



*languages*

# Social and Psychological Factors in Bilingual Speech Production

Edited by

Robert Mayr and Jonathan Morris

Printed Edition of the Special Issue Published in *Languages*

# **Social and Psychological Factors in Bilingual Speech Production**



# Social and Psychological Factors in Bilingual Speech Production

Editors

**Robert Mayr**

**Jonathan Morris**

MDPI • Basel • Beijing • Wuhan • Barcelona • Belgrade • Manchester • Tokyo • Cluj • Tianjin



*Editors*

Robert Mayr	Jonathan Morris
Cardiff Metropolitan University	Cardiff University
UK	UK

*Editorial Office*

MDPI  
St. Alban-Anlage 66  
4052 Basel, Switzerland

This is a reprint of articles from the Special Issue published online in the open access journal *Languages* (ISSN 2226-471X) (available at: <https://www.mdpi.com/journal/languages/special.issues/Speech-Production>).

For citation purposes, cite each article independently as indicated on the article page online and as indicated below:

LastName, A.A.; LastName, B.B.; LastName, C.C. Article Title. <i>Journal Name</i> <b>Year</b> , <i>Volume Number</i> , Page Range.
--

**ISBN 978-3-0365-2277-7 (Hbk)**

**ISBN 978-3-0365-2278-4 (PDF)**

© 2021 by the authors. Articles in this book are Open Access and distributed under the Creative Commons Attribution (CC BY) license, which allows users to download, copy and build upon published articles, as long as the author and publisher are properly credited, which ensures maximum dissemination and a wider impact of our publications.

The book as a whole is distributed by MDPI under the terms and conditions of the Creative Commons license CC BY-NC-ND.

# Contents

<b>About the Editors</b> . . . . .	<b>vii</b>
<b>Robert Mayr and Jonathan Morris</b> Social and Psychological Factors in Bilingual Speech Production: Introduction to the Special Issue Reprinted from: <i>Languages</i> 2021, 6, 155, doi:10.3390/languages6040155 . . . . .	<b>1</b>
<b>Simona Montanari, Robert Mayr and Kaveri Subrahmanyam</b> Maternal Cultural Orientation and Speech Sound Production in Spanish/English Dual Language Preschoolers Reprinted from: <i>Languages</i> 2021, 6, 78, doi:10.3390/languages6020078 . . . . .	<b>7</b>
<b>Tanja Kupisch, Nadine Kolb, Yulia Rodina and Olga Urek</b> Foreign Accent in Pre- and Primary School Heritage Bilinguals Reprinted from: <i>Languages</i> 2021, 6, 96, doi:10.3390/languages6020096 . . . . .	<b>23</b>
<b>Anissa Baird, Angela Cristiano and Naomi Nagy</b> Apocope in Heritage Italian Reprinted from: <i>Languages</i> 2021, 6, 120, doi:10.3390/languages6030120 . . . . .	<b>43</b>
<b>Gisela Tomé Lourido and Bronwen G. Evans</b> Sociolinguistic Awareness in Galician Bilinguals: Evidence from an Accent Identification Task Reprinted from: <i>Languages</i> 2021, 6, 53, doi:10.3390/languages6010053 . . . . .	<b>63</b>
<b>Meinir Williams and Sarah Cooper</b> Adult New Speakers of Welsh: Accent, Pronunciation and Language Experience in South Wales Reprinted from: <i>Languages</i> 2021, 6, 86, doi:10.3390/languages6020086 . . . . .	<b>89</b>
<b>Jonathan Morris</b> Social Influences on Phonological Transfer: /r/ Variation in the Repertoire of Welsh-English Bilinguals Reprinted from: <i>Languages</i> 2021, 6, 97, doi:10.3390/languages6020097 . . . . .	<b>109</b>
<b>Azza Al-Kendi and Ghada Khattab</b> Psycho-Social Constraints on Naturalistic Adult Second Language Acquisition Reprinted from: <i>Languages</i> 2021, 6, 129, doi:10.3390/languages6030129 . . . . .	<b>127</b>
<b>Ulrich Reubold, Sanne Ditewig, Robert Mayr and Ineke Mennen</b> The Effect of Dual Language Activation on L2-Induced Changes in L1 Speech within a Code-Switched Paradigm Reprinted from: <i>Languages</i> 2021, 6, 114, doi:10.3390/languages6030114 . . . . .	<b>149</b>
<b>Denise M. Osborne and Miquel Simonet</b> Foreign-Language Phonetic Development Leads to First-Language Phonetic Drift: Plosive Consonants in Native Portuguese Speakers Learning English as a Foreign Language in Brazil Reprinted from: <i>Languages</i> 2021, 6, 112, doi:10.3390/languages6030112 . . . . .	<b>177</b>
<b>Lisa Kornder and Ineke Mennen</b> Longitudinal Developments in Bilingual Second Language Acquisition and First Language Attrition of Speech: The Case of Arnold Schwarzenegger Reprinted from: <i>Languages</i> 2021, 6, 61, doi:10.3390/languages6020061 . . . . .	<b>203</b>



## About the Editors

**Robert Mayr** (Dr) is a Reader in Linguistics at the Centre for Speech and Language Therapy and Hearing Science, Cardiff Metropolitan University, where he leads the Speech, Hearing and Communication Research Group. His research aims to gain a better understanding of the cognitive, social and interactional factors that affect speech sound development in bilinguals and multilinguals across their lifespan. He has been the principal investigator or a collaborator on various research projects, including a British Academy-funded study on socio-phonetic variation in Welsh–English bilinguals, an FWF-funded study on L1 attrition of speech in native English speakers in Austria and an NIH-funded study on speech perception and production in Spanish–English bilingual children. His work has appeared in international journals, such as *Bilingualism: Language and Cognition*, *International Journal of Bilingualism*, *International Journal of Bilingual Education and Bilingualism*, *Journal of Phonetics*, *Journal of the International Phonetic Association*, *Journal of Child Language and Frontiers in Psychology: Language Sciences*.

**Jonathan Morris** (Dr) is a Senior Lecturer at the School of Welsh, Cardiff University (UK) and leads the Cardiff Multilingualism Research Network. Jonathan’s research focuses on sociolinguistic aspects of bilingualism and aims to shed light on the linguistic consequences of Welsh language revitalisation. Specifically, his work concentrates on socio-phonetic and phonological variation in the bilingual repertoire of Welsh speakers, and the interaction between Welsh and English. He has also published on the sociology of bilingualism in the Welsh context and examined both language use among young people and the intergenerational transmission of Welsh. More recently, his work has investigated the acquisition of Welsh in educational contexts with particular focus on adult learners’ speech and children’s reading skills. He was a co-editor of the volume *Sociolinguistics in Wales* (Palgrave Macmillan, 2016) and his work has appeared in international journals such as *International Journal of Bilingualism*, *Journal of Sociolinguistics* and *Journal of the International Phonetic Association*.





Editorial

# Social and Psychological Factors in Bilingual Speech Production: Introduction to the Special Issue

Robert Mayr <sup>1,\*</sup> and Jonathan Morris <sup>2</sup>

<sup>1</sup> Centre for Speech and Language Therapy and Hearing Science, Cardiff Metropolitan University, Llandaff Campus, Western Avenue, Cardiff CF5 2YB, UK

<sup>2</sup> School of Welsh, Cardiff University, Colum Drive, Cardiff CF10 3EU, UK; morrisj17@cardiff.ac.uk

\* Correspondence: rmayr@cardiffmet.ac.uk

The pronunciation patterns of bilinguals and their development have been investigated in a number of ways. Generally, a distinction is made between internal *linguistic* factors, on the one hand, and external *social and psychological* factors, on the other. Linguistic factors that have been identified as critically affecting L1 and L2 accentual targets include cross-linguistic interactions and language transfer; language universals and markedness; and the similarity of L1 and L2 sounds. The study of linguistic factors contributes significantly to our understanding of the mechanisms that underpin bilingual speech processing and development. This is reflected in the central role that such factors play within current theoretical models in bilingual speech research, such as the revised Speech Learning Model, SLM-r (Flege and Bohn 2021); the Perceptual Assimilation Model of Second Language Learning, PAM-L2 (Best and Tyler 2007); and the Second Language Linguistic Perception Model, L2LP (Escudero 2005).

Importantly, however, these internal factors cannot fully explain the bilingual speech patterns we observe: they are critically mediated by a myriad of extra-linguistic cognitive and psychological factors, such as L1 and L2 use, age of onset of learning, length of residence in an L2-speaking environment or motivation, to name but a few (see, e.g., Piske et al. (2001) for an overview). Moreover, recent variationist and experimental research has documented the significance of social factors, such as peer group identity or cultural and ethnic orientation, on the speech patterns of bilinguals, in particular in the context of long-term contact and minority language bilingualism (Mayr et al. 2017; Nance 2020; Sharma and Sankaran 2011).

The present Special Issue sought to deepen our understanding of such extra-linguistic variables by bringing together ten state-of-the-art articles that investigate the role of one or several social and/or psychological factors on the speech patterns of bilingual speakers. The articles feature a wide range of bilingual populations and contexts—from child and adult bilinguals in heritage language settings to adult new speakers in bilingual societies, and L2 learners and L1 attriters in migration contexts—and encompass a variety of examined languages, including Arabic, English, Galician, German, Italian, Portuguese, Russian, Spanish and Welsh. They are diverse in methodological terms, from auditory and acoustic analyses of bilinguals' speech patterns to accent identification tasks and online questionnaires, and aim to inform current theoretical debates. Contributors are leading experts from universities in Austria, Canada, England, Germany, Italy, Norway, Oman, the United States and Wales.

The ten articles in the Special Issue have been organized thematically. As a result, the order presented here differs somewhat from that of the articles published on the journal website. Specifically, it commences with two articles on child bilinguals in heritage language settings (Montanari, Mayr and Subrahmanyam; Kupisch, Kolb, Rodina and Urek), followed by an article on the speech patterns of adult heritage speakers across generations (Baird, Cristiano and Nagy). The subsequent three papers, in turn, are situated



**Citation:** Mayr, Robert, and Jonathan Morris. 2021. Social and Psychological Factors in Bilingual Speech Production: Introduction to the Special Issue. *Languages* 6: 155. <https://doi.org/10.3390/languages6040155>

Received: 14 September 2021  
Accepted: 24 September 2021  
Published: 28 September 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

within the context of minority languages in bilingual societies (Tomé Lourido and Evans; Williams and Cooper; Morris), with specific focus on the accents of new speakers. Finally, the Special Issue concludes with four articles that focus on adult L2 speech acquisition and/ or the effects of the L2 experience on L1 speech patterns (Al-Kendi and Khattab; Reubold, Diteweg, Mayr and Mennen; Osborne and Simonet; Kornder and Mennen). In what follows, details of each of these are presented in turn.

The study by Simona Montanari (California State University, Los Angeles, Robert Mayr (Cardiff Metropolitan University) and Kaveri Subrahmanyam (California State University, Los Angeles) entitled “Maternal cultural orientation and speech sound production in Spanish/English dual language preschoolers” examines the effect of maternal orientation to Anglo-American (acculturation) and Mexican (enculturation) culture on the speech patterns of Spanish-English bilingual children from low socio-economic backgrounds growing up in Spanish-speaking homes in Southern California. To this end, single-word samples were elicited from the children in Spanish and English, using picture prompts, that together encompass the consonant and vowel categories of the two languages. Mothers’ acculturation and enculturation levels, in turn, were assessed on the basis of a rating scale. The results revealed a significant correlation between maternal acculturation and children’s segmental accuracy in English, that is, the greater the mothers’ orientation towards Anglo-American culture, the more accurately their children produced English consonants and vowels. No such relation was found for children’s accuracy on Spanish segments. The authors interpret these findings as arising from more American-oriented mothers using English more with their children, which may have reinforced their phonological skills in the language. The lack of a correlation for Spanish, on the other hand, is likely due to consistently high levels of Mexican orientation amongst the mothers in the sample or language input differences due to different social practices.

The second article featuring child bilinguals, “Foreign accent in pre- and primary school heritage bilinguals” by Tanja Kupisch (Universität Konstanz and UiT, The Arctic University of Norway), Nadine Kolb (UiT, The Arctic University of Norway), Yulia Rodina (UiT, The Arctic University of Norway) and Olga Urek (UiT, The Arctic University of Norway), investigates the effects of age in the perceived accent of German-Russian bilingual children in Germany alongside native controls in both languages. Using a global accent rating task, samples of speech from the children’s narratives in both languages were rated by native Russian and German listeners. The results revealed that the bilinguals were perceived as foreign-accented in both languages more often than the monolingual controls. Moreover, the older bilinguals’ accent was perceived as more authentic in German than that of the younger bilinguals, while the reverse was true for Russian. These effects were enhanced in families with two Russian parents compared with those with only one Russian-speaking parent. The authors conclude that the primary school years are critical for the maintenance of the heritage language as it comes under pressure from the majority language.

The third article, “Apocope in heritage Italian” by Anissa Baird (University of Toronto), Angela Cristiano (Università di Bologna) and Naomi Nagy (University of Toronto), examines whether three generations of Calabrian Italian heritage speakers in Toronto delete (apocope) or reduce word-final vowels in their heritage language in line with the patterns seen in southern Italian varieties. Using sociolinguistic interviews, word-final vowel productions from the heritage speakers and a homeland comparator group were analyzed auditorily and subsequently related to a large number of predictor variables, including linguistic and social ones, using mixed effects regression models. The results showed that the heritage speakers made use of both apocope and reduction. While their distributions were similar to those of homeland speakers, some inter-generational differences occurred. Both apocope and reduction were found to be primarily influenced by linguistic factors, while surprisingly few social factors mediated the use of these socially marked forms. Together, the findings of the study reveal a complex interplay of social and linguistic factors.

Shifting to studies on the speech patterns of minority language speakers in bilingual societies, the fourth article, “Sociolinguistic awareness in Galician bilinguals: Evidence from an accent identification task” by Gisela Tomé Lourido (Leeds University) and Bronwen G. Evans (University College London), assesses the extent to which a new variety of Galician spoken by Galician new speakers (so-called *neofalantes*) is distinctive enough to be identifiable by listeners in an accent identification task. The results revealed, however, that while Galician-dominant bilinguals and to a lesser extent Spanish-dominant bilinguals could be clearly identified, listeners from Galicia were unable to classify *neofalantes* correctly. Instead, the latter were categorized as both Galician- and Spanish-dominant, suggesting the variety constitutes a hybrid accent. The authors argue that listeners may have a gradient representation of variation in which the accents of Galician-dominant and Spanish-dominant speakers function as anchors, with *neofalantes’* accent situated in between. The results also revealed differences in identification accuracy across listener groups, with *neofalantes* exhibiting particular sensitivity to the accents of Galician-dominant speakers. This may be due to *neofalantes’* heightened awareness of the sociolinguistic landscape in Galicia and their motivation to learn Galician.

The fifth article, “Adult new speakers of Welsh: Accent, pronunciation and language experience in South Wales” by Meinir Williams and Sarah Cooper (Bangor University), investigates the experiences of adult new speakers of Welsh with little to no exposure to the language in childhood, with learning the pronunciation patterns of Welsh. Using an online questionnaire, the study examined how respondents perceived their Welsh accent, which speech sounds they found most challenging and how traditional speakers responded to them. The results revealed that the respondents generally did not perceive their accent as native-like, and that perceptions of their accent depended on their level of competence in Welsh, with beginning and intermediate learners considering their accent as less native, being less proud of their accent and wanting to change their accent more than advanced or fluent learners. They indicated that vowel length and consonants that are not shared with English were most challenging. Finally, the study revealed a range of reactions by traditional speakers, including switching to English, decreasing speaking rate or correcting the respondents’ pronunciations. These findings have important implications for new speakers’ participation in the community and for language learning.

The sixth article, “Social influences on phonological transfer: /r/ variation in the repertoire of Welsh-English bilinguals” by Jonathan Morris (Cardiff University), is also situated within the context of bilingualism in Wales. Specifically, it examines the effects of speaker gender, home language and speech context on the Welsh and English /r/ productions of bilingual speakers from two communities in North Wales, Caernarfon and Mold, that differ in their use of Welsh as a community language. Data collected from young Welsh-English bilinguals via sociolinguistic interviews and word lists revealed cross-linguistic and areal differences. Thus, Welsh /r/ variants, such as the trill and tap, occurred substantially more in the Welsh than the English productions of Caernarfon speakers, and were entirely absent in the English data from Mold. Moreover, in contrast to previous dialectological studies, alveolar approximants were found to be widely used in Welsh in both communities, suggesting cross-linguistic transfer from English. This pattern was more pronounced in new speakers than traditional ones, showing an effect of home language. Effects were also found for task and speaker gender. Together, the study reveals a complex pattern in which the realization of /r/ in the two languages is mediated by a multifaceted interplay of different social variables, reflecting sociolinguistic differences across the two communities and differences in peer group structure.

Shifting to studies examining L1 and L2 speech and its interrelation in bilingual adults, the next article, “Psycho-social constraints on naturalistic adult second language acquisition” by Azza Al-Kendi (Sultan Qaboos University) and Ghada Khattab (Newcastle University), investigates the speech perception and production patterns in L2 Arabic of Foreign Domestic Helpers (FDH) in Omani homes with a range of native languages. As such, it is one of few studies on bilinguals from low educational backgrounds who learn an

L2 predominantly, if not solely, to interact with their employers. The results from an AX discrimination task revealed low sensitivity to Arabic consonantal contrasts that do not exist in the participants' L1 while the results of a production task showed low accuracy on Arabic consonants. Moreover, in an accent rating task, L1 Omani Arabic speakers identified a marked foreign accent in the FDH's speech, in particular in those that were not literate in Arabic. Interestingly, unlike previous studies on L2 speech, length of residence was not found to be a significant predictor of performance in any of the tasks. The authors argue that this is due to the specific social context of FDH and the unequal power relations between them and their employers.

The eighth article, "The effect of dual language activation on L2-induced changes in L1 speech within a code-switched paradigm" by Ulrich Reubold (Karl-Franzens-Universität Graz), Sanne Ditewig (Karl-Franzens-Universität Graz), Robert Mayr (Cardiff Metropolitan University) and Ineke Mennen (Karl-Franzens-Universität Graz), examined the L1 English speech of adult migrants to Austria in a code-switched and monolingual condition alongside that of monolingual English speakers in England. The code-switched materials involved German words with segments known to trigger cross-linguistic interactions being inserted into an otherwise English frame; the monolingual materials, in turn, contained the equivalent segments in wholly English sentences. The sentences were produced in an online reading task and the critical items subsequently analyzed acoustically. The results revealed no differences between the monolingual and bilingual speakers in the monolingual condition. However, on some sounds significant L2-induced shifts in L1 speech production were observed in the code-switched condition. These occurred both before and after the switch. On the other hand, with one exception, i.e., amount of L2 use for [w], none of the predictor variables examined were associated with L2-induced shifts across conditions. These results have significant implications for the role of dual activation in the pronunciation patterns of late bilinguals.

In "Foreign-language phonetic development leads to first-language phonetic drift: Plosive consonants in native Portuguese speakers learning English as a foreign language in Brazil" by Denise M. Osborne (State University of New York at Albany) and Miquel Simonet (University of Arizona), word-initial plosive productions in Portuguese were compared across L1 Brazilian Portuguese speakers learning English as an L2 in Brazil and monolingual native speakers of Brazilian Portuguese. The L2 learners' English plosive productions were also investigated and compared with their L1 categories. The results of an acoustic analysis of the participants' voice onset time (VOT) patterns revealed first that the learners produced Portuguese voiced plosives with a longer voicing lead than the monolinguals. They also showed that the learners' English plosives were similar to their Portuguese ones, showing cross-linguistic interactions but also some phonetic development. Together, the study adds to the growing literature on bidirectional interactions in bilinguals' speech and shows that these can even occur outside L2 immersion contexts.

The final article, "Longitudinal developments in bilingual second language acquisition and first language attrition of speech: The case of Arnold Schwarzenegger" by Lisa Kornder and Ineke Mennen (Karl-Franzens-Universität Graz), investigates the speech productions of the late consecutive bilingual Arnold Schwarzenegger in his L1 (Austrian German) and L2 (English) over a period of 40 years. As such, it is the first longitudinal case study that examines bilingual speech development over several decades in both languages. Specifically, on the basis of speech samples from broadcast interviews, it measured the VOT durations in his plosive productions and the formant frequencies in his vowel productions across different time periods. The results revealed a cross-linguistic merger of /p t k/ in his late productions. His vowel realizations in the two languages also became increasingly more similar. These findings shed new light on bilinguals' reorganization of L1 and L2 sound systems over time and their dynamic interrelations.

Taken together, the ten articles in this Special Issue significantly advance our understanding of the role that different social and psychological factors play in the speech productions of bilinguals across a wide range of settings, and have important theoretical

implications. They also highlight important future directions in this vibrant field of enquiry, which, we hope, will inspire others in their pursuit of new knowledge.

**Acknowledgments:** First and foremost, we would like to thank the authors of the articles in the Special Issue for their fine contributions. It has been a great pleasure to work with each of them. We are also most grateful to the international set of experts who agreed to act as reviewers for the articles in this volume and helped to improve the quality of the contributions. In alphabetical order, these are John Archibald (University of Victoria), Elise Bell (University of California, Los Angeles), Gwen Brekelmans (University College London), Chiara Celata (Università degli Studi di Urbino Carlo Bo), Ulrike Gut (Westfälische Wilhelms-Universität Münster), Mike Hammond (University of Arizona), Michael Hornsby (Adam Mickiewicz University, Poznań), Lisa Kornder (Karl-Franzens-Universität Graz), Rosangela Lai (Università di Pisa), Chia-Cheng Lee (Portland State University), Philip P. Limerick (New Mexico State University), Diarmait Mac Giolla Christ (Cardiff University), Daniela Mereu (Free University of Bozen-Bolzano), Alene Moyer (University of Maryland), Naomi Nagy (University of Toronto), Ramsés Ortín (University of Texas, Rio Grande Valley), Michael Putnam (Pennsylvania State University), Yasaman Rafat (Western University), Anabela Rato (University of Toronto), Xosé-Luis Regueira Fernández (Universidade de Santiago de Compostella), Monika Schmid (University of Essex), Geoffrey Schwartz (Adam Mickiewicz University, Poznań), and Antje Stoehr (Basque Center on Cognition, Brain & Language). Last but not least, we would like to thank the team at *Languages*, in particular Yuki Yu and Milana Arambašić, for their editorial support.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Best, Catherine T., and Michael Tyler. 2007. Nonnative and second-language speech perception: Commonalities and complementarities. In *Language Experience in Second Language Speech Learning, in Honor of James Emil Flege*. Edited by Ocke-Schwen Bohn and Murray J. Munro. Amsterdam: John Benjamins, pp. 13–24.
- Escudero, Paola. 2005. *Linguistic Perception and Second Language Acquisition*. Utrecht: Utrecht University.
- Flege, James E., and Ocke-Schwen Bohn. 2021. The revised Speech Learning Model. In *Second Language Speech Learning, Theoretical and Empirical Progress*. Edited by Rtree Wayland. Cambridge: Cambridge University Press, pp. 3–83.
- Mayr, Robert, Jonathan Morris, Ineke Mennen, and Daniel Williams. 2017. Disentangling the effects of long-term language contact and individual bilingualism: The case of monophthongs in Welsh and English. *International Journal of Bilingualism* 21: 245–67. [[CrossRef](#)]
- Nance, Claire. 2020. Bilingual language exposure and the peer group: Acquiring phonetics and phonology in Gaelic medium education. *International Journal of Bilingualism* 24: 360–75. [[CrossRef](#)]
- Piske, Thorsten, Ian R. A. MacKay, and James E. Flege. 2001. Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics* 29: 191–215. [[CrossRef](#)]
- Sharma, Devyani, and Lavanya Sankaran. 2011. Cognitive and social forces in dialect shift: Gradual change in London Asian speech. *Language Variation and Change* 23: 399–428. [[CrossRef](#)]



Article

# Maternal Cultural Orientation and Speech Sound Production in Spanish/English Dual Language Preschoolers

Simona Montanari <sup>1,\*</sup>, Robert Mayr <sup>2</sup> and Kaveri Subrahmanyam <sup>3</sup>

<sup>1</sup> Child & Family Studies, California State University, Los Angeles, Los Angeles, CA 90032, USA

<sup>2</sup> Centre for Speech and Language Therapy, Cardiff Metropolitan University, Cardiff CF5 2YB, UK; rmayr@cardiffmet.ac.uk

<sup>3</sup> College of Natural and Social Sciences, California State University, Los Angeles, Los Angeles, CA 90032, USA; ksubrah@calstatela.edu

\* Correspondence: smontan2@calstatela.edu

**Abstract:** Empirical work has shown that maternal education is related to children's language outcomes, especially in the societal language, among Spanish-English bilingual children growing up in the U.S. However, no study thus far has assessed the links between maternal cultural orientation and children's speech sound production. This paper explores whether mothers' orientation to American (*acculturation*) and Mexican culture (*enculturation*) and overall linear acculturation are related to children's accuracy of production of consonants, of different sound classes, and of phonemes shared and unshared between languages in both English and Spanish at age 4;6 (4 years and 6 months). The results reveal a link between maternal acculturation and children's segmental accuracy in English, but no relation was found between mothers' enculturation and children's speech sound production in Spanish. We interpreted the results in English as suggesting that more American-oriented mothers may have been using more English with their children, boosting their English production abilities and promoting English speech sound development. At the same time, we speculate that the results in Spanish were possibly due to the high and homogeneous levels of Mexican orientation among mothers, to language input differences attributable to distinct cultural practices, or to the status of Spanish as a minority language.

**Keywords:** maternal acculturation; maternal enculturation; speech sound production; Spanish-English bilingual preschoolers



**Citation:** Montanari, Simona, Robert Mayr, and Kaveri Subrahmanyam. 2021. Maternal Cultural Orientation and Speech Sound Production in Spanish/English Dual Language Preschoolers. *Languages* 6: 78. <https://doi.org/10.3390/languages6020078>

Academic Editors: Juana M. Licerias and Raquel Fernández Fuentes

Received: 18 February 2021

Accepted: 20 April 2021

Published: 22 April 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Considerable empirical work has demonstrated that maternal characteristics such as maternal education are strong predictors of language outcomes in young dual language learners. Specifically, a large body of work has shown that maternal education is related to children's phonological, lexical, and grammatical measures throughout development, especially in the societal language, among Spanish-English bilingual children growing up in the U.S. (Bohman et al. 2010; De Anda et al. 2016; Friend et al. 2017; Hammer et al. 2012; Hoff et al. 2018; Montanari et al. 2020; Place and Hoff 2016). For instance, De Anda et al. (2016) found that maternal educational attainment was significantly related to the English comprehension and production vocabularies of infants growing up in English-dominant bilingual homes, and these results were replicated by Friend et al. (2017) when the same children were tested at 22 months of age. In another study, Place and Hoff (2016) also documented positive correlations between maternal education and English comprehension, productive vocabulary, and, additionally, grammatical skills in Spanish-English bilingual children at 30 months of age. Furthermore, in a more recent investigation comparing the speech and language skills of Spanish-English bilingual children whose mothers had completed secondary school vs. children whose mothers had only attended primary school, Montanari et al. (2020) found that the children of more educated mothers performed



significantly better than children of less educated mothers not only on English lexical and grammatical measures but also on speech sound production at age 3;6 (3 years and 6 months), suggesting that maternal level of schooling affects not only children's language but also speech abilities.

Interestingly, the majority of previous studies have found that maternal level of schooling—measured in terms of the number of years of education attained irrespective of the language—has a different impact on children's English versus Spanish development. Specifically, while research has shown clear links between maternal education and children's English outcomes, mothers' school attainment has not been found to be related to Spanish outcomes, at least in [De Anda et al. \(2016\)](#), [Friend et al. \(2017\)](#), and [Place and Hoff \(2016\)](#). By controlling the language in which mothers received their education, [Hoff et al. \(2018\)](#) found that maternal school attainment in English was significantly related to children's English but not to Spanish vocabulary between 30 and 60 months, and maternal level of education in Spanish predicted children's Spanish but not English lexical skills between the same ages. In contrast, all mothers in [Montanari et al. \(2020\)](#) were educated in Spanish in Mexico, and therefore, the language in which the mothers had received their education did not explain the differences in English speech and language outcomes between children of more and less educated mothers.

[Montanari et al. \(2020\)](#) speculated that characteristics other than the language of mothers' schooling may mediate the relationship between general maternal education and children's English abilities. As noted by the authors (p. 12), "acculturation, language experiences, and literacy practices may differ in homes where mothers have completed more schooling, irrespective of the language in which such schooling has occurred". For example, more educated mothers may be more acculturated to Anglo-American culture, and, therefore, they may engage their children more in language and literacy learning opportunities in English than less educated mothers. Although all mothers in [Montanari et al. \(2020\)](#), p. 5) exhibited a "Very Mexican Orientation" according to the Acculturation Rating Scale for Mexican Americans-II (ARSMA-II, [Cuéllar et al. 1995](#)) and acculturation scores for more and less educated mothers were not statistically different, it is possible that the more educated mothers promoted their children's English skills to a greater extent than the less educated ones.

Unfortunately, very few studies have examined Spanish-English bilingual children's language outcomes as related to mothers' levels of *acculturation* (i.e., orientation to American culture) and *enculturation* (i.e., orientation to Latino culture) ([Gonzales et al. 2004](#)). [Boyce et al. \(2013\)](#) examined the relationship between different maternal factors, including maternal acculturation, and Spanish-English dual language learners' total vocabulary size (i.e., including both English and Spanish words) at 24 and 36 months. The authors found that both maternal acculturation and home language/literacy environment were each related to children's total vocabulary at 24 months. Moreover, a positive significant correlation was found between maternal acculturation measured at 24 months and children's total word knowledge at 36 months. [Cote and Bornstein \(2014\)](#) similarly examined the relationship of maternal acculturation to the productive vocabularies of bilingual American 20-month-old children born to Korean, Japanese, and South American immigrant mothers from Argentina, Colombia, and Peru. The authors found that mothers' acculturation level when their child was five months of age was positively related to his/her vocabulary size in English at 20 months. Path analyses further demonstrated that exposure to English partially mediated the relation between mothers' acculturation level and children's English vocabulary size in that mothers who were more acculturated to American culture exposed their children to more English, which, in turn, resulted in larger English vocabularies for their children. On the other hand, more enculturated immigrant mothers (i.e., more oriented to their heritage culture) exposed their children more to their heritage language, which, in turn, resulted in larger vocabularies in that language for their children. Heritage language exposure fully mediated the effect of acculturation on heritage language vocabulary size.

While these studies have documented a link between maternal cultural orientation and children's language outcomes, no research thus far has examined whether mothers' levels of acculturation and enculturation also affect children's ability to produce speech sounds in both English and the heritage language. As pointed out above, maternal characteristics such as socio-economic status (SES) and maternal education have been found to affect not only language but also speech outcomes in both monolingual and bilingual children. For instance, [Campbell et al. \(2003\)](#) found that English monolingual children whose mothers had attained limited education were 2.5 times more likely to display speech delay compared to children of more educated mothers. Similarly, in a large sample of almost 1500 monolingual English-speaking children from Australia ([Eadie et al. 2015](#)), the prevalence of speech sound disorders was partially predicted by SES and maternal education, and children whose mothers had a college or postgraduate degree were half as likely to have a speech sound disorder compared to children whose mothers had completed less than 12 years of schooling. [Montanari et al. \(2020\)](#) further found that children of more educated mothers had higher English consonant accuracy than children of less educated mothers at age 3;6. However, the English speech sound production abilities of both groups of children were no longer different one year later after one year of exposure to English in the preschool setting, and maternal education was not related to children's Spanish phonological accuracy at either age.

It is the goal of this study to examine the link between children's English and Spanish speech sound production skills during the preschool years and (a) maternal orientation to American culture (acculturation), (b) maternal orientation to Mexican culture (enculturation), and (c) maternal linear acculturation, a single measure of cultural orientation that takes both American and Mexican orientation into account ([Cuéllar et al. 1995](#)). We focus on preschoolers at age 4;6 to explore (1) whether speech sound production in English is related to mothers' American orientation and overall acculturation at an age when prior work suggests that maternal education no longer plays a role and (2) whether maternal orientation to Mexican culture and overall acculturation, unlike educational attainment, are linked to children's ability to produce speech sounds in Spanish at an age when children have begun their own acculturation process through preschool.

In order to assess whether maternal cultural orientation contributes differently to the production of distinct English and Spanish sounds, we base our analysis on the speech production measures that have been widely used to assess bilingual phonological skills in previous studies and to identify typical and atypical development (e.g., [Bunta et al. 2009](#); [Cooperson et al. 2013](#); [Fabiano-Smith and Goldstein 2010a, 2010b](#); [Gildersleeve-Neumann et al. 2008, 2009](#); [Goldstein and Bunta 2012](#); [Goldstein et al. 2010](#); [Keffala et al. 2020](#); [Kehoe and Girardier 2020](#); [Montanari et al. 2018](#); [Ruiz-Felter et al. 2016](#)): (a) overall accuracy of consonant production ("Percentage of Consonants Correct-Revised", PCC-R, [Shriberg et al. 1997](#)), (b) accuracy of production of different sound classes, and (c) accuracy of production of phonemes "shared" and "unshared" between English and Spanish. Shared phonemes include segments that are phonetically similar (despite differences in fine phonetic detail) and present in both languages such as stops, most nasals (/m/ and /n/), and some fricatives (/f/ and /s/). Unshared phonemes include language-specific segments that do not have phonetically similar equivalents in the other language, such as the Spanish trill or English /v/, /θ/, or /z/. Previous research has shown that Spanish L1 children successfully rely on many Spanish consonants in their acquisition of phonetically similar segments in L2 English, producing shared sounds with significantly higher accuracy than unshared sounds ([Montanari et al. 2018](#); [Scarpino 2011](#)). Indeed, according to the *Unified Competition Model* ([MacWhinney 2005](#)), phonological properties that are common across languages—such as shared segments—produce frequent and reliable speech cues, leading to rapid development—or *positive transfer* ([Goldstein and Bunta 2012](#))—of these properties in both languages. On the other hand, properties that are unique across languages—such as unshared sounds—produce less frequent and less strong cues, leading to *negative transfer* and protracted development with these segments ([Goldstein and Bunta 2012](#)). Importantly,

the observed developmental benefits of phonologically shared categories do not preclude cross-linguistic interactions from taking place at the phonetic level. Indeed, there is an abundance of evidence that similar but non-identical L2 sounds are equated with their closest L1 counterparts, resulting in merged representations (e.g., Aoyama et al. 2004; Simonet 2010), in line with Flege's *Speech Learning Model* (SLM, Flege 1995; Flege and Bohn 2021). However, since the present study is situated within an auditory-based approach to children's phonological development and focuses on shared versus unshared *phonemes*, children's realizations in terms of fine phonetic detail are not considered here.

Therefore, based on MacWhinney's (2005) model, we predicted that maternal acculturation—if mediated by language use—may be more strongly correlated with children's accuracy with English-only phonemes than with items shared between Spanish and English since the acquisition of the former depends exclusively on English input. Likewise, mothers' Mexican orientation—and possibly increased Spanish use—should be more strongly related to children's accuracy with Spanish-only phonemes than with common segments based on the premise that the former are uniquely acquired through exposure to Spanish.

## 2. Materials and Methods

### 2.1. Participants

The data for this study come from a larger longitudinal investigation of dual language development in low-SES Spanish L1/English L2 bilingual preschoolers from Southern California. We focused on the same 20 children from Montanari et al. (2020) as well as 17 additional children (for a total of 37 children, 21 girls, and 16 boys) at age 4;6 (range: 4;0–5;2) for whom both speech sound production and maternal acculturation data were available. Information collected through a detailed parent questionnaire administered in Spanish indicated that all children were developing typically and had no hearing, speech, language, cognitive, or neurological deficits. All children were born in the U.S. but came from Spanish-speaking, Mexican-origin families who had been living in Los Angeles County for approximately 14 years (range = 6–24). Spanish was the primary home language, the language spoken by each mother to her child, and the children's native language. However, all participants were born in the U.S.; thus, they had also been exposed to English as a second language through siblings, the media, the Head Start program they had been attending for a year, and the larger community. All children were from low-SES backgrounds as evidenced by their participation in the Head Start program, which is aimed at promoting school readiness among children from low-income families. Indeed, all children's mothers had limited education with one-third of them having attained only six years of schooling (*educación primaria*, "primary education") and the remaining having completed high school (*la escuela preparatoria*) in Spanish in Mexico. All mothers rated their Spanish as "native" and "their strongest language", whereas their proficiency in English was predominantly rated as "limited". The vast majority of the mothers were not employed at the time of the study and thus took care of their child.

### 2.2. Data Collection: Materials

The children's speech sound production abilities were assessed by examining phonological accuracy in single-word samples elicited with the phonology subtest of the Bilingual English Spanish Assessment (BESA; Peña et al. 2014), a standardized, norm-referenced test that has been widely used to assesses phonological skills in Spanish-English dual language learners and identify typical and atypical development (Bunta et al. 2009; Cooperson et al. 2013; Fabiano-Smith and Goldstein 2010a, 2010b; Goldstein and Bunta 2012; Goldstein et al. 2010; Montanari et al. 2018; Ruiz-Felter et al. 2016). The test contains 31 target items for English and 28 target items for Spanish that are elicited through high-quality pictures. These items vary in length and lexical stress pattern and target all English and Spanish consonants (except for /s/ in English) in different word positions, for a total of 47 English consonant tokens (15 plosives, 9 nasals, 12 fricatives, 3 affricates, and 8 approximants) and

54 Spanish consonant tokens (14 plosives, 7 nasals, 17 fricatives, 1 affricate, 3 flaps, 6 trills, and 6 approximants). Note that the Spanish consonantal inventory is more limited than the English one, as it only displays the voiced and voiceless labial, dental and velar stops, three nasals<sup>1</sup> (/m, n, ɲ/), the fricatives /f, s, β<sup>2</sup>, ð, γ, x/, the affricate /tʃ/, the alveolar flap and trill, and the approximants /l, w, j/ (Goldstein 2001).

Mothers' acculturation and enculturation levels were assessed by the Acculturation Rating Scale for Mexican Americans-II (ARSMA-II, Cuéllar et al. 1995), a measure that allows for the independent assessment of an individual's involvement with Anglo-American culture ("acculturation") and Mexican culture ("enculturation"). The ARSMA-II contains two scales that can be used separately. Scale 1 measures Anglo-American orientation (AOS) and Mexican orientation (MOS). Scale 2 measures the concepts of marginality and separation. For the purpose of this study, only Scale 1 was used. This scale includes 30 questions assessing the following cultural domains: (1) language use and preference; (2) ethnic identity and classification; (3) cultural heritage and ethnic behaviors; and (4) ethnic interaction. The AOS contains 13 questions and the MOS 17 questions. Each question is scored on a Likert scale from 1 (*not at all*) to 5 (*extremely often or almost always*). For each participant, a mean AOS score is calculated by adding the scores of the 13 items and dividing the sum by 13. Similarly, a mean MOS score is obtained by adding the scores of the 17 items and dividing it by 17. Higher AOS and MOS scores represent higher orientation to American and Mexican culture, respectively, whereas lower scores represent less cultural orientation. Besides giving an acculturation and enculturation score, the ARSMA-II also generates a linear continuous measure of acculturation that can be obtained by subtracting the mean MOS score from the mean AOS score. This single cultural orientation measure places an individual within one of five acculturative categories along a continuum. Level 1 represents a "Very Mexican Orientation" (mean < -1.33); Level 2 represents "Mexican Oriented to Approximately Balanced Bicultural" individuals (mean ≥ -1.33 and ≤ -0.07); Level 3 represents a "Slightly Anglo Oriented Bicultural" orientation (mean > -0.07 and < 1.19); Level 4 represents a "Strongly Anglo Oriented Bicultural" orientation (mean ≥ 1.19 and < 2.45); and Level 5 represents "Very Assimilated or Anglicized Individuals" (mean > 2.45) (Cuéllar et al. 1995). Reliability and test-retest reliability for ARSMA-II scales are high as indicated by correlations for the AOS and MOS -0.83 and 0.88, respectively (Cuéllar et al. 1995). Correlations between acculturation scores from the original ARSMA and those from the ARSMA-II have further revealed strong construct and concurrent validity as well as high convergent validity for the ARSMA-II, suggesting that it is a valid and reliable measure to assess acculturation among Mexican-Americans (Jimenez et al. 2010).

### 2.3. Data Collection: Procedure

The BESA was administered at the preschool in the fall of children's second year in the Head Start program. Children were asked to name the object depicted in each picture. In the case of no response, children were first given prompts, and then they were allowed to provide a delayed imitation, as in previous studies (e.g., Cooperson et al. 2013; Fabiano-Smith and Goldstein 2010a, 2010b; Gildersleeve-Neumann et al. 2008, 2009; Goldstein and Bunta 2012; Goldstein et al. 2010; Keffala et al. 2020; Montanari et al. 2018; Ruiz-Felter et al. 2016), since only negligible differences have been found between spontaneous and imitated responses in single-word phonology tests (Goldstein et al. 2004). Each session was administered in one language at a time by different research assistants who were native speakers of English and Spanish and only interacted with the child in the language of testing. Half the children were administered the BESA in Spanish first, while the other half were tested in English first.

<sup>1</sup> Although the velar nasal exists in Spanish as a conditional allophone, we followed previous studies (e.g., Montanari et al. 2018; Fabiano-Smith and Goldstein 2010b) and did not code it as a sound shared between English and Spanish because it is phonemic in English but allophonic in Spanish.

<sup>2</sup> Although the spirants /β, ð, γ/ are typically considered allophones of /b, d, g/, we included them in the analysis in line with most recent studies (Montanari et al. 2018; Fabiano-Smith and Goldstein 2010b), since these sounds have been argued to constitute the underlying form and not the phonetic realization of the stops (Barlow 2003).

Samples were recorded using an Edirol R-09HR High-Resolution WAVE/MP3 recorder and a desktop microphone close to the child. Each sample was independently transcribed by the first author and by two separate Spanish/English bilingual research assistants in narrow phonetic transcription, using the conventions of the International Phonetic Alphabet (International Phonetic Association 1999). Inter-rater reliability, calculated for 100% of the target consonants, was 96% for Spanish and 94% for English. Intra-rater reliability was 98% for the first author and 96% for the graduate assistant for Spanish, and 96% for the first author and 95% for the graduate assistant for English. Disagreements on sounds were discussed by listening to the recordings several more times until consensus was reached. The consensus transcriptions were used for the analysis.

The children's mothers were administered the ARSMA-II in Spanish (Cuéllar et al. 1995) at the beginning of the study (when the children were three and a half on average). Trained Spanish-speaking research assistants administered the assessment at the preschool site after having obtained the mothers' consent for their and their child's participation.

#### 2.4. Analyses

Overall consonant accuracy was calculated in terms of Percent of Consonants Correct-Revised (PCC-R) (Shriberg et al. 1997), a widely used and recognized tool in both research and clinical settings to assess phonological skills in dual language learners and differentiate between typical and atypical development (Goldstein and Bunta 2012; Goldstein et al. 2010; Keffala et al. 2020; Montanari et al. 2018; Ruiz-Felter et al. 2016; Scarpino 2011). PCC-R indicates the percentage of consonant sounds that were articulated correctly out of the total number of targeted consonants. However, speech sound distortions and sounds that differ in fine phonetic details, which are common in speech sound development and are not as indicative of speech sound disorder as omissions and substitutions, are not coded as errors as long as they match the adult target in terms of place, manner of articulation, and voicing. The goal of this measure, indeed, is to examine the extent to which children's productions match the adult target on *auditory* grounds—even if fine phonetic details may differ (see Shriberg et al. 1997 and Keffala et al. 2020 for a discussion). PCC-R was also calculated for different sound classes (stops, fricatives, nasals, etc.) and for phonemes shared between Spanish and English, that is, segments that are “phonetically similar” (Flege 1981) and present in both languages (/p, b, t, d, k, g, m, n, f, s, ð, tʃ, l, w, j/), and for unshared phonemes, that is, sounds that exist only in English (/ʃ, v, z, ʒ, θ, h, dʒ, ɹ/) or in Spanish (/ɲ, β, γ, x, r, r<sup>3</sup>/), as in Fabiano-Smith and Goldstein (2010b) and Montanari et al. (2018). As in previous studies, we compared the children's productions to adult targets in the varieties spoken by the children: American English and Mexican Spanish (as outlined in Goldstein 2001).

We first calculated descriptive statistics in order to show the means and standard deviations for children's phonological accuracy scores in English and Spanish and mothers' AOS, MOS, and linear acculturation scores. Then, correlation analyses between children's phonological accuracy measures in English and Spanish and (a) maternal AOS scores, (b) maternal MOS scores, and (c) mothers' linear acculturation scores were run using the Statistical Package for the Social Sciences (SPSS Statistics Version 26). Separate analyses were run for each language, that is, Pearson correlation coefficients were computed for mothers' AOS and linear acculturation scores and children's English PCC-R, their accuracy with English stops, fricatives, nasals, affricates, approximants, and with shared and unshared phonemes. Likewise, separate correlation analyses were run between maternal MOS and linear acculturation scores and children's Spanish PCC-R, their accuracy with Spanish stops, fricatives, nasals, affricates, flaps, trills, approximants, and with shared and unshared phonemes. The alpha level was adjusted for multiple comparisons throughout using the Holm–Bonferroni method (Holm 1979).

<sup>3</sup> Note that although the flap exists in English as an allophone of /t/ and /d/, we coded it as a Spanish-only sound, in line with Fabiano-Smith and Goldstein (2010b) because it is phonemic only in Spanish.

### 3. Results

#### 3.1. Children's Phonological Accuracy in English and Spanish

Table 1 shows children's phonological accuracy in English and Spanish for all consonants, for different sound classes, and for phonemes shared and unshared between English and Spanish. Children produced around 80% of all consonants accurately in both languages, and they displayed higher mean accuracy rates for early-developing sounds (i.e., stops and nasals) than late-developing segments (i.e., fricatives in English, trills and flaps in Spanish) in line with the results of previous studies (Fabiano-Smith and Goldstein 2010b; Montanari et al. 2018; Ruiz-Felter et al. 2016). Accuracy levels were more uniform for early-acquired consonants, whereas, as shown by the large standard deviations, children differed widely in how good they were at producing late-acquired segments. Accuracy rates were significantly higher for phonemes common to both English and Spanish than for unshared segments in both languages, confirming the possible transfer of phonologically similar structures from one language to the other. With the exception of fricatives and affricates, whose accuracy of production was almost 20 percentage points lower in English than in Spanish, speech sound production abilities were comparable in Spanish and English after a year of preschool for this sample of Spanish L1/English L2 bilingual children.

**Table 1.** Children's phonological accuracy in English and Spanish for all consonants (PCC-R), for different sound classes, and for shared and unshared phonemes (in %).

	English		Spanish	
	Mean	SD	Mean	SD
PCC-R	81.3	7.9	80.5	9.2
Stops	91.8	7.8	90.5	10.7
Fricatives	59.8	13.1	76.0	10.3
Nasals	94.0	6.7	98.5	5.6
Affricates	81.5	29.3	100	0.0
Approximants	87.4	15.1	90.9	13.4
Flaps	-	-	78.8	33.0
Trills	-	-	34.2	36.7
Shared Phonemes	86.4	7.5	89.4	8.6
Unshared Phonemes	60.3	12.3	53.2	16.9

#### 3.2. Mothers' Acculturation, Enculturation, and Overall Linear Acculturation

Table 2 reports mothers' mean acculturation (AOS), enculturation (MOS), and overall linear acculturation scores. The acculturation and enculturation scores suggest that mothers were more oriented to Mexican than to Anglo-American culture. The mean linear acculturation score for the sample shows that mothers had, on average, the lowest level of acculturation to Anglo-American culture or a "Very Mexican Orientation" (scores lower than  $-1.33$ , Cuéllar et al. 1995), a categorization that reflects their status as recent immigrants in a region characterized by a large Mexican community. As shown by the standard deviations and ranges, mothers differed substantially in how much they were oriented to American culture, whereas they were more uniform in their Mexican orientation.

**Table 2.** Mothers’ mean Anglo-American orientation (AOS), Mexican orientation (MOS), and linear acculturation scores, standard deviations, and ranges.

	Mean	SD	Range
AOS	2.42	0.88	1.00–4.23
MOS	4.40	0.47	3.18–5.00
Linear Acculturation	−1.98	1.1	−3.69–1.05

3.3. Maternal Acculturation and Children’s English Speech Sound Production

Table 3 shows the results of the correlation analyses between maternal Anglo-American orientation (AOS scores) and children’s phonological accuracy measures in English. The results show moderate positive correlations between maternal American orientation and children’s (1) English PCC-R ( $r = 0.519, p = 0.001, \text{adj. } \alpha = 0.007$ ), (2) accuracy with English stops ( $r = 0.449, p = 0.005, \text{adj. } \alpha = 0.008$ ), (3) accuracy with English approximants ( $r = 0.583, p = 0.000, \text{adj. } \alpha = 0.0056$ ), and (4) accuracy with phonemes unshared between English and Spanish ( $r = 0.521, p = 0.001, \text{adj. } \alpha = 0.006$ ). The magnitude of association was based on Evans’ (1996) account: 00–0.19: “very weak”; 0.20–0.39: “weak”; 0.40–0.59: “moderate”; 0.60–0.79: “strong”; 0.80–1.0: “very strong”. Weak to moderate positive correlations were also found between maternal American orientation and children’s accuracy with fricatives and shared phonemes. However, these correlations were not statistically significant given the Holm–Bonferroni adjusted alpha levels. Overall, the results suggest that mothers with higher Anglo-American orientation had children who were more accurate in the production of overall English consonants, stops, approximants, and sounds specific to English. However, this pattern did not hold across the board, since mothers’ AOS scores were not related to children’s accuracy with the other segments.

**Table 3.** Bivariate correlations between maternal Anglo-American orientation (AOS) scores and children’s English phonological accuracy measures; \*\*\* = significant effect (Holm–Bonferroni adjusted).

	r-Value	p-Value	Adjusted $\alpha$ -Level
PCC-R	0.519	*** 0.001	0.008
Stops	0.449	*** 0.005	0.01
Fricatives	0.395	0.015	0.013
Nasals	−0.004	0.982	0.05
Affricates	0.106	0.533	0.025
Approximants	0.583	*** 0.000	0.006
Shared Phonemes	0.372	0.023	0.017
Unshared Phonemes	0.521	*** 0.001	0.007

Table 4 shows the results of the correlation analyses between maternal linear acculturation scores and children’s phonological accuracy measures in English. Similar to the previous analysis, the results show moderate positive correlations between maternal linear acculturation and children’s (1) overall English consonant accuracy ( $r = 0.471, p = 0.003, \text{adj. } \alpha = 0.008$ ), (2) accuracy with English approximants ( $r = 0.500, p = 0.002, \text{adj. } \alpha = 0.006$ ), and (3) accuracy with unshared phonemes ( $r = 0.502, p = 0.002, \text{adj. } \alpha = 0.007$ ), suggesting, again, that increasing maternal acculturation was related to children’s higher accuracy with English consonants in general, with approximants, and with segments specific to English. Weak to moderate positive correlations were also found between maternal linear acculturation and children’s accuracy with stops, fricatives, and shared phonemes, but again, these correlations did not reach statistical significance due to the Holm–Bonferroni adjusted alpha levels. On the other hand, accuracy with nasals and affricates did not appear at all related to mothers’ overall acculturation level.

**Table 4.** Bivariate correlations between maternal linear acculturation scores and children’s English phonological accuracy measures; \*\*\* = significant effect (Holm–Bonferroni adjusted).

	r-Value	p-Value	Adjusted $\alpha$ -Level
PCC-R	0.471	*** 0.003	0.008
Stops	0.400	0.014	0.01
Fricatives	0.350	0.034	0.013
Nasals	0.046	0.789	0.05
Affricates	0.189	0.262	0.025
Approximants	0.500	*** 0.002	0.006
Shared Phonemes	0.347	0.035	0.017
Unshared Phonemes	0.502	*** 0.002	0.007

### 3.4. Maternal Enculturation and Children’s Spanish Speech Sound Production

Table 5 shows the results of the correlation analyses between maternal Mexican orientation (MOS) and linear acculturation scores and children’s phonological accuracy measures in Spanish. Unlike in English, mothers’ Mexican orientation and their overall acculturation level were not related to any of the children’s phonological accuracy measures in Spanish. These results suggest that maternal cultural orientation and children’s Spanish speech sound production were independent of each other.

**Table 5.** Bivariate correlations between maternal Mexican orientation (MOS) scores and linear acculturation scores and children’s Spanish phonological accuracy measures.

	Mexican Orientation (MOS)		Linear Acculturation	
	r-Value	p-Value	r-Value	p-Value
PCC-R	0.015	0.931	0.262	0.117
Stops	0.016	0.925	0.205	0.225
Fricatives	0.122	0.474	0.115	0.497
Nasals	0.012	0.943	0.049	0.771
Affricates	- 1	-	- 2	-
Approximants	−0.101	0.551	0.153	0.367
Flaps	−0.028	0.871	0.083	0.625
Trills	−0.065	0.704	0.198	0.241
Shared Phonemes	−0.069	0.685	0.248	0.139
Unshared Phonemes	0.022	0.896	0.207	0.219

<sup>1,2</sup> These correlations could not be computed since all children achieved 100% accuracy with Spanish affricates.

## 4. Discussion

The purpose of this study was to examine the link between English and Spanish speech sound production skills at age 4;6 and (a) maternal acculturation, (b) maternal enculturation, and (c) overall maternal linear acculturation, which takes both American and Mexican orientation into account (Cuéllar et al. 1995). Specifically, we focused on (1) whether mothers’ American orientation and overall acculturation were related to children’s accuracy of production of English consonants, of different sounds classes, and of phonemes shared and unshared between English and Spanish, and (2) whether maternal orientation to Mexican culture and overall acculturation were linked to children’s segmental accuracy with Spanish consonants, with different sounds classes, and with shared and unshared phonemes.



Similar to the findings of previous work that documented differences in the speech and language outcomes of children of more vs. less educated mothers (De Anda et al. 2016; Friend et al. 2017; Hammer et al. 2012; Hoff et al. 2018; Montanari et al. 2020; Place and Hoff 2016), we found that children's overall segmental accuracy in English and their production accuracy with English stops, approximants, and English-specific phonemes were positively and moderately correlated with their mothers' levels of American orientation. Likewise, overall English consonant accuracy and accuracy with approximants and with unshared phonemes were also positively and moderately correlated with maternal linear acculturation scores. These results suggest that mothers who were more acculturated when their children began preschool had children with better English speech production abilities one year after the children had begun attending the program. Although the mothers in this study had, on average, a "Very Mexican Orientation" and reported to be using primarily Spanish with their children, their orientation to American culture differed substantially, and it is possible that those mothers who were more American oriented may have socialized their children more into American culture and language, boosting their production abilities and promoting their overall speech sound development in English.

Interestingly, Montanari et al. (2020) found that children of more educated mothers had higher speech sound production accuracy in English than children of less educated mothers only at preschool entry, at age 3;6, but not after one year of preschool attendance. In contrast, in the present study, maternal levels of acculturation were related to children's speech outcomes even at age 4;6, after the children had been exposed to the language and culture of the school for a full year. It is possible that these different findings are due to the studies' differing methodologies and Montanari et al.'s (2020) small sample, as this study compared speech outcomes in children of more and less educated mothers in only 20 participants. In addition, since substantial information on consonant sounds is available even in input that is not particularly frequent or diverse (Dollaghan et al. 1999), one year of preschool may have been enough to close the gap in speech outcomes between children of more and less educated mothers, especially since all mothers had limited education (half completed primary school, the other half secondary school). On the other hand, links between maternal acculturation and children's speech outcomes may persist long-term or emerge with a delay. Indeed, Boyce et al. (2013) documented a positive significant correlation between maternal acculturation measured at 24 months and children's total word knowledge at 36 months. Similarly, Cote and Bornstein (2014) found that mothers' acculturation level when their child was five months of age was positively related to his/her vocabulary size in English at 20 months. In the present study, maternal acculturation was measured when children were at preschool entry, at age 3;6, whereas children's speech sound production was assessed a year later, when they were four and a half. Thus, we speculate that mothers who were more American-oriented at the beginning of preschool possibly socialized their children more into American culture during their children's first year of school, using more English but also exposing them more to English-language activities inside and outside the home. It is also possible that more acculturated mothers had higher levels of English proficiency than less acculturated ones. Taken together, these maternal behaviors and characteristics might have enhanced children's overall production abilities and phonological development in English by the second year of preschool.

It is important to point out, however, that maternal acculturation was not related to all phonological measures in English. For instance, children's accuracy levels with fricatives and shared phonemes were moderately linked to maternal acculturation, but these correlations did not reach statistical significance. On the other hand, the production of nasals and affricates appeared to be completely independent of mothers' American orientation. Accuracy with nasals was overall high and uniform, and ceiling effects may have masked possible relations. On the other hand, children displayed most variation in their production of affricates (some children had 100% accuracy, while others had 0%), and this large variation may have also obscured possible links. Regardless, while our findings

demonstrate a relationship between the majority of children's English segmental accuracy measures and maternal acculturation, we cannot ascertain why this pattern did not hold across the board, and it is possible that a combination of more than one factor came into play.

Most importantly, as we hypothesized, mothers' acculturation scores were particularly related to children's accuracy of production of English-specific phonemes, which require English input for their development. That is, mothers who were more American-oriented had children who were better at producing phonemes that only exist in English, sounds that could not have emerged through the positive transfer of phonetically similar sounds from Spanish and, according to the *Unified Competition Model* (MacWhinney 2005), produce less frequent and less strong cues that lead to negative transfer and protracted development (Goldstein and Bunta 2012). This result suggests, again, that more acculturated mothers possibly created an environment in which children had more access and exposure to English—and to possibly less accented English—both inside and outside the home, promoting their accuracy of production and acquisition of these unique segments.

In contrast to the results for English, none of the children's phonological accuracy measures in Spanish were related to their mothers' Mexican orientation or linear acculturation scores. These results suggest that increased maternal acculturation was not linked to higher speech sound production abilities in Spanish at age 4;6. These findings are interesting as they reflect the same pattern found in studies of maternal education, which has been found to be related to children's English but not to their Spanish speech and language outcomes. Montanari et al. (2020) speculated that maternal education may have a different impact on children's English and Spanish outcomes due to differences in cultural practices. Specifically, maternal education may be more related to children's language skills in individualistic cultures—such as American culture—that value verbal communication and self-expression and where education places an emphasis on the importance of fostering children's language skills (De Anda et al. 2016; Kuchirko and Tami-LeMonda 2019). On the other hand, cultures that value collectivism, cooperation, and obedience—such as Mexican culture—may promote the importance of teaching children about politeness, respect, and collaboration rather than placing emphasis on “intensive” language instruction. While our results do not allow us to draw any conclusions as to how cultural factors may mediate the relationship between maternal cultural orientation and children's speech sound production, we do not exclude the possibility that our contrasting findings for English and Spanish are related to differences in cultural practices. At the same time, it is important to point out that the mothers in our study had overall high and homogenous levels of Mexican orientation, as shown by the standard deviations and ranges in Table 2, and these fairly uniform Mexican orientation scores may have obscured any possible relation with children's segmental accuracy measures in Spanish. After all, larger studies that have examined the link between maternal heritage culture orientation and children's heritage language vocabulary have indeed documented links between the two (Cote and Bornstein 2014). Thus, we do not exclude that this could be the case also for speech outcomes in larger and more heterogeneous samples of Mexican-American mothers and children.

Finally, as we speculated in Montanari et al. (2020), we do not rule out that our findings are due to the role that English plays as the societal language in the context of this study. Recall that children were in their second year of preschool and, therefore, had already begun their own process of acculturation. This means that they were hearing extensive English input in the environment and this may have interacted with maternal acculturation, producing a combined effect on English but not on Spanish speech sound production. Indeed, extensive sociolinguistic work suggests that children's language skills are ultimately more affected by societal language input than maternal input (Labov 2014). Although our participants were only four and a half, one year in the preschool program had dramatically increased their exposure to English as the societal language and, in turn, expanded their English phonological skills and overall language abilities (Montanari et al. 2018). By age 4;6, children had also learned how to use their languages in school and the

community and become aware of the majority status of English and the minority status of Spanish (Montanari et al. 2019). Thus, it is possible that the children's English skills had been affected by societal input in English as the majority language. Given that a link between maternal education and performance in English as a majority language has been documented across a wide range of ages and bilingual groups (Gathercole et al. 2016), we speculate that the same may be true for maternal acculturation.

## 5. Conclusions

In conclusion, this study documented a link between maternal American orientation and overall acculturation and children's speech sound production in English, but it found no relation between mothers' orientation to Mexican culture and children's segmental production in Spanish. We speculated that the results in Spanish may have been due to the high and homogeneous levels of Mexican orientation among mothers, to language input differences attributable to distinct cultural practices, or to the status of Spanish as a minority language. At the same time, we interpreted the results in English as suggesting that more American-oriented mothers may have been socializing their children more to American culture, and they may have been exposing them more to English-language activities inside and outside the home and to native or near-native English input, therefore boosting their English production abilities.

Although this study contributes to the expanding literature on the relationship between maternal cultural orientation and children's linguistic performance, it has several limitations that should be considered when interpreting the results and planning future research. First, our sample was of moderate size, limiting the generalizability of its findings. In particular, our study only focused on mothers with high and homogenous levels of Mexican orientation and with Spanish as their native language, which might have possibly limited the detection of links with children's phonological skills. Since Latinx groups vary tremendously in terms of race, culture, SES, country of origin, patterns of immigration, levels of acculturation and enculturation, proficiency levels in Spanish and English, and parenting and language practices, it is possible that maternal enculturation is related to children's speech sound production in Spanish in some Latinx groups but not in others. Likewise, maternal acculturation was not related to all phonological measures in English, but we do not exclude the possibility that larger and more diverse samples would reveal such links. Thus, future research should explore mothers' cultural orientation and children's speech and language outcomes in larger samples, in different Latinx groups, and among mothers with a wider range of English and Spanish proficiency and acculturation and enculturation levels, possibly from very Anglo-oriented (and dominant in English) to very Mexican-oriented (and dominant in Spanish). A related limitation is that we did not directly assess mothers' Spanish and English proficiency levels. Language competence in the native language and in the language of the host country—in constant fluctuation during immigration—may directly affect maternal cultural orientation and practices. Thus, future studies should directly measure this important variable and examine its link to children's speech and language outcomes. At the same time, since we largely speculated on maternal cultural practices and their possible effects on children's segmental accuracy, future studies should also collect more in-depth ethnographic information through maternal interviews and observations of parent-child interactions in order to reveal whether maternal cultural orientation does indeed affect child socialization patterns.

The second limitation of this study regarded the analyses. First, children's segmental accuracy was assessed via phonetic transcription, a method that does not necessarily capture fine phonetic detail since it relies on the investigator transcribing the speech "as heard". Recall that our study positions itself within the literature that examines phonological skills in moderately sized samples of bilingual children using single-word samples often elicited with the Bilingual English Spanish Assessment (Peña et al. 2014) and transcribed phonetically (e.g., Bunta et al. 2009; Cooperson et al. 2013; Fabiano-Smith and Goldstein 2010a, 2010b; Gildersleeve-Neumann et al. 2008, 2009; Goldstein and Bunta 2012; Goldstein

et al. 2010; Keffala et al. 2020; Montanari et al. 2018; Ruiz-Felter et al. 2016). While the goal of these studies is to examine the extent to which children’s productions match the adult target based on *auditory* grounds, irrespective of fine phonetic details, future studies should increase the sophistication of the analyses by employing instrumental methods such as acoustic analyses. Indeed, acoustic analyses produce an objective, physical measurement of the acoustic signal, and can thus reveal the degree to which children’s productions match the adult targets phonetically rather than phonologically. Furthermore, even if maternal acculturation was related to children’s speech sound production in English, the results do not imply that the former was the cause for the latter. Thus, future research should also expand the types of analyses and explore more in-depth how maternal cultural orientation contributes to bilingual children’s speech outcomes.

Future studies should also focus on a wider range of child language proficiency measures. Recall that previous studies have documented maternal-education-related differences in children’s vocabulary and grammatical abilities—but not in phonological skills—at age 4;6 (Montanari et al. 2020). Thus, it is important to assess whether mothers’ acculturation and enculturation have a different impact on different measures of child language ability—from speech sound production to vocabulary to grammatical skills. Future studies should also be longitudinal to assess whether English exposure and instruction in the school setting ultimately reduce the strength of the link between maternal acculturation and English abilities while strengthening the relationship between mothers’ heritage language orientation and child heritage language skills. Indeed, since most acculturation studies have focused on toddlers (Cote and Bornstein 2014) and preschoolers (Boyce et al. 2013), it is unclear whether the link between maternal cultural orientation and child speech and language outcomes remains steady throughout childhood or whether it decreases over time with increasing societal input (Labov 2014) and at what point this relationship becomes decoupled, perhaps due to ceiling effects in speech and language development.

Despite these limitations, this study makes an original contribution to the understudied topic of maternal cultural orientation and its link to children’s speech and language outcomes, with important educational implications. Since the emergence of preliteracy skills is dependent upon speech perception and production abilities that are developed in early infancy and childhood (Nittrouer and Burton 2005), educators, administrators, and policymakers should make a deliberate effort to obtain information on the degree of acculturation of young dual language learners’ mothers and create interventions that promote English language and literacy among children from less acculturated families. As early differences in oral language skills typically translate into progressively larger differences in language and literacy skills at later ages, intervention should begin as early as possible, possibly in infancy, in order to improve children’s English outcomes by school entry and increase dual language learners’ educational achievement. It is hoped that this study will serve as the springboard for more investigations on the crucial role that immigrant mothers’ identity reshaping process plays on children’s overall linguistic abilities.

**Author Contributions:** Conceptualization, S.M., R.M., and K.S.; methodology, S.M., R.M.; formal analysis, S.M., R.M.; writing—original draft preparation, S.M.; writing—review and editing, S.M., R.M., and K.S.; project administration, K.S.; funding acquisition, K.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the U.S. National Institute of General Medical Sciences, grant number 5SC3GM847583-3 awarded to Kaveri Subrahmanyam, Marlene Zepeda, and Simona Montanari.

**Institutional Review Board Statement:** The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Institutional Review Board of California State University, Los Angeles (IRB 07-22, approved on 2/12/2008).

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** The data presented in this study are not publicly available since IRB and informed consent protocols under which the data were collected do not allow for the data to be publicly shared.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## References

- Aoyama, Katsura, James E. Flege, Susan G. Guion, Reiko Akahane-Yamada, and Tsuneo Yamada. 2004. Perceived phonetic dissimilarity and L2 speech learning: The case of Japanese /r/ and English /l/ and /r/. *Journal of Phonetics* 32: 233–50. [CrossRef]
- Barlow, Jessica. 2003. The stop-spirant alternation in Spanish: Converging evidence from a fortition account. *Southwest Journal of Linguistics* 22: 51–86.
- Bohman, Thomas M., Lisa M. Bedore, Elizabeth D. Peña, Anita Mendez-Perez, and Ronald B. Gillam. 2010. What they hear and what they say: Language performance in young Spanish-English bilinguals. *International Journal of Bilingualism and Bilingual Education* 13: 325–44. [CrossRef] [PubMed]
- Boyce, Lisa K., Sandra L. Gillam, Mark S. Innocenti, Gina A. Cook, and Eduardo Ortiz. 2013. An examination of language input and vocabulary development of young Latino dual language learners living in poverty. *First Language* 33: 572–93. [CrossRef]
- Bunta, Ferenc, Leah Fabiano-Smith, Brian A. Goldstein, and David Ingram. 2009. Phonological whole-word measures in 3-year-old bilingual children and their age-matched monolingual peers. *Clinical Linguistics and Phonetics* 23: 156–75. [CrossRef] [PubMed]
- Campbell, Thomas F., Christine A. Dollaghan, Howard E. Rockette, Jack L. Paradise, Heidi M. Feldman, Lawrence D. Shriberg, Diane L. Sabo, and Marcia Kurs-Lasky. 2003. Risk factors for speech delay of unknown origin in 3-year-old children. *Child Development* 74: 346–57. [CrossRef]
- Cooperson, Soloman J., Lisa M. Bedore, and Elizabeth D. Peña. 2013. The relationship of phonological skills to language skills in Spanish-English-speaking bilingual children. *Clinical Linguistics and Phonetics* 27: 371–89. [CrossRef]
- Cote, Linda R., and Marc H. Bornstein. 2014. Productive vocabulary among three groups of bilingual American children: Comparison and prediction. *First Language* 34: 467–85. [CrossRef]
- Cuéllar, Israel, Bill Arnold, and Roberto Maldonado. 1995. Acculturation Rating Scale for Mexican Americans-II: A revision of the original ARSMA scale. *Hispanic Journal of Behavioral Sciences* 17: 275–305. [CrossRef]
- De Anda, Stephanie, Natalia Arias-Trejo, Diane Poulin-Dubois, Pascal Zesiger, and Margaret Friend. 2016. Minimal second language exposure, SES, and early word comprehension: New evidence from a direct assessment. *Bilingualism: Language and Cognition* 19: 162–80. [CrossRef]
- Dollaghan, Christine A., Thomas F. Campbell, Jack L. Paradise, Heidi M. Feldman, Janine E. Janosky, Dayna N. Pitcairn, and Marcia Kurs-Lasky. 1999. Maternal education and measures of early speech and language. *Journal of Speech, Language, and Hearing Research* 42: 1432–43. [CrossRef]
- Eadie, Patricia, Angela Morgan, Obioha C. Ukoumunne, Kyriaki Ttofari Eecen, Melissa Wake, and Sheena Reilly. 2015. Speech sound disorder at 4 years: Prevalence, comorbidities, and predictors in a community cohort of children. *Developmental Medicine and Child Neurology* 57: 578–84. [CrossRef]
- Evans, James D. 1996. *Straightforward Statistics for the Behavioral Sciences*. Pacific Grove: Brooks/Cole Publishing.
- Fabiano-Smith, Leah, and Brian A. Goldstein. 2010a. Early, middle, and late developing sounds in monolingual and bilingual children: An exploratory study. *American Journal of Speech-Language Pathology* 19: 66–77. [CrossRef]
- Fabiano-Smith, Leah, and Brian A. Goldstein. 2010b. Phonological acquisition in bilingual Spanish-English speaking children. *Journal of Speech, Language, and Hearing Research* 53: 160–78. [CrossRef]
- Flege, James E. 1981. The phonological basis of foreign accent: A hypothesis. *TESOL Quarterly* 15: 443–55. [CrossRef]
- Flege, James E. 1995. Second language speech learning: Theory, findings, and problems. In *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues*. Edited by Winifred Strange. Timonium: York Press, pp. 233–77.
- Flege, James E., and Ocke-Schwen Bohn. 2021. The Revised Speech Learning Model. Unpublished Preprint. Available online: [https://www.researchgate.net/publication/348265785\\_The\\_revised\\_Speech\\_Learning\\_Model\\_SLM-r](https://www.researchgate.net/publication/348265785_The_revised_Speech_Learning_Model_SLM-r) (accessed on 8 April 2021).
- Friend, Margaret, Stephanie DeAnda, Natalia Arias-Trejo, Diane Poulin-Dubois, and Pascal Zesiger. 2017. Developmental changes in maternal education and minimal exposure effects on vocabulary in English- and Spanish-learning toddlers. *Journal of Experimental Child Psychology* 164: 250–59. [CrossRef]
- Gathercole, Virginia C. Mueller, Ivan Kennedy, and Enlli Môn Thomas. 2016. Socioeconomic level and bilinguals' performance on language and cognitive measures. *Bilingualism: Language and Cognition* 19: 1057–78. [CrossRef]
- Gildersleeve-Neumann, Christina E., Elizabeth D. Peña, Barbara L. Davis, and Ellen S. Kester. 2009. Effects on L1 during early acquisition of L2: Speech changes in Spanish at first English contact. *Bilingualism: Language and Cognition* 12: 259–72. [CrossRef]
- Gildersleeve-Neumann, Christina E., Ellen S. Kester, Barbara L. Davis, and Elizabeth D. Peña. 2008. English speech sound development in preschool-aged children from bilingual English-Spanish environments. *Language, Speech and Hearing Services in Schools* 39: 314–28. [CrossRef]
- Goldstein, Brian A. 2001. Transcription of Spanish and Spanish-influenced English. *Communication Disorders Quarterly* 23: 54–60. [CrossRef]

- Goldstein, Brian A., and Ferenc Bunta. 2012. Positive and negative transfer in the phonological systems of bilingual speakers. *International Journal of Bilingualism* 16: 388–401. [\[CrossRef\]](#)
- Goldstein, Brian A., Ferenc Bunta, Jenny Lange, Jenny Rodriguez, and Lauren Burrows. 2010. The effects of measures of language experience and language ability on segmental accuracy in bilingual children. *American Journal of Speech-Language Pathology* 19: 238–47. [\[CrossRef\]](#)
- Goldstein, Brian A., Leah Fabiano, and Aquiles Iglesias. 2004. Spontaneous and imitated productions in Spanish-speaking children. *Language, Speech, and Hearing Services in Schools* 35: 5–15. [\[CrossRef\]](#)
- Gonzales, Nancy A., George P. Knight, Dina Birman, and Amalia A. Siroli. 2004. Acculturation and enculturation among Latino youth. In *Investing in Children, Youth, Families, and Communities: Strengths-Based Research and Policy*. Edited by Kenneth I. Maton, Cynthia J. Schellenbach, Bonnie J. Leadbeater and Andrea L. Solarz. Washington, DC: American Psychological Association, pp. 285–302.
- Hammer, Carol Scheffner, Eugene Komaroff, Barbara L. Rodriguez, Lisa M. Lopez, Shelley E. Scarpino, and Brian Goldstein. 2012. Predicting Spanish–English bilingual children’s language abilities. *Journal of Speech, Language, and Hearing Research* 55: 1251–64. [\[CrossRef\]](#)
- Hoff, Erika, Andrea Burridge, Krystal M. Ribot, and David Giguere. 2018. Language specificity in the relation of maternal education to bilingual children’s vocabulary growth. *Developmental Psychology* 54: 1011–19. [\[CrossRef\]](#) [\[PubMed\]](#)
- Holm, Sture. 1979. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics* 6: 65–70.
- International Phonetic Association. 1999. *Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet*. Cambridge: Cambridge University Press.
- Jimenez, Daniel E, Heather L. Gray, Michael Cucciare, Sheba Kumbhani, and Dolores Gallagher-Thompson. 2010. Using the revised Acculturation Rating Scale for Mexican Americans (ARSM-A-II) with older adults. *Hispanic Health Care International* 8: 14–22. [\[CrossRef\]](#)
- Keffala, Bethany, Shelley Scarpino, Carol Scheffner Hammer, Barbara Rodriguez, Lisa Lopez, and Brian A. Goldstein. 2020. Vocabulary and phonological abilities affect dual language learners’ consonant production accuracy within and across languages: A large-scale study of 3- to 6-year-old Spanish–English dual language learners. *American Journal of Speech-Language Pathology* 29: 1196–211. [\[CrossRef\]](#) [\[PubMed\]](#)
- Kehoe, Margaret, and Chloé Girardier. 2020. What factors influence phonological production in French-speaking bilingual children, aged three to six years? *Journal of Child Language* 47: 945–81. [\[CrossRef\]](#) [\[PubMed\]](#)
- Kuchirko, Yana A., and Catherine S. Tamis-LeMonda. 2019. The cultural context of infant development: Variability, specificity, and universality. *Advances in Child Development and Behavior* 57: 27–63.
- Labov, William. 2014. The sociophonetic orientation of the language learner. In *Advances in Sociophonetics*. Edited by Chiara Celata and Silvia Calamai. Amsterdam: John Benjamins, pp. 17–28.
- MacWhinney, Brian. 2005. A unified model of language acquisition. In *Handbook of Bilingualism: Psycholinguistic Approaches*. Edited by Judith F. Kroll and Annette M. B. De Groot. Oxford: Oxford University Press, pp. 49–67.
- Montanari, Simona, Robert Mayr, and Kaveri Subrahmanyam. 2018. Bilingual speech sound development during the preschool years: The role of language proficiency and cross-linguistic relatedness. *Journal of Speech, Language, and Hearing Research* 61: 2467–86. [\[CrossRef\]](#) [\[PubMed\]](#)
- Montanari, Simona, Robert Mayr, and Kaveri Subrahmanyam. 2020. Speech and language outcomes in low SES Spanish-English bilingual preschoolers: Effects of maternal education. *The International Journal of Bilingual Education and Bilingualism*. [\[CrossRef\]](#)
- Montanari, Simona, Wendy Ochoa, and Kaveri Subrahmanyam. 2019. A longitudinal investigation of language mixing in Spanish-English dual language learners: The role of language proficiency, variability, and socio-linguistic variables. *Journal of Child Language* 46: 913–37. [\[CrossRef\]](#) [\[PubMed\]](#)
- Nittrouer, Susa, and Lisa T. Burton. 2005. The role of early language experience in the development of speech perception and phonological processing abilities: Evidence from 5-year-olds with histories of otitis media with effusion and low socioeconomic status. *Journal of Communication Disorders* 38: 29–63. [\[CrossRef\]](#) [\[PubMed\]](#)
- Peña, Elizabeth D., Vera Gutiérrez-Clellen, Aquiles Iglesias, Brian A. Goldstein, and Lisa M. Bedore. 2014. *The Bilingual English-Spanish Assessment (BESA)*. San Rafael: AR Clinical Publications.
- Place, Silvia, and Erika Hoff. 2016. Effects and noneffects of input in bilingual environments on dual language skills in 2 ½-year-olds. *Bilingualism: Language and Cognition* 19: 1023–41. [\[CrossRef\]](#)
- Ruiz-Felter, Roxanna, Solaman J. Cooperson, Lisa M. Bedore, and Elizabeth D. Peña. 2016. Influence of current input–output and age of first exposure on phonological acquisition in early bilingual Spanish–English-speaking kindergarteners. *International Journal of Language and Communication Disorders* 51: 368–83. [\[CrossRef\]](#)
- Scarpino, Shelley E. 2011. The Effects of Language Environment and Oral Language Ability on Phonological Production Proficiency in Bilingual Spanish-English Speaking Children. Doctoral dissertation, Pennsylvania State University, University Park, PA, USA.
- Shriberg, Lawrence D., Diane Austin, Barbara A. Lewis, Jane L. McSweeney, and David L. Wilson. 1997. The Percentage of Consonants Correct (PCC) metric: Extensions and reliability data. *Journal of Speech, Language, and Hearing Research* 40: 708–22. [\[CrossRef\]](#) [\[PubMed\]](#)
- Simonet, Miquel. 2010. Dark and clear laterals in Catalan and Spanish: Interaction of phonetic categories in early bilinguals. *Journal of Phonetics* 38: 664–79. [\[CrossRef\]](#)



Article

# Foreign Accent in Pre- and Primary School Heritage Bilinguals

Tanja Kupisch <sup>1,2,\*</sup>, Nadine Kolb <sup>2</sup>, Yulia Rodina <sup>2</sup> and Olga Urek <sup>2</sup><sup>1</sup> Department of Linguistics, University of Konstanz, 78464 Konstanz, Germany<sup>2</sup> Department of Language and Culture, UiT The Arctic University of Norway, 9019 Tromsø, Norway; nadine.kolb@uit.no (N.K.); yulia.rodina@uit.no (Y.R.); olga.urek@uit.no (O.U.)

\* Correspondence: tanja.kupisch@uni-konstanz.de

**Abstract:** Previous research has shown that the two languages of early bilingual children can influence each other, depending on the linguistic property, while adult bilinguals predominantly show influence from the majority language to the minority (heritage) language. While this observed shift in influence patterns is probably related to a shift in dominance between early childhood and adulthood, there is little data documenting it. Our study investigates the perceived global accent in the two languages of German-Russian bilingual children in Germany, comparing 4–6-year-old (preschool) children and 7–9-year-old (primary school) children. The results indicate that in German the older children sound less accented than the younger children, while the opposite is true for Russian. This suggests that the primary school years are a critical period for heritage language maintenance.

**Keywords:** global foreign accent; accent rating; heritage language; majority language; preschool children; school children; Russian; German



**Citation:** Kupisch, Tanja, Nadine Kolb, Yulia Rodina, and Olga Urek. 2021. Foreign Accent in Pre- and Primary School Heritage Bilinguals. *Languages* 6: 96. <https://doi.org/10.3390/languages6020096>

Academic Editors: Robert Mayr and Jonathan Morris

Received: 5 March 2021

Accepted: 14 May 2021

Published: 24 May 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Heritage speakers (HSs) are typically characterized as bilinguals growing up with a minority language, passed on by one or both parents, along with the dominant language of the society. Even though heritage speakers acquire two languages during early childhood, the societal (majority) language (ML) typically becomes their dominant language throughout the lifespan (herein, we will use the term “dominance” in terms of proficiency). The term heritage language has traditionally been used in the context of immigrant languages, indigenous languages and colonial languages. In this study, we focus on an immigrant minority language, specifically Russian in Germany, and on children in the second generation. Child (heritage) bilinguals are at the crossroads between first (L1) and second language (L2) learners. They are pre-critical period, L1 acquirers, who develop their phonological representations and their articulatory and acoustic-perceptual skills in tandem. At the same time, they bear resemblance with late L2 learners because they represent cases of language contact within an individual.

Pronunciation is one of the most puzzling phenomena in HSs. With regards to proficiency in their heritage language, HSs are generally said to possess native-like levels of pronunciation and fluency (e.g., Montrul 2008, p. 163; Polinsky and Scontras 2020, p. 8; Rothman 2009, p. 157). Indeed, research has shown that HSs have relatively “authentic” pronunciation, differing in this respect from late L2 learners (Au et al. 2008; Kupisch et al. 2014; Flores and Rato 2016; Kupisch et al. 2020; Chang et al. 2008; Saadah 2011). This is not surprising, since many phonetic and phonological properties are amongst the earliest acquired properties of language and should therefore be relatively resistant to attrition and resilient to cross-linguistic influence (CLI) (Montrul 2008, p. 71; Polinsky and Scontras 2020, p. 8).

In reality, however, HSs are often deemed foreign-sounding in their heritage language, and native speaker raters can easily detect accent features of the speakers’ dominant language (Kupisch et al. 2014; Lloyd-Smith et al. 2020; Kupisch et al. 2020). Benmamoun



et al. (2013, p. 137), while putting forward that HSs tend to retain their native phonological contrasts in the HL, also mention that the phonetic values of vowels and consonants may nevertheless be affected, thus contributing to a non-native accent. Similarly, while pointing out that “aspects of phonetics and phonological competence appear to be robust in heritage languages”, Polinsky and Scontras (2020, p. 10) also observe that HSs are distinguishable from monolingual native speakers because of their “heritage accent”. Thus, HSs benefit from their early exposure to the heritage language (HL), but are nevertheless vulnerable to dominant language influence. Therefore, HS data supports the idea that early exposure is the prerequisite for a speaker’s attainment of monolingual-like pronunciation (e.g., Flege et al. 1997; Abrahamsson and Hyltenstam 2009), but, at the same time, provides us with opportunities to investigate under which circumstances early exposure does not suffice.

While Benmamoun et al. (2013, p. 140) observed that research on HSs’ phonology has “barely scratched the surface”, numerous publications on adult bilingual HSs have since been published (on segmental properties, see, e.g., Amengual 2012, 2016; Nagy and Kochetov 2013; Mayr and Siddika 2018; Elias et al. 2017; Kissling 2018; Einfeldt et al. 2019; on supra-segmental properties, see, e.g., Chang et al. 2011; Colantoni et al. 2016; Henriksen 2016; Kim 2019, 2020). Studies on global<sup>1</sup> accent that have looked at both native languages of adult HSs have shown that HSs are most often perceived as foreign speakers when speaking their HL but as native when speaking their ML (Kupisch et al. 2014, 2020; Lloyd-Smith et al. 2020). In other words, influence in adult HSs is largely unidirectional, although there are some interesting exceptions, typically observed in populations where the HL predominates in the home, e.g., Sylheti-speakers in the UK, Turkish speakers in Germany (e.g., Mayr and Siddika 2018; Kupisch et al. 2020).

There is also a good coverage of studies on developing child bilinguals, i.e., speakers of a heritage language during early childhood (Kehoe 2015, 2018; Lleó 2016, 2018; Lleó and Cortés 2013). In contrast to the studies on adult speakers, these studies leave no doubt that influence—at this early age—is bidirectional, at least when children start acquiring their two languages simultaneously. For example, one and the same individual might show CLI from German (majority language) to Spanish (heritage language) in terms of syllable structure (Lleó et al. 2003), but from Spanish to German in terms of vowel length (Kehoe 2002). The fact that CLI tends to affect both languages in children, while the HL is relatively more affected during adulthood points to a shift in language balance sometime between early childhood and adulthood. Indeed, it has often been mentioned that one decisive period for the HL is when children enter school and are massively exposed to the societal language (Montrul 2008, 2018). However, there are no longitudinal studies linking early childhood and adulthood, and relatively few studies have investigated bilingual children during their early school years.

In the present study, we investigate the perceived foreign accents of German-Russian bilingual children (ages 4–9 years), who grow up as simultaneous or early sequential bilingual children in Germany, in both of their languages. Although German and Russian are typologically distinct, their phonological systems display some similarities: Both languages are stress-timed, where word level prominence is mainly associated with vowel duration in stressed syllables. Word stress and sentence intonation can affect the meaning and interpretation of words and sentences (Hall 2000; Bryzgunova 1963; Odé 1989; Svetozarova 1998). Both Russian and German have word-final devoicing of obstruents. Nevertheless, the languages sound very different, due to their dissimilar phonemic inventories. The phonological system of Russian contains 42 phonemes, with only five vowels (Avanesov 1956; Halle 1959; Panov 1967; Bondarko 1998), and is characterized by a large variety of positional vowel and consonant alternations. Vowel alternations are constrained by stress, which can fall on any syllable and can move. Russian further distinguishes between soft (palatalized) and hard consonants, and voiceless stops are pronounced without aspiration.

<sup>1</sup> When using the term “global” accent, we meant to imply that the accent does not result from one specific accent feature, but a combination of various segmental and suprasegmental features. These may stem from the influence of one or several contact languages or varieties.

German has only 24 consonants and, unlike in Russian, voiceless stops are aspirated, but a comparatively larger vowel inventory with 15 short and long monophthongs and four diphthongs (Hall 2000). German is also subject to substantial regional variation. The varieties relevant to our study are the Alemannic variety spoken in Konstanz and surroundings and the variety in Berlin, which, despite some idiosyncrasies, resembles Standard German more closely. For example, as a result of regional variation, speakers of the Alemannic variety might produce /z/ as voiceless [s] (Beckman et al. 2009) and palatalize /s/ (phonological process /s/ → [ʃ]/\_ [+stop]) (Auer 1990). The variety spoken in Berlin is characterizable as an urban dialect, which means that there are some dialect features (e.g., production of the front vowels [i] and [e] as rounded, i.e., as [y] and [œ]; Wiese and Freywald 2017), but speakers tend to avoid them. Germans can typically identify speakers from other regions of Germany, even though they cannot always determine the exact region. The varieties of Russian are comparatively more homogeneous. There are a few prominent accent features (e.g., fricativization of /g/ in Southern Russian), but variation is much less pronounced than in German. The data in the present study are representative of the Central dialect group, which is characterized by a variety of vowel reduction patterns in contrast to the Northern dialect group where the reduction is weak or lacking (e.g., producing /ə/ vs. /o/ in unstressed positions) (Kasatkin 1989).

Our study compares the accents of Russian-German bilingual preschool and primary school children. We are particularly interested in potential differences between the languages as well as between different age groups (before and after starting primary school). Our results show that, when speaking German, the younger children are perceived as foreign-sounding more often than the older children, while when speaking Russian, it is the older children who tend to be perceived as more foreign-sounding. This suggests that the phonological systems of bilingual primary school children are still malleable and that it is during these years that a shift from sounding accented in the ML to sounding accented in the HL may take place, and could potentially be prevented. Our paper is structured as follows: In the next section, we provide some background on previous research on phonological development in bilingual children, and findings on the perceived accents of early bilingual adults. In Section 3, we formulate our research questions and hypotheses. Section 4 presents the results, which are discussed in Section 5. We conclude with some notes on the societal relevance of our findings.

## 2. Background

The acquisition of phonology is a gradual process, which starts very early but continues into the preschool years. For example, German-learning children produce the consonants [m b d t n] 90% correct between 1;6–1;11, while [j] and [ŋ] are only acquired between 3;0 and 3;5, and the sibilants [s] and [z] and the affricate [ts] are phonetically unstable up to the age of 5;11 (Fox 2007). Similarly, the acquisition of Russian phonology is characterized by early acquisition of the vowel [a] and on the late acquisition of the vowel [i] (e.g., Shvachkin 1948; Gvozdev 1961). The speech of young Russian children is characterized by an overall softness of consonants. According to Zharkova (2005), palatalized consonants are the most frequent substitutes for other consonant phonemes at an early age. In contrast, stressed syllables are shown to be acquired very early, and Russian-speaking children make few mistakes in choosing which syllable to stress, but they can omit unstressed syllables before the age of 3. The fact that the acquisition of phonology, although it starts very early, continues throughout childhood compromises the idea that phonology is generally early acquired. It is likely that the late(r) acquired properties are particularly vulnerable to CLI in bilingual children, in particular in combination with (typically) massive exposure to the ML in the context of entering primary school, if not earlier, when some phonological properties are not yet stable.

### 2.1. Cross-Linguistic Influence in the Phonologies of HSs during Early Childhood

When it comes to the bilingual acquisition of phonology, most research during the past three decades has focused on whether the speech of bilingual children differs in qualitative and quantitative ways from that of monolingual children. Systematic differences would suggest that there is interaction between the two linguistic systems of the bilingual, i.e., CLI. Most of the relevant research in this area has focused on young bilinguals who are either simultaneous bilinguals (with exposure to both languages from birth) or early sequential bilinguals (with exposure to one language from birth and sequential exposure to the other before the age of three). Traditionally, there were two research trends. The first trend represents studies trying to show that bilingual children start out acquiring both of their languages with one unified speech system (Vogel 1975; Volterra and Taeschner 1978). Under the second approach, bilingual children are assumed to build up two systems from the beginning, although these two systems may interact (Paradis and Genesee 1996).

More recent studies within phonology-oriented research, based on in-depth analyses of production data, have tried to model cross-linguistic interaction (see, e.g., Lleó and Cortés 2013; Kehoe 2015; Lleó 2018). These detailed linguistic studies suggest that the acquisition of phonology may be accelerated or delayed depending on the specific structures under investigation. For example, findings on the same group of German-Spanish bilingual children revealed different patterns of interaction: acceleration of codas in Spanish, deceleration of the vowel length distinction in German, and transfer of long lag voicing into Spanish (Kehoe 2015). The use of the term “transfer” in this approach refers specifically to the presence of a non-native sound or structure in one of the bilingual’s languages, which comes from its presence in the other language, thus differing from its more traditional use, where it is synonymous with cross-linguistic interaction. It can be considered a qualitative difference to monolinguals. By contrast, acceleration and deceleration refer to quantitative differences between monolinguals and bilinguals, which implies that bilinguals undergo qualitatively similar processes as monolinguals but they do so earlier/faster or later/more slowly. In the following, the three types are illustrated on the basis of studies on German-Spanish children in Germany (see Kehoe 2015 for an excellent overview with more examples).

A clear example of acceleration has been provided by Lleó et al. (2003), who showed that German-Spanish bilinguals (HSs of Spanish) produced syllable-final codas in Spanish earlier than Spanish monolinguals, arguably due to experience with more complex codas in German. However, with regard to the same phenomenon in comparable language combinations other authors found deceleration or no influence (e.g., Gildersleeve-Neumann et al. 2008; Almeida et al. 2012; Ezeizabarrena and Alegria 2015), suggesting that language external factors also play a role. An example of deceleration comes from a study by Kehoe (2002) on the acquisition of vowel length in German-Spanish bilinguals. The German vowel system is more complex than the Spanish one and Kehoe (2002) findings suggest indeed that bilingual children experienced difficulty acquiring the more marked system of German vowels. Interestingly, German was the ML in this study, which means that at least at an early age, CLI can also affect the majority language. Moreover, the children did not differ from monolinguals in their acquisition of the five-vowel system of Spanish, their HL, which implies that in some cases, the HL is not affected by influence even if there are differences between the ML and the HL. An example of transfer was provided by Kehoe et al. (2004), who have shown that one of the four German-Spanish children they investigated transferred long lag VOT from German to Spanish. Both German and Spanish have the voiceless stops /p, t, k/, but the phonetic basis underlying the voicing distinction is different in the two languages, as they are realized with long lag in German and with short lag in Spanish. Since transfer of long lag was also found in a study on English-Japanese bilinguals with Japanese as their HL (Johnson and Wilson 2002), it is possible that transfer only affects the HL, while quantitative effects (acceleration and deceleration) can be bidirectional. Finally, note that evidence for language interaction can still be found at primary school ages. For example, Lleó (2018) investigated assimilation of

place in coda nasals in the Spanish productions of German-Spanish bilingual children in Germany. Comparisons to monolinguals and bilinguals growing up in Spain showed that in the productions of the bilinguals living in Germany the phenomenon was still unstable at the age of 7;0. In summary, research on young bilingual children has provided plenty of evidence for language interaction. The influence can be quantitative or qualitative, it can affect the HL and the ML, and it does not necessarily cease when children enter primary school.

There are very few accent rating studies with children, looking at accent globally. “Global” means that there is no focus on any particular phonetic or phonological feature, because an accent can result from a combination of different segmental or suprasegmental features. Accent rating studies with children are particularly challenging because, due to language development, children’s “accents” are different from those of adults for an extended period of time, yet without being foreign. To date, only three accent rating studies have been carried out with children (Asher and García 1969; Snow and Hoefnagel-Höhle 1977; Wrembel et al. 2019): The former two explored accentedness in bilingual children’s ML, and the latter in their HL Polish. Asher and García (1969) analyzed the accentedness of Spanish-English bilinguals aged 7–19 years compared to English monolinguals in their ML English in the US. Ratings were based on four English sentences rated by high school students. Findings show that all bilinguals were rated as accented; however, with an early age of onset upon arrival and with increasing length of exposure the amount of children rated as near-native increased. Snow and Hoefnagel-Höhle (1977) conducted two studies on accentedness to compare the abilities of children and adults (L1 English) to pronounce words in a new foreign language (Dutch): In study 1, three raters judged 136 English-speaking children and adults. In study 2, one rater judged 47 children and adults. The authors found that the pronunciation of the older participants was initially more target-like than that of the younger ones, and that after about a year of exposure this trend started to reverse. Wrembel et al. (2019) combined a detailed phonetic analysis with accent ratings investigating the speech of Polish-English bilinguals in their HL Polish. The rating study was based on data from a sentence repetition task and complemented by data from a narrative task. The bilinguals were preschool and primary school children (mean age 5.79) who were exposed to English (the ML) before the age of three and had at least one Polish-speaking parent. The raters for this study were teachers and teacher trainees, who assessed the degree of accentedness on a 7-point Likert scale. The Polish-English bilingual children were perceived to be significantly more accented than monolingual Polish children, and the amount and quality of the HL Polish input was found to be the main predictor. The relationship in global accentedness between the children’s two languages (HL and ML) during childhood has not been explored, so far. An open question concerns the adequacy of raters for accent rating studies with bilingual children. There has been some discussion in the context of late second language learners on whether it is crucial to work with experienced (i.e., phonetically trained) raters, but with no clear advantage pointing to either of the two types (see Jesney 2004). Therefore, in a study on early bilinguals, Kupisch et al. (2014) engaged 50% of each type, finding no difference between those groups. An additional challenge pertaining to child speech is the fact that children’s sound systems are still in the process of developing and that it may be difficult to distinguish “developmental” accents from “foreign” accents. This would justify prioritizing phonetically untrained raters, as long as they are familiar with child language, provided raters with the combined expertise (phonetics and child speech) are unavailable.

## 2.2. Global Accent in Adult Early Bilinguals

A foreign accent has been defined as any perceived divergence from a native speaker, resulting not from the presence of regional varieties but from the influence of another language (Derwing and Munro 2009). HSs are native speakers by definition, but they may nevertheless be perceived as foreign-sounding. Often, a foreign accent results from a number of features, including segmental and suprasegmental properties as well as

phonological rules, which is why relevant studies have often used the term “global” (“global foreign accent”). At the same time, a single sound that is pronounced in a divergent fashion may be sufficient to contribute to a perceived foreignness.

While many accent rating studies have focused on late L2 learners (see [Jesney 2004](#) for an excellent overview), only a few studies have assessed the perceived accents of early bilinguals and have included ratings for both languages. [Kupisch et al. \(2014\)](#) compared HSs when speaking their HL to (i) monolingual speakers of the same language, (ii) late L2 learners and (iii) bilinguals with the same language combination but an inverse dominance relation between ML and HL. For example, HSs of Italian in Germany were compared to (i) monolingual controls from Italy, (ii) L2 speakers of Italian and (iii) HSs of German in Italy speaking Italian as their ML. In addition, there was a (mirror) experiment for German, where the same two early bilingual groups were compared to monolinguals and L2ers. Crucially, all heritage bilinguals in this study had been exposed to both of their languages from birth so that potential effects of Age of Onset (AoO) could be excluded when comparing the ML and HL. The results showed monolingual-like ratings for HSs when speaking their ML and advantages over late L2ers when speaking their HL. Nevertheless, HSs were more often than not deemed foreign speakers of their native language. The study also included experiments with French HSs in Germany and German HSs from France with similar outcomes. Finally, individual speakers were compared in their two languages, and the comparison indicated that a native accent in the ML did not coincide with a foreign accent in the HL, or vice versa, as there were also speakers who were perceived to be native in their two languages.

In a follow-up study with HSs of Italian in Germany ([Lloyd-Smith et al. 2020](#)), HSs with sequential exposure to German were studied. The results confirmed those reported above, i.e., HSs performed on a par with monolingual speakers when speaking their ML, while in their HL they outperformed late L2ers but were nevertheless often perceived to have a foreign accent. It was also shown that a later AoO in German neither had a negative effect on the speakers’ accents when speaking German, nor a positive effect on their accent when speaking Italian. The latter might have been expected because an earlier AoO in German implies less time for the HL to develop independently. However, speakers who had reported using Italian more were perceived to sound more native-like when speaking Italian, which parallels [Wrembel et al. \(2019\)](#) study with Polish-speaking children. This suggests that language use is at least as important a factor for HL development and maintenance, as AoO is (said to be) an important factor for the development of the ML. Another interesting aspect comparing this study with [Kupisch et al. \(2014\)](#) was that the rate of perceived foreignness was lower: 49% in this study as compared to 70% in the Italian HSs of the previous study. This could be related to the different geographical settings, as [Lloyd-Smith et al. \(2020\)](#) study was situated in the South where density of the Italian populations is higher than in the region where [Kupisch et al. \(2014\)](#) had tested their Italian speakers. It could also be related to the higher number of sequential bilinguals in the latter study, i.e., speakers with a later AoO in German who spoke only Italian in the home. In another study with HSs of Turkish in Germany ([Kupisch et al. 2020](#)), these findings were confirmed once more, except that in this study the HSs were often perceived to be mildly accented when speaking their ML. The authors argued that what the raters have perceived might have been features of a variety of German, “Kiezdeutsch” (see [Wiese 2012](#)), that is often spoken by young immigrants and native Germans in cities with larger linguistic minorities.

In summary, accent rating studies with HSs during adulthood show that these speakers are often, though not always, perceived to be foreign speakers of their HL but more rarely of their ML. Amount of language use seems to be a crucial factor for the development of a native-sounding accent in the HL, while it is somewhat controversial whether an early AoO in the ML is necessary for developing a native-sounding accent in the ML, as long as exposure happens during the preschool years. Other factors, such as the density of

the population, which comes with more opportunities to use the language, could play an additional role.

### 3. Research Questions and Predictions

Based on what we know from previous studies, we have formulated the following questions and predictions (P):

RQ1. How often are monolingual children perceived to have a foreign accent?

RQ2. Are bilingual children perceived to have a foreign accent more often than monolingual children when speaking the majority language?

RQ3. Are younger bilingual children perceived to have a foreign accent more often than older children when speaking the majority language?

RQ4. Are older bilingual children perceived to have a foreign accent more often than younger children when speaking the heritage language?

P1. Monolingual children will occasionally be rated as accented, but the incidence of perceived foreignness will fade in the older children. This hypothesis is motivated by the fact that monolingual children develop their phonological systems gradually and that some phonological phenomena are late acquired. Given individual variation, not all adult-listeners might be familiar with all accent features, even if they are used to hearing child speech and sensitized prior to the experiment.

P2. When speaking German, the bilingual children will be perceived as more accented than the monolingual children due to cross-linguistic influence from Russian. This hypothesis is motivated by the literature on young bilingual children having shown that the two languages of bilingual children can mutually influence each other. Of course, a later age of onset in German (in the case of children who speak only Russian at home) and schooling in German might play an additional role.

P3. Perceived accentedness in German, the majority language, will decrease over time. The motivation for this is twofold. First, children's speech becomes more adult-like with increasing age (see P1). Second, rating studies of early bilinguals during adulthood have shown that adult early bilinguals are rarely considered to be foreign-sounding.

P4. Perceived accent in Russian, the children's heritage language, will increase in the older bilingual children. The motivation for this prediction is again twofold. Research on child bilinguals has shown that the two languages of bilingual children show cross-linguistic influence. However, there is no indication that this influence is unidirectional from the majority language in the heritage language or vice versa. During adulthood, by contrast, heritage speakers tend to be accented in their heritage language and not in their majority language (Kupisch et al. 2014; Lloyd-Smith et al. 2020). Thus, a shift from bidirectional to (more) unidirectional influence must take place during their development in later childhood, catalyzed by a change in exposure patterns. We assume that this period coincides with entry to primary school where exposure and use of the majority language steadily increases.

### 4. Study

#### 4.1. Child Participants and Preparation of Materials

We recorded German-Russian simultaneous and sequential bilingual children when narrating picture-based stories in German and in Russian, using the German and Russian versions of the Multilingual Assessment Instrument for Narratives (MAIN) task (Gagarina et al. 2012)<sup>2</sup>. Narratives of 12 children aged 4–6 years (henceforth Group I, mean age 5;4), and of 12 children aged 7–9 years (henceforth Group II, mean age 8;4) were selected based on sound quality and availability of recordings of the same child in both languages. The reason for choosing these ages was as follows: In Germany, children start school at the age of 6 years. Thus, children aged 4–6 will have had exposure to German but presumably not

<sup>2</sup> The MAIN task was developed specifically to assess narrative abilities in the two languages of bilingual preschoolers and school-aged children up to the age of ten. In the production modus of the task, the children are asked to tell a story based on a sequence of six color pictures. The narrative data of the L1 Russian children is from Rodina (2017).

as much as with the beginning of school. At the age of 6, exposure to German is supposed to increase substantially in both quantity and quality, because alphabetization starts as well. Moreover, the educational setting might trigger changes in the children’s identity (self-concept) and attitudes.

All children were born and raised in Germany (either in the Konstanz or the Berlin area<sup>3</sup>). Most of the children (19/24) were growing up in families in which both parents were Russian-speaking and in which only Russian was spoken. These children could be characterized as early sequential bilinguals, since their parents generally reported that their children had their first (intensive) experience with German around the age of three years when entering kindergarten. It needs to be acknowledged, however, that, since the children were growing up in Germany, they had most likely had some exposure to German before, so that the reported AoO of three years can only be approximate. The remaining children (5/24) grew up in families in which one parent was Russian and the other German, and in which both languages were spoken. These children could be characterized as early simultaneous bilinguals, as they had been exposed to both languages from birth with their parents generally following the one parent–one language strategy. Since the information on AoO is more problematic, as mentioned above, we prefer to refer to the groups in terms of family type (“Russian” vs. “mixed”) rather than AoO. Although we have not collected detailed information on input quantity and quality, it seems reasonable to say that the children from Russian families have been exposed to Russian more substantially.

Sound files of 30 s ( $\pm 10\%$ ) per child in each language were extracted from the MAIN narratives (based on two of the MAIN stories). The length of the samples varied slightly in order to avoid interruptions within a sentence. In some cases, edits were required due to the lack of uninterrupted 30-s speech samples. The quality of the files was double-checked by native speakers of each language.

We further included control data from 12 monolingual Russian and 12 monolingual German age-matched children. Fewer control samples were included to avoid a bias towards more unaccented samples (we assumed that not all bilinguals would sound accented). The mean ages of Group I and II for the Russian controls were 4;8 and 8;6 respectively. The mean ages of Group I and II for the German controls were 5;1 and 8;1 respectively. The control data were included to ensure that the raters were able to identify a monolingual accent in German or Russian respectively (see Table 1 for an overview).

**Table 1.** Participating children.

	Group I (4–6-Year-Olds)		Group II (7–9-Year-Olds)	
	Bilinguals	Monolinguals	Bilinguals	Monolinguals
Russian	12	6	12	6
German		6		6

#### 4.2. Rating Task and Raters

The sound files were judged by native speakers of Russian (Russian experiment) and German (German experiment) who lived in Russia and Germany, respectively. The main inclusion criterion for the raters was being familiar with child speech by spending time with 4–9-year-old children on a regular basis. Most of the participating raters ( $n = 36$ ) had children, grandchildren<sup>4</sup> or younger siblings in this age range. This choice of raters had the disadvantage that they were phonetically untrained (unlike in Wrembel et al. 2019), but it came with the advantage that they know what children at comparable ages sound like. We opted for this choice because there is mixed evidence concerning the need of phonetically

<sup>3</sup> There was no intrinsic motivation for this choice.

<sup>4</sup> In Russia, older siblings and grandparents often live together and if grandparents live separately they take care of their grandchildren several days a week.

trained raters, while it was crucial that the raters would be sensitive to developmental aspects of child speech.

The sound files were randomized and presented to the raters by means of a PowerPoint Presentation. The raters' task was to determine for each sample whether the child has "no accent", a "weak foreign accent" or a "strong foreign accent".<sup>5</sup> The rating task consisted of two parts: preschoolers were rated in the first part, school children in the second part. Before the experiment, the raters were familiarized with the procedure and sensitized to the fact that they would listen to children. The raters were instructed to listen carefully to the children's short narratives, and to determine whether the specific child had a (weak/strong) foreign accent or not. The raters could listen to each recording twice. We explained that a foreign accent corresponds to a pronunciation in German or Russian respectively that is influenced by characteristics of another language. In addition, the raters were asked to ignore dialectal features and grammatical errors, and they were reminded of the fact that child speech differs from adult speech as young children often use shorter sentences and have a smaller vocabulary. The German raters participated in the experiment via internet in a Zoom meeting (due to the pandemic). The sound was checked carefully before starting the experiment.

Twenty-two Russian raters (16 females) with a mean age of 44 years (age range 18–69) participated in the Russian experiment. Most of them were mothers and grandmothers (age range 31–69) and one was an 18-year-old sister of a preschool boy. All raters were born and have lived most of their lives in Ivanovo, a city in Central Russia. They all spoke standard Russian without particular dialectal features. Our monolingual controls were from the same city. Twenty-one German raters (13 females) with a mean age of 44 years (age range 22–72) participated in the German experiment. The majority of the raters were parents of preschool and/or school children ( $n = 14$ ), while others were primary school teachers ( $n = 3$ ), grandparents ( $n = 2$ ) and babysitters ( $n = 2$ ) of 4–9-year-old children. At the time of testing, most of the raters ( $n = 14$ ) were living in the Berlin area; the remainder came from other areas in Germany including Western and Southern Germany, often coinciding with the area in which they had grown up.

#### 4.3. Results

In Figure 1, we plotted the distribution of foreign accent ratings ("weak accent", "strong accent" or "no accent") in Russian (left) and German (right) by the family type of the children: heritage mixed (family in which German and Russian are spoken), heritage Russian (family in which both parents speak Russian), monolingual Russian family in Russia or monolingual German family in Germany. In order to calculate the mean accent-ness of speakers (see Figures 2 and 3), accent ratings were converted into numeric scores, where "no foreign accent" corresponded to 0, "weak foreign accent" corresponded to 1, and "strong foreign accent" corresponded to 2. The measure of internal consistency reliability of the ratings in each language, Cronbach's alpha (lrm package, Rizopoulos 2006), was 0.899 for Russian and 0.703 for German data.

<sup>5</sup> The choice of rating options might be criticized on the grounds that we primed the raters by asking them to judge *foreign* accentedness. However, we knew from previous research that early bilinguals are sometimes perceived as foreign-sounding (see literature reviewed above), and our goal was to find out how often this was the case in our data in relation to language, language use at home and age. It is possible that the degree of perceived foreignness would have been lower if we had not primed the raters towards foreignness.



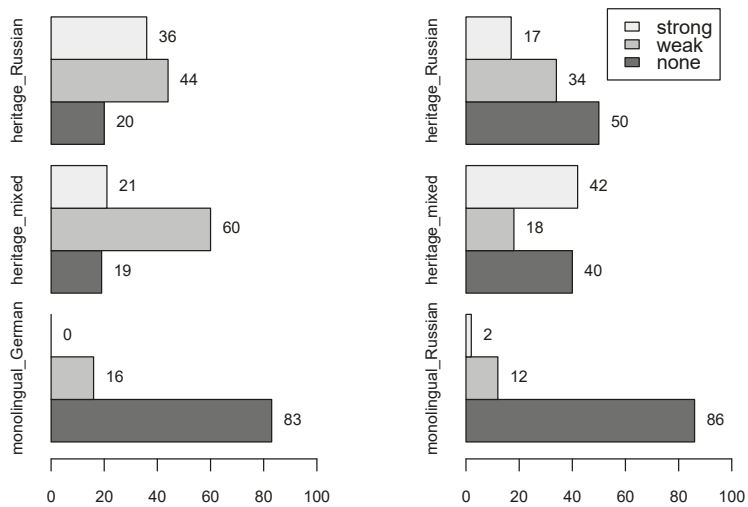


Figure 1. Distribution of accent ratings in German (left) and Russian (right) by language background.

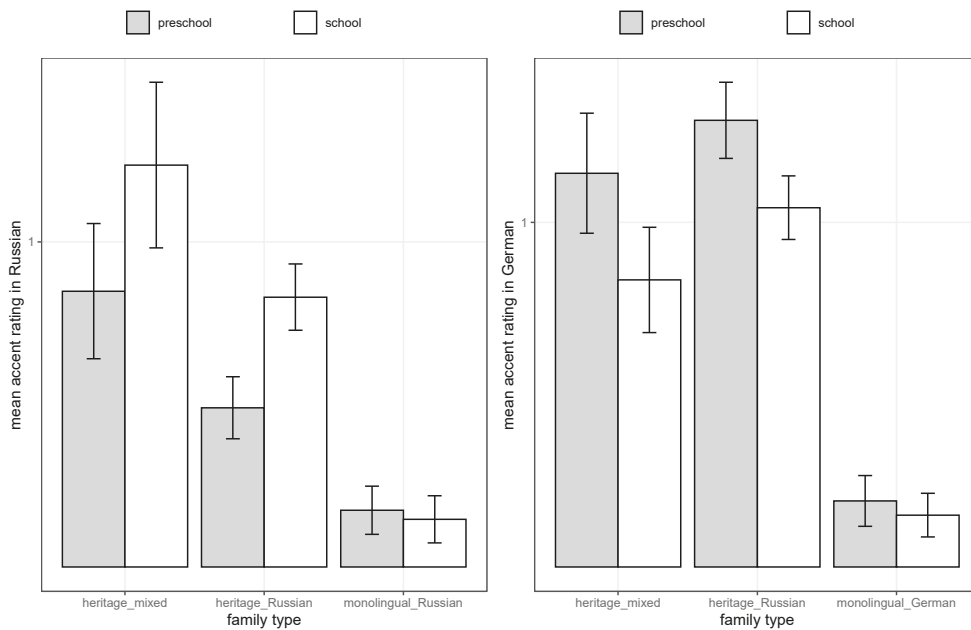
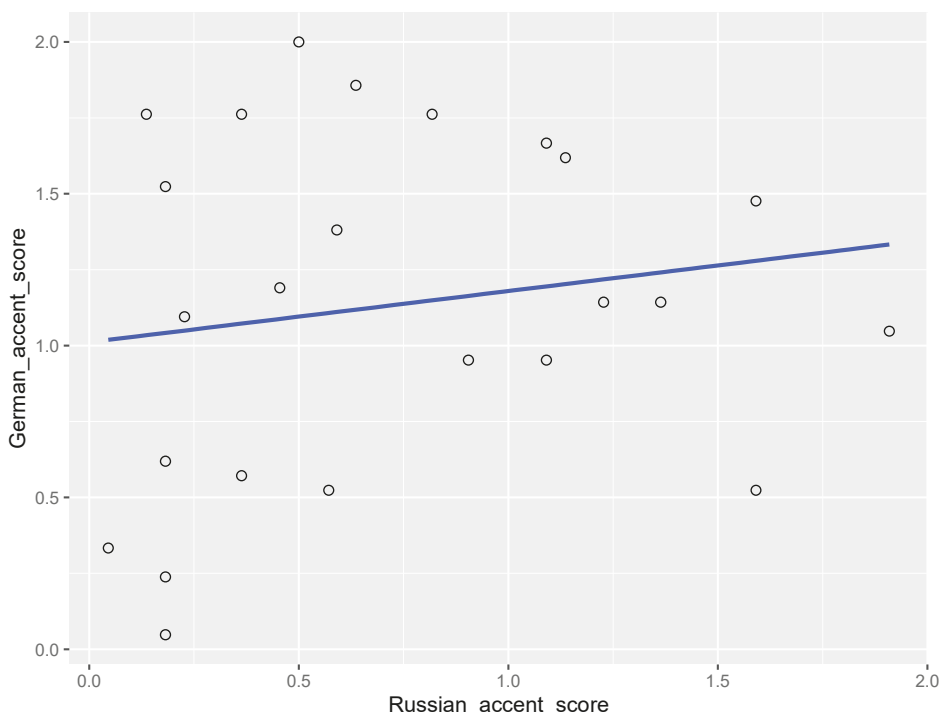


Figure 2. Mean accent ratings in Russian (left) and German (right) by age group and family type (error bars indicate a 95% confidence interval).



**Figure 3.** Mean accent score in German by mean accent score in Russian.

As expected, monolingual Russian and German children were most often rated as unaccented (86% and 83% respectively), although they were perceived as having a weak foreign accent, respectively, 12% and 16% of the time. Bilingual children from families with two Russian-speaking parents were rated as unaccented 50% of the time in Russian, but only 20% of the time in German. In this type of family, accent strength in Russian was mostly perceived to be “weak” (34% of the time) and less often as “strong” (17% of the time), while in German, the accent was perceived to be weak 44% of the time and strong 36% of the time. Bilingual children with one German-speaking parent were perceived as unaccented 40% of the time when speaking Russian, and only 19% of the time when speaking German. In this family type accents tended to be perceived as strong in Russian and as weak in German: 42% of the cases in Russian and 21% in German were rated as strongly accented; 18% of the cases in Russian and 60% of the cases in German were rated as weakly accented.

The effect of family type and age group on the perceived accentedness of the children was investigated with an ordinal mixed-effects logistic model in R (Ordinal package, Christensen and Christensen 2021), where accent score (ordinal variable taking values 0, 1 or 2) was predicted from family type, age group and their interaction, and “rater” was included as a random intercept term. For Russian, the model indicates a significant effect of family type, and a significant interaction between family type and age group, suggesting that the relation between perceived accentedness and age depends on the family type (Tables 2 and 3). For German, the interaction term was not included due to the incomplete information problem (Field et al. 2012, p. 322), resulting from the absence of “strong accent” ratings in German monolingual school children. The results are presented in Tables 4 and 5. For German, both the family type and age group are significantly associated with perceived accentedness.

**Table 2.** Mixed-effects model for perceived accentedness in Russian: random-effects factor.

Groups	Name	Variance	Standard Deviation
ID_rater	(Intercept)	0.4018	0.6338

**Table 3.** Mixed-effects model for perceived accentedness in Russian: fixed-effects factors (\* significant, \*\*\* highly significant).

	Estimate	Std. Error	z-Value	Pr (>  z )	
FamilyBilingual (Mixed)	2.34694	0.34335	6.835	$8.17 \times 10^{-12}$	***
FamilyBilingual (Russian)	1.39138	0.28358	4.906	$9.28 \times 10^{-7}$	***
AgeGroupschool	-0.09154	0.37177	-0.246	0.8055	
FamilyBilingual (Mixed)	1.14514	0.53114	2.156	0.0311	*
FamilyBilingual (Russian)	1.03717	0.41993	2.470	0.0135	*
AgeGroupschool					

**Table 4.** Mixed-effects model for perceived accentedness in German: random-effects factor.

Groups	Name	Variance	Standard Deviation
ID_rater	(Intercept)	0.09011	0.3002

**Table 5.** Mixed-effects model for perceived accentedness in German: fixed-effects factors. (\*\*\*) highly significant).

	Estimate	Std. Error	z-Value	Pr (>  z )	
FamilyBilingual (Mixed)	2.8406	0.2566	11.07	$<2 \times 10^{-16}$	***
FamilyBilingual (Russian)	3.3350	0.2169	15.38	$<2 \times 10^{-16}$	***
AgeGroupschool	-0.6756	0.1557	-4.34	$1.42 \times 10^{-5}$	***

We conducted post hoc pairwise comparisons (Lenth 2016) in order to compare mean accentedness scores between different participant groups (see Appendix A for detailed results). The data are illustrated in Figure 2, which plots mean accent ratings by age group in the different populations (error bars indicate a 95% confidence interval). The analysis shows that for Russian, both family type and age significantly affect the perceived accentedness of the children. Thus, monolingual Russian preschoolers were rated as significantly less accented than their peers from bilingual families with two Russian-speaking parents ( $z$  ratio =  $-4.907$ ,  $p < 0.0001$ ), who, in turn, are rated as less accented than children coming from mixed heritage families ( $z$  ratio =  $-3.408$ ,  $p = 0.0086$ ). The same step-wise distribution is also evident for school-age children, where monolingual Russian children are perceived to be less accented than bilingual children from families where both parents speak Russian ( $z$  ratio =  $-7.696$ ,  $p < 0.0001$ ), who again are rated as less accented than children from families with one Russian- and one German-speaking parent ( $z$  ratio =  $-3.310$ ,  $p = 0.012$ ).

For Russian, the effect of age is evident in that preschoolers are perceived as less accented than school children. The effect holds for the bilingual children from Russian bilingual families ( $z$  ratio =  $-4.844$ ,  $p < 0.0001$ ), and for the children coming from mixed bilingual families, it is marginally significant ( $z$  ratio =  $-2.780$ ,  $p = 0.06$ ). At the same time, no difference in accentedness scores is found between monolingual preschoolers and school-age children. For German, we find an overall effect of family type, such that monolingual German children are perceived as less accented than bilingual children with one Russian- and one German-speaking parent ( $z$  ratio =  $-11.068$ ,  $p < 0.0001$ ), who, in turn, are rated as less accented than children with two Russian-speaking parents ( $z$  ratio =  $-2.414$ ,  $p < 0.04$ ).

The effect of age is significant overall, such that school-age children are perceived as less accented than preschoolers ( $z$  ratio =  $-4.340$ ,  $p < 0.0001$ ).

Figure 3 plots, for each bilingual child, the mean accent score in German against the mean accent score in Russian. A Pearson's test did not show a significant correlation between the two scores ( $r = 0.16$ ,  $p = 0.45$ ). The visual inspection of the scatter plot indicates that there are children who are perceived as not having an accent in any language and other children who are perceived as foreign-accented in one of their languages, while very few children are perceived as strongly accented in both of their languages.

## 5. Discussion

### 5.1. Perceived Accent in Monolingual Children

We predicted (P1) that monolingual children will occasionally be rated as accented but that the incidence of perceived foreignness will fade in the older children. This prediction was motivated by the fact that monolingual children develop their phonological systems gradually and, given individual variation, even listeners sensitized to hearing child speech may sometimes mistake a native accent for a foreign accent.

The monolingual children in our study were perceived to be (weakly) accented 12% (Russian) and 16% (German) of the time, respectively. Whether this small percentage of misclassifications is due to the raters' lack of familiarity with the dialectal features in the speech of individual children, or with difficulties distinguishing the idiosyncrasies of child speech from foreign accents, is hard to tell. The inter-rater reliability in our Russian study suggests that the ratings are reliable. However, they are somewhat less reliable for German, where dialectal variation is more typical, both in the language per se and also in our sample. Therefore, although we did not find that raters from specific areas (e.g., Berlin) systematically misclassified particular children (e.g., from Southern Germany), an overall effect of dialectal variation might be at play. Moreover, in previous studies with monolingual raters for various target languages, the highest rate of "misclassifications" (20%) was found in a study on German (Kupisch et al. 2014).<sup>6</sup> Future studies might want to avoid such effects by including homogeneous speaker samples, and raters who are trained phonetically, familiar with the dialects of the target system and with developmental features of child speech. However, such studies would come at the cost of ecological validity, because in real life situations the two target systems of bilinguals typically do show regional variation, and those people who judge their speech, even if unconsciously, rarely ever happen to be linguists.

### 5.2. Perceived Accent in the Majority Language of Bilingual Children

Our second prediction (P2) was that, when speaking German, the bilingual children would be perceived as being more accented than monolingual children due to CLI from Russian. This prediction is motivated by the literature on young bilingual children showing cross-linguistic interaction, i.e., influence into both the minority and the majority languages. Moreover, although AoO effects are not always visible in HSs during adulthood (see Stangen et al. 2015), they may (still) be visible during childhood. Given that the youngest children in our study were only four years old and partially from monolingual Russian families, their exposure to German had, in the most extreme cases, only started 1–2 years before the data were collected. The results show that, when speaking German, the bilingual children tended to be perceived as accented (79%), compared to only 16% of the time in monolingual children, although in both groups the accents were mostly perceived to be mild. Strong accents were the exception and occurred more often in children from exclusively Russian speaking families with a later AoO in German. Thus, both monolingual and bilingual children are sometimes perceived as accented, but the incidence of perceived

<sup>6</sup> In this specific case, some of the raters commented on one particular speaker, who was indeed the only Southern German speaker in the experiment, and whom they perceived to be foreign.

foreignness is higher in the latter population. This indicates that developmental features, “foreign” features and, potentially, dialectal features are not always easy to tease apart.

We further assumed (P3) that perceived accentedness in German would decrease over time, since rating studies of adult early bilinguals had shown that adult early bilinguals are rarely ever deemed foreign-accented. It was indeed the case that the older children were perceived to be less accented than the younger children, and this trend occurred in both family types. While we need to be cautious in not over-interpreting our data, these results could be taken to suggest that speaking (only) the HL in the family does not harm the development of the ML, at least not as far as accentedness is concerned. Of course, strictly speaking, our study does not allow us to conclude that a foreign accent in the ML decreases over time, because we have not looked at the same children longitudinally. While we can only speculate that by the time these children are adults, they will sound just like monolinguals, previous studies with adult bilinguals (e.g., Kupisch et al. 2014; Lloyd-Smith et al. 2020) suggest that this is likely.

### 5.3. Perceived Accent in the Minority Language of Bilingual Children

We finally predicted (P4) that the incidence of perceived accents in Russian, the children’s HL, would increase in the older bilingual children. To begin with, we found that bilingual children from Russian families or families in which both German and Russian were spoken were perceived to be accented more often (50% and 60% of the time, respectively) than monolingual Russian children (12%). However, they were less likely to be perceived as accented when speaking Russian than when speaking German. Their foreign accents, if perceived, were more likely to be strong when they came from families with one German parent (42%) than when coming from Russian-speaking families (17%). Since the children were exposed to Russian from birth, their perceived accents must be due to the influence of German, which is arguably stronger if German is spoken in the family. These findings are unexpected under the assumption that early exposure is sufficient to attain a native-like accent, but they are expected given previous findings of CLI in early bilingual children. We further observed that the likelihood of a perceived foreign accent was higher for school children than for preschoolers. This was also expected given the high incidence of perceived foreignness in adult HSs in previous studies (e.g., Kupisch et al. 2014, 2020; Lloyd-Smith et al. 2020), as well as Wrembel et al. (2019) finding that Polish-English bilinguals in the UK are perceived to be more accented than their monolingual peers from Poland. However, it is somewhat counterintuitive that older children, despite having had more language experience with the target language than younger children, do not move closer to the target but further away, at least in terms of pronunciation.

Finally, although one might have expected an inverse correlation in accentedness in the children’s two languages, such that children who sound native in German would sound foreign in Russian and vice versa, no such correlation was found. There are children who sound like monolinguals in both languages and children who are perceived as foreign-accented in one of their languages. The number of children being perceived as strongly accented in both languages is small. This is in line with previous studies on adult HSs in which both of the bilinguals’ languages were compared.

## 6. Conclusions and Relevance

In summary, German-Russian children were deemed foreign-sounding significantly more often than monolingual children. This was the case in both languages. There was a difference between the majority language and the heritage language, such that a perceived accent was more likely in German. When speaking Russian, children from German-Russian families were more likely to be perceived as foreign than bilingual children with two Russian-speaking parents, while the opposite was found for German. The incidence of a perceived foreign accent decreased from younger (preschool) to older (primary school) children in German, while increasing for Russian. The latter may be caused by a gradual change in the relative exposure to both languages, as the beginning of school comes

with more German input and qualitatively different input, including formal and written registers. This change in linguistic experience may be accompanied by changes in the HS's attitudes, caused by the prevalent ideologies in the educational settings of in countries with only one national language (e.g., [Hornberger and Wang 2008](#); [Valdéz et al. 2008](#)). Germany represents such a setting.

The observed increase in accent over time in Russian, and its decrease in German must be seen with caution. The first and foremost reason is that we did not investigate individual children longitudinally. Although unlikely, it is possible that we have included school children who happened to be generally more accented in Russian, and preschoolers who happened to be generally more accented in German. Second, our study leaves open which properties exactly are subject to CLI and the related question what exactly the raters perceived to be "different" from monolinguals. Anecdotally, the raters have commented on the pronunciation of /r/, vowel quality and rhythm in German, vowel quality, lack of palatalization and stress in Russian. However, we cannot exclude that the raters also based their judgment on morpho-syntactic properties, even if we explicitly told them not to do so. This might have been the case especially for the younger children when speaking German, as they sometimes omitted articles and produced gender and case errors. Such properties are developmental, i.e., also found in monolingual German children, but in monolingual acquisition they typically occur at an earlier age. Lastly, there was a lot of individual variation, suggesting that factors beyond family type and age were at play. From a methodological perspective, these might include the small sample size as well as a higher degree of misclassifications (compared to previous studies with adults) due to less familiarity with dialectal or developmental features.

Finally, one could question the relevance of such a study, asking why it is important that bilinguals sound "monolingual-like", given that being bilingual is part of their identity and having an accent is part of being bilingual, just as it is typical to speak a standard language with a regional dialect if one comes from a region in which a dialect is (still) spoken. In other words, do we, by studying bilingual children's accents, promote the idea of a "monolingual habitus" ([Gogolin 2008](#))? It is, of course, not our intention to promote a "monolingual habitus", nor do we subscribe to the idea that there is only one "correct" way of speaking a language. On the other hand, it is a fact that a foreign accent is a feature that can reveal a person's origin and identity. If there are positive associations with this origin, this is beneficial. However, an accent may also have negative associations. As shown in a large-scale study by [Gärtig et al. \(2010\)](#), Germans show positive attitudes towards French-, Italian-, English- or Spanish-sounding foreign accents, but not necessarily for certain other accents, and a third of the respondents mentioned comprehension difficulties in conversations with migrants. Such attitudes can affect people's well-being. One example is that young people with Turkish citizenship sometimes report that people in Turkey identify them as coming from Germany, while in Germany they count as "the Turkish", which could make them feel like strangers in both countries ([Kupisch et al. 2020](#)). As to the majority language, potential employers may unfairly discriminate against applicants with non-native sounding accents due to stereotyping or cultural biases ([Munro 2003](#); [Hosoda and Stone-Romero 2008](#)). From a linguistic perspective, it will be crucial to investigate in future studies whether differences in phonological production are mirrored in perception. Perception differences might have implications for the acquisition of morpho-syntactic properties, e.g., gender and number, which are typically marked by word-final sounds ([Colantoni et al. 2020](#)). If these are perceived differently, they may also be produced differently. Taking these points together, the choice about whether or not to consciously develop one's accents, e.g., by increasing use and/or exposure, should be made by the speakers themselves, and a first step towards this is to gain knowledge about when a foreign accent emerges.

**Author Contributions:** Methodology, all authors; formal analysis, O.U.; data collection, N.K. and Y.R.; writing and draft preparation, T.K. (60%), O.U. (20%), N.K. (10%), Y.R. (10%). All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was partially supported by a grant from the Research Council of Norway for the project Microvariation in Multilingual Acquisition & Attrition Situations (MiMS), project code 250857.

**Institutional Review Board Statement:** The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Institutional Review Board (or Ethics Committee) of NSD Norwegian Centre for research data (protocol code: 51174, date of approval: 18 November 2016).

**Informed Consent Statement:** Informed consent was obtained from the parents of all children and all adults who have participated in the study.

**Acknowledgments:** We wish to thank Anne-Sophie Hufer and Ksenia Mack for their help in preparing the stimuli and all children, parents and raters for their participation. We thank Bernhard Brehmer, Henrik Gyllstad and Martin Schweinberger for comments on individual aspects of the paper. Special thanks the Kifu Akademie in Singen for supporting the data collection.

**Conflicts of Interest:** The authors declare no conflict of interest.

### Appendix A

**Table A1.** Russian experiment: pairwise comparisons, age group by family type (subset).

Contrast	Estimate	SE	df	z. Ratio	p-Value
Monolingual (RU) preschool—Bil. (Mixed) preschool	−2.34694	0.343	Inf	−6.835	<0.0001
Monolingual (RU) preschool—Bil. (RU) preschool	−1.39138	0.284	Inf	−4.906	<0.0001
Monolingual (RU) preschool—Mon. (RU) school	0.09154	0.372	Inf	0.246	0.9999
Monolingual (RU) preschool—Bil. (Mixed) school	−3.40054	0.389	Inf	−8.738	<0.0001
Monolingual (RU) preschool—Bil. (RU) school	−2.33701	0.278	Inf	−8.404	<0.0001
Bilingual (Mixed) preschool—Bil. (RU) preschool	0.95556	0.280	Inf	3.408	0.0086
Bilingual (Mixed) preschool—Mon. (RU) school	2.43848	0.374	Inf	6.516	<0.0001
Bilingual (Mixed) preschool—Bil. (Mixed) school	−1.05360	0.379	Inf	−2.780	0.0607
Bilingual (Mixed) preschool—Bil. (RU) school	0.00993	0.269	Inf	0.037	1.0000
Bilingual (RU) preschool—Mon. (RU) school	1.48292	0.320	Inf	4.628	0.0001
Bilingual (RU) preschool—Bil. (Mixed) school	−2.00916	0.333	Inf	−6.034	<0.0001
Bilingual (RU) preschool—Bil. (RU) school	−0.94563	0.195	Inf	−4.844	<0.0001
Monolingual (RU) school—Bil. (Mixed) school	−3.49208	0.417	Inf	−8.375	<0.0001
Monolingual (RU) school—Bil. (RU) school	−2.42855	0.316	Inf	−7.695	<0.0001
Bilingual (Mixed) school—Bil. (RU) school	1.06353	0.321	Inf	3.310	0.0120

**Table A2.** German experiment: pairwise comparisons, age group.

Contrast	Estimate	SE	df	z. Ratio
preschool—school	0.676	0.156	Inf	4.340

**Table A3.** German experiment: pairwise comparisons, family type.

Contrast	Estimate	SE	df	z. Ratio
Monolingual (German)—Bilingual (Mixed)	−2.841	0.257	Inf	−11.068
Monolingual (German)—Bilingual (Russian)	−3.335	0.217	Inf	−15.379
Bilingual (Mixed)—Bilingual (Russian)	−0.494	0.205	Inf	−2.414

### References

Abrahamsson, Niclas, and Kenneth Hyltenstam. 2009. Age of onset and nativelikeness in L2: Listener perception versus linguistic scrutiny. *Language Learning* 59: 249–306. [CrossRef]

Almeida, Letícia, Yvan Rose, and João Freitas. 2012. Prosodic influence in bilingual phonological development: Evidence from a Portuguese-French first language learner. In *Proceedings of the 36th Annual Boston University Conference on Language Development*. Edited by Alia Biller, Chung Esther and Kimball Amelia. Somerville: Cascadilla Press, pp. 42–52.

- Amengual, Mark. 2012. Interlingual influence in bilingual speech: Cognate status effect in a continuum of bilingualism. *Bilingualism: Language and Cognition* 15: 517–30. [CrossRef]
- Amengual, Mark. 2016. Acoustic correlates of the Spanish tap-trill contrast: Heritage and L2 Spanish speakers. *Heritage Language Journal* 13: 88–112. [CrossRef]
- Asher, James, and Ramiro García. 1969. The Optimal Age to Learn a Foreign Language. *The Modern Language Journal* 53: 334–41. [CrossRef]
- Au, Terry Kit-Fong, Janet S. Oh, Leah M. Knightly, Sun-Ah Jun, and Laura F. Romo. 2008. Salvaging a childhood language. *Journal of Memory and Language* 58: 998–1011. [CrossRef]
- Auer, Peter. 1990. *Phonologie der Alltagssprache: Eine Untersuchung zur Standard/Dialekt-Variation am Beispiel der Konstanzer Stadtsprache*. *Studia Linguistica Germanica* 28. Berlin and New York: Walter de Gruyter.
- Avanesov, Ruben Ivanovich. 1956. *Fonetika Sovremenno Russkogo Literaturnogo Jazyka [Phonetics of Modern Russian Literary Language]*. Moscow: Moscow University Press.
- Beckman, Jill, Michael Jessen, and Catherine Ringen. 2009. German fricatives: Coda devoicing or positional faithfulness? *Phonology* 26: 231–68. [CrossRef]
- Benmamoun, Ellabas, Silvina Montrul, and Maria Polinsky. 2013. Heritage languages and their speakers: Opportunities and challenges for linguistics. *Theoretical Linguistics* 39: 29–181. [CrossRef]
- Bondarko, Liya V. 1998. *Fonetika Sovremenno Russkogo Jazyka [Phonetics of Modern Russian Language]*. St. Petersburg: St. Petersburg University.
- Bryzgunova, Elena A. 1963. *Prakticheskaya Fonetika i Intonatsiya Russkovo Yazyka [Practical Phonetics and Intonation of the Russian Language]*. Moscow: Moskovskiy Gosudarstveniy Universitet.
- Chang, Charles B., Erin Haynes, Russel Rhodes, and Yao Yao. 2008. *A Tale of Two Fricatives: Consonant Contrast in Heritage Speakers of Mandarin*. University of Pennsylvania Working Papers in Linguistics. Philadelphia: University of Pennsylvania, vol. 15, pp. 37–43.
- Chang, Charles B., Yao Yao, Erin F. Haynes, and Russel Rhodes. 2011. Production of phonetic and phonological contrast by heritage speakers of Mandarin. *The Journal of the Acoustical Society of America* 129: 3964–80. [CrossRef]
- Christensen, Rune, and Bojesen Christensen. 2021. Analysis of Ordinal Data with Cumulative Link Models—Estimation with the R-Package Ordinal. Available online: [https://cran.r-project.org/web/packages/ordinal/vignettes/clm\\_article.pdf](https://cran.r-project.org/web/packages/ordinal/vignettes/clm_article.pdf) (accessed on 19 May 2021).
- Colantoni, Laura, Alejandro Cuza, and Natalia Mazzaro. 2016. Task-related effects in the prosody of Spanish heritage speakers and long-term immigrants. In *Intonational Grammar in Ibero-Romance: Approaches across Linguistic Subfields*. Edited by E. Meghan, E. Armstrong and Nicholas Henriksen. Amsterdam: John Benjamins, pp. 4–23.
- Colantoni, Laura, Ruth Martínez, Natalia Mazzaro, Ana T. Pérez-Leroux, and Natalia Rinaldi. 2020. A phonetic account of Soanish-English bilinguals' divergence with agreement. *Languages* 5: 58. [CrossRef]
- Derwing, Tracey. M., and Murray J. Munro. 2009. Putting accent in its place: Rethinking obstacles to communication. *Language Teaching* 42: 476–90. [CrossRef]
- Einfeldt, Marieke, Tanja Kupisch, and Joost van de Weijer. 2019. The production of geminates in Italian-dominant bilinguals and heritage speakers of Italian. *Language, Interaction, and Acquisition* 10: 177–203. [CrossRef]
- Elias, Vanessa, Sean McKinnon, and Angel Milla-Muñoz. 2017. The Effects of Code-Switching and Lexical Stress on Vowel Quality and Duration of Heritage Speakers of Spanish. *Languages* 2: 29. [CrossRef]
- Ezeizabarrena, Maria-José, and Alaitz Alegria. 2015. Early coda production in bilingual Spanish and Basque. In *The Acquisition of Spanish in Understudied Language Pairings*. Edited by Tiffany Judy and Silvia Perpiñán. Amsterdam: John Benjamins, pp. 75–105.
- Field, Andy, Jeremy Miles, and Zoe Field. 2012. *Discovering Statistics Using R*. Thousand Oaks: SAGE Publications.
- Flege, James Emil, Elaina M. Frieda, and Takeshi Nozawa. 1997. Amount of native-language (L1) use affects the pronunciation of an L2. *Journal of Phonetics* 25: 169–86. [CrossRef]
- Flores, Cristina, and Anabela Rato. 2016. Global Accent in the Portuguese Speech of Heritage Returnees. *Journal of Heritage Language* 13: 161–83. [CrossRef]
- Fox, Anette V. 2007. *Kindliche Aussprachestörungen*. Idstein: Schulz-Kirchner.
- Gagarina, Natalia, Daleen Klop, Sari Kunnari, Koula Tantele, Taina Välimaa, Ingrida Balčiūnienė, Ute Bohnacker, and Joel Walters. 2012. *Multilingual Assessment Instrument for Narratives (MAIN)*. ZAS Papers in Linguistics 56. Berlin: ZAS, Available online: <http://www.zas.gwz-berlin.de/zaspil56.html> (accessed on 13 May 2020).
- Gärtig, Anne-Kathrin, Albrecht Plewnia, and Astrid Rothe. 2010. Wie Menschen in Deutschland über Sprache denken: Ergebnisse einer bundesweiten Repräsentativerhebung zu aktuellen Spracheinstellungen. In *Arbeitspapiere und Materialien zur Deutschen Sprache* 40. Mannheim: Institut für Deutsche Sprache.
- Gildersleeve-Neumann, Christina E., Ellen S. Kester, Barbara L. Davis, and Elizabeth D. Peña. 2008. English speech sound development in preschool-aged children from bilingual English-Spanish environments. *Language, Speech, and Hearing Services in Schools* 39: 314–28. [CrossRef]
- Gogolin, Ingrid. 2008. *Der monolinguale Habitus der Multilingualen Schule*. Münster: Waxmann.
- Gvozdev, Aleksandr Nikolaevich. 1961. *Formirovanie u Rebenka Grammatičeskogo Stroja Russkogo Jazyka*. Moscow: APN RSFSR.
- Hall, Tracy Alan. 2000. *Phonologie: Eine Einführung*. Berlin: DeGruyter.



- Halle, Morris. 1959. *The Sound Pattern of Russian: A Linguistic and Acoustical Investigation*. The Hague: Mouton & Co.
- Henriksen, Nicholas. 2016. Convergence effects in Spanish-English bilingual rhythm. *Speech Prosody*, 721–25.
- Hornberger, Nancy, and Shuhan C. Wang. 2008. Who are our heritage language learners? Identity and biliteracy in heritage language education in the United States. In *Heritage Language Education*. Edited by Donna M. Brinton, Olga Kagan and Susan Bauckhus. New York: Routledge, pp. 3–35.
- Hosoda, Megumi, and Eugene Stone-Romero. 2008. The effects of foreign accent on employment-related decisions. *Journal of Managerial Psychology* 25: 113–32. [CrossRef]
- Jesney, Karen. 2004. *The Use of Global Foreign Accent Rating in Studies of L2 Acquisition*. Annotated Bibliography. Calgary: Calgary, Language Research Centre, University of Calgary.
- Johnson, Carolyn E., and Ian L. Wilson. 2002. Phonetic evidence for early language differentiation: Research issues and some preliminary data. *The International Journal of Bilingualism* 6: 271–89. [CrossRef]
- Kasatkin, Leonid Leonidovich. 1989. *Russkaja Dialektologija [Russian Dialectology]*. Moscow: Prosveshchenie.
- Kehoe, Margaret. 2002. Developing vowel systems as a window to bilingual phonology. *International Journal of Bilingualism* 6: 315–34. [CrossRef]
- Kehoe, Margaret. 2015. Cross-linguistic interaction: A retrospective and prospective view. In *Proceedings of the International Symposium on Monolingual and Bilingual Speech*. Edited by Elena Babatsouli and David Ingram. Chania: Institute of Monolingual and Bilingual Speech, pp. 141–67.
- Kehoe, Margaret. 2018. The development of rhotics: A comparison of monolingual and bilingual children. *Bilingualism: Language and Cognition* 21: 710–31. [CrossRef]
- Kehoe, Margaret, Conxita Lleó, and Martin Rakow. 2004. Voice onset time in bilingual German-Spanish children. *Bilingualism: Language and Cognition* 7: 71–88. [CrossRef]
- Kim, Ji-Young. 2019. Heritage speakers' use of prosodic strategies in focus marking in Spanish. *International Journal of Bilingualism* 23: 986–1004. [CrossRef]
- Kim, Ji Young. 2020. Discrepancy between heritage speakers' use of suprasegmental cues in the perception and production of Spanish lexical stress. *Bilingualism* 23: 233–50. [CrossRef]
- Kissling, Elizabeth M. 2018. An Exploratory Study of Heritage Spanish Rhotics: Addressing Methodological Challenges of Heritage Language Phonetics Research. *Heritage Language Journal* 15: 25–70. [CrossRef]
- Kupisch, Tanja, Dagmar Barton, Ewgenia Klaschik, Tatjana Lein, Ilse Stangen, and Joost van de Weijer. 2014. Foreign accent in adult simultaneous bilinguals. *The Heritage Language Journal* 11: 123–50. [CrossRef]
- Kupisch, Tanja, Anika Lloyd-Smith, and Ilse Stangen. 2020. Foreign accent in Turkish heritage speakers in Germany. In *Turkish as a Heritage Language in Europe*. Edited by Fatih Bayram. SiBil Series; Amsterdam: John Benjamins, pp. 207–28.
- Lenth, Russell V. 2016. Least-Squares Means: The R Package 'lsmeans'. *Journal of Statistical Software* 69. [CrossRef]
- Lleó, Conxita. 2016. *Bilingualism and Child Phonology*. New York: Oxford Handbooks Online.
- Lleó, Conxita. 2018. Aspects of the phonology of Spanish as a heritage language: From incomplete acquisition to transfer. *Bilingualism: Language and Cognition* 21: 732–47. [CrossRef]
- Lleó, Conxita, and Susana Cortés. 2013. Modeling the outcome of language contact in the speech of German-Spanish and Catalan-Spanish bilingual children. *International Journal of the Sociology of Language* 221: 101–25. [CrossRef]
- Lleó, Conxita, Imme Kuchenbrandt, Margaret Kehoe, and Cristina Trujillo. 2003. Syllable final consonants in Spanish and German monolingual and bilingual acquisition. In *(In)vulnerable Domains in Multilingualism*. Edited by Natascha Müller. Amsterdam: John Benjamins, pp. 191–220.
- Lloyd-Smith, Anika, Marieke Einfeldt, and Tanja Kupisch. 2020. Italian-German bilinguals: The effects of HL use on accent in the early-acquired languages. *International Journal of Multilingualism* 24: 289–304. [CrossRef]
- Mayr, Robert, and Aysha Siddika. 2018. Inter-generational transmission in a minority language setting: Stop consonant production by Bangladeshi heritage children and adults. *International Journal of Bilingualism* 22: 255–284. [CrossRef]
- Montrul, Silvina. A. 2008. *Incomplete Acquisition in Bilingualism: Re-Examining the Age Factor*. Amsterdam: John Benjamins.
- Montrul, Silvina A. 2018. Heritage language development: Connecting the dots. *International Journal of Bilingualism* 22: 530–46. [CrossRef]
- Munro, Murray J. 2003. A primer on accent discrimination in the Canadian context. *TESL Canada Journal* 20: 38–51. [CrossRef]
- Nagy, Naomi, and Alexei Kochetov. 2013. VOT across the generations: A cross-linguistic study of contact-induced change. In *Multilingualism and Language Contact in Urban Areas: Acquisition-Development-Teaching-Communication*. Edited by Peter Siemund, Ingrid Gogolin, Monika Schulz and Julia Davydova. Amsterdam: John Benjamins, pp. 19–38.
- Odé, Cecilia. 1989. *Russian Intonation: A Perceptual Description*. Amsterdam: Rodopi.
- Paradis, Johanne, and Fred Genesee. 1996. Syntactic acquisition in bilingual children: Autonomous or interdependent? *Studies in Second Language Acquisition* 18: 1–25. [CrossRef]
- Panov, Mikhail Viktorovich. 1967. *Russkaya Fonetika [Russian Phonetics]*. Moscow: Prosveshchenie.
- Polinsky, Maria, and Gregory Scontras. 2020. Understanding heritage languages. *Bilingualism: Language and Cognition* 23: 4–20. [CrossRef]
- Rizopoulos, Dimitris. 2006. ltm: An R package for latent variable modelling and item response theory analyses. *Journal of Statistical Software* 17: 1–25. Available online: <http://www.jstatsoft.org/v17/i05/> (accessed on 15 May 2021).

- Rodina, Yulia. 2017. Narrative abilities of preschool bilingual Norwegian-Russian children. *International Journal of Bilingualism* 21: 617–35. [CrossRef]
- Rothman, Jason. 2009. Understanding the nature of early bilingualism: Romance languages as heritage languages. *International Journal of Bilingualism* 13: 155–63. [CrossRef]
- Saadah, Eman. 2011. The Production of Arabic Vowels by English L2 Learners and Heritage Speakers of Arabic. Ph.D. dissertation, University of Illinois at Urbana-Champaign, Champaign, IL, USA.
- Shvachkin, Nikolai Hristoforovich. 1948. Razvitije fonematičeskogo vosprijatija reči v rannem vozraste. *Izvestija Akademii Pedagogičeskikh Nauk RSFSR* 13: 101–32.
- Snow, Catherine E., and Marian Hoefnagel-Höhle. 1977. Age differences in the pronunciation of foreign sounds. *Language and Speech* 20: 357–65. [CrossRef]
- Stangen, Ilse, Tanja Kupisch, Anna Lia Proietti Ergün, and Marina Zielke. 2015. Foreign accent in heritage speakers of Turkish in Germany. In *Transfer Effects in Multilingual Language Development*. Edited by Peukert Hagen. Amsterdam: John Benjamins, pp. 87–108.
- Svetozarova, Natalia. 1998. Intonation in Russian. In *Intonation Systems: A Survey of Twenty Languages*. Edited by Daniel Hirst and Albert di Christo. Cambridge: Cambridge University Press, pp. 261–74.
- Valdéz, Guadalupe, Sonia V. González, Dania López García, and Patricio Márquez. 2008. Heritage language ideologies of language. In *Heritage Language Education*. Edited by Donna M. Brinton, Olga Kagan and Susan Bauckhus. New York: Routledge, pp. 107–30.
- Vogel, Irene. 1975. One system or two: An analysis of a two-year-old Romanian-English bilingual's phonology. *Papers and Reports on Child Language Development* 9: 43–62.
- Volterra, Virginia, and Traute Taeschner. 1978. The acquisition and development of language by bilingual children. *Journal of Child Language* 5: 311–26. [CrossRef]
- Wiese, Heike. 2012. *Kiezdeutsch: Ein Neuer Dialekt Entsteht*. München: C.H. Beck.
- Wiese, Heike, and Ulrike Freywald. 2017. "Könn'Se berlinern?"—Dialektgrammatik im Deutschunterricht. In *Sprache-Literatur-Region im Deutschunterricht: Fachliche Grundlagen und Unterrichts Anregungen*. Edited by Elisabeth Berner. Potsdam: Universitätsverlag Potsdam, pp. 127–47.
- Wrembel, Magdalena, Marta Marecka, Jakub Szewczyk, and Agnieszka Otwinowska. 2019. The predictors of foreign-accentedness in the home language of Polish–English bilingual children. *Bilingualism: Language and Cognition* 22: 383–400. [CrossRef]
- Zharkova, Natalia. 2005. Strategies in the acquisition of segments and syllables in Russian-speaking children. In *Developmental Paths in Phonological Acquisition. Special issue of Leiden Papers in Linguistics 2.1*. Edited by Marina Tzakosta, Claartje Levelt and Jeroen van de Weijer. Leiden: University of Leiden, pp. 189–213.



Article

# Apocope in Heritage Italian

Anissa Baird <sup>1</sup>, Angela Cristiano <sup>2</sup> and Naomi Nagy <sup>1,\*</sup>

<sup>1</sup> Department of Linguistics, University of Toronto, Toronto, ON M5S 3G3, Canada; anissa.baird@mail.utoronto.ca

<sup>2</sup> Department of Classical Philology and Italian Studies, Università di Bologna, 40126 Bologna, Italy; angela.cristiano989@gmail.com

\* Correspondence: naomi.nagy@utoronto.ca

**Abstract:** Apocope (deletion of word-final vowels) and word-final vowel reduction are hallmarks of southern Italian varieties. To investigate whether heritage speakers reproduce the complex variable patterns of these processes, we analyze spontaneous speech of three generations of heritage Calabrian Italian speakers and a homeland comparator sample. All occurrences (N = 2477) from a list of frequent polysyllabic words are extracted from 25 speakers' interviews and analyzed via mixed effects models. Tested predictors include: vowel identity, phonological context, clausal position, lexical frequency, word length, gender, generation, ethnic orientation and age. Homeland and heritage speakers exhibit similar distributions of full, reduced and deleted forms, but there are inter-generational differences in the constraints governing the variation. Primarily linguistic factors condition the variation. Homeland variation in reduction shows sensitivity to part of speech, while heritage speakers show sensitivity to segmental context and part of speech. Slightly different factors influence apocope, with suprasegmental factors and part of speech significant for homeland speakers, but only part of speech for heritage speakers. Surprisingly, for such a socially marked feature, few social factors are relevant. Factors influencing reduction and apocope are similar, suggesting the processes are related.



**Citation:** Baird, Anissa, Angela Cristiano, and Naomi Nagy. 2021.

Apocope in Heritage Italian.

*Languages* 6: 120. <https://doi.org/10.3390/languages6030120>

Academic Editors: Robert Mayr and Jonathan Morris

Received: 25 February 2021

Accepted: 7 July 2021

Published: 13 July 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Keywords:** heritage language; apocope; vowel centralization; vowel reduction; variationist sociolinguistics; Calabrese; Italian

## 1. Introduction

Heritage language (HL) sociolinguistics generally targets “conflict sites” (Poplack and Meechan 1998), that is, portions of a language that differ from the majority language (English, in many HL studies), in order to determine the degree of influence of English on the given heritage language. Frequent examples are pro-drop (cf. Otheguy et al. 2007; Silva-Corvalán 1994) and voice onset time (Nagy and Kochetov 2013). Sometimes the investigations are framed in terms of “simplification” or “attrition” of the HL grammar, but that often coincides with becoming more English-like.

Instead, this paper examines an indigenous pattern of variation, one that appears in non-heritage varieties and, thus, is not attributable directly to contact effects of the majority language in the heritage context. In this first empirical sociolinguistic investigation of the two related phenomena of word-final vowel reduction and apocope (word-final vowel deletion) in Italian, our primary goal is to establish the distribution of contexts in which full, reduced and deleted word-final vowels occur. That is, we first establish the relative frequency of each variant in each context in which it occurs, and what factors, linguistic and social, condition the variability. This contributes to understanding how and if the two processes are related to each other—i.e., if apocope is the final stage of a reduction process (cf. Sections 2.1 and 6). We first describe the diatopic variation in Italy and Calabria, describing the conversational speech we recorded as ‘Regional Italian’. Section 2.1 discusses two possible changes to vowels—reduction and deletion, focusing on word-final vowels in Calabria.

Our second goal is to determine the extent to which Heritage Italian speakers, living in Toronto, replicate the grammar that conditions this variability, including sensitivity to both linguistic and social factors. In addition to social factors such as age and gender that are predicted to influence homeland speakers, we consider generation (since immigration) and an array of ethnic orientation measures for the heritage speakers. It is important to note that, for heritage-language speakers, use of the language is, generally, more restricted to casual (at home) contexts than is the case for speakers in Italy. Hjelde et al. (2019, p. 186) note that heritage language speakers often are “able to use the heritage language but with a reduced register and with lexical and grammatical reductions, and with deviations compared to the baseline,” likely resulting from the restricted contexts of use, see Section 2.2.

## 2. Language Varieties in Italy

The Italian linguistic repertoire is complex and multidimensional. Therefore, systematic descriptions, based on spontaneously produced speech (talk-in-interaction), of the main aspects of this system are important for understanding the factors that influence Italian speech.

The diatopic axis plays a primary role in the Italian picture: Berruto (2006) sets geographic variation as the basis for all the other types of variation, with which it intertwines. Some scholars claim that certain patterns only emerge in spontaneous speech (cf. Abete 2011; Sornicola 2002). Local and Walker (2005) set the use of ecologically valid data, drawn from talk-in-interaction, such as the interviews we examine, as imperative for the study of the phonetic organization of spontaneous speech.

An important consequence is that every geographic variety has its own diaphasic (stylistic) variation. Moreover, each axis of variation represents a continuum where every variety dissolves into another (Berruto 2007). Since the axes intersect, the Italian repertoire is usually considered a multidimensional continuum of continua (Grassi et al. 1997). This multifaceted situation can be explained historically (cf. Cerruti et al. 2017). The prolonged contact between Italo-Romance dialects (derived from Vulgar Latin) and what became Italian, created varieties of Italian on the diatopic axis (Crocco 2017, p. 91): Regional Italians (henceforth RI) originated. RIs are varieties of Italian spoken in specific geographical areas of the peninsula, whose differentiating features mainly come from local dialects (Grassi et al. 1997). RIs are, on the one hand, the consequence of what Berruto (2005, p. 83) calls “dialectization of Italian”; on the other hand, the result of the convergence from dialects to Italian (Cerruti 2011). Poggi Salani (2010) suggests that RIs constitute the current spoken Italian. In contrast, Standard Italian is an abstraction without native speakers.

RIs are divided into northern, central, southern and Sardinian varieties (Pellegrini 1977). The southern area is further grouped into high-southern and low-southern, with Calabria overlapping both. Our speakers come from parts of Calabria in the low-southern region. Some phonetic features of Calabrian RI are (cf. De Blasi 2014, pp. 110–11):

- Aspiration of -t-, especially following -r- or -n- (*certamente* [tʃɛrtʰa'mɛntʰe] ‘surely’);
- Retroflexion of the cluster -tr-;
- Gemination of intervocalic -b- and -g- (*cabina* [kab'bi:na] ‘cabin’, *cugino* [kud'dʒi:no] ‘cousin’);
- Voicing of plosives after nasals (*tanto* [‘tando] a lot);
- Epenthesis in consonant clusters (*aritmetica* [arite'me:tika] ‘arithmetic’).
- Stressed mid vowels pronounced mid-low, [ɛ] [ɔ] (*certamente* [tʃɛrtʰa'mɛntʰe] ‘surely’).

The interviews with speakers of Calabrian origin contain these traits, together with other phonetic, morphosyntactic and lexical traits featured in Calabrian dialects (e.g., *ci attualizzante* or ‘emphatic *ci*–*ci* ‘there’ before the verb *avere* ‘to have’ for emphasis; dialectal articles; dialectal words such as *ammucciare* ‘to hide’; and confusion between *imparare* ‘to learn’ and *insegnare* ‘to teach’). Examples 1–3, below, from our corpus, illustrate phonetic details:

1. *Una volta una signora ha detto: “Voglio assaggiare il pranzo che avete fatto.”*

- [ 'u:na 'vɔltʰə nə sɪn'jɔ:ra a d' dɛttʰu 'vɔll assad' dʒa:re i p'pʰranʃs ke a' ve:te 'fatto]  
 'Once a lady told me: "I want to taste the meal you have prepared."' [I1F73A, 18:42]<sup>1</sup>
2. *I genitori avevano un pezzo di terra, la campagna, ma io stavo al paese, andavo a insegnarmi l'arte di falegname, di muratore.*  
 [i dʒeni'tɔ:rə ə'və:vno um 'pɛts di 'tɛrrə la kəm'pʰaɲɲa ma 'i:ɔ 'sta:vo al pʰa'e:sə ən'da:va a iŋʃɛn'jarmi,l 'artʰe di faleɲ'ja:me di mura'tɔ:re]  
 'His parents had a piece of land, the farmland, but I stayed in town, I was learning the craft of woodworking, of bricklaying.' [I1M62A, 21:28]
3. *Diciamo che mi sono . . . oltre allo studio fatto un pochino di casini là, no, mi sono occupato di vita d'associazione, ne avev-ho creato un'associazione studentesca.*  
 [di'dʒa:mɔ kə m' 'sɔ:nɔ 'ɔlʃs 'allo 'stu:djɔ o f'fatt m po'ki:nə di: di ka'si:ni' lla nɔ mɔ 'sɔ:nɔ okkʰ'pa:to di 'vi:ta d as,soʃ'at'tsjone n a've:ɔ o kkrɛ'a:t 'u:na assodʒə'zjɔ:nɔ studen'tʰeska]  
 'Let's say I . . . In addition to studying, I've made a bit of a mess, there, you know, I've handled club life, I had- I've created a student club.' [IXM35A, 20:44]

The convergence of a group of dialectal traits onto a base still recognizable as Italian allows us to identify what we have elicited as Calabrian RI, although these traits are present to different degrees across speakers. As the literature claims (cf. Grassi et al. 1997; Cerruti 2011; Crocco 2017; Section 2.2), the rate of perceived 'regionality' varies within the RI. One conditioning factor is how much 'regionality' is valued across the community of speakers. We examined speakers from southern Calabria (primarily Vibo Valentia), and their local varieties remain subject to stigma today, affecting this sense of belonging. Nodari's (2017) review of linguistic autobiographies of Calabrian teenagers shows that parents today do not directly teach dialects and often scold their children for using them. Several participants describe their local variety negatively or report that teachers punish those who use it at school. Some teens feel shame in admitting knowing their dialect. Despite this, dialect is used among peers in informal, everyday contexts.

### 2.1. Final Vowel Phenomena

We distinguish two phenomena that affect final vowels in Italian varieties: reduction and apocope, or deletion. An Italian word with a full-timbre final vowel, like [bam'bi:ɲo] 'child', can be produced with the final vowel reduced, [bam'bi:nə], or deleted, [bam'bi:n]. Some studies claim that reduction and apocope are stages of a single process, with the former being a step toward the latter (cf. Section 2.1.2). We summarize each phenomenon in turn. Final vowel reduction is reported to be typical of high-southern Italian varieties, with some research mentioning deletion there too. As studies attest it also in low-southern varieties (cf. Section 2.1.1), we expect our speakers to show variability.

#### 2.1.1. Reduction

'Reduction' can include any type of decrease in the difficulty, quantity, or size of the articulatory movements required for the realization of a phone. The relevant type of reduction is centralization: the articulation of a vowel in a more central area of the vowel space.

While final vowels in Standard Italian are traditionally considered to be full-timbre, many studies demonstrate the contrary, showing centralization and reduction of duration in every RI and style (cf. Albano Leoni et al. 1995; Fava and Caldognetto 1976; Savy and Cutugno 1997). In fact, the tendency that these studies attest may well be universal (Vayra 1991). Nevertheless, Marotta and Soriano (1998) claim that final vowel reduction is subject to diatopic and substrate factors that modify its intensity in different parts of Italy. In other words, every speaker tends to reduce when using the everyday variety (RI), some speakers more than others, and this may be due to the influence of the local dialect. In fact, descriptions of RIs (De Blasi 2014; Vietti 2019) note reduction only in RIs whose corresponding dialect also has it. Considering the corresponding dialect to glean

information about the RI is justified, since, according to [Telmon \(1993, p. 96\)](#), the dialect is the clearest means to understand and interpret the RI. Even though we cannot know *a priori* which dialectal features will be found in the corresponding RI, we hypothesize, based on this claim, that higher rates of final vowel reduction will be observable in the RI of speakers whose dialect presents the phenomenon.

Therefore, let us turn to Calabrian low-southern dialects, to see if traces of final vowel reduction, potentially transmittable to RI, can be found. While Calabrian low-southern dialects do not traditionally present centralized unstressed vowel systems ([Loporcaro 2009](#)), experimental work shows otherwise. [Trumper et al. \(2001\)](#) assert that Calabrese presents both centralization and notable duration reduction in post-tonic vowels. [Romito et al. \(1997\)](#) and [Loporcaro et al. \(1998\)](#) experimentally attest final atonic vowel centralization for the dialects of the extreme South of Calabria. Romito et al.'s study Crotonese and postulate a synchronic variable rule that operates in that dialectal system: final atonic vowels /i, a, u/ → [ə], with the centralization varying depending on speech style. Loporcaro et al. describe this rule as a step towards the creation of a phonological /ə/ in Southern Calabrese.

Dialect may not be the only source of reduction in Calabrian RI: "Regiolects may develop linguistic innovations of their own which have no basis in the standard variety, nor in the dialects" ([Auer 2005, p. 31](#)) and "The dialectal substrate is not the only element capable of explaining RI" ([Cerruti 2009, p. 31](#)). Furthermore, despite the aforementioned studies, the status of final vowel reduction in Calabria is still ambiguous, empirically. More diachronic and/or apparent time data on both dialect and RI are needed, and there is no clear empirical evidence to determine whether the phenomena described in the literature are ongoing innovations or established patterns (and, if so, since when). Thus, we are interested in seeing whether Calabrian RI presents final vowel reduction, leaving to other works the problem of understanding the origins of the phenomenon.

Considering reduction more broadly, some factors known to influence it are discussed: [Van Bergem \(1994\)](#) mentions lexical frequency and the phonemic context of the vowel. These factors are significant in [Romano and Manco's \(2003\)](#) study of vowel reduction in the varieties of Bari and Lecce (Apulia), and [Bybee \(2003\)](#) notes that high lexical frequency can lead to reduction phenomena. In [Beckman's \(1996\)](#) work on Montreal French, reduction of high vowels occurs especially after fricatives. With regard to vowel identity, [Farnetani and Busà \(1999\)](#) test Italian final vowel reduction and show that /a/ and /o/ tend to maintain a longer duration than (in order of length) /e/, /u/ and /i/. Additionally, [Romito et al.'s \(1997\)](#) study of the Crotonese dialect points out a smaller degree of centralization for /a/—which remains distinct from schwa—than for /i/ and /u/. [Silvestri's \(2018\)](#) analysis of north-western Calabrian varieties also shows that final atonic /a/ is the only vowel not neutralized. Finally, [Voghera \(2001\)](#) notes that vowel reduction can also affect segments that are morphological markers: she reports a study by [Savy \(1999a, 1999b\)](#) which highlights the deletion of morphological suffixes. While this list is not exhaustive, it pertains to factors that can specifically be tested in a corpus of spontaneous speech (i.e., excluding stylistic variation).

### 2.1.2. Deletion

We first distinguish variable apocope from categorical apocope. The latter is a process that eliminates sequences of vowels across a word boundary. It affects all Italian varieties including Standard Italian. Categorical apocope, according to [Nespor \(1990\)](#), includes elision (*la elica* > *l'elica*, 'the propeller'), vowel degemination (*erano orridi* > *èranòrriidi*, 'they were horrible') and specifier vowel deletion (*quell'albero* (« *quello*», 'that tree'). All such contexts are excluded from our dataset so that we can focus on variable apocope in other (preconsonantal) contexts.

Turning to variable apocope, [Russo and Barry \(2002\)](#) assert that the deletion of the centralized vowel corresponds to the final stage of a weakening process. [Farnetani and Busà \(1999\)](#) similarly see apocope in Italian as "the endpoint of gradient articulatory reduction rather than the outcome of phonological rules". If we once again seek clues

in the dialect for what may happen in RI, we find some studies claiming that apocope has started to affect dialects that previously only showed reduction. Examples are the dialects of Naples (cf. Albano Leoni 2015; Maturi and Mastantuoni 2012; Cristiano 2019; Radtke 1997; Rohlf 1966), Apulia and Lucania (Loporcaro 1988; Romano 2020), Ischia and Pozzuoli (Russo and Barry 2002) and North-Western Calabria (Silvestri 2018). We can thus conclude that centralization and apocope might well be connected. The former affects the quality of vowels, while the latter affects their duration. We expect that dialects such as the low-southern Calabrian ones, which already feature centralization and reduction in atonic vowels according to studies cited in Section 2.1.1, may be inclined to (variably) delete the same atonic vowels, and that they may transfer this inclination to their RI.

Let us now list some conditioning factors for apocope. The position of the token seems to be particularly influential. Russo and Barry (2002) include some “standard” contexts for the phenomenon: intonation-phrase final and intermediate phrase boundary. They also link apocope to speaking rate. Loporcaro’s (1988) claim (regarding the Altamura dialect) is that final vowel apocope only happens pre-pausally, as the step following reduction. Romano and Manco (2003) hypothesize syntactic and intonational boundaries to be influential for final-vowel deletion too. Maturi and Mastantuoni (2012) also mention pre-pausal position, while adding to the list of influencing factors sociolinguistic features such as age, the relationship of speakers with their variety and communicative context.

Conditioning factors proposed in literature are hence summarized as follows:

- Speaking rate;
- Intonational and syntactic context;
- Social variables.

## 2.2. Stylistic Variation

In spoken language, features from a range of (diastratic and diatopic) varieties may mix, making it difficult to provide exact matches between extracts of spontaneous speech and the named varieties of spoken language reported to exist in Italy. A quick look at the examples from our corpus (1–3, above) show that they contain features connected to both Standard Italian and RI or even dialect. Bell (1984) pointed out that individuals’ stylistic variation mirrors interspeaker variation defined by social status. Thus, we expect to find, in a corpus of spontaneous speech recorded in casual contexts, variation that includes forms associated with varieties of lower status. If we compare the descriptions of standard Italian, RIs and dialects, we see that reduction and deletion are associated with the latter but not the former. For heritage speakers, who virtually only use their heritage language in casual contexts, we might expect a higher concentration of these non-standard variations than in a sample elicited in the same way from homeland speakers, who have more exposure to (and, in some cases, use of) Standard Italian. Indeed, Cerruti (2011) proposes the existence of “folk” RI, educated RI, and standard RI, which can presumably all be found in our corpus and likely even produced by a single individual.

As Grassi et al. (1997) and Crocco (2017) state that the rates of local traits in RI depend on social and contextual factors, we anticipate variability in their rates of use across speakers and across linguistic contexts, even when all speakers are speaking the same RI and all speech is elicited in the same context. Thus, our task is to describe how these social and linguistic context factors predict (or are indexed to) the variation between full, reduced and deleted word-final vowels. Careful analysis will reveal whether these processes reflect an ongoing innovation or something already established and, if so, since when.

## 3. Research Questions

The questions we aim to answer are:

- Can we establish systematic patterns of variation in spontaneous speech concerning vowel reduction and apocope?
- What linguistic and social factors constrain the process? Do these differ between homeland (in Calabria) and heritage (in Toronto) speakers?



- Can we find evidence that relates the processes of vowel reduction and apocope? To probe these questions, the following hypotheses are tested:
  1. We hypothesize that there are differences between older and younger speakers in both the rate and the factors which condition selection between full, reduced and deleted word-final vowels. This is motivated by sociolinguistic expectations that younger speakers exhibit more use of innovative variants than older speakers, allowing us to understand linguistic change through an apparent time lens.
  2. Additionally, we predict that heritage speakers exhibit more reduction and more deletion than homeland speakers because their use of Italian is, generally, more restricted to casual (at home) contexts than is the case for speakers in Italy.
  3. We further anticipate effects of generation and/or ethnic orientation, as heritage speakers who use Italian more often (than English) or are more oriented toward Italian culture (than Canadian culture) are expected to more robustly maintain features of the homeland variety. Each successive generation of heritage speakers has more opportunities to be immersed in English than the previous one.
  4. We also test a series of hypotheses, based on the literature discussed in Section 2, that various linguistic factors (listed in Table 2) affect patterns of deletion and reduction. It is critical to control for variation across linguistic contexts in order to properly measure social factor effects.

**4. Materials and Methods**

The data come from the Heritage Language Documentation Corpus (Nagy 2009, 2011), developed by the Heritage Language Variation and Change in Toronto Project (HLVC, <http://ngn.artsci.utoronto.ca/HLVC>, accessed on 6 July 2021). We examined 25 recorded and transcribed interviews in Calabrian RI. Speakers are distributed as in Table 1. In the HLVC project, speakers are categorized as follows (Nagy 2015, p. 313):

- Homeland: speaker has always lived in southern Calabria, and has southern Calabrian parents<sup>2</sup>;
- Generation 1 (Gen1): speaker has lived at least their first 18 years in Calabria and has lived in the Greater Toronto Area (henceforth Toronto) for at least 20 years;
- Generation 2 (Gen2): parents qualify as Generation 1 (although they might not be in the corpus) and speaker was either born in Toronto or arrived before age six;
- Generation 3 (Gen3): parents qualify as Generation 2 and speaker was born in Toronto.

**Table 1.** Speaker distribution.

Gender	Homeland	Gen1	Gen2	Gen3	Total
Female	5	2	4	1	12
Male	6	4	2	1	13

Interviews were conducted according to standard sociolinguistic interview protocol (Labov 1984). The interviews were conducted in spaces selected by the participants, often at home, to encourage relaxed conversational speech. Recordings used a Zoom H4n digital recorder (44.1 kHz sampling rate) with an Audio Technica lavalier microphone. Elan (Wittenburg et al. 2006) was then used to create time-aligned transcriptions and code contextual factors.

From these interviews, we analyzed 2477 tokens. To select tokens, a concordance was constructed from the 25 interviews. From this word list, 18 frequent polysyllabic words were selected. The five most frequent nouns after these were selected as well. All tokens of these words were selected. From this set, contexts of elision and obligatory apocope were excluded. Mixed-effects models were fitted to the remaining data set.

The dependent variable in these models is the form of the final vowel. This variable was impressionistically coded with three levels: Full, Reduced and Deleted. The coders

are all Italian speakers. Unclear cases of deletion were confirmed by visual examination in Praat. Ambiguous coding decisions were resolved collaboratively by both coders.

We conducted two analyses, each handling these three levels differently, in order to compare the distributions relevant to the two (putatively connected) processes of reduction and deletion. As multivariate analysis using logistic regression requires binary variables, we first combined Deleted and Reduced tokens to create a binary variable: Full vs. Reduced&Deleted, to examine reduction patterns. In the second analysis, Full tokens were excluded to create a binary variable: Reduced vs. Deleted, to examine deletion patterns.

The linguistic factors we predict to influence this variable, based on the literature discussed in Section 2, are FinalVowel, NumberOfSyllables, Stress, ClausalPosition, WordFrequency, PartOfSpeech, TokenOnsetManner, TokenOnsetPlace, NextWordOnsetManner and NextWordOnsetPlace. The latter four were reduced to ResyllabificationPossible (of the two newly adjacent consonants post-deletion), after examination of their behavior. The levels for each predictor are listed in Table 2.

**Table 2.** Linguistic factors analyzed and their variants.

Linguistic Factor	Variants	Example
<b>Segmental</b>		
FinalVowel	a e i o	[bam.'bi.na] '(female) child' [a.'mi.ke] '(female) friends' [bam.'bi.ni] 'children' [bam.'bi.no] '(male) child'
TokenOnsetManner	voiceless stop voiceless fricative/affricate sonorant	[a.'mi.ke] '(female) friends' [,pja.tʃe] 'like it' [bam.'bi.na] '(female) child'
TokenOnsetPlace	alveolar non-alveolar	[bam.'bi.na] '(female) child' [a.'mi.ke] '(female) friends'
NextwordOnsetManner	voiced/voiceless stop voiced/voiceless fricative voiceless affricate or pause nasal liquid pause	<i>erano calabresi</i> 'they were Calabrian' <i>proprio sopra</i> 'right above' <i>mia famiglia</i> . 'my family' <i>diciamo mio</i> 'we say my' <i>siamo rimasti</i> 'we remained' <i>mia famiglia</i> . 'my family'
ResyllabificationPossible	yes no following pause geminate created	<i>adess(o) l'uomo</i> 'now the man' [sl] <i>abbiamo visto</i> 'we saw' [*mv] <i>abbiamo</i> . 'we have [pause].' <i>diciam(o) mio</i> 'we say my' ([mm])
<b>Suprasegmental</b>		
Stress	antepenultimate penultimate	[ 'e.ra.no] '(they) were' [bam.'bi.na] '(female) child'
ClausalPosition	final nonfinal	<i>mia famiglia</i> . 'my family' <i>mia famiglia è ...</i> 'my family is'
<b>Lexical</b>		
WordFrequency	continuous: 34–228 occurrences	<i>proprio</i> 'proper/exact/correct'
PartOfSpeech	adjective adverb noun verb	<i>veramente</i> 'really' <i>bambina</i> '(female) child' <i>siamo</i> 'we are'

Some factors need explanation regarding their purpose in our analyses. The first segmental factor is FinalVowel, the identity of the word-final vowel (that is, the target of apocope and reduction), included due to claims that some vowels are more susceptible to reduction than others (see Section 2.1.1). Place and Manner of both the onset of the final syllable and the onset of the following syllable (in the next word) were coded to investigate

whether sonority influences the processes (cf. Section 2.1). After examination, these four were reduced to ResyllabificationPossible, a factor that distinguished sequences that form a licit onset cluster if the intervening vowel is deleted vs. those that do not form a licit cluster.

Two suprasegmental factors are considered. Stress indicates which syllable of the token word receives primary stress to determine whether foot structure influences the variable processes. ClausalPosition indicates where the token appears in its clause because, as mentioned in Section 2.1.2, boundaries (both syntactic and intonational) may influence apocope.

Two factors relating to the lexical item are considered. PartOfSpeech of the word was included because studies such as Voghera (2001) and Savy (1999a, 1999b) note the role of morphological markers for reduction. This allows us to investigate if the morphological markers of certain word types are more prone to reduction and/or deletion than others. Word frequency has frequently been proposed to affect reduction processes.

Reduction of levels in many of these factors are required by the analyses, described in Section 4, to produce robust models. This was done by combining levels that are most similar phonologically and in terms of rates of deletion or reduction. The results in the tables in Section 5 indicate the final decisions about which levels are distinguished.

To determine whether the variable pattern of deletion is stable or undergoing change, and whether homeland and heritage speakers apply apocope similarly, we included in the model the social factors generation, age, and gender of the speaker. For the heritage speakers, we also considered a selection of ethnic orientation (EO) scores, each corresponding to averaged, quantified responses from one section of the HLVC Ethnic Orientation Questionnaire. The categories are: Ethnic Identity, Language Choices, Cultural Environment, Language Use, Cultural Choices and Experiences of Discrimination. These cover questions such as how the speakers identify themselves ethnically—in the case of Heritage Italian speakers, whether they consider themselves Italian, Canadian, or Canadian-Italian; if most of their neighbors or their friend group is Italian; what languages they speak and in what contexts; their parents' and grandparents' heritage; and their experiences with Italian culture and discrimination. The questionnaire is posted at [http://ngn.artsci.utoronto.ca/pdf/HLVC/short\\_questionnaire\\_English.pdf](http://ngn.artsci.utoronto.ca/pdf/HLVC/short_questionnaire_English.pdf) (accessed on 6 July 2021). EO responses are scored on a scale in which "0" indicates orientation toward English or Canada, "2" indicates orientation toward Italian or Italy, and "1" indicates a "both" or "mixed" response. We predict that heritage speakers with higher scores will produce patterns more like homeland speakers.

Homeland speakers' age range is 19–61, with an average age of 45. The heritage speakers are grouped into three generations (see Table 1). Generations are not defined by age, but age ranges for each generation differ: Generation 1, 61–73 years; Generation 2, 44–57; Generation 3, 21–22. Because age and generation are collinear, separate models are compared, one with generation and the other with age (as a continuous factor).

The tokens, which we code in ELAN, were consolidated with the coding for the speakers' characteristics in a dataframe for analysis. Mixed effects multivariate analyses performed in Rbrul (Johnson 2009) measured the weight of each independent variable, or predictor, on the selection between Full and Reduced or Deleted, and then between Reduced and Deleted, final vowels. Step-up/step-down comparison of models determined the best-fitting models. Models were constructed for the dataset as a whole, and for each generation. These allowed us to compare the factors that predict the reduction and deletion processes in each group.

After testing the linguistic factors and age, gender and generation, the effects of EO were explored. Because EO is not independent of generation and not relevant to homeland speakers, we tested its effect using Spearman ranked-correlation tests to see whether speakers with higher scores also have higher deletion or reduction rates. For this, we used the estimates for Speaker as a random effect from the reported mixed effects models, in order to make comparisons with the effects of differing contextual distributions taken out.

### 5. Results

We first present the distribution of the three variants: Full, Reduced and Deleted word-final vowels. Figure 1 illustrates the stability of the system: each generational group produces about two-thirds of the word-final vowels as Full, about one-quarter Reduced and 1/10 Deleted. Individual are reasonably similar, in terms of use of full forms: rates range from 81% to 100% full forms. Of (the small portion of) non-full forms, speakers range from 47% to 86% reduced (deleting the balance of the vowels).

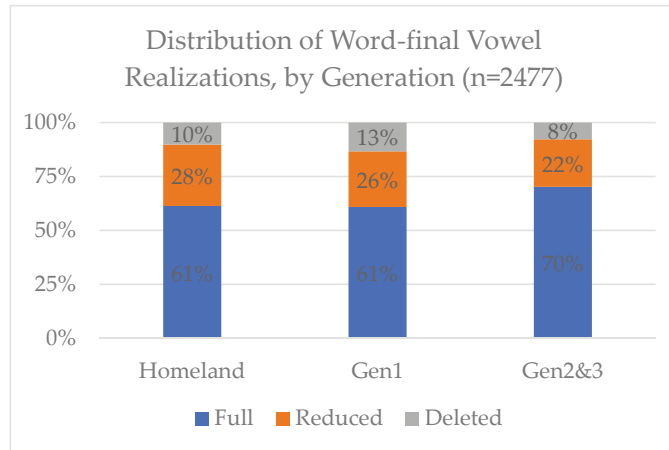


Figure 1. Distribution of vowel realizations by generation.

As confirmed statistically below, there is an age effect for deletion, but not reduction: younger speakers are less likely to delete final vowels. This is illustrated in Figure 2, which shows a difference of 11% in the rate of full forms (from 55% for the oldest group to 66% for the youngest), but only an 8% difference for the rate of Deleted forms. We are not sure why these putatively related processes don't change over apparent time in a more synchronized way, but perhaps the second stage (deletion) is too recent and not yet regularized in parallel with reduction. Alternatively, our models may be missing relevant predictors.

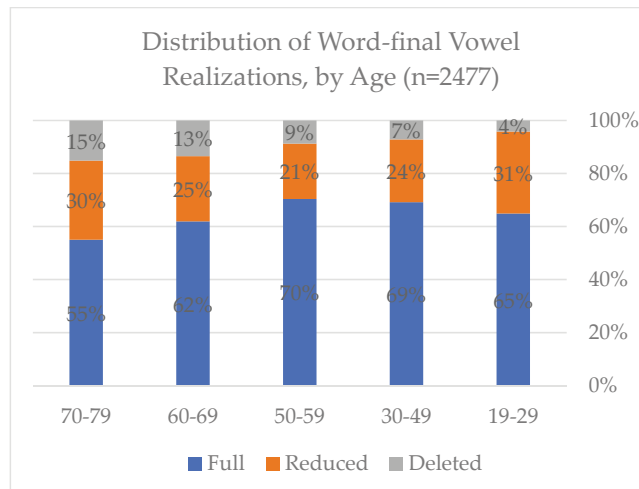


Figure 2. Distribution of vowel realizations by age.

5.1. Full vs. Reduced&Deleted Comparisons

We next compare mixed effects models with various predictors of the form of the final vowel, presenting the models that best fit the distribution of the data. We first consider binary models that contrast the Full form (application value) against the Reduced and Deleted forms (non-application values), to examine the reduction process. Models fit via step-up/step-down comparison indicate that the predictors of production of Full vowels relate to the resyllabifiability of the newly-formed cluster (if the vowel indeed deletes), Stress, Part of Speech and Age. The strongest effect (range = 31) is that Full forms are favoured unless a geminate is created. Full forms are favoured more by Antepenult-stressed words than Penult-stressed, and more by Adverbs, Adjectives, and Nouns than by Verbs. The effect of age is that the older the speaker, the less they favour full forms, as seen in Figure 3. All other factors listed in Table 2 were tested but not found significant in this best-fitting model. Notably, generation did not emerge as significant in models that included it in place of age. Table 3 reports a model with centered factor weights, and range calculated as the difference between the highest and lowest factor weight ( $\times 100$ , by convention). Speaker is included as a random effect and shows that individuals vary from 80–96% Full forms (after excluding five speakers who only produced Full forms).

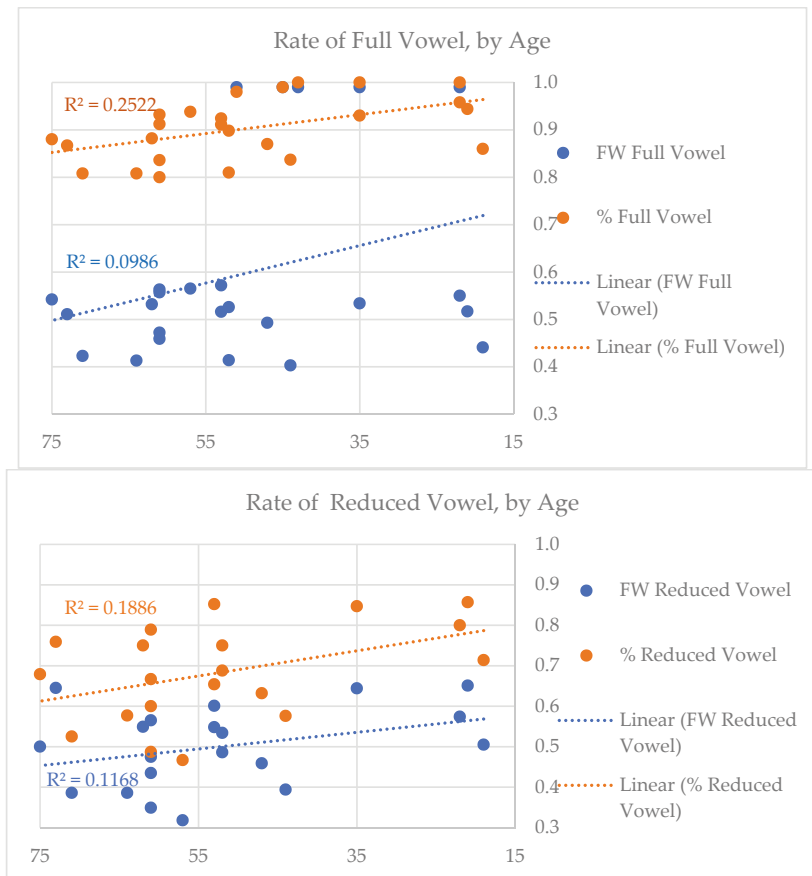


Figure 3. Scatterplots of Speakers for rate of Full vs. Reduced&Deleted (top, n = 2477, of which 2189 included in model) and Reduced vs. Deleted (bottom, N = 815). Speakers' age is on the x-axis.

**Table 3.** Best-fitting mixed-effects model with Full Vowel as application value and Reduced&Deleted as non-application value, for all speakers (N = 2189).

Predictor	Logodds	n	% Full Vowel	Factor Weight	Range
<b>Resyllabification Possible</b> ( $p < 0.001$ )					
Pause follows	0.412	455	91%	0.60	31
Yes	0.349	71	92%	0.59	
No	0.111	1568	88%	0.53	
Geminate created	-0.872	95	76%	0.30	
<b>Stress</b> ( $p < 0.001$ )					
Antepenult	0.299	514	92%	0.57	15
Penult	-0.299	1675	87%	0.43	
<b>Part of Speech</b> ( $p < 0.001$ )					
Adverb	0.207	450	90%	0.55	14
Adjective or Noun	0.157	898	91%	0.54	
Verb	-0.364	841	85%	0.41	
<b>Age</b> ( $p = 0.04$ )					
continuous	-0.015				

As Farnetani and Busà’s (1999) descriptions of Italian apocope report patterns depending of the identity of the final vowel, we report this distribution in Table 4, though it does not have a significant effect in our analyses. (Vowel emergence as a significant predictor of Reduction if Stress is excluded from the model, but Stress produces a better-fit to the data.)

**Table 4.** Distribution of surface forms by final-vowel identity (N = 2477).

Dependent Variable (3 Levels)	i	e	o	a
Deleted	11%	13%	12%	6%
Reduced	19%	21%	19%	44%
Full vowel	70%	65%	70%	50%

We next compare models for the homeland, Gen1 and Gen2/Gen3 speakers in order to determine whether the same grammar determines the selection of Full vs. Reduced word-final vowels in each. (One of our two Gen3 speakers categorically produces full vowels, so the one remaining Gen3 speaker is combined with Gen2.) In these models with fewer tokens than the model of all speakers shown in Table 1, fewer factors emerge as significant. Rather than showing the full models, we note only which factors are significant for each speaker group. Table 5 shows an increase in the complexity of the factors governing reduction: homeland speakers are governed by one (lexical) factor, while heritage speakers are governed by that factor plus a phonological factor. Levels are ranked similarly across models for each generation.

**Table 5.** Significant predictors (indicated with “✓”) in MEMs comparing Full vs. Reduced&Deleted vowels, for each speaker group.

Predictor	All Speakers	Homeland	Gen 1	Gen 2/3
n	2189	644	860	685
Part of Speech	✓	✓	✓	✓
Resyllabification possible	✓		✓	✓
Stress	✓			
Age	✓			

5.2. Reduced vs. Deleted Comparisons

We turn next to a comparison of reduction vs. deletion. In the following models, tokens with full vowels are excluded in order to focus on the alternation between reduction and deletion (which were combined in the non-application value in Section 5.1). The application value is Reduction. Again, we first consider the patterns in the full speaker sample, shown in Table 6. The same three linguistic factors that were significant in the choice between Full vs. Reduced&Deleted are significant here, with Resyllabification again playing a large role. Reduction is disfavoured (in comparison to Deletion) if the deletion introduces a licit cluster or geminate. However, differing from the models of Full vs. Reduced&Deleted, age is not significant here.

Table 6. MEM for all speakers together, comparing Reduction to Deletion (N = 815).

Predictor	Logodds	n	% Reduced	Factor Weight	Range
<b>Resyllabification possible</b> ( $p < 0.001$ )					
Pause follows	0.539	147	74%	0.63	32
No	0.467	607	69%	0.62	
Yes	-0.225	17	65%	0.44	
Geminate created	-0.781	44	48%	0.31	
<b>Part of speech</b> ( $p < 0.001$ )					
Adverb	0.543	191	76%	0.63	32
Adjective or Noun	0.259	344	77%	0.56	
Verb	-0.802	280	54%	0.31	
<b>Stress</b> ( $p < 0.001$ )					
Antepenult	0.348	186	76%	0.59	17
Penult	-0.348	629	67%	0.41	

Next, we compare models for each speaker group, to determine how the grammars governing reduction vs. deletion differ across generations, summarized in Table 7.

Table 7. Significant predictors (indicated with “✓”) in MEMs comparing Reduced vs. Deleted vowels, for each speaker group.

Predictor	All Speakers	Homeland	Gen 1	Gen 2/3
n	815	261	337	217
Part of speech	✓	✓	✓	✓
Resyllabification possible	✓	✓		
Stress	✓			
Age				

Unlike the pattern in Table 5, for Reduction vs. Deletion, we see a simplification of the grammar when we compare heritage speakers (governed only by one lexical factor) to homeland speakers (governed by that same factor plus the Resyllabification factor). This may be indicative of some reanalysis of how both processes work. Alternatively, these smaller samples of available tokens may not support robustly representative models. However, within each significant factor, levels are identically ordered.

5.3. Age Effects in Reduction and Deletion

Let us consider the effect of Age more closely. Recall that Age is significant for the analysis of Full vs. Reduced&Deleted forms, but not for the analysis of Reduced vs. Deleted forms. This is unexpected if Deletion is the final stage of the reduction process. Figure 3 contains scatterplots of individual rates for both processes. In the top plot, the orange dots indicate the percentage of tokens that are Full (vs. Reduced&Deleted). In the bottom plot, they indicate the percentage of tokens that are Reduced (vs. Deleted). Blue dots indicate

Speaker random effect estimates for these same application values. Speakers are arranged from oldest to youngest (left to right). Speakers who categorically produce the application value are shown at 100% and with a factor weight of 1.0 (but they were not included in the regression analyses). Including them here provides a more realistic picture of the age effect. The slopes for the two processes are similar: younger speakers produce more Full as well as more Reduced forms. This fuller picture, including categorically Full-vowel speakers, contradicts the models described above, where Age was a significant predictor only for Full vs. Reduced&Deleted. The lack of a significant age effect for the second stage of the process may be due to the smaller sample size. In “real life,” speakers have much larger samples to calibrate to, and this difference between the two processes would be less stark.

#### 5.4. Ethnic Orientation (EO)

As quite a bit of inter-speaker variation is unaccounted for in the models above, we seek explanation in “ethnic orientation”. As described in Section 4, the Ethnic Orientation Questionnaire elicited self-reports on how speakers orient to their Italian ethnicity, their language preferences and practices, their cultural preferences and practices and whether they have been subject to discrimination as Italians. As these responses are not independent of each other, we separately examine the correlation of each factor to speakers’ rates of reduction and deletion, using the estimates calculated for Speaker as a random effect in the models for all speakers reported above (Tables 3 and 6). For the choice between Full vs. Reduced&Deleted, one set of responses is correlated both significantly and strongly: Discrimination. This is a combined score for the following questions, where higher scores reflect experience of discrimination of this type:

- Have you ever had a problem getting a job because you’re Italian?
- What about renting an apartment or buying a house?
- Were you treated differently by your teachers in school?
- Have you ever been treated badly because you’re Italian?
- Is there a lot of discrimination against Italians?

For the choice between Reduced vs. Deleted, a different factor is correlated significant: speakers’ orientation. This factor is composed of responses to these five questions, where higher scores reflect orientation toward their Italian ethnicity:

- Do you think of yourself as Italian, Canadian or Italian-Canadian?
- Are most of your friends Italian?
- Are people in your neighborhood Italian?
- Are the people you work with Italian?
- When you were growing up, were the kids in your school Italian? Were your friends? The kids in your neighborhood?

This suggests that speakers’ orientation toward their culture, and resulting discrimination, are more connected to their speech production than the expected factors, such as how often they use the language or actually engage in cultural activities with other Italians.

These analyses, using Spearman’s rank correlation  $\rho$ , are summarized in Table 8. As has been reported for other variables examined in this same corpus of Heritage Italian data (Nagy 2018), there is little relationship between ethnic orientation factors and linguistic variation.



**Table 8.** (Non-)correlation of ethnic orientation to apocope and reduction in Heritage Italian (13 speakers).

	Ethnic Identity	Lg. Choices	Cult. Env.	Lg. Use	Cult. Choices	Discrimination
<b>Full vs. Reduced&amp;Deleted</b>						
<i>p</i> -value	0.21	0.41	0.74	0.70	0.96	0.01
$\rho$	0.39	0.26	0.11	0.12	<b>0.01</b>	<b>−0.70</b>
<b>Reduced vs. Deleted</b>						
<i>p</i> -value	<b>0.03</b>	0.29	0.65	0.22	0.11	0.48
$\rho$	<b>0.66</b>	−0.35	−0.15	−0.40	0.51	−0.24

### 6. Discussion

Surprisingly, our expectation of differences between younger and older speakers is only partially supported. For the rates of both reduction and deletion, we see that age is significant for the alternation between full vs. non-full forms. This is apparent-time evidence of a change in progress.<sup>3</sup> Age, however, may not be a conditioning effect for the alternation between reduction and deletion.

Interpreting the factors conditioning the two processes, we detect small changes from generation to generation in the grammar governing patterns of vowel reduction. We see an increase in the number of constraints affecting the alternation between Full and Reduced&Deleted forms, but a reduction for the constraints governing Reduced vs. Deleted forms. Several factors motivated by the literature turned out not to significantly affect this variation: Vowel Identity, ClausalPosition and Lexical Frequency. Place and Manner of the onset consonants turn out not to model the data as well as a combined Resyllabification factor. However, it remains to be explained why Full forms are favored similarly by contexts that do and do not create licit resyllabified clusters, and only disfavored if a geminate would be the result.

Turning to our second hypothesis, we expected the heritage speakers to reduce/delete more than homeland speakers because of assumed restrictions on their use of Italian to informal contexts. Since we do not, in fact, find inter-generational differences in frequency, our results might be interpreted in different ways. We could conclude that Reduction&Deletion rates in one fixed context (the sociolinguistic interview) are not influenced by the overall frequency of use of the variety.

We must also, however, consider the differences in the conditioning effects. Homeland speakers have more linguistic constraints on Deletion (vs. Reduction) than heritage speakers. We are tempted to suggest that this reflects social differences. As noted in Section 2, Calabrian varieties can be subject to negative social evaluation. In order to avoid stigma, homeland speakers may orient their speech away from dialectal variants, using them in a more restricted number of contexts—both linguistic and extralinguistic. That is, there are contexts to which homeland speakers limit their deletions, but these constraints are not transferred to heritage speakers. Such restrictions by the homeland might persist unconsciously even in informal contexts in which dialectal variants are expected. Heritage speakers may have a different attitude towards dialects and dialectal variants, showing appreciation for the dialect as part of their unique cultural identity. Consequently, they do not manifest the same preference for prestigious variants, resulting in less restrictions and more “freedom” in the way they speak, and therefore, in a less constrained grammar of Reduction&Deletion. However, it is hard to align this with the lack of difference in rates between these groups.

We not only anticipated effects linked to generations among the heritage speakers, but also to ethnic orientation. However, the alternation between Full vs. non-full forms is governed by only one EO factor, as is the alternation between Reduced vs. Deleted forms. This suggests that attitudes and beliefs towards the community might indeed be connected to speech production, providing additional support for our interpretation of the decrease in constraints discussed in the previous paragraph.

Critically, our data do not support any attrition-based account, as we do not see significant rate differences based on generation or the EO scores related to language use—two factors that should correlate to deterioration of the grammar as a result of decreasing exposure. Since later generations have more exposure to English (and in more and more contexts), one could expect English to influence their heritage language production increasingly. We must equally reject any account attributing variation to English influence. That is, if any English pattern of reducing unstressed vowels were being adapted by Heritage Italian speakers, then we would expect to see this more frequently in the speakers with more contact with English, but we do not.

Our fourth set of hypotheses about linguistic factors affecting patterns of deletion and reduction have been partially confirmed, with a consistent set of factors affecting both processes. PartOfSpeech has a surprisingly consistent effect, and Resyllabification a partially unexplainable one, for now. The consistent effect of more Full forms, and then more Reduction, for Antepenultimate-stressed words than Penult-stress words indicates that foot structure must be considered in any phonological analysis of the processes, and more broadly, that these are indeed phonological processes.

We consider next the lack of effect of the identity of the final vowel. This may be due to an accident in the structure of our dataset: all words ending in /e/ or /i/ have penultimate stress. This means that Stress and VowelIdentity cannot be included in the same model. While Stress produces a better-fitting model, it does not mean that there is no secondary role for the identity of the vowel. Indeed, Table 4 shows sizable differences among vowels, particularly for reduction rates. If the results actually reflect a situation where vowels undergo reduction or deletion depending on their identity, the greater resistance of /a/ to reduction and deletion phenomena described in Section 2.1 would be partially confirmed, since our dataset only shows 6% deletion for /a/ (Table 4). However, we find 44% reduction for /a/, countering Silvestri's (2018) report for North-Western Calabrese, which says there is no centralization for this vowel.

Having considered the effects of the linguistic predictors, we can consider a final goal of this study: to establish an empirical link between reduction and deletion, as proposed in the literature (Section 2.1). The outcome is ambiguous. We expected the conditioning effects of the two phenomena to be similar, since the linguistic frame of the process should not change. The more similarities between the patterns, the more connected reduction and deletion would be. Comparing Tables 5 and 7 shows that the significant predictors are the same for the two phenomena. The hierarchy of constraints (ranking of factors by effect size and ranking of levels within each factor by rate of favouring the process) is quite similar. If we consider only linguistic factors, we can confirm that reduction and deletion are part of the same process. What is interesting is that social factors suggest otherwise. If apocope is the endpoint of the reduction process, it should be subject to the same social constraints as reduction. We would expect older speakers to exhibit less reduction, just as they exhibit less Reduction&Deletion, than younger speakers. However, Table 5 shows, this is not the case: age is not a significant factor in the alternation between Reduced vs. Deleted forms. Perhaps the second stage of the process—deletion—is too recent and not yet regular.

Finally, this study has the added bonus of shedding light on previous findings on apocope in Faetar. Faetar is a Francoprovençal isolate spoken in Apulia, a province not far from Calabria geographically and sharing many linguistic features including both apocope and centralization (De Blasi 2014; Marcato 2007; Loporcaro 2009). Nagy and Bill (1997) illustrate that Faetar is undergoing change in apocope and conclude that it is due to the influence of Italian stress patterns competing with Francoprovençal tonic-final phonotactics (cf. Francoprovençal tonic final syllable, e.g., *pan* [pa.n] 'bread' vs. Italian (ante-) penultimate tonic pattern, e.g., *pane* [pa.ne] 'bread'). Until now, no quantitative studies of apocope in Italian varieties in or near Apulia were available for comparison. The present report shows a similar pattern of phonological conditioning and age as a significant predictor, in a variety without a Francoprovençal substrate. Thus, it may be a more widespread Italian variable, rather than conflict between Francoprovençal and

Italian phonotactics, or perhaps both in tandem, that best account for the distribution reported for Faetar. We hope, in future research, to show that the OT analysis with “floating constraints” (Reynolds 1994) can apply to both these varieties. An important test of a theoretical construct developed for one data set is that it can equally account for another.

## 7. Conclusions

This paper has examined two stigmatized variable phonological processes associated with a regional variety. This study provides empirical evidence that Reduction and Deletion both exist in the RI spoken in southern Calabria. It is the first sociolinguistic analysis that applies variationist methods to establishing the relative effect of multiple linguistic and social factors on these processes.

While the literature suggests that, over time, reduction (vowel centralization) is leading to apocope (deletion of these same word-final vowels), we find an apparent-time effect only for the rate of Reduction&Deletion (vs. Full), but not for Reduction vs. Deletion, and we find an age effect only for speakers living in Calabria, not for their descendants living in Toronto and speaking Heritage Italian. We find different linguistic effects governing the heritage vs. the homeland variety. In spite of the reduced set of registers in which heritage speakers use their heritage language, we do not find a difference in the rate of use of the less-standard variants either between homeland and heritage or among heritage generations.

However, our models of the two processes provide empirical support for the connection between reduction and deletion: the same set of factors, ranked in virtually the same constraint hierarchies, govern the two processes. The distribution also supports several claims about environmental conditioning that had not previously been quantitatively tested in spontaneous speech.

Many analyses of heritage languages, when based on spontaneously-produced speech and multivariate analysis techniques, show strong similarities, or even identity, with homeland varieties (Nagy 2018). When it comes to Italian apocope and reduction, we see an extension of this trend, in terms of the rates of each variant of final atonic vowels: generation does not emerge as a significant predictor in any model. However, as with the one other variable with socioindexicality in the homeland variety that has been examined in this corpus, VOT of unstressed syllable onsets (Nagy and Kochetov 2013), we find small differences in the constraint hierarchies between heritage and homeland speakers, suggesting that socially-indexed aspects of the language weaken in successive heritage generations. The virtually complete lack of social conditioning for apocope and deletion add to this trend. However, we must keep in mind that, in spite of the stigma associated with southern Italian varieties, we were not able to establish social conditioning in the homeland variety either, with the exception of an effect of age.

**Author Contributions:** Conceptualization and methodology were developed by all authors. Data was coded by A.B. and A.C. and formal analysis conducted by N.N. The three authors collaboratively drafted and edited all sections of the paper, under N.N.’s supervision; funding acquisition, N.N. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by SSHRC, grants number 410-2009-2330 and 435-2016-1430.

**Institutional Review Board Statement:** The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Institutional Review Board of The University of Toronto (protocol code 24041, approved 30 April 2009).

**Informed Consent Statement:** Informed consent was obtained from all participants involved in the study.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author. The data are not publicly available due to confidentiality restrictions.

**Acknowledgments:** We are extremely grateful to the many students and research assistants who collected, transcribed, analyzed and interpreted the data used in this study and particularly to

Rachel Keir who co-authored the pilot investigation on which this study is based. They are listed at [http://ngn.artsci.utoronto.ca/HLVC/3\\_2\\_active\\_ra.php](http://ngn.artsci.utoronto.ca/HLVC/3_2_active_ra.php) and [http://ngn.artsci.utoronto.ca/HLVC/3\\_3\\_former\\_ra.php](http://ngn.artsci.utoronto.ca/HLVC/3_3_former_ra.php) (accessed on 6 July 2021). We thank the editors and reviewers for help clarifying our findings.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Notes

- <sup>1</sup> Speakers are identified by speakercode. “I” identifies the Italian HLVC corpus. The second character identifies the speaker’s generation (see Table 1), with “X” marking Homeland speakers. The third character identifies the speaker’s sex followed by the speaker’s age. The final character is a unique identifier.
- <sup>2</sup> Seven speakers live in Vibo Valentia and one in Zumpano (CS).
- <sup>3</sup> It might, alternatively, be age-grading, but then we expect effects tied to age divisions like in- vs. post-workforce age for a stylistically-important variable.

## References

- Abete, Giovanni. 2011. *I Processi di Dittongazione nei Dialetti dell’Italia Meridionale. Un Approccio Sperimentale*. Roma: Aracne.
- Albano Leoni, Federico. 2015. <Carmniell o’srngar>. Osservazioni sulla ortografia ingenua del napoletano e sulle sue possibili implicazioni fonetiche. In *Elaborazione Ortografica delle Varietà non Standard. Esperienze Spontanee in Italia e all’estero*. Edited by Silvia Dal Negro, Federica Guerini and Gabriele Iannacarro. Bergamo: Bergamo University Press/Edizioni Sestante, pp. 51–78.
- Albano Leoni, Federico, Francesco Cutugno, and Renata Savy. 1995. The Vowel System of Italian Connected Speech. *Proceeding of the XIIIth International Congress of Phonetic Sciences* 4: 396–99.
- Auer, Peter. 2005. Europe’s Sociolinguistic Unity, or: A Typology of European Dialect/Standard Constellations. In *Perspectives on Variation. Sociolinguistic, Historical, Comparative*. Edited by Nicole Delbecque, Johan van der Auwere and Dirk Geeraerts. Berlin and New York: de Gruyter, pp. 7–42.
- Beckman, Mary E. 1996. When a syllable is not a syllable? In *Phonological Structure and Language Processing, Cross-Linguistic Studies*. Edited by Takeshi Otake and Anne Cutler. Berlin: Mouton de Gruyter, pp. 95–123.
- Bell, Allan. 1984. Language Style as Audience Design. *Language in Society* 13: 145–204. [CrossRef]
- Berruto, Gaetano. 2005. Dialect/Standard Convergence, Mixing, and Models of Language Contact: The Case of Italy. In *Dialect Change. Convergence and Divergence in European Languages*. Edited by Peter Auer, Frans Hinskens and Paul Kerswill. Cambridge: Cambridge University Press, pp. 81–97.
- Berruto, Gaetano. 2006. *Sociolinguistica dell’italiano Contemporaneo*. Roma: Carocci.
- Berruto, Gaetano. 2007. *Fondamenti di Sociolinguistica*. Roma and Bari: Laterza.
- Bybee, Joan. 2003. *Phonology and Language Use*. Cambridge: Cambridge University Press.
- Cerruti, Massimo. 2009. *Strutture dell’italiano Regionale: Morfosintassi Di Una Varietà Diatopica in Prospettiva Sociolinguistica*. Bern: Peter Lang, vol. 7.
- Cerruti, Massimo. 2011. Regional Varieties of Italian in the Linguistic Repertoire. *International Journal of the Sociology of Language* 210: 9–28. [CrossRef]
- Cerruti, Massimo, Claudia Crocco, and Stefania Marzo, eds. 2017. *Towards a New Standard: Theoretical and Empirical Studies on the Restandardization of Italian*. Language and Social Life 6. Berlin and Boston: De Gruyter Mouton.
- Cristiano, Angela. 2019. *Fenomeni di Riduzione Fonetica nel Napoletano Parlato*. Napoli: Università degli studi di Napoli Federico II.
- Crocco, Claudia. 2017. Everyone Has an Accent. Standard Italian and Regional Pronunciation. *Towards a New Standard: Theoretical and Empirical Studies on the Restandardization of Italian*, 89–117. [CrossRef]
- De Blasi, Nicola. 2014. *Geografia e Storia dell’italiano Regionale*. Bologna: Il Mulino.
- Farnetani, Edda, and Maria Grazia Busà. 1999. Quantifying the Range of Vowel Reduction in Italian. Paper presented at the 14th International Congress of Phonetic Sciences, San Francisco, CA, USA, August 1–7; vol. 1, pp. 491–94.
- Fava, Elisabetta, and Emanuela Magno Caldognetto. 1976. Studio sperimentale delle caratteristiche elettroacustiche delle vocali toniche e atone in bisillabi italiani. In *Studi di Fonetica e Fonologia*. Edited by Raffaele Simone, Ugo Vignuzzi and Giulianella Ruggiero. Roma: Bulzoni, pp. 35–80.
- Grassi, Corrado, Alberto Sobrero, and Tullio Telmon. 1997. *Fondamenti Di Dialettologia Italiana*. Bari: Laterza.
- Hjelde, Arnstein, Bjørn Harald Kvitte, Linda Evenstad Emilsen, and Ragnar Arntzen. 2019. Processability Theory as a Tool in the Study of a Heritage Speaker of Norwegian. In *Widening Contexts for Processability Theory: Theories and Issues*. Edited by Lenzing Anke, Howard Nicolas and Jana Roos. Amsterdam and Philadelphia: John Benjamins Publishing Company, pp. 185–206.
- Johnson, Daniel E. 2009. Getting off the GoldVarb Standard: Introducing Rbrul for Mixed-Effects Variable Rule Analysis. *Language and Linguistics Compass* 3: 359–83. [CrossRef]
- Labov, William. 1984. Field Methods of the Project on Linguistic Change and Variation. In *Language in Use: Readings in Sociolinguistics*. Edited by John Baugh and Joel Scherzer. Englewood Cliffs: Prentice Hall, pp. 28–53.
- Local, John, and Gareth Walker. 2005. Methodological Imperatives for Investigating the Phonetic Organization and Phonological Structures of Spontaneous Speech. *Phonetica* 62: 120–30. [CrossRef] [PubMed]

- Loporcaro, Michele. 1988. *Grammatica storica del Dialetto di Altamura*. Pisa: Giardini.
- Loporcaro, Michele, Luciano Romito, Antonio Mendicino, and Tizia Turano. 1998. La neutralizzazione delle vocali finali in crotonese: Un esperimento percettivo. In *Unità Fonetiche e Fonologiche: Produzione e Percezione, Atti delle 8 Giornate di Studio del Gruppo di Fonetica Sperimentale (A.I.A.)*, Pisa, 17–19 December 1997 (Collana degli Atti dell'Associazione Italiana di Acustica). Edited by Pier Marco Bertinetto and Lorenzo Cioni. Pisa: Scuola Normale Superiore, pp. 91–100.
- Loporcaro, Michele. 2009. *Profilo Linguistico dei Dialetti Italiani*. Rome and Bari: Laterza.
- Marcato, Carla. 2007. *Dialetto, Dialetti e Italiano*. Bologna: Il Mulino.
- Marotta, Giovanna, and Patrizia Soriano. 1998. Vocali contigue a confine di parola. In *Unità fonetiche e fonologiche: Produzione e percezione. Atti delle VIII Giornate di Studio del Gruppo di Fonetica Sperimentale, Scuola Normale Superiore di Pisa, Pisa, 18–20 December 1997*. Edited by Pier Marco Bertinetto and Lorenzo Cioni. Pisa: Stamperia SNS, pp. 101–13.
- Maturi, Pietro, and Edoardo Mastantuoni. 2012. La variabile schwa nei dialetti campani. In *La Variazione nell'italiano e nella sua Storia. Varietà e Varianti Linguistiche e Testuali*. Edited by Patricia Bianchi, Nicola De Blasi, Chiara De Caprio and Francesco Montuori II. Firenze: Franco Cesati, pp. 869–78.
- Nagy, Naomi. 2009. Heritage Language Variation and Change in Toronto. Available online: <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.591.912&rep=rep1&type=pdf> (accessed on 6 July 2021).
- Nagy, Naomi. 2011. A Multilingual corpus to explore variation in language contact situations. *Rassegna Italiana di Linguistica Applicata* 43: 65–84.
- Nagy, Naomi. 2015. A Sociolinguistic View of Null Subjects and VOT in Toronto Heritage Languages. *Lingua* 164: 309–27. [CrossRef]
- Nagy, Naomi. 2018. Linguistic Attitudes and Contact Effects in Toronto's Heritage Languages: A Variationist Sociolinguistic Investigation. *International Journal of Bilingualism* 22: 429–46. [CrossRef]
- Nagy, Naomi, and Alexei Kochetov. 2013. VOT across the Generations: A Cross-Linguistic Study of Contact-Induced Change. In *Multilingualism and Language Contact in Urban Areas: Acquisition-Development-Teaching-Communication*. Edited by Peter Siemund, Ingrid Gogolin, Monika Schulz and Julia Davydova. Amsterdam: John Benjamins, pp. 19–38.
- Nagy, Naomi, and Reynolds Bill. 1997. Optimality Theory and variable word-final deletion in Faetar. *Language Variation and Change* 9: 37–55. [CrossRef]
- Nespor, Marina. 1990. Vowel Deletion in Italian: The Organization of the Phonological Component. *The Linguistic Review* 7: 375–98. [CrossRef]
- Nodari, Rosalba. 2017. Autobiografie linguistiche di adolescenti calabresi. *Studi Italiani di Linguistica Teorica e Applicata* XLVI: 519–40.
- Otheguy, Ricardo, Ana-Celia Zentella, and Daniel Livert. 2007. Language and Dialect Contact in Spanish in New York: Toward the Formation of a Speech Community. *Language* 83: 770–802. [CrossRef]
- Pellegrini, Giovan Battista. 1977. *Carta dei dialetti d'Italia*. Pisa: Pacini.
- Poggi Salani, Teresa. 2010. Italiano regionale. *Enciclopedia dell'italiano* I: 726–29.
- Poplack, Shana, and Marjory Meechan. 1998. How Languages Fit Together in Codemixing. *International Journal of Bilingualism* 2: 127–38. [CrossRef]
- Radtke, Edgar. 1997. *I Dialetti della Campania*. Roma: Il Calamo.
- Reynolds, Bill. 1994. Variation and Phonological Theory. Doctoral dissertation, University of Pennsylvania, Philadelphia, PA, USA.
- Rohlf, Gerard. 1966. *Grammatica Storica della Lingua Italiana e dei suoi Dialetti*. Torino: Einaudi.
- Romano, Antonio. 2020. Vowel Reduction and Deletion in Apulian and Lucanian Dialects with Reference to Speech Rhythm. *Italian Journal of Linguistics* 32: 85–102.
- Romano, Antonio, and Francesca Manco. 2003. Incidenza di Fenomeni di Riduzione Vocalica nel Parlato Spontaneo a Bari e a Lecce. In *Il Parlato Italiano. Atti del Convegno Naz., Napoli, 13–15 Febbraio 2000*. Edited by Federico Albano Leoni, Francesco Cutugno, Massimo Pettorino and Renata Savy. CD-rom. Napoli: D'Auria.
- Romito, Luciano, Tiziana Turano, Michele Loporcaro, and Antonio Mendicino. 1997. Micro e macrofenomeni di centralizzazione nella variazione diafasica: Rilevanza dei dati fonetico-acustici per il quadro dialettologico calabrese. In *Fonetica e Fonologia Degli stili dell'italiano Parlato. Atti delle VII Giornate di Studio del GFS, Napoli 14–15.XI.1996*. Edited by Francesco Cutugno. Roma: Esagrafica, pp. 157–75.
- Russo, Michela, and William John Barry. 2002. Gradi di elisione dello schwa finale nelle varietà di Ischia e Pozzuoli: Un'analisi strumentale e implicazioni ritmiche. In *Atti di XII Giornate di Studio del Gruppo di Fonetica Sperimentale (GFS), Macerata*. Edited by Agostino Regnicoli. Roma: Il Calamo, pp. 13–15.
- Savy, Renata. 1999a. *Riduzioni Foniche nel Parlato Spontaneo: Il Ruolo della Morfologia Nell'interpretazione del Messaggio e nella Comunicazione*. Roma: Università da Roma Tre.
- Savy, Renata. 1999b. Riduzioni foniche nella morfologia del sintagma nominale nel parlato spontaneo: Indagine quantitativa e aspetti strutturali. In *Fonologia e Morfologia dell'italiano e dei Dialetti d'Italia, Atti del XXXI Congresso della SLI*. Edited by Paola Benincà, Alberto Mioni and Laura Vanelli. Roma: Bulzoni, pp. 201–21.
- Savy, Renata, and Francesco Cutugno. 1997. Ipoarticolazione, Riduzione Vocalica, Centralizzazione: Come Interagiscono Nella Variazione Diafasica? *Fonetica e Fonologia Degli Stili dell'italiano Parlato* XI: 177–94.
- Silva-Corvalán, Carmen. 1994. *Language Contact and Change*. New York: Oxford University Press.
- Silvestri, Giuseppina. 2018. Word-Final Schwa Licensed by Prosody and Syntax: Evidence from Southern Italian Dialects. *Research in Generative Grammar* 40. Available online: <http://hdl.handle.net/11707/7779> (accessed on 11 July 2021).

- Sornicola, Rosanna. 2002. La variazione dialettale nell'area costiera napoletana. Il progetto di un archivio di testi dialettali parlati. *Bollettino Linguistico Campano* 1: 131–55.
- Telmon, Tullio. 1993. Varietà regionali. In *Introduzione all'italiano Contemporaneo. La Variazione e gli usi*. Edited by Alberto Sobrero. Rome and Bari: Laterza, pp. 93–149.
- Trumper, John, Luciano Romito, and Antonio Mendicino. 2001. Fonetica, fonologia e dialettologia: Un rapporto complesso. In *La Dialettologia oggi fra Tradizione e Nuove Metodologie*. Edited by Alberto Zamboni, Alberto Del Puente and Maria Teresa Vigolo. Pisa: Edizioni ETS, pp. 93–116.
- Van Bergem, Dick R. 1994. Perceptual and acoustic aspects of lexical vowel reduction, a sound change in progress. *Speech Communication* 16: 329–58. [[CrossRef](#)]
- Vayra, Mario. 1991. Un'interfaccia Tra Fonetica e Fonologia: Declination Intonativa e Altre 'Declinazioni' Nel Parlato. In *L'interfaccia Tra Fonologia e Fonetica. Atti Del Convegno Di Padova*. Edited by Magno Caldognetto and Paola Benincà. Padova: Unipress, vol. XII, pp. 37–154.
- Vietti, Alessandro. 2019. Phonological Variation and Change in Italian. In *Oxford Research Encyclopedia of Linguistics*. Oxford: Oxford University Press. [[CrossRef](#)]
- Voghera, Miriam. 2001. Teorie linguistiche e dati di parlato. In *Dati Empirici e Teorie Linguistiche, Atti del XXXIII Congresso Internazionale di Studi della Società di Linguistica Italiana SLI*. Edited by Federico Albano Leoni, Eleonora Stenta Krosbakken, Rosanna Sornicola and Carolina Stromboli. Roma: Bulzoni, pp. 75–95.
- Wittenburg, Peter, Hennie Brugman, Albert Russel, Alex Klassmann, and Han Sloetjes. 2006. Elan: A Professional Framework for Multimodality Research. Paper presented at the 5th International Conference on Language Resources and Evaluation, Genoa, Italy, May 22–28.



Article

# Sociolinguistic Awareness in Galician Bilinguals: Evidence from an Accent Identification Task

Gisela Tomé Lourido <sup>1,\*</sup> and Bronwen G. Evans <sup>2</sup>

<sup>1</sup> Department of Linguistics and Phonetics, University of Leeds, Leeds LS2 9JT, UK

<sup>2</sup> Department of Speech, Hearing & Phonetic Sciences, University College London, London WC1N 1PF, UK; bronwen.evans@ucl.ac.uk

\* Correspondence: g.tomelourido@leeds.ac.uk

**Abstract:** The inclusion of European minority languages in public spaces such as education, administration and the media has led to the emergence of a new profile of speakers, “new speakers”, who typically acquire a minority language through education, but vary in terms of their language experience and use. The present study investigated whether a distinctive variety spoken by Galician new speakers (neofalantes) has emerged in the community and whether listeners’ language background influences accent identification abilities and patterns. Galician-Spanish bilingual listeners completed an accent identification task and were asked to comment on factors influencing their decision. Results demonstrated that all listeners could identify Galician-dominant better than Spanish-dominant bilinguals but could not identify neofalantes. Neofalantes were categorised as both Spanish- and Galician-dominant, supporting the idea that neofalantes have a hybrid variety. This finding suggests that listeners have a gradient representation of language background variation, with Galician-like and Spanish-like accents functioning as anchors and the neofalantes’ accent situated somewhere in the middle. Identification accuracy was similar for all listeners but neofalantes showed heightened sensitivity to the Galician-dominant variety, suggesting that evaluation of sociophonetic features depends on the listener’s language and social background. These findings contribute to our understanding of sociolinguistic awareness in bilingual contexts.



**Citation:** Tomé Lourido, Gisela, and Bronwen G. Evans. 2021.

Sociolinguistic Awareness in Galician Bilinguals: Evidence from an Accent Identification Task. *Languages* 6: 53. <https://doi.org/10.3390/languages6010053>

Received: 25 January 2021

Accepted: 8 March 2021

Published: 18 March 2021

**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Keywords:** new speakers; accent identification; sociolinguistic awareness; bilingual speech processing; Galician phonetics; minority languages

## 1. Introduction

When we receive a phone call from an unknown number, if it is a person we know we can often recognise their voice even if we only hear the word ‘hello’. When we do not know the person, we are still able to infer some of their characteristics, e.g., gender, geographical origin, language background, based on their speech. Extensive research in sociolinguistics, phonetics and speech perception over the last few decades has confirmed our intuition that listeners are sensitive to accent variation (e.g., Giles 1970; Lambert et al. 1960; Preston 1989) and has provided evidence that we use accent variation to understand speech (e.g., Niedzielski 1999; Strand 1999; Strand and Johnson 1996). However, the process of how listeners extract indexical information from the speech signal and use it in speech processing is not yet fully understood.

An interesting context to investigate how a set of phonetic features may become associated with a particular group of speakers is the emergence of a new accent in a community, i.e., how linguistic features become ‘enregistered’ as a variety (Agha 2003; Johnstone et al. 2006; Silverstein 2003). In minority language communities in Europe, a new profile of speakers has emerged as a result of the inclusion of minority languages in public spaces such as education, administration and the media. These changes in the sociolinguistic landscape have also led to changes in the symbolic value and transmission of minority languages (Ramallo 2013), with some speakers learning them through schooling



or immersion programmes for the first time. These speakers are known as ‘new speakers’ (O’Rourke et al. 2015). Although the bilingual experience of new speakers varies widely in terms of language exposure, use and proficiency, this new profile of speakers typically has little or no home exposure to the minority language, and instead typically acquire the language through education (O’Rourke et al. 2015).

In the bilingual community of Galicia, in the north west of the Iberian peninsula, the ‘new speaker’ (neofalante, in Galician) label is used within the community to designate early bilinguals who learn Spanish at home, but switch language dominance to Galician in adolescence for ideological reasons (O’Rourke and Ramallo 2011, 2015; Ramallo 2013; Tomé Lourido and Evans 2019). Unlike new speakers in other bilingual contexts, neofalantes usually have early exposure to and high competence in Galician, which do not necessarily come exclusively from schooling; but also from acquiring the language from the environment (Ramallo and O’Rourke 2014), e.g., through grandparents, the wider community. Previous research has investigated the consequences of the switch in language dominance on their speech production and found that neofalantes pattern with Galician-dominant speakers in the production of certain phonetic variables, but with Spanish-dominant speakers in the production of others, exhibiting a hybrid variety (Tomé Lourido and Evans 2019). The current study aims to investigate whether the neofalantes’ accent is sufficiently distinct for listeners in the community to recognise it, and therefore emerging as a variety, and whether the listener’s language background influences the patterns of accent identification. To our knowledge, this is the first experimental study investigating whether the variety used by new speakers can be identified by listeners in the community and contributes to our understanding of sociolinguistic awareness in bilingual contexts.

## 2. Literature Review

### 2.1. Identifying Accents and Talkers

Language attitude studies have investigated how listeners use the indexical information embedded in the speech signal to draw inferences about speakers’ regional or social background (Giles 1970, 1971a, 1971b; Giles and Powesland 1975). Work in perceptual dialectology has provided further evidence that listeners are sensitive to regional variation by examining naive listeners’ perception of dialect boundaries. In a seminal study, Preston (1986, 1989) gave American English speakers maps of the United States and asked them to label the places where they judged people to speak differently. This technique also enabled elicitation of attitudes towards the selected accents (see also Preston 1996, 1999). Crucially, this work showed that, in general terms, listeners agree on the attitudes and stereotypes associated with the accents. However, more recent research has revealed that the social meaning of accent features emerges in the context of language use: the particular accent features listeners tune into and how these are evaluated depend on other perceived characteristics of the speaker (Campbell-Kibler 2011; Levon 2014; Montgomery and Moore 2018; Pharaoh et al. 2014) and the background of the listener (Jaeger and Weatherholtz 2016). Other studies have shown that listeners can group speakers according to regional accent but that this is affected by listeners’ own accent background and their experience with a given accent. In a series of studies, Clopper and Pisoni (2004, 2006) presented American listeners with sentences read by talkers from six different American English dialects in a forced-choice categorisation task and found that listeners were able to distinguish broad dialect categories (New England, South and South Midland and North Midland and West). Performance in these tasks was modulated by participants’ background: listeners who had lived in different areas performed better than those who had only lived in one area and, additionally, listeners who lived in a particular region performed better with the accent from that region. These results were taken to mean that greater exposure to linguistic variation and specific experience with one variety benefit accent categorisation. Similar results have been found using free classification tasks (Clopper 2008; Clopper and Pisoni 2007).

Listeners are not only sensitive to variation that signals geographical origin but also other social factors, including ethnicity. Using a matched-guise technique, [Purnell et al. \(1999\)](#) showed that landlords discriminated against prospective tenants based on the inferences they made about the speaker's ethnicity from hearing their accent on the phone. Baugh referred to this as 'linguistic profiling', a process "based upon auditory cues that may be used to identify an individual or individuals as belonging to a linguistic subgroup within a given speech community, including a racial subgroup" ([Baugh 2000](#), p. 363). There is extensive evidence that listeners are sensitive to variation and use it to evaluate speakers, but less is known about the levels of processing involved in the extraction and use of indexical information. Using a neuroimaging technique, magnetoencephalography (MEG), [Scharinger et al. \(2011\)](#) presented listeners with the sentence-initial 'hello' tokens from [Purnell et al. \(1999\)](#) to investigate when the change in accents was detected. Results from the mismatch negativity (MMN) response to accent changes showed that the extraction of accent features occurs very rapidly and is pre-attentive, categorical and speaker-independent. The authors propose that, given that the stimuli presented were acoustically variable, accent extraction involves a process of abstraction by which low-level acoustic information is mapped to a memory trace associated with a phonetic feature which is linked to a social category, in this case, accent background. Another important finding from this study is that accent information appears to be processed in the same way as speaker voice information. A recent study has provided further evidence that indexical information is processed at a relatively early stage. Although research that presented listeners with synthetic speech had suggested that non-linguistic information is ignored at early stages of processing, [Tuninetti et al. \(2017\)](#) found that when presented with natural speech, listeners are sensitive to indexical information (gender and regional background) at an unattended low level of processing.

An interesting question that emerges from this research is concerned with when in development listeners start to acquire the sociolinguistic competence that enables them to identify the regional and social background of talkers by associating a set of phonetic features with a social category. Studies using free classification tasks have shown that non-native listeners ([Clopper and Bradlow 2009](#)), and children, some as early as the age of 4–5 years old ([Jones et al. 2017](#)), are also able to group speakers into broader accent categories, although they are less accurate than adult native listeners. These results suggest that indexical and phonological categories are acquired together in first (L1) and second language (L2) acquisition ([Clopper and Bradlow 2009](#)).

One category that listeners learn to discriminate very early on is that of their native language. [Nazzi et al. \(2000\)](#) used a head-turn preference procedure to show that 5-month-old American infants could always discriminate between languages either when their native language was one of the two languages presented or when the two foreign languages belonged to different rhythmic classes (e.g., Japanese vs. Italian), but not when the two foreign languages belonged to the same rhythmic class (e.g., Italian vs. Spanish). In a similar study, [Butler et al. \(2011\)](#) showed that 5-month-old infants were able to discriminate between their native (South-West English) accent and an unfamiliar regional accent (Welsh English), but were unable to differentiate two unfamiliar regional accents (Welsh English and Scottish English).

Indeed, other research suggests that the ability to discriminate unfamiliar accents does not develop until later in life. [Girard et al. \(2008\)](#) showed that 5–6-year-old French-speaking children distinguished their own accent from a foreign accent, but could still not discriminate between different regional varieties of French. These findings indicate that, at this age, young children have not yet developed fine-grained perceptual representations for regional accents, at least based on the varieties tested here. [Flocchia et al. \(2009\)](#) replicated this result in a similar study with British children and suggested that the acoustic distance between the accents could have played a role in children's discrimination patterns. They demonstrated that consonant differences between the native and the foreign accent were larger and interpreted this finding to mean that foreign accents introduce greater distortions to the signal than regional accents making the accent itself more distinctive. Similar results

were found for American children, aged 5–6 years old, who were able to discriminate their native accent from an L2 accent (Indian English, produced by speakers who acquired English as an L2), but who were unable to discriminate between their native and a regional accent, or a regional vs. L2 accent (Wagner et al. 2014). Based on these findings, it has been hypothesized that children have a gradient representation of dialect variation with representations organised relative to the native accent, such that those a greater distance apart are easier to discriminate (Wagner et al. 2014).

Much less research has examined accent identification in the context of bilingual communities. Evans and Lourido (2019) replicated Wagner et al. (2014)'s study with monolingual children in London, U.K., but also showed that bilingual children were able to discriminate talkers in all three conditions (native vs. regional, regional vs. L2 and native vs. L2), suggesting that early experience with variation benefits identification of talkers from different language backgrounds. Arguably, bilingual children had more exposure to variation in a community where that variation is useful in identifying talkers and navigating relationships (Evans and Lourido 2019, p. 156), and this most likely led to an earlier development of sociolinguistic awareness in comparison to monolingual peers.

Studies with adult bilingual listeners additionally show that identification is affected by listeners' identity as well as experience. Tan (2012) investigated whether Singaporean bilingual listeners were able to identify the ethnicity of English-Chinese, English-Malay and English-Tamil bilingual speakers. The results showed that listeners identified Chinese speakers more accurately than Malay and Indian speakers, in this order. The author argues that the findings could be explained by the amount of exposure listeners had to the different accents; Singaporean-Chinese speakers make up most of the population and, therefore, listeners in the community are likely to hear this variety more frequently. There was also a significant effect of age; younger Singaporeans were less accurate than older Singaporeans. The author suggests that the younger group may have a more national-based, rather than ethnic-based identity, compared to the older group and their performance may reflect this link between their own identity and perception. In a minority language context, Mayr et al. (2020) showed that both Welsh-English bilinguals and English monolingual listeners from Wales were able to identify whether someone can speak Welsh on the basis of their accent in English above chance level, although performance was lower than in similar studies with L2 speakers. Listeners performed better with talkers from the same area of Wales as them, but there was no difference between bilingual and monolingual listeners (Mayr et al. 2020, p. 752).

In the context of the current study, in which all listener groups are bilingual in Galician and Spanish, it is possible that, differences in language background will lead to differences in accent identification patterns. Given that the degree of distinctiveness will likely be more similar to that of regional than foreign accents as regardless of language dominance, all speakers will likely have a Galician accent (e.g., in contrast with L2 Galician speakers from a different part of Spain), how 'Galician' a speaker sounds will vary as a function of their language dominance (Amengual and Chamorro 2015; Tomé Lourido and Evans 2019; Aguete Cajiao 2019), whether they come from an urban or rural environment (Mayr et al. 2019; Tomé Lourido and Evans 2019; Regueira and Fernández Rei 2020) and other factors. As well as greater exposure with a given variety leading to better identification, the participants' social background and aims may also influence identification patterns.

How might listeners store and consequently access indexical information during speech processing to enable them to group talkers into different social categories? As mentioned above, recent work has proposed that accent information is processed in the same way as speaker voice information (Scharinger et al. 2011). Such work has highlighted the likely contribution of episodic memory in models of speech processing (e.g., Nygaard and Pisoni 1998). Episodic models of lexical access propose that phonetic variation in the speech signal, such as indexical or talker information, is not discarded in speech perception, but instead is retained and stored in memory (Docherty and Foulkes 2014; Goldinger 1998). Indeed, it has been shown that listeners can use fine-grained phonetic

information, such as VOT, to identify talkers (Allen and Miller 2004). Additionally, work on talker identification has consistently shown a Language Familiarity Effect (LFE), i.e., listeners are better at identifying talkers in their native language (e.g., Fleming et al. 2014; Goggin et al. 1991; Perrachione et al. 2011; Thompson 1987; Levi 2019). For example, Goggin et al. (1991) showed that monolingual English listeners were better at identifying English voices than German ones, and German listeners exhibited the opposite pattern. Similarly, English monolinguals were better at identifying English voices when compared to Spanish voices, with intermediate performance with Spanish-accented voices, but the pattern did not hold for English-Spanish bilinguals. One possible interpretation of these findings is that language familiarity is beneficial for voice recognition. However, whether this effect is related to language comprehension or familiarity with the phonological structure of the language is unclear.

Perrachione et al. (2011) examined whether knowledge of phonology played a role in voice recognition. In this experiment, dyslexic listeners, who have impaired phonological processing, identified voices in English (native language) and Chinese (unfamiliar language). Whilst the monolingual English control group were more accurate with the English voices, displaying a language familiarity effect, dyslexic listeners were no better able to identify English than Chinese talkers. These results led the authors to suggest that phonological representations are important for recognising speakers and that the process of voice recognition functions by comparing the segments in the input voice with the listener's own phonological representations. Thus, voice recognition is more difficult when listeners cannot relate the speaker's segments to their own representations because they are either missing (when they hear an unfamiliar language) or impaired (in the case of dyslexic listeners). On the other hand, Fleming et al. (2014) have argued that as the LFE is already apparent in 7–8-month-old infants (Johnson et al. 2011; Nazzi et al. 2000), who cannot understand speech, the effect could also be driven by experience with native phonological categories. Fleming et al. (2014) presented English and Chinese adult listeners with unintelligible time-reversed sentences in English and Mandarin, which they argued preserved phonological information but meant that the speech was unintelligible. Both listener groups rated pairs of native-language speakers as more dissimilar than foreign-language speakers, suggesting that the LFE is based not on comprehension, but on familiarity with the native language phonological system. With the aim of elucidating the underlying cause of the LFE, Johnson et al. (2018) claim that relative familiarity with a variety, i.e., the frequency of encountering talkers from that linguistic background, is not enough to account for the LFE, which is instead driven by 'attunement' to the underlying phonological structure. They tested this hypothesis by asking English listeners to identify talkers with a familiar and unfamiliar variety of English (Australian and North American English). They found no differences in performance between the two varieties, which supports the idea that familiarity alone does not account for the LFE. The authors argue that Australian and North American English share the same underlying abstract phonology and propose that it is the listeners' 'attunement' to the phonology that drives this effect. However, they also point out that it is not clear whether this would be the case for other varieties differing in their phonological structure, e.g., syllable structure, rhythm and that further research is needed to 'map the boundaries of phonological attunement' (Johnson et al. 2018, p. 643).

In sum, although the ability to identify accents develops relatively late in life and at different rates in monolingual and bilingual communities, listeners use the indexical information embedded in the speech signal to draw inferences about speakers' regional, social and language background. Additionally, listener's ability to categorise talkers is likely affected by their own language background, experience and possibly even attitude towards a given variety. Finally, the ability to identify accents may function in a similar way to voice identification with both familiarity and 'attunement' to the phonological system playing a role.

## 2.2. The Neofalantes' Accent as an Emerging Variety

New speakers in minority language communities have been defined as “individuals with little or no home or community exposure to a minority language but who instead acquire it through immersion or bilingual educational programs, revitalization projects or as adult language learners” (O'Rourke et al. 2015, p. 1). They have been documented in most minority language communities in Europe: Ireland (O'Rourke and Ramallo 2010; Walsh and O'Rourke 2014), Wales (Robert 2009), Scotland (McLeod and O'Rourke 2015; Nance et al. 2016; O'Rourke and Walsh 2015), Isle of Man (Ó hÍfearnáin 2015), Provence (Costa 2015), Brittany (Hornsby 2005, 2009, 2015), Corsica (Jaffe 2015), Galicia (O'Rourke and Ramallo 2010, 2013a, 2013b, 2015; Ramallo 2013; Ramallo and O'Rourke 2014; Tomé Lourido and Evans 2015, 2017, 2019; Agüete Cajiao 2019; Regueira and Fernández Rei 2020), Catalonia (Pujolar and Puigdevall 2015; Woolard 2011) and the Basque Country (Ortega et al. 2014; Ortega et al. 2015). Though this label is particularly useful in examining their sociolinguistic ideologies and practices, it is important to understand that they are a heterogeneous group from the point of view of language acquisition, ranging from early bilinguals with great exposure to the minority language to L2 learners with varying degrees of proficiency.

There is limited experimental research investigating the phonetics and phonology of new speakers of minority language communities. Nance (2013, 2015) and Nance et al. (2016) investigated the speech of Gaelic speakers in Scotland. Nance (2015) compared the speech of young adults attending Gaelic-medium secondary schools in Glasgow, an area with low numbers of Gaelic speakers, young adults attending Gaelic-medium secondary schools in the Isle of Lewis, an area with the densest concentration of Gaelic speakers and a group of older adults from Lewis who were considered ‘traditional speakers’. Young speakers from Glasgow differed from both young and older speakers on Lewis in the three phonetic variables investigated, the high back vowel /u/, the lateral system and intonation, suggesting that the new speakers' variety is different from that of previous generations. However, when comparing the production of word-final rhotics by highly proficient urban adult new speakers and ‘traditional speakers’, Nance et al. (2016) found that some new speakers distinguished traditional Gaelic rhotic categories, but others did not. The variation in the new speaker group was not only accounted for by L1 interference, but also how they constructed their identity as Gaelic speakers.

Nance (2015, p. 556) states that the ‘new speaker’ label is not used by New Gaelic speakers themselves, but is instead an analytical label which has emerged from the minority language revitalisation literature. However, this is not the case in all communities. In Galicia, a bilingual community situated in the north west of the Iberian Peninsula, the new speakers' group has become socially salient within certain spheres of Galician society, and the ‘neofalante’ label has been used beyond academia to designate the social group (O'Rourke and Ramallo 2011, 2015; Ramallo 2013; Tomé Lourido and Evans 2019) such that it is sometimes used as a self-defining category by neofalantes (O'Rourke et al. 2015, p. 13). For example, there is a Twitter account named ‘O neofalante’, ‘The new speaker’ (Neofalante 2021). Most Galician neofalantes are bilinguals who learn Spanish at home, but have early exposure to Galician and high competence in both languages. O'Rourke and Ramallo describe neofalantes as “individuals for whom Spanish was their language of primary socialization, but who at some stage in their lives (usually early to late-adolescence) have adopted Galician language practices and on occasions displaced Spanish all together” (O'Rourke and Ramallo 2015, p. 148, see also O'Rourke and Ramallo 2010, 2011, 2013a, 2013b; Ramallo 2010, 2013; Ramallo and O'Rourke 2014). O'Rourke and Ramallo (2015) and Ramallo (2010) suggest that neofalantes' linguistic behaviour can contribute to the transformation of the sociolinguistic reality and characterise these speakers as proponents of social change, arguing for ‘neofalantismo’ as a social movement, with neofalantes an active minority, one in which “individuals or groups [ . . . ] through their behaviour attempt to influence both the attitudes and practices of the majority and in doing so, bring about social change” (O'Rourke and Ramallo, p. 151).

Impressionistic descriptions of neofalantes' speech have proposed that they use a Spanish-accented variety of Galician (Freixeiro Mato 2014; González González 2008; Ramallo 2010), which has been referred to as 'New Urban Galician' (Novo galego urbano, Dubert García 2002; González González 2008; Regueira 1999a; Vidal Figueroa 1997). Tomé Lourido and Evans (2019) were the first to provide a detailed acoustic description of the variety of Galician used by neofalantes and also to examine potential differences in their perception of Galician with respect to other bilingual groups. Neofalantes in this study were early bilinguals who changed from being dominant in Spanish to speaking Galician almost exclusively in adolescence for ideological, political or socio-cultural reasons. A series of studies examined three variables which differ in Galician and Spanish: Galician mid-vowel contrasts / $\epsilon$  e/ and / $\text{o}$  o/, which are not contrastive in Spanish; the Galician contrast sibilant fricative contrast / $\text{s}$   $\text{ʃ}$ /, where Spanish only has / $\text{s}$ /; and the reduction of word-final vowels, a Galician-specific feature. Neofalantes were compared to two early bilingual groups of Galician-dominant and Spanish-dominant speakers. For vowels, the perception tasks revealed that neofalantes' performance on a mid-vowel identification task was not different from that of Spanish-dominants and was poorer than that of Galician-dominant listeners. For the fricative contrast, though the three groups had a categorical contrast between the two sibilants, Galician-dominants had an earlier boundary than both neofalantes and Spanish-dominant groups. In production, neofalantes also patterned with Spanish-dominant speakers in their realisation of mid vowels, neutralising the contrast, and sibilant fricatives, producing a smaller contrast than that of Galician-dominants. However, they patterned with Galician-dominants in the production of reduced final vowels, exhibiting a hybrid variety made up of a combination of traditional Galician and Spanish features.

What is yet to be established is whether Galician listeners can identify the neofalantes' accent as a distinctive variety in the community, i.e., whether a particular set of linguistic features have become associated with the label. Agha (2003, p. 231) proposed the term 'enregisterment' to describe the "processes through which a linguistic repertoire becomes differentiable within a language as a socially recognized register of forms" (see also Silverstein 2003). Since then, this term has been also used to describe the emergence of new accents. For example, Johnstone et al. (2006) and Johnstone and Kiesling (2008) investigated how a set of linguistic features which were not noticed by listeners at first, became linked to socio-economic class, then associated with a region and 'enregistered' as a dialect called 'Pittsburghese', spoken in the United States. In this case, the linguistic features associated with 'Pittsburghese' were highly enregistered, as they were overtly linked to specific sociolinguistic spaces and discussed in metalinguistic commentary. Although Tomé Lourido and Evans (2019) found no evidence that neofalantes produced phonetic features which were distinctively different from those of Galician- and Spanish- dominant bilinguals, it is possible that listeners in the community use other features not measured in that study to identify the neofalantes variety. The current study sets out to investigate this question using an accent identification task.

### 2.3. The Current Study

The study aims to investigate whether a distinctive variety spoken by Galician neofalantes has emerged in the community and whether listeners' language background influences accent identification abilities and patterns. To address these questions, Galician-Spanish bilingual listeners completed an accent identification task and were asked to comment on factors influencing their decision.

Based on the research reviewed, we hypothesise that all Galician listeners will be able to categorise talkers from a Galician-dominant and Spanish-dominant background. A question that remains though, is whether listeners are able to recognise the neofalantes' accent. In the study, listeners heard sentences produced by bilingual speakers belonging to three groups (neofalantes, Galician-dominant and Spanish-dominant speakers) and categorised them according to their language background to address two research questions:

1. Are neofalantes' shifts in production sufficient for listeners in the community to identify their accent?
2. Does identification ability depend on listeners' language background?

If listeners are able to recognise the neofalantes' accent, this would indicate that it has become enregistered as a variety, one that has become associated with a set of linguistic features and is recognisable as a distinctive variety in the community. Nevertheless, if listeners are not able to link the neofalantes' accent with the sociolinguistic label, whether they classify neofalantes as Spanish-dominant or Galician-dominant speakers would be informative of whether neofalantes' speech production patterns have changed after the language dominance switch. Language ability, language familiarity and attunement to the phonological system have been shown to be beneficial for talker identification (Fleming et al. 2014; Goggin et al. 1991; Johnson et al. 2018; Levi 2019; Perrachione et al. 2011; Thompson 1987) and experience with a particular variety appears to enhance the accuracy of identification of that variety (Clopper and Pisoni 2004, 2006). If accent categorisation ability relies on similar mechanisms to talker identification skills, it might be influenced by similar factors. It is unclear whether Galician- and Spanish-dominant varieties would be considered to have a similar or different underlying phonology in the 'phonological attunement' account and therefore it is difficult to use this to inform the predictions. However, an effect of language ability, or more specifically more robust phonological and phonetic representations of the language, would predict an advantage in accent identification for Galician-dominant listeners. In contrast, a LFE would predict similar performance for all listener groups, as they live in a bilingual environment where they listen to both Galician and Spanish on a daily basis.

### 3. Methods

#### 3.1. Participants

This study set out to test the wider community and therefore, the sample is formed of a pool of varied participants from different backgrounds and professions. A total of 162 participants took part in the online task; 20 participants were excluded because they did not meet the criteria. The remaining 142 participants were raised in Galicia, had not lived anywhere else for more than seven years and were bilingual in Galician and Spanish. Their age ranged between 18–54 years old (median = 27). After the experiment, they completed the language background questionnaire used in Tomé Lourido and Evans (2019). The questionnaire was used to classify participants into the three groups of interest, following the criteria in Tomé Lourido and Evans (2019):

- Neofalantes: raised predominantly in Spanish (i.e., their parent(s) used to speak to them in Spanish), but decided to adopt Galician as their dominant language in adolescence (14–20 years old, median = 16) for ideological or cultural reasons. Since this switch, they have mainly spoken Galician.
- Galician-dominant bilinguals: raised predominantly in Galician (i.e., their parent(s) spoke Galician to them) and have always spoken mainly Galician.
- Spanish-dominant bilinguals: raised predominantly in Spanish (i.e., their parent(s) spoke Spanish to them) and have always spoken mainly Spanish.

This resulted in 13 neofalantes (6 female, 7 male), 58 Galician-dominants (34 female, 24 male) and 61 Spanish-dominants (34 female, 24 male). The remaining 10 participants did not belong to any of these three groups, but were included in the first set of analyses, as these were focussed on whether the three groups of speakers were correctly identified, regardless of listeners' language background. The second set of analyses examined specifically whether listeners' language background played a role in identification, and therefore those 10 participants were excluded. Two pilot participants completed the experiment before data collection took place; their data were not included in any of the analyses. None of the subjects reported any speech, hearing or language disorders at the time of testing.

### 3.2. Stimuli

The stimuli consisted of the first sentence of ‘The north wind and the sun’ passage in Galician: *O vento do norte e mais o sol porfiaban sobre cal deles era o máis forte* (The North Wind and the Sun were disputing which was the stronger). This sentence was selected because it includes key phonetic variables which have been shown to differ in the speech of Galician- and Spanish-dominant speakers: mid vowels in stressed position (Amengual and Chamorro 2015; Tomé Lourido and Evans 2019) e.g., *norte*, unstressed word-final vowels, e.g., *vento*, and the voiceless alveolar fricative (Tomé Lourido and Evans 2019), e.g., *sobre*. The sentence also includes other Galician-specific features, such as the voiced velar nasal in syllable final position (Freixeiro Mato 2006; Regueira 1999b), e.g., *porfiaban* and connected speech processes (Freixeiro Mato 2006; Regueira 1999b), e.g., *norte e, mais o, era + o*. The sentence was extracted from recordings of the passage used in Tomé Lourido and Evans (2019) produced by 56 speakers: 14 neofalantes (7 female, 7 male), 22 Galician-dominant speakers (12 female, 10 male) and 20 Spanish-dominant speakers (12 female, 8 male), classified following the same method used for listeners. The speakers were early bilinguals in Galician and Spanish recruited from the University of Santiago de Compostela who grew up in Galicia, had not lived anywhere else for more than a year and were 18–30 years old at the time of the recording. They came from both urban and rural backgrounds (neofalantes: 8 urban, 6 rural; Galician-dominant: 5 urban, 17 rural; Spanish-dominant: 11 urban, 9 rural). Speakers raised in one of the main 7 Galician cities (A Coruña, Pontevedra, Ourense, Lugo, Santiago de Compostela, Vigo and Ferrol) were considered to come from an urban background. Speakers raised in smaller towns, villages or smaller areas within villages (e.g., A Baña, Aguiño, Noia, Porto do Son, Silleda) were considered to come from rural backgrounds. The stimuli were scaled for intensity to 65 dB and 50ms silence was added at the beginning and end of each file. The duration of the stimuli ranged from 3.001 s to 5.510 s ( $M = 4.038$  s). All processing was done using Praat (Boersma and Weenink 2015). Stimuli were presented in a random order.

### 3.3. Procedure

Participants completed the accent identification task online using Qualtrics (2015). All the instructions were written in Galician. The definitions and the illustration of the trial procedure presented below correspond to English translations (for the Galician version see Appendix A). Before the task started, definitions for the three different groups were provided as follows:

- (the speaker) Usually speaks Galician: This person speaks Galician in their daily life and has always spoken more Galician than Spanish.
- (the speaker) Usually speaks Spanish: This person speaks Spanish in their daily life and has always spoken more Spanish than Galician.
- (the speaker) Is a new speaker: This person used to speak more Spanish, but now they speak Galician in their daily life.

These definitions were provided in case listeners were unfamiliar with the neofalantes label; although the label is widely used, listeners were recruited to be from a diverse set of backgrounds and not all may have been familiar with it. The trial procedure is illustrated in Figure 1 (for the Galician version, see Figure A1). Participants were instructed to listen to each sentence over headphones and indicate to which group the speaker belonged. The sentence was played only once. Participants were subsequently asked to comment on whether particular factors had influenced their decision (see Section 5. Discussion). In this case, they were allowed to listen to the audio again with no limit on the number of times. Although the experiment was distributed online, it was only advertised through friends and acquaintances of the first author to give some control over who participated and seek to guarantee that participants listened to the stimuli over headphones in a quiet environment. In fact, participants overall spent a considerable amount of time completing the task (mean experiment duration = 65.22 min), which indicates that they took the time to provide detailed comments. Given that the recruitment method was through friends



of friends, and that this was also the case for recruiting the speakers who produced the recordings, participants were asked whether they knew the speaker. Participants indicated that they knew the talker in 114 trials (1.56% of the total number of trials); these trials were excluded from further analysis. Finally, participants completed a language background questionnaire which elicited demographic data and information about their residential history and language background, including how they acquired and use their languages.

What group does this speaker belong to?

(the speaker) usually speaks Galician

(the speaker) usually speaks Spanish

(the speaker) is a new speaker

Next

Why do you think this person usually speaks Spanish?

Give an answer as specific as possible. For example, is there a particular sound or word which makes you think that this speaker usually speaks Spanish? You don't have to use technical language and don't worry about how you explain it. Just use your intuition.

If you would like to, you can listen to the audio again.

Do you know this person?

No Yes

Next

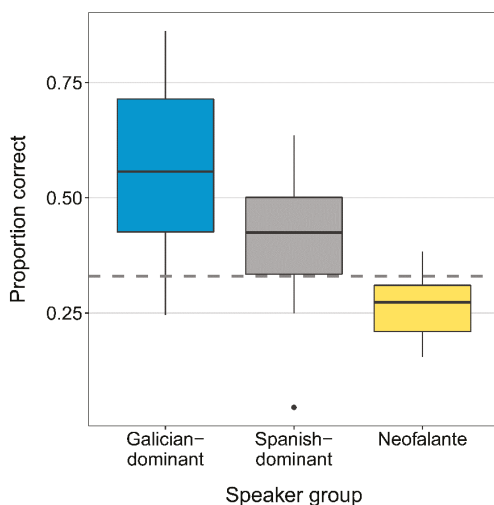
**Figure 1.** Representation of the procedure. First, participants identified to which group they thought the speaker belonged. Then, they provided comments about what influenced their decision. They also indicated whether they thought they knew the speaker.

#### 4. Results

##### 4.1. Can Listeners Identify the Neofalantes' Accent?

Figure 2 shows the identification score (proportion correct) for each of the speaker groups averaged over listeners ( $N = 142$  listeners). The data is available at <https://osf.io/4nwpv> (Supplementary Materials). To investigate which accents were identified at above chance level, the real data were compared to randomly generated data of corresponding dimensions. This method was selected instead of scoring the dependent variable as correct or incorrect and comparing the intercept to chance, because the experiment was a three-way discrimination task, and therefore chance level was not 50%. Three separate logistic regression models were fit to the real and fake data for each of the groups. The dependent variable was the binomial response (correct/incorrect) and the only predictor variable was

type of data (fake or real). Participant and item were included as crossed random effects. Table 1 shows the results of each of the models. Both Galician-dominant [Mean proportion correct (MProp) = 0.57] and Spanish-dominant speakers (MProp = 0.41) were identified at above chance level, but neofalantes were identified systematically worse than chance (MProp = 0.26).



**Figure 2.** Boxplot showing accent identification scores (proportion correct) for all listeners. The three boxplots represent speaker group: Galician-dominant on the left, Spanish-dominant in the centre and Neofalantes on the right. The dashed line represents chance level performance.

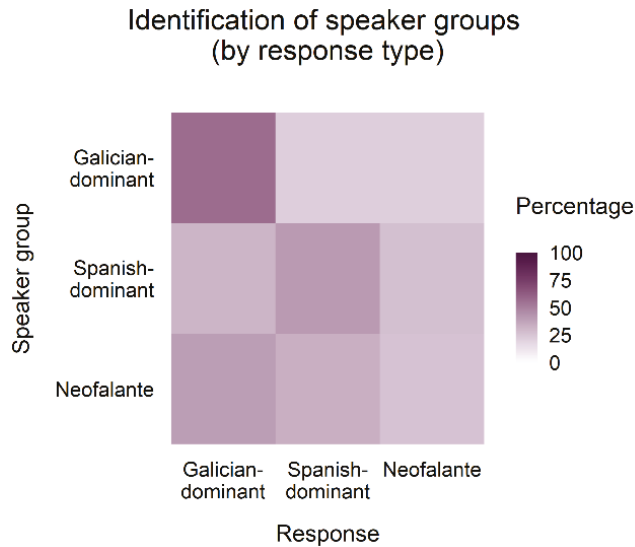
**Table 1.** Summary of the results of the regression models for each speaker group compared to a random baseline. The baseline for the categorical predictor variable was the fake data. Numbers represent Estimates ( $\beta$ ), Standard Errors (SE), Wald statistics (z-values) and p-values.

	$\beta$	SE	z-Value	p-Value
<b>Model 1: Galician-Dominant Speakers</b>				
Intercept	-0.659	0.096	-6.839	<0.001
Real data	0.965	0.057	17.046	<0.001
<b>Model 2: Spanish-Dominant Speakers</b>				
Intercept	-0.704	0.085	-8.307	<0.001
Real data	0.329	0.059	5.582	<0.001
<b>Model 3: Neofalantes</b>				
Intercept	-0.608	0.063	-9.617	<0.001
Real data	-0.438	0.073	-5.963	<0.001

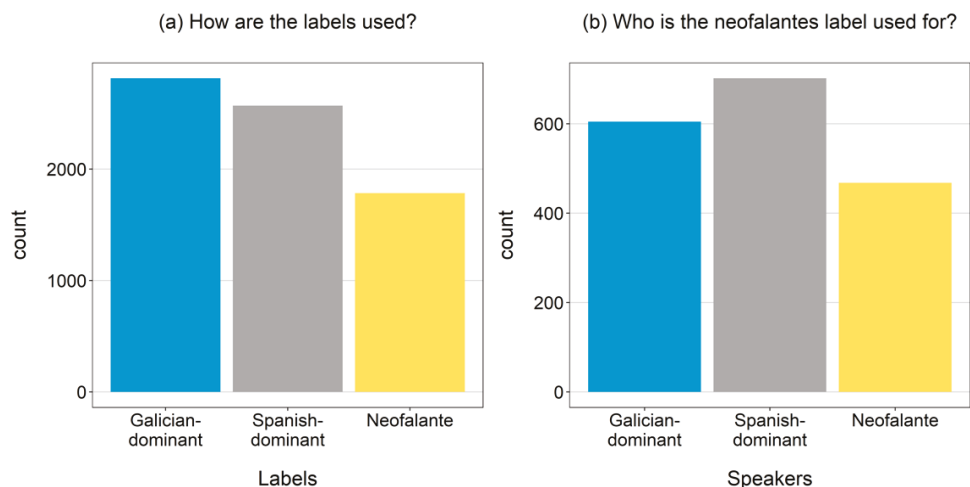
To further investigate whether there were any differences between the two groups of speakers that were identified above chance a separate regression model was fit to the binomial response (correct/incorrect) for Galician-dominant and Spanish-dominant speaker groups in the real data. Speaker group was included as the predictor variable, with Galician-dominant as the baseline. Participant and item were included as crossed random effects. The model revealed a significant difference in identification of Galician-dominant speakers when compared to Spanish-dominant speakers (Intercept:  $\beta = 0.343$ , SE = 0.169,  $z = 2.029$ ,  $p = 0.042$ ; Speaker group:  $\beta = -0.774$ , SE = 0.241,  $z = -3.210$ ,  $p = 0.001$ ); listeners were more accurate in identifying Galician-dominant speakers.

It is clear from these results that listeners could not recognise neofalantes based on their accent. Figure 3 displays the pattern of responses for each speaker group. The confusion matrix shows that neofalantes were not only identified as Spanish-dominant, but also as Galician-dominant speakers. To further explore this question, the responses that corresponded to when neofalantes were misidentified were analysed. An intercept-only logistic regression model was fitted to the categorical response (Galician-dominant vs. Spanish-dominant) when the neofalantes speaker group was misidentified. The model showed that the intercept is significantly different from zero ( $\beta = 0.163$ ,  $SE = 0.055$ ,  $z = 2.945$ ,  $p = 0.003$ ), which implies that the event probability is different from 0.5. This suggests that there is a bias in classifying neofalantes as Galician-dominant; they were classified as Galician-dominants 54% of the time and as Spanish-dominant 46% of the time (see Figure 3).

One possible explanation for the consistent misidentification of neofalantes would be the existence of a bias against choosing the neofalantes label. However, it was not the case that listeners did not choose this label. The left panel on Figure 4 illustrates the counts for each of the speaker labels and shows that all three labels were used for classification. Although the neofalantes label was used the least, the use of labels reflects the distribution of speakers: there were more Galician-dominant ( $N = 22$ ) and Spanish-dominant speakers ( $N = 20$ ) than neofalantes ( $N = 14$ ). Given that the neofalantes label was indeed used, but not for categorising the correct speakers, the question then remains as to which speakers were assigned this label. The right panel on Figure 4 shows counts of the use of the neofalantes label and reveals that it was used to identify Spanish-dominant and Galician-dominant speakers more often than neofalantes themselves.



**Figure 3.** Confusion matrix showing the identification of speaker groups by response type. The *y*-axis represents the speaker group (Galician-dominant, Spanish-dominant, and Neofalantes and the *x*-axis represents the response all listeners gave per speaker group. The darker the colour the higher the percentage of responses in that category.



**Figure 4.** Barplots showing (a) counts for each of the three speaker labels and (b) counts for the Neofalantes label. The plot on the left shows how often each of the three speaker groups labels was selected, with the speaker group labels on the x-axis (Galician-dominant, Spanish-dominant and Neofalantes and frequency counts on the y-axis. The plot on the right shows how often each speaker group (Galician-dominant, Spanish-dominant and Neofalantes) was identified as Neofalantes.

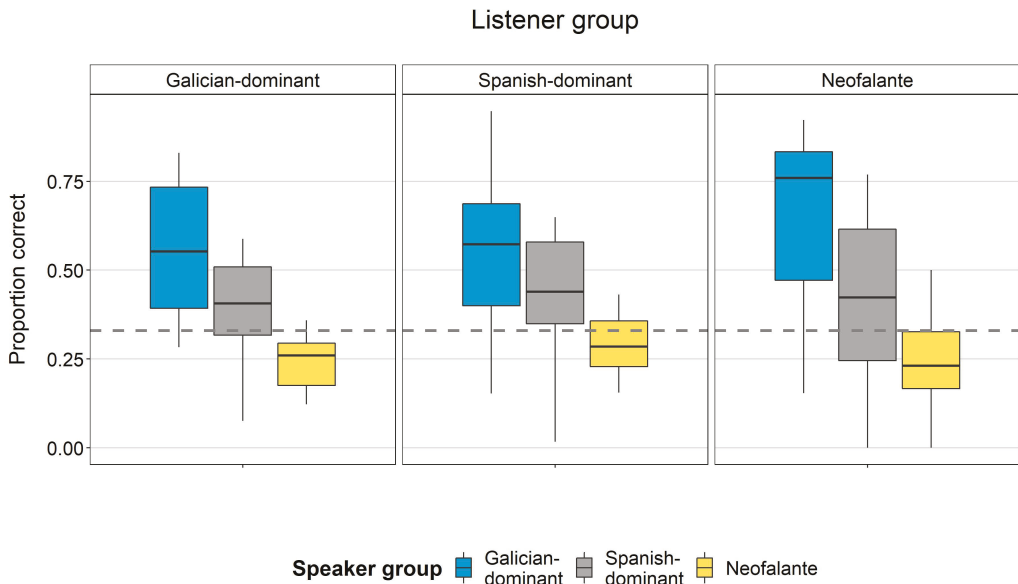
#### 4.2. Does Identification Ability Depend on Listeners’ Language Background?

To investigate whether identification ability depended on listeners’ language background, only data from the three groups of interest was included in the analyses. A logistic mixed effect regression was fitted on the binomial response (correct/incorrect), speaker group and listener group were included as fixed factors. Participant and speaker were included as crossed random effects. The main effects from this model were interpreted using Wald  $\chi^2$  tests, as reported by the Anova() function in the car package (Fox and Weisberg 2011) in R (R Core Team 2013);  $p$ -value  $< 0.001 = ***$ ,  $p$ -value  $< 0.01 = **$ ,  $p$ -value  $< 0.05 = *$ ,  $p$ -value  $> 0.05 = n.s.$  The main effect of speaker group was highly significant [ $\chi^2 (2) = 34.8393 ***$ ]. As discussed in the previous section, Galician-dominant speakers were identified more accurately ( $M = 57%$ ) than Spanish-dominant speakers ( $M = 42%$ ), and both groups were identified more accurately than neofalantes, for whom identification was below chance ( $M = 27%$ ). The effect of listener group was not significant [ $\chi^2 (2) = 4.5787 n.s.$ ], suggesting that language background did not affect overall identification. This can be seen in Figure 5, which shows the accent identification scores, and which illustrates that the pattern of identification was very similar for all three listener groups.

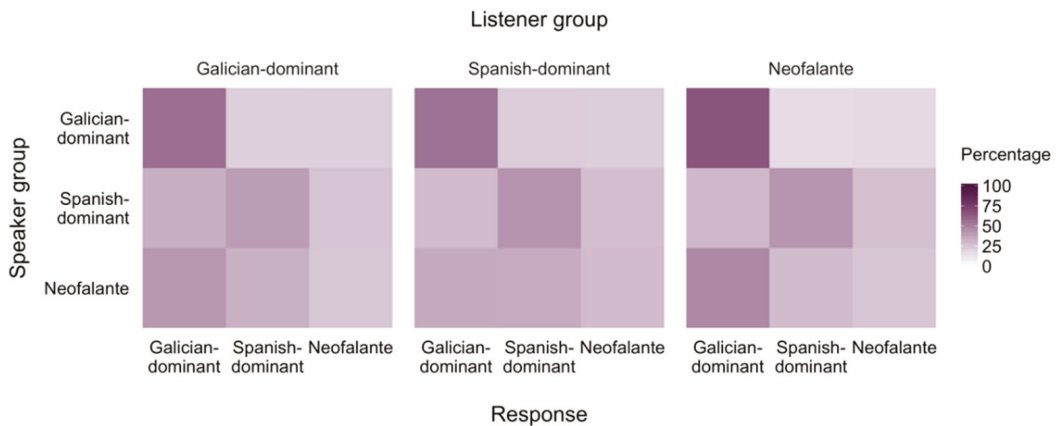
The analysis also showed a significant interaction between speaker group and listener group [ $\chi^2 (4) = 12.4894 *$ ]. To follow up this interaction, pairwise post-hoc tests were carried out using the lsmeans package (Lenth 2016) in R, adjusting for multiple comparisons using the Tukey method. The interaction appeared to be driven by the identification of Galician-dominant speakers by neofalantes listeners when compared to both Galician-dominant (GD vs. NF:  $\beta = -0.446$ ,  $SE = 0.160$ ,  $z = -2.774$ ,  $p = 0.015$  and Spanish-dominant listeners (SD vs. NF:  $\beta = -0.504$ ,  $SE = 0.159$ ,  $z = -3.161$ ,  $p = 0.004$ ). No other interactions were significant. This indicates that neofalantes were better ( $M = 66%$ ) than the other two listener groups (GD:  $M = 56%$ , SD:  $M = 55%$ ) at identifying Galician-dominant speakers.

This effect is illustrated in Figure 6, which displays the identification of speaker groups by response type and listener group. The graph shows that the cell with the darkest colour (i.e., highest number of accurate responses) corresponds to the identification of Galician-dominant speakers by neofalantes listeners (matrix on the right), indicating that neofalantes were more accurate than Galician-dominant and Spanish-dominant listeners at identifying Galician-dominant speakers, as revealed by the significant interaction

between speaker and listener groups in the regression model. Another apparent difference in the classification pattern concerns which listener groups classified neofalantes as Galician-dominant speakers. To investigate if groups differed in their classification of neofalantes, a mixed-effect logistic regression model was fit to the binomial response (Galician-dominant/Spanish-dominant) from the subset of data where neofalantes were identified incorrectly. Listener group was included as a fixed factor in the model, with neofalantes as a baseline, and participant was included as a crossed random effect. The model (Intercept:  $\beta = 0.529$ ,  $SE = 0.210$ ,  $z = 2.511$ ,  $p = 0.012$ ) revealed that Galician-dominant listeners did not differ from neofalantes listeners when labelling neofalantes speakers as Galician-dominant ( $\beta = -0.273$ ,  $SE = 0.233$ ,  $z = -1.176$ ,  $p = 0.239$ ), but Spanish-dominant listeners did differ from neofalantes listeners when labelling neofalantes speakers as Galician-dominant ( $\beta = -0.533$ ,  $SE = 0.232$ ,  $z = -2.295$ ,  $p = 0.022$ ). This suggests that neofalantes were identified as Galician-dominant more frequently by Galician-dominant listeners (56% of the time) and neofalantes themselves (62% of the time), than by Spanish-dominant listeners, who identified them as Galician-dominant 50% of the time and as Spanish-dominant 50% of the time.



**Figure 5.** Boxplot showing accent identification scores (proportion correct) by the three listener groups: Galician-dominant (left rectangle), Spanish-dominant (middle rectangle) and neofalantes (right rectangle). Boxplots represent speaker group: Galician-dominant on the left, Spanish-dominant in the centre and Neofalantes on the right. The dashed line represents chance level performance. The accent identification pattern was very similar for all three listener groups.



**Figure 6.** Confusion matrices showing the identification of speaker groups by response type and listener group (Galician-dominant, Spanish-dominant and Neofalantes). The y-axis represents the speaker group (Galician-dominant, Spanish-dominant, and Neofalantes) and the x-axis represents the response each listener group gave per speaker group. The matrix on the left corresponds to Galician-dominant listeners, the one in the centre to Spanish-dominant listeners and the one on the right to Neofalantes. The darker the colour the higher the percentage of responses in that category. Neofalantes listeners were better than the two other listener groups at identifying Galician-dominant speakers and neofalantes speakers were identified as Galician-dominant more frequently by Galician-dominant listeners and neofalantes themselves than by Spanish-dominant listeners.

## 5. Discussion

### 5.1. The Neofalantes' Accent as an Emerging Variety

Listeners in the Galician community, regardless of language background, can identify Galician-dominant and Spanish-dominant bilinguals above chance and perform better with the former group. However, they cannot identify the neofalantes' accent. Although there are differences in how individual neofalantes are classified, with some speakers more often classified as Galician-dominant and others more often classified as Spanish-dominant, overall, neofalantes speakers are not only confused with Spanish-dominant but also with Galician-dominant speakers. This result suggests that their accent contains features used by both Galician-dominant and Spanish-dominant speakers. There are also differences in categorization patterns according to listener background; neofalantes listeners show heightened sensitivity to the Galician-dominant variety, in comparison to the other two groups, classifying Galician-dominant speakers more accurately than Spanish-dominant and Galician-dominant listeners. Despite the frequent use of the neofalantes label to designate this social group (O'Rourke and Ramallo 2011, 2015; Ramallo 2013; Tomé Lourido and Evans 2019), the results of this study indicate that Galician listeners are unable to recognise the variety used by neofalantes, that is, they are not able to associate the label with a set of phonetic features, whereas they can do so for Galician-dominant or Spanish-dominant speakers. One possibility is that some participants in the experiment might not have been familiar with the existence of neofalantes as a social group. This study deliberately set out to test the wider community and selected a pool of participants from all backgrounds and professions to investigate whether a neofalantes accent had emerged as a new variety in the community as whole, rather than in only particular areas of society (e.g., those related with language planning and revitalisation or Galician linguistics). However, it seems unlikely that participants did not understand the label, as they were provided with definitions for each group before starting the experiment and the results showed that participants used all three labels. Besides, even though they might not use the label themselves, Galician listeners are often aware that individual speakers may switch language dominance during their lives. In fact, some of the comments they provided

to justify their choice when they identified a speaker as neofalante illustrate this point (participants' quotes were translated by the first author):

- |   |  |
|---|--|
| <p>[1] <i>Fala galego habitualmente pero non parece que sempre fora así, como se pensara en castelán.</i></p> <p>[2] <i>A entoación segue sendo lixeiramente castelá. Tenta falar galego, pero lle queda ese acento castelanfalante.</i></p> <p>[3] <i>Este chico non falou galego ata que chegou a universidade.</i></p> <p>[4] <i>Prosodia e pronuncia "aprendida", non soa "natural".</i></p> <p>[5] <i>Boa fonética, mais penso que adquirida a posteriori.</i></p> | <p>[1] '(The speaker) usually speaks Galician, but it doesn't seem like it has always been like this, as if (they) thought in Spanish'</p> <p>[2] 'The intonation continues to be slightly Spanish. (The speaker) tries to speak Galician, but (they) are left with that Spanish-speaking accent.'</p> <p>[3] 'This guy didn't speak Galician until he got to university.'</p> <p>[4] "'Learnt' prosody and pronunciation, it doesn't sound "natural".'</p> <p>[5] 'Good phonetics, but I think it was acquired a posteriori.'</p> |
|---|--|

These comments suggest that listeners were aware that the definition of a neofalante involved a long-term language switch. Therefore, it seems unlikely that the reason why neofalantes were not identified as such was related to not understanding the label.

A question that then arises is in what ways the 'neofalante' label is becoming associated with a particular set of linguistic features. It is possible that listeners have not yet tuned into the phonetic forms produced by neofalantes to be able to link them with the social group to which they belong. However, this interpretation would assume that the changes after the neofalantes' language switch are sufficiently phonetically distinct to constitute an identifiable variety. To evaluate this assumption, it is worth considering that listeners were less accurate in identifying Spanish-dominant than Galician-dominant speakers. Spanish-dominant speakers are not L2 learners and thus, are likely to have a certain type of Galician accent, both in Galician and in Spanish. Therefore, variation due to language background differences could be organised along a continuum with Galician-dominant speakers at one end and L2 Galician speakers at the other end (e.g., a person from Madrid). The accent of Spanish-dominant speakers then, which would fall in the middle of the continuum, but towards the L2 accent side, would not be as distinctive as the Galician-dominant one. Recent work on Galician and Galician Spanish also supports the idea of an existing continuum of varieties, with more traditional Galician varieties, typically produced by rural Galician-dominant speakers at one end and varieties which are more influenced by Spanish, typically produced by urban Spanish-dominant speakers at the other end (e.g., [Regueira 2019](#); [Regueira and Fernández Rei 2020](#)). Regarding variation within Galician-dominant speakers, [Aguete Cajiao \(2019, 2020\)](#) proposes the existence of two models of stressed vowel systems in Galician: a conservative model with seven vowels (see also [de la Fuente Iglesias and Castillejo 2020a, 2020b](#)) and an innovative model with five vowels, as a result of both language internal and language contact factors. The latter model, with merged mid vowel contrasts is associated with urban and semi-urban areas, where Spanish is more widespread ([Aguete Cajiao 2019, 2020](#); [Tomé Lourido and Evans 2019](#); [Mayr et al. 2019](#)), and also with neofalantes and Spanish-dominant speakers ([Amengual and Chamorro 2015](#); [Tomé Lourido and Evans 2019](#); [Regueira 2019](#); [Regueira and Fernández Rei 2020](#)). [Regueira and Fernández Rei \(2020\)](#) examined the stressed and unstressed vowels systems and intonation patterns of six Galician bilingual speakers from different language backgrounds. As well as confirming the patterns found in previous studies for stressed vowels in Galician ([Tomé Lourido and Evans 2019](#); [Aguete Cajiao 2019](#); [Amengual and Chamorro 2015](#)), they found that for unstressed final vowels Galician-dominant, neofalantes and rural Spanish-dominant speakers used reduced vowels, a traditional Galician feature. However, the urban Spanish speaker used an unstressed vowel system that was closer to Castilian Spanish and different from the rest of the participants, providing further support for the existence of a continuum, but also illustrating that individual phonetic variables may behave differently.

The existence of a continuum would also explain why the neofalantes' accent was not accurately identified. These speakers would be situated between Galician-dominant and Spanish-dominant bilingual speakers, and thus, it might not be possible for this accent to emerge as distinctive, due to the degree of overlap with the other two varieties. This idea is similar to explanations of how children develop awareness of regional accent variation. Wagner et al. (2014) argue that children have a gradient representation of accent variation in which the native accent forms the core set of experience and other accents are categorised in relation to that core (see also Evans and Lourido 2019). One possibility is that such gradient representations not only form the basis of adult representations, but that they continue to be used in adulthood. In our case, it is possible that a prototypical Galician-like accent and a prototypical Spanish-like accent function as anchors at both ends of a continuum, and that other language backgrounds are identified relative to these. In fact, some comments that participants made when identifying neofalantes' speakers support this idea:

- |  |   |
|--|---|
| [1] <i>Non vexo claro se é máis galego ou máis castelán.</i>   | [1] 'It is not clear to me if it is more Galician or more Spanish.'   |
| [2] <i>Os enes e a articulación das consoantes son casteláns, pero semella polo ton e as vogais que fala galego normalmente.</i>   | [2] 'The "n"s and the articulation of consonants are Spanish, but in terms of the tone and the vowels, it seems that (the speaker) usually speaks Galician.'  |
| [3] <i>Hai moita variabilidade entre rasgos de pronuncia típicamente galegos e outros moi alleos.</i>  | [3] 'There is a lot of variability between typically Galician pronunciation features and very alien ones.'  |
| [4] <i>Ten unha mezcla de pronunciacións.</i>  | [4] '(The speaker) has a mixture of pronunciations.'  |
| [5] <i>Ten un amago de sete vogais, pero non tan claras como nos galegofalantes. Transmítete sensación de inseguridade, como se non soubese exactamente como ten que dicir cada palabra. Podería vir xusto desa condición de neofalante.</i> | [5] '(The speaker) has something like seven vowels, but they are not as clear as those of Galician speakers. It conveys to me a feeling of insecurity, as if (they) didn't know how exactly (they) have to say each word. It could come from precisely that condition of neofalante.' |
| [6] <i>Ten un bo acento galego pero algunhas trazas son do castelán.</i>   | [6] (The speaker) has a good Galician accent, but some features are Spanish.'   |

These comments also reveal that neofalantes were described as using a mix of features, some of which were identified as Spanish and others which were identified as Galician, indicating that neofalantes use a hybrid variety. Tomé Lourido and Evans (2019) investigated the production of three segmental variables by these three bilingual groups and showed that neofalantes pattern with Spanish-dominant speakers for mid-vowel and fricative contrasts, but with Galician-dominant speakers for reduced word-final vowels. These findings showed that neofalantes did not produce categories that were distinctive from the other two groups. Likewise, the accent identification study showed that neofalantes patterned with both bilingual groups, as they were not only identified as Spanish-dominant but also as Galician-dominant, specifically by Galician-dominant listeners and neofalantes themselves. This is in contrast to impressionistic descriptions of neofalantes' varieties that have suggested that these speakers have a Spanish-accented variety of Galician (Freixeiro Mato 2014; González González 2008; Ramallo 2010) and are speakers of 'New Urban Galician' (*Novo galego urbano*, Dubert García 2002; Regueira 1999a). However, the phonetic variables examined in Tomé Lourido and Evans (2019) represent only a limited number of features—these represent only part of their accent—and it is possible that listeners are sensitive to other segmental or suprasegmental features. In sum, it appears that neofalantes use a mixture of Galician- and Spanish-like variables, including the phonetic features examined in Tomé Lourido and Evans (2019) but that they may also use other variables that have not yet been explored. It is possible then, that listeners in the community are sensitive not only to the Spanish-like variables, but also to the Galician-like features that neofalantes acquire after their switch, and that this leads them to categorise neofalantes speakers as both Spanish- and Galician-dominant speakers.



## 5.2. Accent Identification and Listeners' Language Background

Our second research question examined whether identification ability depended on listeners' language background. Overall, identification accuracy was similar for the three listener groups. These results do not provide full support for the idea that language ability facilitates identification of the speakers' language background (see [Perrachione et al. 2011](#), for effects of language ability on voice identification), as an effect of language ability would predict better performance in Galician-dominant listeners. Although all bilingual groups were familiar with the phonological system of Galician, neofalantes and Spanish-dominant listeners likely perceive the sounds of Galician through their native Spanish categories ([Tomé Lourido and Evans 2019](#); [Iverson et al. 2003](#); [Pallier et al. 1997](#)). [Tomé Lourido and Evans \(2019\)](#) showed that neofalantes and Spanish-dominant listeners' accuracy when identifying the mid-vowel contrasts in minimal pairs in a word identification task was not as good as that of Galician-dominant listeners', who performed at ceiling. Other studies have found similar results when comparing Spanish- and Galician-dominant listeners ([Amengual and Chamorro 2015](#); [Aguete Cajiao 2019](#)). However, many participants from the neofalantes and Spanish-dominant groups claimed to use the mid-vowel contrasts to categorise speakers:

- |   |   |
|---|---|
| [1] Spanish-dominant listener (SD): <i>Boa distinción entre vogais medias abertas e pechadas.</i>   | [1] SD: 'Good distinction between open and close mid vowels.'   |
| [2] SD: <i>Todas as vogais me sonan igual de pechadas. Creo que as non logra diferenciar con facilidade.</i>  | [2] SD: 'All the vowels sound equally close to me. I think (the speaker) can't differentiate them easily.'  |
| [3] Neofalante (NF): <i>Véxolle seguridade na fala e non emprega as vogais abertas, que para min é algo moi característico para saber quen é galego falante e quen non.</i> | [3] NF: 'I see that (the speaker) is confident when speaking and doesn't use the open vowels, which for me is something very characteristic to know who is a Galician speaker and who isn't.' |
| [4] NF: <i>Só lle escoito cinco vogais.</i>   | [4] NF: 'I can only hear five vowels.'  |

One possibility is that listeners believe they use certain phonetic features, such as mid vowels, to classify speakers when they might be, in fact, using different variables. This would imply a mismatch between what they think they use and what they actually use. Mid vowels could be considered a sociolinguistic stereotype, one which forms part of the knowledge of members of the bilingual community, even though it may not conform to an objective fact ([Labov 1972](#)). There is a high degree of awareness among individuals in the community about the fact that one of the differences between Galician and Spanish is the vowel system. This is particularly true for younger listeners, who have been taught the Galician language at school. Besides, there is a widespread belief that a 'good speaker' of Galician must have all seven vowels. This may also be why listeners are able to report specific phonetic features, often vowels, using linguistic terminology in the comments above. However, it seems rather unlikely that Spanish-dominant and neofalantes listeners who were not always able to identify the mid-vowel contrast in a vowel identification task ([Tomé Lourido and Evans 2019](#)) would be able to use this contrast in accent categorisation. It is likely that instead, they use other phonetic features, such as unstressed word-final vowels, a feature that has been claimed to be easily perceptible and distinctive ([Regueira 2012](#)), but that they believe they use mid vowels. Indeed, there were remarkably fewer comments highlighting the influence of word-final vowels in participants' decisions, and those comments were expressed in less explicit ways. For example, in comments (1) and (2) the participants represent in spelling the reduction of unstressed word-final vowels by writing 'norti' instead of *norte*, 'mailu' instead of *mailo*, and 'du' instead of *do*. In comment (3) the listener refers to this feature by saying that the final vowel is almost not pronounced.

- |  |  |
|--|--|
| [1] <i>Clarísima galego polo acento. O vento do 'norti' e 'mailu' sol ...</i>                                  | [1] 'Clearly Galician because of the accent. O vento do "norti" e "mailu" sol ...'                           |
| [2] <i>A contracción 'do' pronúnciaa como unha persoa que fala en galego normalmente. Case pronuncia 'du'.</i> | [2] '(The speaker) pronounces "do" like a person who usually speaks Galician. (They) almost pronounce "du".' |
| [3] <i>Casi no pronuncia la &lt;-e&gt; final.</i>  | [3] '(The speaker) almost doesn't pronounce the final <-e>.'   |

Listeners also made references to other segmental features such as the pronunciation of /l/, /s/ and /ŋ/ and liaison processes, e.g., 'era o' [l̪ɛrə] and 'máis o' as [l̪majlo\_] (transcription of individual words follows the transcription system proposed in Regueira's *Diccionario de pronuncia da lingua galega* (Regueira 2010)). The phonemes /l/ and /s/ exist in both languages, but the realisation of the /s/ has been found to be different for Galician-dominant and Spanish-dominant speakers (Tomé Lourido and Evans 2019). Additionally, there is individual and regional variation in the realisation of /s/ in Galician (Regueira and Ginzo 2018; Regueira 2014; Labraña Barrero 2009, 2014). In contrast, the phoneme /ŋ/ and the liaison processes that occur in the sentence are characteristic of Galician and do not exist in Spanish (see Fernández Rei 2005; Regueira et al. 1998; for vowel elision in Galician). Suprasegmental features, such as rhythm, intonation and prosody, which are typically different in both languages, were consistently mentioned (see Fernández Rei 2005, 2016; Fernández Rei et al. 2014; for Galician prosody).

Assuming that a similar mechanism underlies voice and accent identification skills, the result that all listener groups showed a similar level of accuracy in identifying talkers supports an account in which familiarity with a phonological system, rather than more robust phonological representations, facilitates talker identification (Fleming et al. 2014; Goggin et al. 1991; Thompson 1987). In this context, all three listener groups live in a bilingual community where they have everyday exposure to all the accents. These findings are in line with other work in the area of voice identification. Bregman and Creel (2014) showed that listeners learnt to recognise talkers faster in their L1 than in their L2, but that early bilinguals learnt voices equally quickly in both of their languages. They suggest that one way to account for the difference between early and late learners is that languages or cultures differ in terms of the features that are used to differentiate between talkers. As they acquire the sound inventory of their L2, early bilinguals, unlike late learners, are also thought to acquire the 'talker-varying characteristics unique to a particular culture' (Bregman and Creel 2014, p. 94). In the case of Galician bilinguals, it is possible that from an early age, they gain sensitivity to the phonetic cues that help to identify the speaker's language background and that's why no overall difference in accuracy was found between the three groups. Clopper and Pisoni (2004, 2006) found that performance in accent categorisation tasks appears to be modulated by participants' background: listeners who had lived in different areas performed better than those who had only lived in one area and, additionally, listeners who lived in a particular region performed better with the accent from that region. The authors proposed that greater exposure to linguistic variation and specific experience with one variety benefits accent categorisation. The results of the current study do not contradict Clopper and Pisoni's findings. All listeners had been exposed to all the accents presented here, at least to Galician-dominant and Spanish-dominant varieties. and although listeners did not show an advantage for their own accent, this may be because of their frequent exposure to all accents. This is similar to findings in other bilingual contexts (Mayr et al. 2020; Tan 2012). Mayr et al. (2020) found that monolingual and bilingual listeners in Wales were able to identify whether a person was able to speak Welsh based on their accent in English. The accuracy rate in their study was similar to that for Galician- and Spanish-dominant speakers in our study; above chance, but not exceptional. Likewise, although the listener groups in both studies are not fully comparable, they also found no difference in performance between their English monolingual and Welsh-English bilingual listener groups. Note that the sociolinguistic situation in Wales and Galicia is different in this regard, as it would be rare to find monolingual speakers of Galician or Spanish born

and raised in Galicia within the age range tested (18–54 years old), at least at the time when the study was carried out; this may change in the future.

However, identification accuracy was not exactly the same for all listener groups; neofalantes showed heightened sensitivity to one of the accents, the Galician-dominant variety. This could be due to neofalantes' increased metalinguistic awareness about Galician. Neofalantes are typically very aware of the way they speak and the fact that their accent is different from that of Galician-dominant speakers. They are usually very motivated to learn Galician and invest time and effort in doing so. O'Rourke and Ramallo (2013b, 2015) argue that neofalantes have a heightened sense of awareness about their own sociolinguistic reality and the sociolinguistic context in Galicia. Taking all these factors into consideration, it seems reasonable to hypothesise that neofalantes would be more sensitive to phonetic features in the Galician variety, as that is likely the model most of them follow after they switch languages. In sociolinguistics, listeners' sensitivity to a particular phonetic cue or awareness of a sociolinguistic variable has been related to the concept of 'salience' (Drager and Kirtley 2016; Jaeger and Weatherholtz 2016; Montgomery and Moore 2018; Nycz 2016; Rącz 2013). Jaeger and Weatherholtz (2016) distinguish between the 'initial salience' of a novel feature a listener experiences for the first time and salience at a later stage, i.e., the cumulative exposure the individual has had to the variant. A feature is perceived to have initial salience when it is unexpected in relation to the listener's previous language experience and, therefore, varies between individuals and communities. One could hypothesise that neofalantes were more sensitive to Galician-specific features because these were not part of their phonetic repertoire, or at least not before the language dominance switch.

Previous work has also shown that associations between phonetic variables and social meanings may not be the same for all listeners in the community. Eckert (2008) argues that variables do not have fixed and static meanings, but instead that they acquire that meaning in a particular context. Identifying Galician-dominant speakers or monitoring their speech might not be so important for Spanish-dominant listeners or Galician-dominant listeners themselves, whilst it might be particularly relevant for neofalantes. This explanation is consistent with Evans and Lourido (2019) findings for monolingual and bilingual children in London; bilinguals were able to differentiate talkers with a foreign, regional and their home accents, whilst monolinguals were only able to differentiate the foreign accent from their own. Like the bilingual children, it is possible that neofalantes develop and then benefit from the accent identification skills needed to navigate the relationships within their community. One important caveat is that the listener groups were not balanced for sample size. Whilst there were 58 Galician-dominant and 61 Spanish-dominant listeners, there were only 13 neofalantes, due to the difficulties in recruiting this group of bilinguals (Tomé Lourido and Evans 2019, p. 645). One possibility is that this result is due to variability in the neofalantes group, and replication of this effect is thus needed to ensure its validity.

## 6. Conclusions

In sum, this study showed that although neofalantes are a distinct social group that acquire and use both their languages in a different way to Galician- and Spanish-dominant bilinguals, the emergence of this profile of speakers has not led to the creation of a distinct neofalantes variety (see also Nance et al. 2016) that is recognised by Galician listeners. Instead, listeners categorise neofalantes as both Spanish- and Galician-dominant, supporting findings from production studies that show that neofalantes use a variety containing a mix of Spanish and Galician features (Tomé Lourido and Evans 2019; Regueira and Fernández Rei 2020). One possibility is that listeners have a gradient representation of variation, with Galician-like accents and Spanish-like accents functioning as anchors and the neofalantes' accent situated somewhere in the middle. There was also evidence to support the view that familiarity with a phonological system, rather than more robust phonological representations, benefits accent identification; the overall identification accuracy was similar

for bilinguals from the three language backgrounds, suggesting that the three groups are sensitive to the phonetic cues that are used to identify the background of a speaker in a community and likely acquire them early in life. However, the differences in the patterns of identification indicate that listeners did not weigh phonetic features in the same way. These findings suggest that representations of accent variation vary according to language background and provides further evidence that the evaluation of phonetic features not only varies as a function of context, but also depends on the social and language experience of the individual.

**Supplementary Materials:** The following are available online at <https://osf.io/4nwpv>, Dataset.

**Author Contributions:** Conceptualisation, methodology, analysis and writing: G.T.L. and B.G.E. Both authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Ethical approval was obtained from the Division of Psychology and Language Sciences at UCL (SHaPS-2014-BE-004).

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Acknowledgments:** We are grateful to Xosé Luís Regueira and Elisa Fernández Rei for their help with participant recruitment. We also thank Ewan Dunbar for his help with data visualisation and statistical analysis.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A. Materials related to the Procedure in Galician

Definitions:

- Normalmente fala galego: Esta persoa fala galego no seu día a día e sempre falou máis galego que castelán.
- Normalmente fala castelán: Esta persoa fala castelán no seu día a día e sempre falou máis castelán que galego.
- E neofalante: Esta persoa sempre falaba máis castelán, pero agora fala galego no seu día a día.

A que grupo pertence este falante?

normalmente fala galego

normalmente fala castelán

é neofalante

Por que cres que esta persoa normalmente fala castelán?

Dá unha resposta tan específica como che sexa posible. Por exemplo, hai algún son ou palabra en particular que che faga pensar que este falante normalmente fala castelán? Non tes por que utilizar linguaxe técnica e non te preocupes por como o explicas. Simplemente utiliza a tua intuición.

Se queres, podes volver escoitar o audio.

Cofiecces a esta persoa?

Non  Si

**Figure A1.** Representation of the procedure. First, participants identified to which group they thought the speaker belonged. Then, they provided comments about what influenced their decision. They also indicated whether they thought they knew the speaker.

## References

- Agha, Asif. 2003. The Social Life of Cultural Value. *Language and Communication* 23: 231–73. [CrossRef]
- Aguete Cajiao, Alba. 2019. Vocalismo galego. Motivacións fonolóxicas do cambio. Doctoral thesis, Universidade de Santiago de Compostela, Santiago, Spain.
- Aguete Cajiao, Alba. 2020. Contacto de linguas e variedade na emerxencia de novos modelos de vocalismo galego. In *Variedades lingüísticas en contacto na Península Ibérica*. Edited by Francisco Dubert-García, Vitor Míguez and Xulio Sousa. Santiago de Compostela: Consello da Cultura Galega. [CrossRef]
- Allen, J. Sean, and Joanne L. Miller. 2004. Listener Sensitivity to Individual Talker Differences in Voice-Onset-Time. *The Journal of the Acoustical Society of America* 115: 3171–83. [CrossRef] [PubMed]
- Amengual, Mark, and Pilar Chamorro. 2015. The Effects of Language Dominance in the Perception and Production of the Galician Mid Vowel Contrasts. *Phonetica* 72: 207–36. [CrossRef] [PubMed]
- Baugh, John. 2000. Racial Identification by Speech. *American Speech* 75: 362–64. [CrossRef]
- Boersma, Paul, and David Weenink. 2015. Praat: Doing Phonetics by Computer [Computer Program]. Version 6.0. Available online: <http://www.praat.org/> (accessed on 13 November 2015).
- Bregman, Micah R., and Sarah C. Creel. 2014. Gradient Language Dominance Affects Talker Learning. *Cognition* 130: 85–95. [CrossRef]

- Butler, Joseph, Caroline Floccia, Jeremy Goslin, and Robin Panneton. 2011. Infants' Discrimination of Familiar and Unfamiliar Accents in Speech. *Infancy* 16: 392–417. [\[CrossRef\]](#) [\[PubMed\]](#)
- Campbell-Kibler, Kathryn. 2011. Intersecting Variables and Perceived Sexual Orientation in Men. *American Speech* 86: 52–68. [\[CrossRef\]](#)
- Clopper, Cynthia G. 2008. Auditory Free Classification: Methods and Analysis. *Behavior Research Methods* 40: 575–81. [\[CrossRef\]](#)
- Clopper, Cynthia G., and Ann R. Bradlow. 2009. Free Classification of American English Dialects by Native and Non-Native Listeners. *Journal of Phonetics* 37: 436–51. [\[CrossRef\]](#) [\[PubMed\]](#)
- Clopper, Cynthia G., and David B. Pisoni. 2004. Homebodies and Army Brats: Some Effects of Early Linguistic Experience and Residential History on Dialect Categorization. *Language Variation and Change* 16: 31–48. [\[CrossRef\]](#)
- Clopper, Cynthia G., and David B. Pisoni. 2006. Effects of Region of Origin and Geographic Mobility on Perceptual Dialect Categorization. *Language Variation and Change* 18: 193–221. [\[CrossRef\]](#)
- Clopper, Cynthia G., and David B. Pisoni. 2007. Free Classification of Regional Dialects of American English. *Journal of Phonetics* 35: 421–38. [\[CrossRef\]](#)
- Costa, James. 2015. New Speakers, New Language: On Being a Legitimate Speaker of a Minority Language in Provence. *International Journal of the Sociology of Language* 231: 127–45. [\[CrossRef\]](#)
- de la Fuente Iglesias, Mónica, and Susana Pérez Castillejo. 2020a. L1 Phonetic Permeability and Phonetic Path towards a Potential Merger The Case of Galician Mid Vowels in Bilingual Production. *Linguistic Approaches to Bilingualism*, 1–29. [\[CrossRef\]](#)
- de la Fuente Iglesias, Mónica, and Susana Pérez Castillejo. 2020b. Phonetic Interactions in the Bilingual Production of Galician and Spanish /e/ and /o/. *International Journal of Bilingualism* 24: 305–18. [\[CrossRef\]](#)
- Docherty, Gerard J., and Paul Foulkes. 2014. An Evaluation of Usage-Based Approaches to the Modelling of Sociophonetic Variability. *Lingua* 142: 42–56. [\[CrossRef\]](#)
- Drager, Katie, and M. Joelle Kirtley. 2016. Awareness, Salience, and Stereotypes in Exemplar-Based Models of Speech Production and Perception. In *Awareness and Control in Sociolinguistic Research*. Edited by Anna M. Babel. Cambridge: Cambridge University Press, pp. 1–24. [\[CrossRef\]](#)
- Dubert García, Francisco. 2002. Os sociolectos galegos. *Cadernos de lingua* 24: 5–27.
- Eckert, Penelope. 2008. Variation and the Indexical Field. *Journal of Sociolinguistics* 12: 453–76. [\[CrossRef\]](#)
- Evans, Bronwen G., and Gisela Tomé Lourido. 2019. Effects of Language Background on the Development of Sociolinguistic Awareness: The Perception of Accent Variation in Monolingual and Multilingual 5: The 7-Year-Old Children. *Phonetica* 76: 142–62. [\[CrossRef\]](#) [\[PubMed\]](#)
- Fernández Rei, Elisa. 2005. O estándar prosódico. In *Norma lingüística e variación*. Edited by Rosario Álvarez and Henrique Monteagudo. Santiago de Compostela: Consello da Cultura Galega and Instituto da Lingua Galega, pp. 97–125.
- Fernández Rei, Elisa. 2016. Dialect, Historical and Sociolinguistic Aspects of Galician Intonation. *Dialectologia* 6: 147–69.
- Fernández Rei, Elisa, Lurdes de Castro Moutinho, and Rosa Lúcia Coimbra. 2014. As entoacións galega e portuguesa: A fronteira á luz da dialectometría e da percepción. In *Lingua e identidade na fronteira galego-portuguesa*. Edited by Xulio Sousa, Marta Negro Romero and Rosario Álvarez. Santiago de Compostela: Consello da Cultura Galega, pp. 115–41.
- Fleming, David, Bruno L. Giordano, Roberto Caldara, and Pascal Belin. 2014. A Language-Familiarity Effect for Speaker Discrimination without Comprehension. *Proceedings of the National Academy of Sciences* 111: 13795–98. [\[CrossRef\]](#) [\[PubMed\]](#)
- Floccia, Caroline, Joseph Butler, Frédérique Girard, and Jeremy Goslin. 2009. Categorization of Regional and Foreign Accent in 5- to 7-Year-Old British Children. *International Journal of Behavioral Development* 33: 366–75. [\[CrossRef\]](#)
- Fox, John, and Sanford Weisberg. 2011. *An R Companion to Applied Regression*, 2nd ed. Thousand Oaks: Sage.
- Freixeiro Mato, Xosé Ramón. 2006. *Gramática da lingua galega I. Fonética e fonoloxía*, 2nd ed. Vigo: A Nosa Terra.
- Freixeiro Mato, Xosé Ramón. 2014. Lingua oral, calidade da lingua e futuro do galego. In *Modelos de lingua e compromiso*. Edited by Xosé Manuel Sánchez Rei. A Coruña: Baía Edicións, pp. 13–76.
- Giles, Howard. 1970. Evaluative Reactions to Accents. *Educational Review* 22: 211–27. [\[CrossRef\]](#)
- Giles, Howard. 1971a. Ethnocentrism and the Evaluation of Accented Speech. *British Journal of Social and Clinical Psychology* 10: 187–88. [\[CrossRef\]](#)
- Giles, Howard. 1971b. Patterns of Evaluation to R.P., South Welsh and Somerset Accented Speech. *British Journal of Social and Clinical Psychology* 10: 280–81. [\[CrossRef\]](#) [\[PubMed\]](#)
- Giles, Howard, and Peter F. Powesland. 1975. *Speech Style and Social Evaluation*. London and New York: Academic Press.
- Girard, Frédérique, Caroline Floccia, and Jeremy Goslin. 2008. Perception and Awareness of Accents in Young Children. *British Journal of Developmental Psychology* 26: 409–33. [\[CrossRef\]](#)
- Goggin, Judith P., Charles P. Thompson, Gerhard Strube, and Liza R. Simental. 1991. The Role of Language Familiarity in Voice Identification. *Memory & Cognition* 19: 448–58. [\[CrossRef\]](#)
- Goldinger, Stephen D. 1998. Echoes of Echoes? An Episodic Theory of Lexical Access. *Psychological Review* 105: 251–79. [\[CrossRef\]](#)
- González González, Manuel. 2008. O novo galego urbano. In *Cada palabra pesaba, cada palabra medía. Homenaxe a Antón Santamarina*. Edited by Mercedes Brea, Francisco Fernández Rei and Xosé Luís Regueira. Santiago de Compostela: Universidade de Santiago de Compostela, pp. 363–74.
- Hornsby, Michael. 2005. Néo-Breton and Questions of Authenticity. *Estudios de Sociolingüística* 6: 191–218. [\[CrossRef\]](#)
- Hornsby, Michael. 2009. Globalisation Processes and Minority Languages: Linguistic Hybridity in Brittany. Doctoral thesis, University of Southampton, Southampton, UK.

- Hornsby, Michael. 2015. The “new” and “traditional” Speaker Dichotomy: Bridging the Gap. *International Journal of the Sociology of Language* 231: 107–25. [CrossRef]
- Iverson, Paul, Patricia K. Kuhl, Reiko Akahane-Yamada, Eugen Diesch, Yoh’ich Tohkura, Andreas Kettermann, and Claudia Siebert. 2003. A Perceptual Interference Account of Acquisition Difficulties for Non-Native Phonemes. *Cognition* 87: 47–57. [CrossRef]
- Jaeger, T. Florian, and Kodi Weatherholtz. 2016. What the Heck Is Salience? How Predictive Language Processing Contributes to Sociolinguistic Perception. *Frontiers in Psychology* 7: 1–5. [CrossRef]
- Jaffe, Alexandra. 2015. Defining the New Speaker: Theoretical Perspectives and Learner Trajectories. *International Journal of the Sociology of Language* 231: 21–44. [CrossRef]
- Johnson, Elizabeth K., Ellen Westrek, Thierry Nazzi, and Anne Cutler. 2011. Infant Ability to Tell Voices Apart Rests on Language Experience. *Developmental Science* 14: 1002–11. [CrossRef] [PubMed]
- Johnson, Elizabeth K., Laurence Bruggeman, and Anne Cutler. 2018. Abstraction and the (Misnamed) Language Familiarity Effect. *Cognitive Science* 42: 633–45. [CrossRef] [PubMed]
- Johnstone, Barbara, and Scott F. Kiesling. 2008. Indexicality and Experience: Exploring the Meanings of /Aw/-Monophthongization in Pittsburgh. *Journal of Sociolinguistics* 12: 5–33. [CrossRef]
- Johnstone, Barbara, Jennifer Andrus, and Andrew E. Danielson. 2006. Mobility, Indexicality, and the Enregisterment of “Pittsburghese”. *Journal of English Linguistics* 34: 77–104. [CrossRef]
- Jones, Zack, Qingyang Yan, Laura Wagner, and Cynthia G. Clopper. 2017. The Development of Dialect Classification across the Lifespan. *Journal of Phonetics* 60: 20–37. [CrossRef]
- Labov, William. 1972. *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.
- Labraña Barrero, Sabela. 2009. Las consonantes fricativas de la lengua gallega. *Estudios de Fonética Experimental* 18: 193–213.
- Labraña Barrero, Sabela. 2014. Parámetros acústicos dos sons fricativos da lingua galega. *Estudios de Fonética Experimental* 23: 203–43.
- Lambert, Wallace E., Richard C. Hodgson, Robert C. Gardner, and Samuel Fillenbaum. 1960. Evaluational Reactions to Spoken Languages. *The Journal of Abnormal and Social Psychology* 60: 44–51. [CrossRef] [PubMed]
- Lenth, Russell V. 2016. Least-Squares Means: The R Package lsmeans. *Journal of Statistical Software* 69: 1–33. [CrossRef]
- Levi, Susannah V. 2019. Methodological Considerations for Interpreting the Language Familiarity Effect in Talker Processing. *WIREs Cognitive Science* 10: 1–15. [CrossRef] [PubMed]
- Levon, Erez. 2014. Categories, Stereotypes, and the Linguistic Perception of Sexuality. *Language in Society* 43: 539–66. [CrossRef]
- Mayr, Robert, Laura López-Bueno, Martín Vázquez Fernández, and Gisela Tomé Lourido. 2019. The Role of Early Experience and Continued Language Use in Bilingual Speech Production: A Study of Galician and Spanish Mid Vowels by Galician-Spanish Bilinguals. *Journal of Phonetics* 72: 1–16. [CrossRef]
- Mayr, Robert, Llian Roberts, and Jonathan Morris. 2020. Can You Tell by Their English If They Can Speak Welsh? Accent Perception in a Language Contact Situation. *International Journal of Bilingualism* 24: 740–66. [CrossRef]
- McLeod, Wilson, and Bernadette O’Rourke. 2015. ‘New Speakers’ of Gaelic: Perceptions of Linguistic Authenticity and Appropriateness. *Applied Linguistics Review* 6: 151–72. [CrossRef]
- Montgomery, Chris, and Emma Moore. 2018. Evaluating S(c)illy Voices: The Effects of Salience, Stereotypes, and Co-Present Language Variables on Real-Time Reactions to Regional Speech. *Language* 94: 629–61. [CrossRef]
- Nance, Claire. 2013. *Phonetic Variation, Sound Change, and Identity in Scottish Gaelic*. Doctoral Thesis, University of Glasgow, Glasgow, Scotland. [CrossRef]
- Nance, Claire. 2015. ‘New’ Scottish Gaelic Speakers in Glasgow: A Phonetic Study of Language Revitalisation. *Language in Society* 44: 553–79. [CrossRef]
- Nance, Claire, Wilson McLeod, Bernadette O’Rourke, and Stuart Dunmore. 2016. Identity, Accent Aim, and Motivation in Second Language Users: New Scottish Gaelic Speakers’ Use of Phonetic Variation. *Journal of Sociolinguistics* 20: 164–91. [CrossRef]
- Nazzi, Thierry, Peter W. Jusczyk, and Elizabeth K. Johnson. 2000. Language Discrimination by English-Learning 5-Month-Olds: Effects of Rhythm and Familiarity. *Journal of Memory and Language* 43: 1–19. [CrossRef]
- Neofalante, O. 2021. Available online: <https://twitter.com/unneofalante> (accessed on 15 March 2021).
- Niedzielski, Nancy. 1999. The Effect of Social Information on the Perception of Sociolinguistic Variables. *Journal of Language and Social Psychology* 18: 62–85. [CrossRef]
- Nycz, Jennifer. 2016. Awareness and Acquisition of New Dialect Features. In *Awareness and Control in Sociolinguistic Research*. Edited by Anna M. Babel. Cambridge: Cambridge University Press, pp. 62–79. [CrossRef]
- Nygaard, Lynne C., and David B. Pisoni. 1998. Talker-Specific Learning in Speech Perception. *Perception and Psychophysics* 60: 355–76. [CrossRef]
- Ó hIfearnaín, Tadhg. 2015. Sociolinguistic Vitality of Manx after Extreme Language Shift: Authenticity without Traditional Native Speakers. *International Journal of the Sociology of Language* 231: 45–62. [CrossRef]
- O’Rourke, Bernadette, and Fernando Ramallo. 2010. Los nuevos hablantes de lenguas minoritarias: Una comparación entre Irlanda y Galicia. Paper present at the Actas del IX Congreso Internacional de Lingüística General, Valladolid, Spain, June 21–23; pp. 1684–700.
- O’Rourke, Bernadette, and Fernando Ramallo. 2011. The Native-Non-Native Dichotomy in Minority Language Contexts: Comparisons between Irish and Galician. *Language Problems and Language Planning* 35: 139–59. [CrossRef]

- O'Rourke, Bernadette, and Fernando Ramallo. 2013a. "A miña variedade é defectuosa": A lexitimidade social das neofalas. *Estudos de lingüística galega* 5: 89–103.
- O'Rourke, Bernadette, and Fernando Ramallo. 2013b. Competing Ideologies of Linguistic Authority amongst New Speakers in Contemporary Galicia. *Language in Society* 42: 287–305. [CrossRef]
- O'Rourke, Bernadette, and Fernando Ramallo. 2015. Neofalantes as an Active Minority: Understanding Language Practices and Motivations for Change amongst New Speakers of Galician. *International Journal of the Sociology of Language* 231: 147–65. [CrossRef]
- O'Rourke, Bernadette, and John Walsh. 2015. New Speakers of Irish: Shifting Boundaries across Time and Space. *International Journal of the Sociology of Language* 231: 63–83. [CrossRef]
- O'Rourke, Bernadette, Joan Pujolar, and Fernando Ramallo. 2015. New Speakers of Minority Languages: The Challenging Opportunity—Foreword. *International Journal of the Sociology of Language* 231: 1–20. [CrossRef]
- Ortega, Ane, Estibaliz Amorrortu, Jone Goirigolzarri, Jacqueline Urla, and Belen Uranga. 2014. New Basque Speakers: Linguistic Identity and Legitimacy. *Digithum* 16: 47–58. [CrossRef]
- Ortega, Ane, Jacqueline Urla, and Estibaliz Amorrortu. 2015. Linguistic Identity among New Speakers of Basque. *International Journal of the Sociology of Language* 231: 85–105. [CrossRef]
- Pallier, Christophe, Laura Bosch, and Núria Sebastián-Gallés. 1997. A Limit on Behavioral Plasticity in Speech Perception. *Cognition* 64: B9–B17. [CrossRef]
- Perrachione, Tyler K., Stephanie N. Del Tufo, and John D. E. Gabrieli. 2011. Human Voice Recognition Depends on Language Ability. *Science* 333: 5–8. [CrossRef]
- Pharao, Nicolai, Marie Maegaard, Janus Spindler Moller, and Tore Kristiansen. 2014. Indexical Meanings of [S+] among Copenhagen Youth: Social Perception of a Phonetic Variant in Different Prosodic Contexts. *Language in Society* 43: 1–31. [CrossRef]
- Preston, Dennis R. 1986. Five Visions of America. *Language in Society* 15: 221–40. [CrossRef]
- Preston, Dennis R. 1989. *Perceptual Dialectology*. Providence: Foris Publications.
- Preston, Dennis R. 1996. Where the Worst English Is Spoken. In *Focus on the USA*. Edited by E. W. Schneider. Amsterdam: John Benjamins, pp. 297–360.
- Preston, Dennis R. 1999. *Handbook of Perceptual Dialectology*. Amsterdam: John Benjamins.
- Pujolar, Joan, and Maite Puigdevall. 2015. Linguistic Mudes: How to Become a New Speaker in Catalonia. *International Journal of the Sociology of Language* 231: 167–87. [CrossRef]
- Purnell, Thomas, William J. Idsardi, and John Baugh. 1999. Perceptual and Phonetic Experiments on American English Dialect Identification. *Journal of Language and Social Psychology* 18: 10–30. [CrossRef]
- Qualtrics. 2015. Qualtrics. Provo, UT, USA. Available online: <http://www.qualtrics.com> (accessed on 8 December 2015).
- R Core Team. 2013. R: A Language and Environment for Statistical Computing. Available online: <http://www.r-project.org/> (accessed on 8 December 2015).
- Rácz, Péter. 2013. *Saliency in Sociolinguistics. A Quantitative Approach*. Berlin and New York: De Gruyter Mouton.
- Ramallo, Fernando. 2010. Cara a unha tipoloxía sociolingüística dos falantes de galego. In *Educación e linguas en Galicia*. Edited by Beito Silva Valdivia, Xesús Rodríguez Rodríguez and Isabel Vaquero Quintela. Santiago de Compostela: Universidade de Santiago de Compostela, pp. 15–37. [CrossRef]
- Ramallo, Fernando. 2013. Neofalantismo. In *Contacto de linguas, hibrididade, cambio: Contextos, procesos e consecuencias*. Edited by Eva Gugenberger, Henrique Monteagudo and Gabriel Rei-Doval. Santiago de Compostela: Consello da Cultura Galega and Instituto da Lingua Galega, pp. 245–58.
- Ramallo, Fernando, and Bernadette O'Rourke. 2014. Perfiles de neohablantes de gallego. *Digithum* 16: 98–105.
- Regueira, Xosé Luís. 1999a. Estándar oral e variación social da lingua galega. In *Cinguidos por unha arela común: Homenaxe ó profesor Xesús Alonso Montero*. Edited by Rosario Álvarez Blanco and Dolores Vilavedra Fernández. Santiago de Compostela: Universidade, pp. 855–75.
- Regueira, Xosé Luís. 1999b. Galician. *Journal of the International Phonetic Association* 26: 119–22. [CrossRef]
- Regueira, Xosé Luís. 2010. *Diccionario de pronuncia da lingua galega*. A Coruña: Real Academia Galega.
- Regueira, Xosé Luís. 2012. Autenticidade e calidade da lingua: Purismo e planificación lingüística no galego actual. *Estudos de lingüística galega* 4: 187–201. [CrossRef]
- Regueira, Xosé Luís. 2014. Mantenimiento y cambio lingüístico: Fricativas sibilantes en gallego. *Revista de Filología Románica* 31: 151–68. [CrossRef]
- Regueira, Xosé Luís. 2019. Variación fonética, a lingua da esfera pública e estándar oral: Entre a lingüística e a política. *Revista Galega de Filoloxía* 20: 119–47. [CrossRef]
- Regueira, Xosé Luís, and Elisa Fernández Rei. 2020. The Spanish Sound System and Intonation in Contact with Galician. In *Spanish Phonetics and Phonology in Contact. Studies from Africa, the Americas, and Spain*. Edited by Rajiv Rao. Amsterdam and Philadelphia: John Benjamins Publishing Company, pp. 328–62. [CrossRef]
- Regueira, Xosé Luís, and María José Ginzo. 2018. A Cross-Linguistic Study of Voiceless Fricative Sibilants in Galician and European Portuguese. In *Romance Phonetics and Phonology*. Edited by Mark Gibson and Juana Gil. Oxford: Oxford University Press, pp. 62–76. [CrossRef]
- Regueira, Xosé Luís, Francisco Dubert García, Edelmira Parga Valiña, and Xulio C. Sousa Fernández. 1998. *Os sons da lingua*. Vigo: Edicións Xerais de Galicia.



- Robert, Elen. 2009. Accommodating “new” Speakers? An Attitudinal Investigation of L2 Speakers of Welsh in South-East Wales. *International Journal of the Sociology of Language* 195: 93–116. [[CrossRef](#)]
- Scharinger, Mathias, Philip J. Monahan, and William J. Idsardi. 2011. You Had Me at “Hello”: Rapid Extraction of Dialect Information from Spoken Words. *NeuroImage* 56: 2329–38. [[CrossRef](#)] [[PubMed](#)]
- Silverstein, Michael. 2003. Indexical Order and the Dialectics of Sociolinguistic Life. *Language and Communication* 23: 193–229. [[CrossRef](#)]
- Strand, Elizabeth A. 1999. Uncovering the Role of Gender Stereotypes in Speech Perception. *Journal of Language and Social Psychology* 18: 86–100. [[CrossRef](#)]
- Strand, Elizabeth A., and Keith Johnson. 1996. Gradient and Visual Speaker Normalization in the Perception of Fricatives. In *Natural Language Processing and Speech Technology. Results of the 3rd KOVENS Conference, Bielefeld, Germany, October 1996*. Edited by Dafydd Gibbon. Berlin: Mouton de Gruyter, pp. 14–26. [[CrossRef](#)]
- Tan, Ying Ying. 2012. Age as a Factor in Ethnic Accent Identification in Singapore. *Journal of Multilingual and Multicultural Development* 33: 569–87. [[CrossRef](#)]
- Thompson, Charles P. 1987. A Language Effect in Voice Identification. *Applied Cognitive Psychology* 1: 121–31. [[CrossRef](#)]
- Tomé Lourido, Gisela, and Bronwen G. Evans. 2015. Switching Language Dominance for Ideological Reasons: A Study of Galician New Speakers’ Speech Production and Perception. In *Proceedings of the International Congress of Phonetic Sciences*. Glasgow: University of Glasgow.
- Tomé Lourido, Gisela, and Bronwen G. Evans. 2017. Os neofalantes galegos no contexto europeo: A produción e a percepción da fala. In *Cambio lingüístico no galego actual*. Edited by Xosé Luís Regueira and Elisa Fernández Rei. Santiago de Compostela: Instituto da Lingua Galega and Consello da Cultura, pp. 99–127.
- Tomé Lourido, Gisela, and Bronwen G. Evans. 2019. The Effects of Language Dominance Switch in Bilinguals: Galician New Speakers’ Speech Production and Perception. *Bilingualism: Language and Cognition* 22: 637–54. [[CrossRef](#)]
- Tuninetti, Alba, Kateřina Chládková, Varghese Peter, Niels O. Schiller, and Paola Escudero. 2017. When Speaker Identity Is Unavoidable: Neural Processing of Speaker Identity Cues in Natural Speech. *Brain and Language* 174: 42–49. [[CrossRef](#)] [[PubMed](#)]
- Vidal Figueroa, Tiago. 1997. Estructuras fonéticas de tres dialectos de Vigo. *Verba* 24: 313–32.
- Wagner, Laura, Cynthia G. Clopper, and John H. Pate. 2014. Children’s Perception of Dialect Variation. *Journal of Child Language* 41: 1062–84. [[CrossRef](#)] [[PubMed](#)]
- Walsh, John, and Bernadette O’Rourke. 2014. Becoming a New Speaker of Irish: Linguistic Mudes throughout the Life Cycle. *Digithum* 16: 67–74. [[CrossRef](#)]
- Woolard, Kathryn. 2011. Is There Linguistic Life after High School? Longitudinal Changes in the Bilingual Repertoire in Metropolitan Barcelona. *Language in Society* 40: 617–48. [[CrossRef](#)]

Article

# Adult New Speakers of Welsh: Accent, Pronunciation and Language Experience in South Wales

Meinir Williams \* and Sarah Cooper

College of Arts, Humanities and Business, Bangor University, Bangor LL57 2DG, UK; s.cooper@bangor.ac.uk  
\* Correspondence: weu423@bangor.ac.uk

**Abstract:** This study examines the experiences of adult new speakers of Welsh in Wales, UK with learning pronunciation in Welsh. Questionnaire data were collected from 115 adult L2 speakers with English as an L1 located in South Wales. We investigated self-reported perceptions of accent and pronunciation as well as exploring which speech sounds were reported to be challenging for the participants. We also asked participants how traditional native speakers responded to them in the community. Perceptions of own accent and pronunciation were not rated highly for the participants. We found that speaker origin affected responses to perceptions of accent and pronunciation, as well as speaker learning level. In terms of speech sounds that are challenging, the results show that vowel length as well as the consonants absent in the L1 (English) were the most common issues reported. A range of responses from traditional native speakers were reported, including speaking more slowly, switching to English, correcting pronunciation or not responding at all. It is suggested that these results indicate that adult new speakers of Welsh face challenges with accent and pronunciation, and we discuss the implications of this for language teaching and for integration into the community.

**Keywords:** second language acquisition; speech production; accent; pronunciation; new speakers; minority language bilingualism



**Citation:** Williams, Meinir, and Sarah Cooper. 2021. Adult New Speakers of Welsh: Accent, Pronunciation and Language Experience in South Wales. *Languages* 6: 86. <https://doi.org/10.3390/languages6020086>

Academic Editors: Robert Mayr and Jonathan Morris

Received: 5 March 2021  
Accepted: 11 May 2021  
Published: 13 May 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

In the field of L2 pronunciation teaching and learning, intelligibility and comprehensibility are often regarded the most important considerations (Derwing and Munro 2015). Whilst models of L2 speech learning (Best 1995; Tyler 2019; Flege 1995; Escudero and Boersma 2004; Van Leussen and Escudero 2015) do allow for 'native like' production by individuals who learn a language during adulthood, this is seen as an unusual outcome rather than the expectation for most individuals (Best and Tyler 2007). As such, direct imitation of the accents, dialects or other pronunciation traits of native speakers is not seen as being particularly desirable for L2 speakers, especially in contexts where they would be using the target language mostly with other non-native speakers (Jenkins 2002). However, whilst this may be true in the case of international languages, such as English, there is evidence that this is not the case for smaller, regional or minority languages. New speakers of many, such as Breton (Hornsby 2015a), Sami (Jonsson and Rosenfors 2017), and Corsican (Jaffe 2013), face the additional challenges of navigating the questions of identity and legitimacy posed by the perceived 'nativeness' in their speech. Indeed, the concept of new speakerhood is intrinsically linked to the context of minority language endangerment and the will to protect and transmit these languages.

The issue of the perceived degree of 'nativeness' is reflected in the terms used to describe speakers and their roles in these contexts. Research in the field of language acquisition and research regarding widely spoken, global languages, such as English and Spanish, refer to first and second language (L1 and L2) or 'native' and 'non-native' speakers. This is viewed as problematic in the context of minority languages (O'Rourke and Pujolar 2013) as it is considered that these terms encourage the perception of a hierarchy of speaker legitimacy, which undermines individuals who have acquired the language

outside familial settings (Hornsby 2019). The terms ‘new’ and ‘traditional’ speakers are used to better reflect the complex social realities attached to speakerhood of non-majority languages. Hornsby (2015b) states that new speakers are often defined as individuals who have attained the minority language through formal education, as opposed to familial transmission, which means that every new speaker has invested their time and resources to attain language. This does not mean that there are not those who acquire minority languages through community transmission or any other linguistic context, but these are exceptions and most have made significant efforts.

The current study considers the experiences of individuals who are, or have, acquired the language as adults. This is opposed to individuals who are acquiring Welsh through Welsh medium education at primary or secondary levels (Mayr et al. 2017), or those learning Welsh in English medium primary or secondary schools (Selleck 2018). We choose to refer to individuals who have made this effort as ‘new’ speakers as opposed to ‘learners’. The term ‘new speaker’ has been used in the Welsh context, often to refer to individuals accessing the language through Welsh medium education (Selleck 2018; Robert 2009) but is arguably confined to academic discourse. The alternatives used in non-academic discourse in Wales, such as ‘dysgwyr’ (learners), can exclude competent users who are able and wish to use the language (Hornsby and Vigers 2018). We also use the terms L1 and L2 when discussing the process of acquiring language competence. Understanding the motivations and perceptions of individuals who often go to great lengths to attain a language is of vital importance to support language transmission outside of the classroom setting, both for new and traditional speakers.

### 1.1. The Welsh Context

Welsh is a Brittonic Celtic language spoken by 562,000 (19%) of Wales’ 3.1 million inhabitants (Welsh Government 2011). All speakers also speak English and both languages have equal legal status, with education at all levels, local and national government, and some media available in Welsh. There was a marked decline in speaker density during the 19th century, a period of rapid industrialization which led to great population growth and movement within the country, especially towards the coal mines of the southern valleys. These social changes, combined with the negative attitudes towards bilingualism commonly held at the time, led to many parents choosing not to transmit the language to their children. There are now many individuals who are reclaiming the language despite the broken chain of transmission within their families. Hodges (2010) refers to this as the ‘cenhedlaeth goll’ (lost generation) of speakers and in the context of individuals who desire to interact with the Welsh language and how this can influence choices regarding Welsh medium education.

There is variety in terms of speaker density within the population in different areas, with as much as 80% of some communities speaking Welsh in the north West, whilst many communities in the south and east have considerably lower percentages. This means that new speakers in the various regions can face different challenges in their linguistic journeys. Whilst an individual in Gwynedd, a county in the north west with the highest percentage of Welsh speakers, might expect to find the language in daily use in the community, individuals in the more urban southern regions would have to seek out opportunities to gain access to the language. The present study focuses on individuals currently living in South Wales. The average population density over the 14 counties of South Wales reporting ability to speak Welsh was 13.6%, ranging from 43.9% in the county of Carmarthenshire to 7.8% in the county of Blaenau Gwent (Welsh Government 2011). This suggests that opportunities for Welsh to be transmitted within the community may be more limited in South Wales and that individuals who choose to acquire the language may be forced to search for more formalised opportunities to use Welsh.

In recent years, there has been an increase in the interest and support for individuals choosing to learn Welsh as adults. These individuals residing in Wales may come from Wales or from further afield. Individuals who were educated in Wales will almost always

have had some Welsh lessons in school as a second language. Therefore, speakers of Welsh English who are learning Welsh as adults are likely to have had some contact with the language in childhood whilst those from outside of Wales often have no previous experience. In 2017, the Welsh Government published the current Welsh language strategy, *Cymraeg 2050: A million Welsh speakers*, which outlines the aim to increase the number of speakers of Welsh to one million by 2050. The Welsh Government recognises the importance and contribution of the Welsh for Adults sector in achieving opportunities for adults to learn Welsh and improve their Welsh language skills. Data from providers of language teaching, such as the National Centre for Learning Welsh, indicate that growing numbers of people are choosing to learn, with 13,260 individuals accessing their Welsh for Adults provision in 2018–2019 (National Centre for Learning Welsh 2020). Online providers, such as Duolingo, have also noted a marked increase in uptake, with Welsh being the fastest growing language on the platform (Watkins 2020). Understanding experiences of learning pronunciation, as well as how new speakers interact with traditional speakers, will not only allow providers to tailor their teaching to the needs of their clients but also allow for greater understanding of the challenges that individuals face with integrating into communities.

### 1.2. *Sounding 'Native'*

Passing for a native speaker is often considered to be the ultimate mark of success in language learning, and pronunciation is one of the most obvious elements (Gnevshева 2017). However, it is commonly accepted that native-like acquisition is not the norm for most individuals learning languages, especially those learning as adults. Whilst there is evidence that some individuals do indeed acquire 'native-like' proficiency (Birdsong 2003) these are generally accepted to be the exception and not the rule. Most models of second language perception and production do allow for the possibility of 'native-like' perception and production in the target language despite cross-linguistic influences from the first language. The Second Language Linguistic Perception model (Escudero 2005; Van Leussen and Escudero 2015) also posits that dialects within the L1 can influence the perception of the target language, meaning that individuals who speak the same language might face differing challenges, especially in terms of vowel perception and production (Escudero and Williams 2012). This is of particular interest in the context of Welsh new speakers as the linguistic background of the individuals choosing to learn and use the language can vary greatly, from Welsh English to other varieties of British English (Müller and Ball 1999), to an array of languages from all around the world.

In addition, the concept of adopting the accents and pronunciation of traditional speakers may not be desirable for some new speakers, given that they have connotations of links to geographical locations and socio-economic status. In many instances, traditional varieties of minority languages are linked to rural life, with the speech of older, traditional speakers being viewed by some as more 'authentic' (Bucholtz 2003). This concept of an inherent bond between specific ideals of language, history and location may not reflect the modern, urban realities of many new speakers across many language contexts. This was reported for Scottish Gaelic (McLeod and O'Rourke 2015; Nance et al. 2016), where individuals displayed very different aspirations in terms of constructing speaker identities around more or less traditional varieties. In both papers, whilst some speakers felt strongly that they wished to use localised dialects, especially those used by family members, others felt that they had no desire to emulate the speech of traditional linguistic heartlands to which they had no connection. Trosset (1986) observed similar attitudes by some new speakers of Welsh, especially those who had no connection to particular areas, who felt that using more traditional varieties could become performative rather than allowing for greater ease of communication.

This calls into question the assumption that 'nativeness' is the ultimate goal in minority language learning. Whilst some do aspire to pass as native speakers, assuming that this is a universal phenomenon for all new speakers risks adding pressure to conform with identities that may not feel relevant to their experiences.

### 1.3. Sounding ‘Welsh’

Despite having a standard written form, Welsh does not have a standard spoken variety, and as such, the phonetic inventory of the language varies between dialects (Ball and Williams 2001; Awbery 2009). These are generally discussed as northern and southern varieties, though there are numerous local variants within this divide. While this study considers individuals currently living in South Wales, we provide a general overview of the phonetic inventory of northern and southern varieties as it is possible that some tutors use northern variants when teaching, and there are some aspects of vowel pronunciation that differ between the varieties that can be ambiguous to those learning the language.

Welsh has 29 distinctive consonants (Ball and Jones 1984; Ball and Williams 2001; Hannahs 2013). These include the plosives /b, d, g, p, t, k/, the nasals /m, n, ŋ/, the trills /r, r<sup>h</sup>/, the fricatives /f, v, θ, ð, s, ʃ, χ, h/. The lateral fricative /ɬ/, the approximant /j/ and the lateral approximant /l/ are also present. The affricates /tʃ/ and /dʒ/ are present in some loan words. Welsh has initial consonant mutation in certain lexical and syntactic environments, and the voiceless nasals /m<sup>h</sup>, n<sup>h</sup>, ŋ<sup>h</sup>/ occur in word initial position in nasal mutated forms. The orthography of Welsh consonants is generally agreed to be fairly transparent, and the consonants are similarly pronounced across dialects, though the identity of certain phonemes in southern dialects is contested (Ball and Williams 2001). The identity of the uvular fricative, for example, /χ/, has been argued by some to be velar in some varieties (Fynes-Clinton 1913; Watkins 1961). Generally, there is a 1:1 relationship between the orthographic representation of consonants and their pronunciation, with most phonemic units being represented by a single grapheme. However, certain phonemic units are represented with digraphs, such as <ll> and <ch>, representing /tʃ/ and /χ/.

The vowels of Welsh vary more between dialects, though there are many similarities (Ball and Williams 2001; Hannahs 2013; Mayr and Davies 2011). All monophthongs, except for schwa, have both a short and long form. As shown in Table 1 below, southern varieties contain 11 monophthongs, while northern varieties have an additional pair /i, i:/. This difference influences the number of diphthongs in both dialects: 13 for northern varieties and 8 for southern varieties. Earlier accounts (Ball and Williams 2001) suggested that the northern pairs are only differentiated in length whilst the southern pairs differ in terms of vowel quality as well as duration. However, acoustic analyses of the vowel inventories of both dialects (Mayr and Davies 2011) suggest that these distinctions are present in both varieties.

When it comes to orthography, in minimal pairs where vowel length is the differentiating factor, e.g., ‘tan’ (until) and ‘tân’ (fire), the circumflex is used to indicate the long vowel. However, when there is no minimal pair, there is no indication of vowel length, e.g., ‘nos’ (night) having a long vowel /no:s/ but no orthographic information to show this. There are several rules which dictate vowel length in various contexts, but this is not always reflected in the orthography (Morris-Jones et al. 1928; Ball and Williams 2001). This is confounded by the vowels in certain words, e.g., ‘pell’ (far) being pronounced differently according to differing dialects. The standard pronunciation in northern varieties would feature a short vowel [peɨ], but speakers of southern varieties would more likely produce [pe:ɨ]. The orthography, therefore, gives much less information regarding the pronunciation of monophthongs than consonants and could potentially pose difficulty to new speakers, particularly if they lack access to spoken input.

Diphthongs are represented by a combination of the two vowel graphemes but are not considered as unit graphemes in the way that consonants, such as <ch> and <ll>, are. Some diphthongal graphemes have more than one realisation. For example, <wy> represents /wɨ/ but also a combination of consonant plus vowel /wɨ/ or /wə/ and <yw> represents /ɨw/ and in some cases /əw/ (Ball and Williams 2001). The ambiguities around vowel length and diphthong identity are considered in this study as they may pose challenges for new speakers due to the fact that there are several options for pronunciation, especially as these are not present in English L1.

**Table 1.** The vowels of northern and southern Welsh. Adapted from [Mayr and Davies \(2011\)](#).

Northern Welsh		Southern Welsh	
Short vowels	Long vowels	Short vowels	Long vowels
i i: u	i: i: u:	ɪ ʊ	i: u:
e ə o	ə: ɔ:	ɛ ə ɔ	e: ɔ:
a	a:	a	a:
13 diphthongs		8 diphthongs	
Front closing	Back closing	Front closing	Back closing
aɪ ɔɪ eɪ	ɪʊ ʊʊ aʊ əʊ iʊ	aɪ ɔɪ ʊɪ eɪ	ɪʊ ɛʊ aʊ əʊ
Central closing			
aɨ aɨ ɔɨ ʊɨ əɨ			

There is evidence of the influence of long-term language contact on the pronunciation of Welsh and English in Wales. There have been several investigations of language contact and (the lack of) differences between the pronunciation of bilingual speakers of Welsh and English, compared to monolingual speakers of English who live in the same communities. For example, [Mayr et al. \(2017\)](#) investigated the effects of long-term language contact on the production of monophthongs by bilingual and monolingual speakers in south Wales. They found strong convergence between both languages and that the majority of cross-linguistic vowel pairs were produced identically in Welsh and English by the participants. This convergence has also been reported in prosodic features by [Mennen et al. \(2020\)](#) who investigated the production of lexical stress in the same speakers as [Mayr et al. \(2017\)](#). They reported that Welsh and Welsh English have become alike in their realisation of lexical stress which they suggest indicates convergence between the two languages in the community. This convergence in segmental and prosodic features is important for this study because we investigate people learning Welsh who are originally from Wales and have lived in the community for many years, as well as participants with English as an L1 from elsewhere in the British Isles. This may have implications for participants’ perception of their own pronunciation in the present study.

The production of vowels was acoustically investigated in adults learning Welsh by [Müller and Ball \(1999\)](#). They compared the production of Welsh and English monophthongs and diphthongs by the individuals attending a Welsh language class in South Wales and their tutor. They found that speakers of ‘Welsh English’ were more likely to produce vowels similar to those of the tutor, a native speaker of Welsh, whilst those who spoke ‘non-Welsh English’ showed a tendency to diphthongise monophthongs. As mentioned above, we are interested in the question of speaker origin in the present study, aiming to investigate whether speakers from Wales who have decided to learn Welsh have differing perceptions of their own accent and pronunciation to speakers who have moved to South Wales from further afield.

[Rees and Morris \(2018\)](#) considered the challenge of pronunciation from the point of view of Welsh for Adults tutors. By using a questionnaire and conducting focus groups they were able to gain understanding of the elements of pronouncing Welsh which tutors saw as challenging for individuals at various points of their language learning trajectories. The input from language tutors is of particular interest, as these individuals teach a wide range of students and are often very experienced, but generally have very little phonetic training. The tutors generally agreed that the vowel sounds posed greater challenges in the longer term, especially the diphthongs. In addition, they suggest that further pronunciation training across all levels, from beginners to advanced speakers, could be beneficial. However, they also note that developing opportunities for new and traditional speakers to interact within their communities is vital.

#### 1.4. Interactions between New and Traditional Speakers

The interactions between new and traditional speakers of minority languages are of great importance when considering the trajectories of new speakers. There is evidence of hierarchization, and even friction, between these groups in many settings, including Breton (Hornsby 2019) and Scottish Gaelic (McLeod et al. 2014; McEwan-Fujita 2010). In some cases, it is even claimed that the differences between the varieties spoken by both groups are so vast that most new speakers do not understand traditional varieties (Hewitt 2020), but this is strongly refuted by others, who view this as “actively [engaging] in the creation of such divisions, through a rhetoric of failure to reach ‘authentic speakerhood’” (Hornsby 2019, p. 395). In the case of Welsh, the experiences of individuals starting to learn Welsh by attending Welsh for Adults courses for beginners in North Wales have been examined from different perspectives in two large-scale studies. Baker et al. (2011) considered the implications of differing motivation to learn in the context of language planning, noting that individuals who had long-term integrative motivation, such as wanting to speak Welsh with their children, generally had more favourable outcomes in terms of language learning. This was echoed by Andrews (2011) who saw that integrative motivations, such as using the language in the community, were considered to be of greater importance than instrumental reasons, such as improved employment opportunities. It is, however, worth remembering that these studies were carried out in areas where there are more opportunities to use Welsh in the community than the areas considered in the current study.

Indeed, in areas where the language is only spoken by a small percentage of the population, access and opportunities to use the language are often considered to pose some of the greatest challenges. Mac Giolla Chríost et al. (2012) suggest that this lack of opportunities to communicate is partly due to the tendency of traditional speakers to switch to English when speaking with new speakers. This may happen for many reasons, which can be positive in their intention, e.g., wanting to facilitate communication (Trosset 1986), but is discouraging for individuals wishing to integrate into Welsh-speaking communities. This challenge is even greater in communities where the language is not used by the majority of the population as casual contact with traditional speakers, e.g., interactions in shops or cafés, is unlikely. This would suggest that individuals are forced to search for more formal opportunities to interact with other speakers. However, there is very little evidence available directly from the new speakers in communities in South Wales that are not in the traditional heartlands of the language. Hornsby and Vigers (2018) investigated the experiences of five ‘new’ speakers in the traditional Welsh-speaking areas of North Pembrokeshire and South Ceredigion. They found that, despite having attended Welsh-medium education and having a high level of competence in the language, the new speakers did not always feel that traditional speakers treated them as valid speakers and therefore that “a linguistic repertoire that includes Welsh competence does not automatically confer legitimacy as a speaker.” (p. 425) Indeed, they mentioned that they encountered individuals who chose not to speak Welsh with them, despite being aware of their ability to use the language. This raises questions regarding the ways in which new speakers define themselves within their communities, but also the ways in which the communities define them.

#### 1.5. The Present Study

There are many considerations when discussing adult new speakers’ accent and pronunciation as they are linked to questions of identity. Whilst studies have been conducted regarding the perception of language tutors of the challenges faced by individuals in their classes, no large-scale study has been conducted with the adult new speakers themselves. Therefore, very little is known about adult new speakers’ perception of their own pronunciation. Studies in minority language teaching and learning are moving away from the ‘native’ speaker as a model (O’Rourke and Pujolar 2013). It is still assumed that more

“native-like” pronunciation would lead to increased confidence in speaking and using Welsh, despite new speakers not having been surveyed.

Based on the considerations presented above, the present study sought to investigate aspects of adult new speakers’ perceptions of their accent and pronunciation in Welsh. Specifically, we sought to investigate whether there is a relationship between speaker background and the stage of their learning journey in perceptions of their own pronunciation in Welsh. Based on previous work on contact between Welsh and English in Wales, as well as the findings of Müller and Ball (1999), we are interested in comparing the experiences of participants with English as an L1 from Wales with speakers from elsewhere in the British Isles. We also sought to identify some of the specific segmental aspects that adult new speakers perceive as challenging. Finally, we sought to investigate the reported responses of traditional ‘native’ speakers of Welsh to L2 accent and pronunciation encountered by participants learning or who have learnt Welsh.

## 2. Method

An online questionnaire was implemented, which asked adult new speakers of Welsh about the challenges that they face in learning to pronounce and use Welsh as new speakers in their communities.

### 2.1. Participants

114 (77 female (68.4%); 36 male (31.6%)) adult new speakers completed an online questionnaire about their experiences of learning and using Welsh. They were recruited online through Welsh language social media, advertising by language course providers and by word of mouth. This allowed for individuals who have a wide range of contact with the Welsh language.

All were living in South Wales at the time of responding, with 55 originally from Wales (47.8%) and 59 from outside of Wales (52.2%) (England  $n = 53$ , Scotland,  $n = 3$  Northern Ireland = 2, Republic of Ireland  $n = 1$ ). The participants’ ages ranged from 21–82 ( $M = 52.65$ ,  $SD = 15.37$ ). All reported being L1 speakers of English and 65 participants reported that they were familiar with languages other than Welsh and English. When asked to describe their abilities in Welsh 52 (49.56%) identified as ‘beginners’, 37 (33.04%) as ‘intermediate speakers’, 18 (15.65%) as ‘advanced speakers’ and 7 (6.09%) as ‘fluent speakers’. Twenty-one speakers had been learning Welsh for less than 1 year, 48 had been learning Welsh for between 1 and 4 years, 20 for between 5 and 10 years and 25 for greater than 11 years. Fourteen participants reported having started learning before the age of 10, but six of these individuals reported a prolonged period of not interacting with the language and noted that they had been learning for 2 years or less. None of the 14 reported being fluent speakers.

### 2.2. The Questionnaire

The questionnaire was designed using JISC Online surveys to enable the collection of a broad range of qualitative and quantitative data about the experiences of L2 speakers of Welsh. The questionnaire design was influenced by earlier research in the field, both in terms of speaker identity and pronunciation. The questionnaire was presented bilingually, and participants were free to answer in either language, or a combination of both. A variety of types of questions were used, from open-ended questions allowing free text answers to statements that were responded to on Likert scales. This was done to collect a broad range of data in a relatively short amount of time, but also to allow for the expression of experiences and attitudes that would not be possible in closed questions alone.

Firstly, we asked participants to respond to a series of statements on a 9-point Likert scale about whether individuals sound like, or wish to sound like, ‘native’ speakers. We also asked whether speakers were proud of their accent, and whether they wanted to change their accent. By considering perceptions of accent as well as the perceived difficulty of various speech sound, we are able to see general trends of self-evaluation. Full details of



these items can be found in Table A1 in the Appendix A. Whilst this method is notorious for providing unreliable data in terms of pronunciation accuracy (Mitterer et al. 2020) it reveals how individuals feel about their own difficulties.

The section of the questionnaire that asked about specific speech sounds that were challenging was designed to mirror the questions asked by Rees and Morris (2018) of Welsh for Adults tutors. By considering the similarities and differences in responses it is hoped that a fuller picture will emerge of the challenges of perception and pronunciation which new speakers face.

The participants were asked which sounds were difficult for them by responding to several examples of individual segments, presented in quotation marks and with example words, e.g., *Ynganu/pronouncing 'll' e.e. lle, pell* (full details of these items can be found in Table A2 in Appendix A). Participants could choose as many of the speech sounds that they found challenging. They were also encouraged to leave further comments regarding elements that they find, or found, challenging. Nine individual aspects of pronunciation were discussed. They were chosen based on the questions posed by Rees and Morris (2018) and their absence from Standard Southern British English. These included three consonants: the lateral fricative <ll>/ɬ/, the velar/uvular fricative <ch>/x/ or /χ/ and the trilled <r, rh>/r/. It should be noted that there are options for other r-realizations in Welsh. For example, intervocally, a tap [ɾ] would be expected. Some native speakers may also produce an approximant [ɹ] or a uvular trill [ʀ] (Morris 2013) but we focus on the “rolling” of the /r/ following Rees and Morris (2018). We were particularly interested in vowel length, as well as the vowel /o/ in word-final position. We also asked participants about their production of the diphthongs <ae>/aɪ/, <wy>/ʊɪ/ and <yw>/ɪʊ/. For vowel length, we asked participants about their ability to distinguish between long and short vowels with and without the circumflex, with examples to illustrate the distinction *tân/tan* and *bys/byr*.

Finally, we asked participants how traditional “native” speakers responded to their accent or pronunciation (Table A3 in Appendix A). Participants could tick as many options as were applicable and were asked to expand if they clicked “other”.

### 2.3. Analysis

The questionnaire featured several questions investigating Likert responses to questions about accent and pronunciation, as well as tick box questions about challenging pronunciation features and responses of ‘native’ speakers. In order to investigate the effect of speaker background in the analysis of the four questions about accent and pronunciation, we used statistical analysis. We were investigating a multinomial dependent variable with ordinal values from 1–9, where ‘9’ is higher than ‘8’ which is higher than ‘7’, etc. We conducted an Ordinal Logistic Regression which is specifically designed for ordinal data analysis (Baayen 2008, p. 208; Endresen and Janda 2016). This approach produces an analysis model, similar to a conventional multiple regression where there is one dependent variable and one or more independent variables. However, the dependent variable is treated as an ordered categorical variable and it is not assumed that there are equal intervals between categories on the response scale. In this analysis, we used the PLUM procedure in SPSS version 27. We report the Wald chi-square test statistics along with the corresponding *p*-values of the overall omnibus tests for each independent variable. In order to investigate how levels of the independent variables Origin (Welsh, Other) and Level (Beginner, Intermediate, Advanced, Fluent) contribute to the model, we also examine the coefficients (Appendix B, Tables A4–A7). We report deviance goodness-of fit test results for the models where a non-significant *p*-value indicates a good fit. We also report the pseudo R<sup>2</sup> values (Nagelkerke) to provide an indication of the overall performance of the model by indicating the proportion of the variance explained by the model.

For the questions about challenging features of pronunciation, we present percentage responses to tick box questions, and proportion of overall responses for the question about how native speakers respond for each level (Beginner, Intermediate, Advanced and Fluent).

### 3. Results

#### 3.1. Accent and Pronunciation

The participants were asked to rate how far they agreed or disagreed with four statements about their accent and pronunciation when speaking Welsh. We presented the four statements on a nine-point Likert scale from 1 = not at all, 9 = completely. We asked how far they agreed with the following statements:

- My pronunciation in Welsh is like a ‘native’ speaker’s.
- I’m proud of my accent when speaking Welsh.
- I want to change my accent in Welsh.
- I want to sound like a ‘native speaker’.

The median scores for each of the statements above is presented in Figure 1 below.

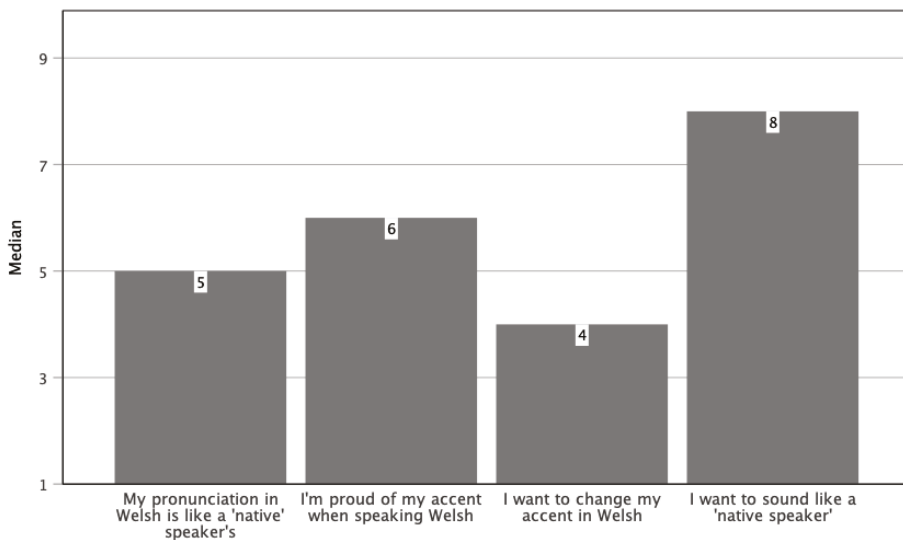


Figure 1. Median scores for statements on accent and pronunciation (1 = Not at all, 9 = Completely).

The median scores across the dataset illustrated in Figure 1 show that participants generally did not strongly agree that their pronunciation was like that of a native speaker, nor that they were particularly proud of their accent. Interestingly, however, participants tended to disagree with the statement “I want to change my accent in Welsh” but strongly agreed that they wanted to sound native. In order to investigate whether responses to these statements were affected by social variables, we explored the impact of four predicting factors.

In order to investigate the effect of speaker background variables on the responses presented above, we conducted ordinal logistic regression analyses on the four statements separately. We were primarily interested in whether participants from Wales had different perceptions of their accent and pronunciation as participants from outside Wales and included the variable Origin (Wales vs. other) in the models. As we had collected data from participants at different levels on their language learning journey, we also included Level in the models (Beginner, Intermediate, Advanced and Fluent) as well as gender (Male, Female) and age. We report the final and most optimal ordinal regression models below containing statistically significant predictors. In the ordinal regression analyses for all four statements, gender and age were not found to be significant predictors in the models. Figures 2 and 3 below illustrate the median rating for each question based on the predictors Origin and Level.

### 3.1.1. My Pronunciation in Welsh Is Like a ‘Native’ Speaker’s

The final and most optimal model included Origin and Level as statistically significant predictors of rating on sounding more like a native speaker (Origin Wald  $\chi^2(1) = 20.412$ ,  $p < 0.001$ ; Level Wald  $\chi^2(3) = 27.095$ ,  $p < 0.001$ ). The deviance goodness-of-fit test indicated that the model was a good fit to the observed data  $\chi^2(52) = 46.092$ ,  $p = 0.704$ . The pseudo R2 value for this model is 0.366, suggesting that the model explains around 36.6% of the variance in responses. Inspection of Figure 2 and inspection of the coefficients in Table A4 indicate that participants from Wales considered their pronunciation to be more similar to that of a native speaker than participants from outside Wales. For level, Figure 3 illustrates that, in general, Advanced and Fluent participants rated their pronunciation as more similar to that of a native speaker than the Beginner and Intermediate groups. The coefficients in Table A4 for Beginner, Intermediate and Advanced are negative, indicating that, compared the Fluent group, these groups rated their pronunciation as less native-like. Consultation of Table A4 suggests that, as the level increases, the likelihood of rating pronunciation to be more native-like increases.

### 3.1.2. I’m Proud of My Accent When Speaking Welsh

Both Origin and Level were statistically significant predictors for pride in accent (Origin Wald  $\chi^2(1) = 1.435$ ,  $p < 0.001$ ; Level Wald  $\chi^2(3) = 12.455$ ,  $p = 0.006$ ). The deviance goodness-of-fit test indicated that the model was a good fit to the observed data  $\chi^2(52) = 59.095$ ,  $p = 0.232$ . The pseudo R2 value for this model is 0.226, suggesting that the model explains around 22.6% of the variance in responses. Inspection of Figure 2 and the coefficients in Table A5 illustrates that participants from Wales were prouder of their accent than participants from outside Wales. For level, Figure 3 illustrates that participants of a higher level (Advanced and Fluent) were prouder of their accent than Beginner or Intermediate participants. This is confirmed in the coefficients in Table A4, which are negative and significant for Beginner and Intermediate compared to the Fluent Group. However, the comparison between the Advanced and the Fluent group is not significant. Overall, this suggests that participants of a higher level (Advanced and Fluent) were prouder of their accent when speaking Welsh than Beginner or Intermediate participants.

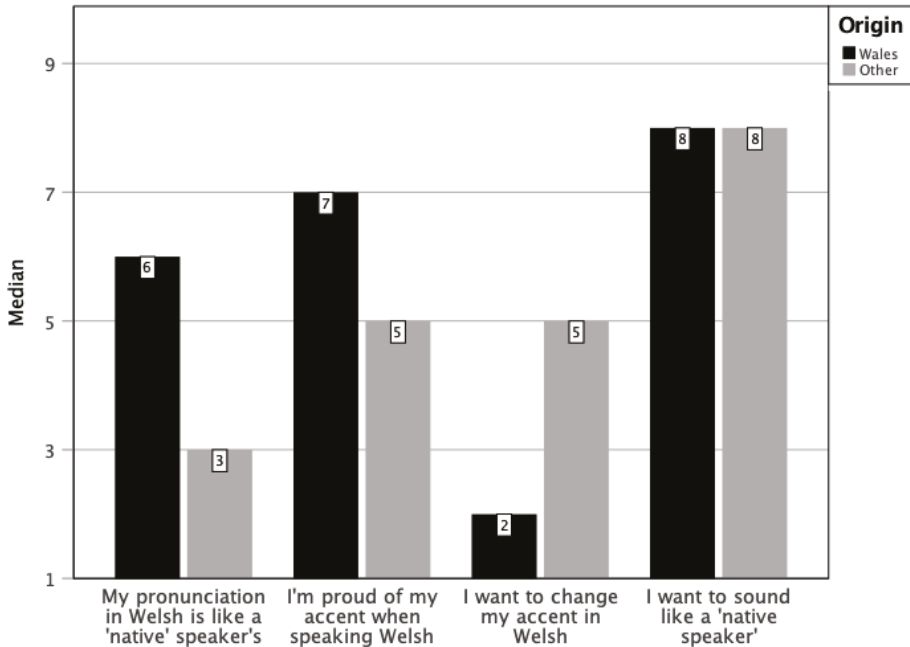
### 3.1.3. I Want to Change My Accent in Welsh

Origin and Level were also included as statistically significant predictors of wanting to change accent in Welsh (Origin Wald  $\chi^2(1) = 16.184$ ,  $p < 0.001$ ; Level Wald  $\chi^2(3) = 8.852$ ,  $p = 0.031$ ). The deviance goodness-of-fit test indicated that the model was a good fit to the observed data  $\chi^2(52) = 49.464$ ,  $p = 0.574$ . The pseudo R2 value for this model is 0.210, suggesting that the model explains around 21% of the variance in responses. Figure 2 illustrates that participants from outside of Wales were more likely to want to change their accent in Welsh than speakers who were from Wales, which is reflected by the positive coefficient in Table A6. The effect of level is illustrated in Figure 3, suggesting that lower-level participants were more likely to want to change their accent than higher level participants. Inspection of Table A6 indicates positive and significant coefficients for Beginner and Intermediate groups compared to the Fluent group, indicating that Beginner and Intermediate groups were more likely to agree that they wanted to change their accent than the Fluent group. The comparison between the Advanced and Fluent group is not significant. Taken with Figure 3, this suggests that participants in the lower groups were more likely to want to change their accent than participants in the Advanced and Fluent groups.

### 3.1.4. I Want to Sound Like a ‘Native Speaker’

For the final question about accent and pronunciation, speakers were asked to what extent they wanted to sound like a native speaker. Neither variable was significant in the model, suggesting that neither of the factors influenced responses to this question (Origin Wald  $\chi^2(1) = 0.425$ ,  $p = 0.514$ ; Level Wald  $\chi^2(3) = 3.103$ ,  $p = 0.376$  coefficients also appear

in Table A7). The deviance goodness-of-fit test indicated that the model was a good fit to the observed data  $\chi^2(52) = 42.137, p = 0.804$ . The pseudo R2 value for this model is 0.036, suggesting that the model only explains around 3.6% of the variance in responses. From Figures 2 and 3, responses across groups for both origin and level are similar for this question, indicating that, in general, participants wanted to sound native to a similar extent.



**Figure 2.** Median scores for statements on accent and pronunciation for speakers from Wales, and from outside of Wales (1 = Not at all, 9 = Completely).

Summing up these results, the ordinal logistic regression analyses indicate that speaker origin and level are statistically significant predictors of ratings on three of the four questions about accent and pronunciation. In general, speakers from outside of Wales hold more negative views of their own accent and pronunciation and want to change their accent more than speakers from Wales. We also discovered that, as level increased, perceived similarity to native speakers, as well as pride in accent was greater, and desire to change accent was less strong. In general, respondents strongly agreed that they wanted to sound native, but no speaker factors predicted responses.

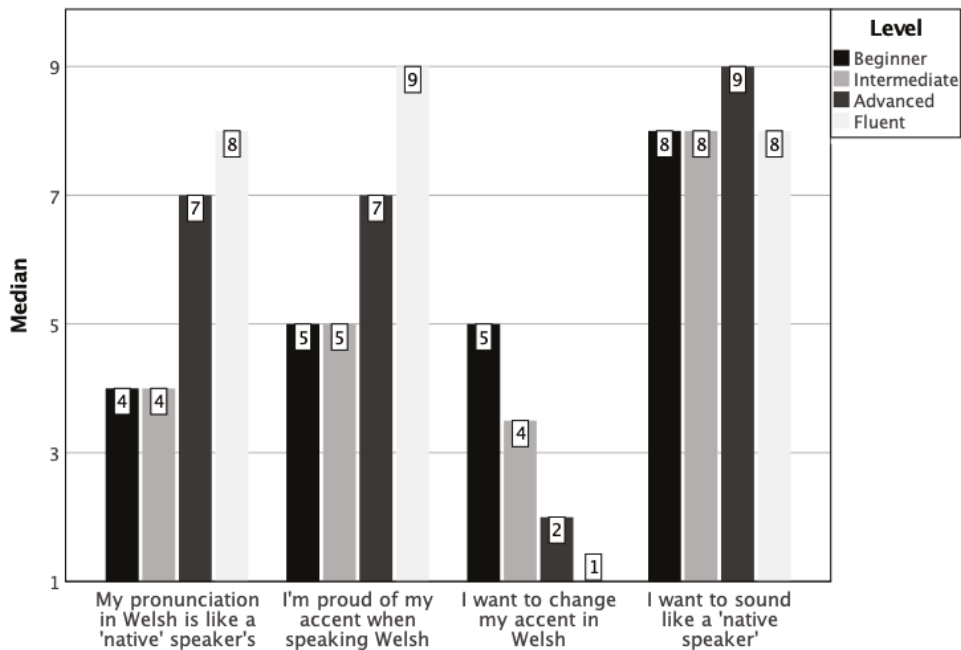


Figure 3. Median scores for statements on accent and pronunciation for the four levels (1 = Not at all, 9 = Completely).

### 3.2. Questions About Specific Speech Sounds

We asked several questions on which speech sounds were challenging for the participants: consonants that are absent in the L1 (English), vowel length with and without the diacritical mark of the circumflex, the pronunciation of three diphthongs and the pronunciation of /o/ at the ends of words. Participants responded by ticking features that were “challenging for me”. Figure 4 below presents the percentage of participants ( $n = 115$ ) who ticked each feature.

With regard to the consonants, inspection of Figure 4, above, shows that the pronunciation of <r> was reported as challenging by around 22% ( $n = 26$ ) of the respondents. Challenges with the trilled /r/ were reflected in the comments made by participants, with several noting that they find “rolling the r sound” difficult, with some also mentioning that they find perceiving and producing the difference between <r>/r/ and <rh>/r<sup>h</sup>/challenging. The pronunciation of the voiceless fricatives <ll> and <ch> were reported to be less challenging than <r>. The voiceless lateral fricative <ll> was reported to be challenging by 10% ( $n = 12$ ) of respondents. Only 8% ( $n = 9$ ) of respondents reported that the voiceless velar/uvular fricative <ch> was challenging. Some participants commented that the <ll> sound was challenging in certain contexts, especially word-finally, and others mentioned that they found it difficult to distinguish between <ll> and <ch>.

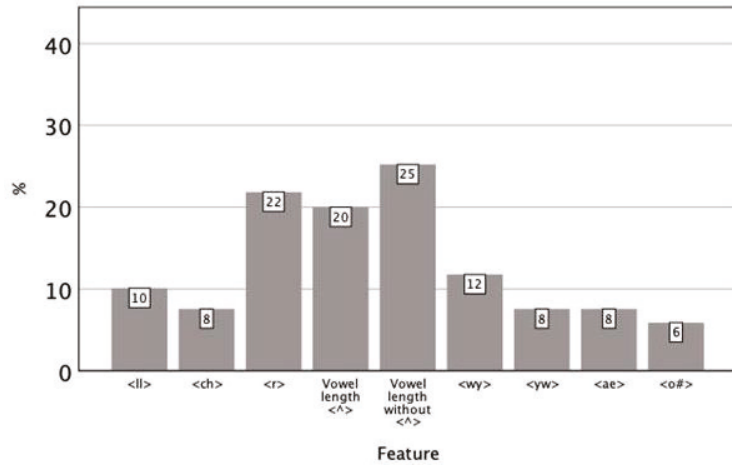
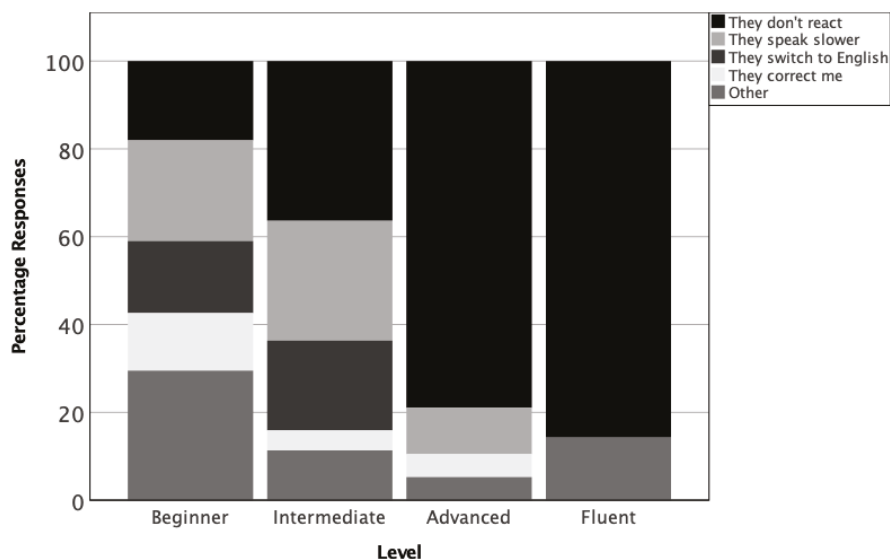


Figure 4. Percentage of speaker responses about individual features of their own pronunciation.

Inspection of Figure 4, above, indicates that the most reported difficulties that individual participants faced with vowels was vowel length. This was reported to be difficult for 25% ( $n = 29$ ) of individuals when the distinction was not signalled by a circumflex in the orthography, but 20% ( $n = 23$ ) of respondents also found vowel length difficult even when the circumflex was present. In terms of the diphthongs, participants generally reported that these were not as challenging as vowel length distinctions. For <wy>, 12% ( $n = 14$ ) found the pronunciation difficult. Several individuals mentioned the difficulties posed by the orthographic ambiguity of <wy>, which can represent a pure diphthong, or a combination of consonant [w] plus vowel. For example, “Knowing which way some sounds need to be pronounced in a word if there are two alternatives, e.g., wy can be either ‘oi’ or ‘wee’”, “gwahaniaethu gwahanol ynganiadau wy” (“differentiating between different pronunciations of ‘wy’”). Eight percent ( $n = 9$ ) responded that they struggled with the pronunciation of <yw>. This was reflected in a comment from one respondent, who stated that they faced challenges with “Vowels, especially groups of vowels—w and y most of all. I understand there are meant to be rules that govern this, but it seems like there must be quite a lot of variations depending on the exact letter combinations.” Only 8% of respondents indicated that <ae> was challenging. The final vowel feature we asked participants about was the production of the monophthong /o/ in word final position based on tutors’ responses in the Rees and Morris (2018) paper that this vowel is often diphthongized. Very few respondents (5.8%,  $n = 7$ ) said that this was challenging to them. We therefore see that some individual speakers are aware of the difficulties that can arise from the ambiguity with vowel length and the digraphs, but participants did not comment specifically about /o/ being challenging to pronounce at the ends of words.

### 3.3. Responses of Traditional ‘Native’ Speakers

Finally, we asked participants how traditional ‘native’ Welsh speakers responded to their accent/pronunciation. Participants were provided with a several response options and were able to tick multiple answers. Figure 5 below illustrates the proportion of responses for each level.



**Figure 5.** Proportion of responses per level for “How do ‘native’ speakers react to your accent/pronunciation”? Beginner  $n = 52$ ; Intermediate  $n = 38$ ; Advanced  $n = 18$ ; Fluent  $n = 7$ .

As illustrated by Figure 5, as the level increased, the proportion of responses that native speakers did not respond to their accent increased. In fact, most Advanced and Fluent speakers reported that native speakers did not respond to their accent/pronunciation at all. Beginner- and intermediate-level speakers, on the other hand, experienced native speakers responding to their accent by speaking slower, switching to English or correcting their pronunciation. Interestingly, all the “other” responses in the beginner group were related to speakers not having an opportunity to speak Welsh with a ‘native’ speaker outside the classroom. Responses in the “other” category for the intermediate level indicated that native speakers responded differently depending on the situation. One speaker reported “most Welsh speakers seem pleased that I’ve tried” whilst another speaker reported that native speakers “talk to me like I’m a baby/small kid”. Some advanced speakers experienced native speakers speaking slower and correcting their pronunciation but did not report native speakers switching to English. Participants in the advanced and fluent groups who selected “other” reported that native speakers were surprised that they had learnt Welsh, and that they were asked if they were from North Wales. Overall, it seems that there were a range of responses by native speakers reported by the participants in the study, but that native speakers responded less to a participant’s accent/pronunciation in more advanced speakers.

#### 4. Discussion

The results reported within this paper demonstrate that in producing Welsh speech, adult new speakers face a range of challenges. Exploring the experiences of adult new speakers has revealed that pronunciation and accent is an important consideration for individuals at all stages in their language learning journeys.

There is evidence that the individuals perceived their accent and pronunciation in general to be not like that of a native speaker, but that most individuals felt strongly that they wanted to sound similar to a ‘native speaker’. This is important in terms of understanding the motivations and aspirations of the new speakers and their interactions with traditional speakers. The nature of ‘native’ speech is vague, especially when some of the participants noted that they had no contact with traditional speakers outside the

classroom setting. However, despite the native-like target, participants did not indicate a strong desire to change their own accents. This may be due to participants seeing native-like pronunciation as unattainable. The questions of the perceived links between native-like pronunciation and speaker legitimacy have been discussed in Welsh and in other contexts. For example, [Nance et al. \(2016\)](#) explored the accent aims of new speakers of Scottish Gaelic. They found that some speakers wished to sound like native-speakers, but other speakers preferred production in line with a new-speaker model and considered a native-speaker target as inauthentic. In our study, participants strongly agreed that sounding like a native speaker was the ideal, but in the future, we may want to explore what this ideal means, how new speakers construct their identity and where they fit into their communities.

Feelings about one's own accent and pronunciation changed as the stage of the learning journey increased. That is, beginner and intermediate individuals reported sounding less native, were less proud of their accent and wanted to change their accent more than advanced and fluent speakers. Individuals reported becoming prouder and more native as the level of ability increased. However, we also found that speakers who were from Wales, and identified as being Welsh, were less likely to want to change their accent, were prouder, and reported having pronunciation that more closely resembled that of a native speaker than individuals who were from outside Wales. This finding is in line with the previous research by [Müller and Ball \(1999\)](#), who found that the variety of English spoken as the L1 affected the production of the vowels in Welsh in the late adult bilinguals. This was explained due to different varieties of English having certain sounds that are closer or further from those found in Welsh. In particular, they found that speakers who had Welsh English as their L1 were more likely to pronounce monophthongs without glide, whereas the non-Welsh English speakers tended to diphthongize the front and back mid vowels. They note that, for all speakers, this reflected their use in English and resulted in an "English" accent in Welsh for the non-Welsh English group. Furthermore, [Mayr et al. \(2017\)](#) and [Mennen et al. \(2020\)](#) investigated language contact and the convergence in segmental and prosodic features seen between Welsh and English in Ammanford in Carmarthenshire. Our finding that speakers from outside Wales are less proud of their accent and believe that they sound less native suggests that the variety of the L1 is important. That is, individuals from Wales may already have the 'new' phonemes in their repertoire from the community or in Welsh lessons in early education. Others from outside of Wales have to learn to distinguish the production of vowels from those in their L1 variety.

This has important implications for language teachers in the classroom, who may have speakers with both differing abilities when it comes to pronunciation, but also different levels of self-confidence in their accent and pronunciation. Indeed, the comparison between the results from the current study and [Rees and Morris \(2018\)](#) provide an insight into the language learning process from both sides. Both groups agree that the unfamiliar consonant sounds can initially pose difficulties, and that orthographic ambiguities could also prove challenging. However, the tutors also saw the diphthongization of monophthongs, especially /o/, as features that could pose problems, with 76% of participants (the highest of any of the sounds) stating that it was 'challenging for some', with a further 10% stating that it was challenging for beginners only. This is in contrast with the participants of the current study where very few reported that this was a challenging feature. This poses questions regarding vowel perception, especially in terms of individuals who have very limited input. It could be that the word-final /o/ is assimilated into an existing vowel category in the English L1 /oo/ and that the speakers do not perceive the difference. These differences and their links to different variants or dialects of the L1 could be an important direction for future research. For example, this is a key point in the L2LP model of speech perception and learning ([Van Leussen and Escudero 2015](#)) which posits that the L1 dialect can affect the perception of the target language. Future research on the perception (and production) of speech by adults who are learning Welsh could consider the effects of the L1 variety. This would have implications for the theoretical understanding of perceiving and



learning new sounds and for the development of pedagogy in the field where speakers with different dialects of the L1 learn in the same classrooms.

An interesting situation arises from the findings of how traditional ‘native’ speakers respond to new speaker accent and pronunciation. We found that new speakers at different levels on their learning journey received different responses from native speakers. Beginners reported a range of responses from native speakers: no reaction, speaking more slowly, switching to English, correcting pronunciation and several responded that they did not have an opportunity to speak with traditional speakers outside of the classroom, which highlights the challenges of learning a language in a minority setting. This may have been caused in part by the fact that data collection was conducted during the COVID-19 pandemic, which meant that opportunities to interact with others more widely were limited. Intermediate speakers also reported a range of responses, but a greater proportion reported no response from native speakers than in the Beginner group. On the other hand, the majority of more advanced participants (Advanced, Fluent) did not note any reaction at all from native speakers. This suggests a more complicated picture than the reported tendency of native speakers switching to English with adult new speakers in general (Trosset 1986; Mac Giolla Chríost et al. 2012). Further research in this area may be beneficial in highlighting the interactions between new and traditional speakers in the community. Increasing the use of Welsh for new speakers, especially outside traditional heartland areas, will be important for achieving the Welsh Government’s ambitious aim of reaching 1 million speakers of Welsh by 2050 (Welsh Government 2017).

This study has highlighted some of the challenges that adult new speakers face in learning the pronunciation of Welsh. Whilst some findings do echo those previously reported by language tutors by Rees and Morris (2018), many elements differ, suggesting that ‘learners’ may not perceive differences in their speech that are important for traditional native speakers. One important limitation of our study is that the data were self-reported, which rely on the participants’ awareness of their own pronunciation and traditional ‘native’ pronunciation. We noted that several participants did not have a chance to interact with native speakers outside of the classroom, which may have implications for what they perceive as ‘native’. Furthermore, the focus of this study was to explore some of the segmental aspects of Welsh pronunciation in south Wales, but it should be recognized that it is not only segmental aspects that can be challenging for people learning a language. Future research on the pronunciation of Welsh by adults who are learning the language may want to consider differences in lexical stress placement and realization, as well as the implementation of intonational tunes. Similarly, we have concentrated on data from south Wales. Future research may want to consider comparisons with communities with larger proportions of Welsh speakers, for example in north west Wales, where there may be more opportunities to use Welsh in the community.

In addition, the participants of this study were mostly less experienced speakers, potentially due to a lack of opportunities to use the language. Future research may also consider the perception and production of Welsh speech by fluent new speakers who are using the language in their community on a regular basis, as well as the reception of their accent and pronunciation by traditional speakers. This would be useful in highlighting non-native features of L2 speech in order to inform language tutors and providers to support new speakers in achieving competence or confidence in pronunciation that allows them to integrate into Welsh speaking communities and use Welsh beyond the classroom.

**Author Contributions:** Conceptualization, M.W. and S.C.; Data curation, M.W. and S.C.; Formal analysis, M.W. and S.C.; Funding acquisition, S.C.; Investigation, M.W. and S.C.; Methodology, M.W. and S.C.; Supervision, S.C.; Writing—original draft, M.W. and S.C.; writing—review and editing, M.W. and S.C.; visualization, M.W. and S.C.; supervision, S.C.; project administration, M.W.; funding acquisition, S.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was part of a doctoral project funded by the Coleg Cymraeg Cenedlaethol, Bangor University and the National Centre for Learning Welsh.

**Institutional Review Board Statement:** The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Ethics Committee of Bangor University (protocol code 2020-MW1), date of approval 16 October 2020.

**Data Availability Statement:** The data presented in this study are not publicly available since the Ethics Committee and informed consent processes do not allow for the raw data to be publicly shared.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results. The National Centre for Learning Welsh advertised the study in newsletters and via social media.

## Appendix A

**Table A1.** Accent and pronunciation items in the questionnaire.

I ba raddau mae'r datganiadau isod yn wir i chi?   To What Extent Are the Following Statements True for You	1 = Ddim o gwbl/Not at all 9 = Yn llwyr/Completely								
	1	2	3	4	5	6	7	8	9
Mae fy ynganiad wrth siarad Cymraeg yr un peth â siaradwyr 'brodorol'   My pronunciation in Welsh is like a 'native' speaker's	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Rydw i'n falch o fy acen wrth siarad Cymraeg   I'm proud of my accent when speaking Welsh	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Rydw i eisiau newid fy acen yn Gymraeg   I want to change my accent in Welsh	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Rydw i eisiau swnio fel 'siaradwr brodorol'   I want to sound like a 'native speaker	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**Table A2.** Individual aspects of pronunciation items in the questionnaire.

Pa mor heriol ydi'r rhain i unigolion sy'n dysgu Cymraeg?   How Challenging Are the Following for Individuals Learning Welsh?	Heriol i fi/Challenging for me
Ynganu/pronouncing 'll' e.e. lle, pell	<input type="checkbox"/>
Ynganu/pronouncing 'ch' e.e. chi, cwch	<input type="checkbox"/>
Rolio/rolling 'r' e.e. oren, dŵr	<input type="checkbox"/>
Ynganu 'o' ar ddiwedd geiriau/pronouncing 'o' at the end of words e.e. nofio, eto	<input type="checkbox"/>
Ynganu/pronouncing 'ae' e.e. mae	<input type="checkbox"/>
Ynganu/pronouncing 'wy' e.e. mwy	<input type="checkbox"/>
Ynganu/pronouncing 'yw' e.e. byw	<input type="checkbox"/>
Gwahaniaethu rhwng llafariad hir a byr gyda to bach/Distinguishing between long and short vowels with a circumflex e.e. tân/tan	<input type="checkbox"/>
Gwahaniaethu rhwng llafariad hir a byr heb do bach/Distinguishing between long and short vowels without a circumflex e.e. bys/byr	<input type="checkbox"/>

**Table A3.** ‘native’ speaker reaction to accent and pronunciation item in the questionnaire.

Sut mae siaradwyr ‘brodorol’ yn ymateb i’ch acen/ynghanid?   How Do ‘Native’ Speakers React to Your Accent/Pronunciation?	
Dydyn nhw ddim yn ymateb/They don’t react	<input type="checkbox"/>
Mae nhw’n siarad yn arafach/They speak more slowly	<input type="checkbox"/>
Mae nhw’n troi i’r Saesneg/They switch to English	<input type="checkbox"/>
Mae nhw’n ‘cywirio’/They ‘correct’ me	<input type="checkbox"/>
Arall/Other	<input type="checkbox"/>
Os dewisoch chi Arall, rhowch fanylion os gwelwch yn dda   If other, please give details:	

**Appendix B**

**Table A4.** Coefficients for My pronunciation in Welsh is like a ‘native’ speaker’s.

Factor	Coef	S.E.	Wald	p-Value
[LikeNative = 1]	-4.925	0.862	32.683	<0.001
[LikeNative = 2]	-4.166	0.843	24.446	<0.001
[LikeNative = 3]	-3.203	0.830	14.902	<0.001
[LikeNative = 4]	-2.833	0.826	11.747	0.001
[LikeNative = 5]	-1.862	0.817	5.190	0.023
[LikeNative = 6]	-1.477	0.812	3.307	0.069
[LikeNative = 7]	-0.378	0.791	0.229	0.632
[LikeNative = 8]	1.549	0.775	3.997	0.046
[Origin = Welsh]	1.634	0.362	20.412	<0.001
[Level = Beginner]	-3.987	0.850	22.031	<0.001
[Level = Intermediate]	-3.464	0.852	16.539	<0.001
[Level = Advanced]	-2.284	0.880	6.730	0.009

**Table A5.** Coefficients for I’m proud of my accent when speaking Welsh.

Factor	Coef	S.E.	Wald	p-Value
[ProudAccent = 1]	-3.652	0.796	21.052	<0.001
[ProudAccent = 2]	-3.259	0.778	17.559	<0.001
[ProudAccent = 3]	-2.209	0.752	8.641	0.003
[ProudAccent = 4]	-1.900	0.747	6.462	0.011
[ProudAccent = 5]	-1.066	0.740	2.074	0.150
[ProudAccent = 6]	-0.559	0.737	0.575	0.448
[ProudAccent = 7]	0.227	0.734	0.096	0.757
[ProudAccent = 8]	1.227	0.740	2.751	0.097
[Origin = Welsh]	1.435	0.356	16.232	<0.001
[Level = Beginner]	-2.014	0.757	7.068	0.008
[Level = Intermediate]	-2.015	0.771	6.824	0.009
[Level = Advanced]	-0.757	0.821	0.848	0.357

**Table A6.** Coefficients for I want to change my accent in Welsh.

Factor	Coef	S.E.	Wald	p-Value
[ProudAccent = 1]	0.018	0.838	0.000	0.983
[ProudAccent = 2]	0.714	0.842	0.718	0.397
[ProudAccent = 3]	1.077	0.845	1.625	0.202
[ProudAccent = 4]	1.588	0.849	3.497	0.061
[ProudAccent = 5]	2.311	0.857	7.266	0.007
[ProudAccent = 6]	2.630	0.862	9.298	0.002
[ProudAccent = 7]	3.297	0.8579	14.074	<0.001
[ProudAccent = 8]	3.568	0.889	16.117	<0.001
[Origin = Welsh]	-1.446	0.360	16.184	<0.001
[Level = Beginner]	2.302	0.871	6.991	0.008
[Level = Intermediate]	1.831	0.879	4.340	0.037
[Level = Advanced]	1.420	0.935	2.306	0.129

**Table A7.** Coefficients for I want to sound like a ‘native speaker’.

Factor	Coef	S.E.	Wald	p-Value
[ProudAccent = 1]	−3.391	0.861	15.513	<0.001
[ProudAccent = 2]	−2.964	0.811	13.355	<0.001
[ProudAccent = 3]	−2.526	0.776	10.597	<0.001
[ProudAccent = 4]	−2.207	0.758	8.479	0.004
[ProudAccent = 5]	−1.031	0.724	2.029	0.154
[ProudAccent = 6]	−0.648	0.719	0.812	0.368
[ProudAccent = 7]	−0.452	0.718	0.396	0.529
[ProudAccent = 8]	0.224	0.717	0.098	0.755
[Origin = Welsh]	−0.227	0.348	0.425	0.514
[Level = Beginner]	0.017	0.731	0.001	0.982
[Level = Intermediate]	−0.238	0.747	0.101	0.750
[Level = Advanced]	0.735	0.826	0.791	0.374

## References

- Andrews, Hunydd. 2011. Llais y Dysgwyr: Profiadau Oedolion Sydd Yn Dysgu Cymraeg Yng Ngogledd Cymru. *Gwerddon* 9: 37–58.
- Awbery, Gwen. 2009. Welsh. In *The Celtic Languages*, 2nd ed. Edited by Martin J. Ball and Nicole Müller. London: Routledge, pp. 359–426.
- Baayen, Harald. 2008. *Analyzing Linguistic Data: A Practical Introduction to Statistics Using R*. Cambridge: Cambridge University Press.
- Baker, Colin, Hunydd Andrews, Ifor Gruffydd, and Gwyn Lewis. 2011. Adult Language Learning: A Survey of Welsh for Adults in the Context of Language Planning. *Evaluation & Research in Education* 24: 41–59. [[CrossRef](#)]
- Ball, Martin J., and Glyn E. Jones. 1984. *Welsh Phonology: Selected Readings*. Cardiff: University of Wales Press.
- Ball, Martin J., and Briony Williams. 2001. *Welsh Phonetics*. Lampeter: Edwin Mellen Press.
- Best, Catherine. 1995. A Direct Realist View of Crosslanguage Speech Perception. In *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*. Edited by W. Strange. Timonium: York Press, pp. 171–204.
- Best, Catherine. T., and Michael. D. Tyler. 2007. Non-Native and Second Language Speech Perception: Commonalities and Complementarities. In *Language Experience in Second Language Learning: In Honour of James Emil Flege*. Edited by Ocke-Schwen Bohn and Murray J. Munro. Amsterdam: John Benjamins, pp. 13–34.
- Birdsong, David. 2003. Authenticité de prononciation en français L2 chez des apprenants tardifs anglophones: Analyses segmentales et globales. *Acquisition et Interaction en Langue Étrangère* 18: 17–36. [[CrossRef](#)]
- Bucholtz, Mary. 2003. Sociolinguistic nostalgia and the authentication of identity. *Journal of Sociolinguistics* 7: 398–416. [[CrossRef](#)]
- Derwing, Tracey M., and Murray J. Munro. 2015. *Pronunciation Fundamentals: Evidence-Based Perspectives for L2 Teaching and Research*. Amsterdam: John Benjamins. [[CrossRef](#)]
- Endresen, Anna, and Laura A. Janda. 2016. Five statistical models for Likert-type experimental data on acceptability judgments. *Journal of Research Design and Statistics in Linguistics and Communication Science* 3: 217–50. [[CrossRef](#)]
- Escudero, Paola. 2005. *Linguistic Perception and Second Language Acquisition: Explaining the Attainment of Optimal Phonological Categorization*. LOT Dissertation Series 113; Utrecht: Utrecht University.
- Escudero, Paola, and Paul Boersma. 2004. Bridging the Gap between L2 Speech Perception Research and Phonological Theory. *Studies in Second Language Acquisition* 26: 551–85. [[CrossRef](#)]
- Escudero, Paola, and Daniel Williams. 2012. Native dialect influences second-language vowel perception: Peruvian versus Iberian Spanish learners of Dutch. *The Journal of the Acoustical Society of America* 131: 406–12. [[CrossRef](#)] [[PubMed](#)]
- Flege, James. 1995. Second Language Speech Learning: Theory, Findings and Problems. In *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*. Edited by W. Strange. Timonium: York Press, pp. 229–73.
- Fynes-Clinton, Osbert Henry. 1913. *The Welsh Vocabulary of the Bangor District*. Oxford: Oxford University Press.
- Gnevshva, Ksenia. 2017. Within-speaker variation in passing for a native speaker. *International Journal of Bilingualism* 21: 213–27. [[CrossRef](#)]
- Hannahs, S. J. 2013. *The Phonology of Welsh*. Oxford: Oxford University Press.
- Hewitt, Steve. 2020. The Problem of Neo-Speakers in Language Revitalization: The Example of Breton. Paper presented at FEL24: Teaching and Learning Resources for Endangered Languages, UCL, London, UK, September 23–25.
- Hodges, Rhian Siân. 2010. Tua'r Goleuni: Rhesymau Rhieni Dros Ddewis Addysg Gymraeg i'w Plant Yng Nghwm Rhymni. *Gwerddon* 6: 9–33.
- Hornsby, Michael. 2015a. *Revitalizing Minority Languages: New Speakers of Breton, Yiddish and Lemko*. London: Palgrave Macmillan. [[CrossRef](#)]
- Hornsby, Michael. 2015b. The “New” and “Traditional” Speaker Dichotomy: Bridging the Gap. *International Journal of the Sociology of Language* 231: 107–25. [[CrossRef](#)]
- Hornsby, Michael. 2019. Positions and Stances in the Hierarchization of Breton Speakerhood. *Journal of Multilingual and Multicultural Development* 40: 392–403. [[CrossRef](#)]

- Hornsby, Michael, and Dick Vigers. 2018. “New” Speakers in the Heartlands: Struggles for Speaker Legitimacy in Wales. *Journal of Multilingual and Multicultural Development* 39: 419–30. [CrossRef]
- Jaffe, Alexandra. 2013. Minority Language Learning and Communicative Competence: Models of Identity and Participation in Corsican Adult Language Courses. *Language & Communication* 33: 450–62. [CrossRef]
- Jenkins, Jennifer. 2002. A Sociolinguistically Based, Empirically Researched Pronunciation Syllabus for English as an International Language. *Applied Linguistics* 23: 83–103. [CrossRef]
- Jonsson, Carla, and Mona Rosenfors. 2017. “I Have Struggled Really Hard to Learn Sami”: Claiming and Regaining a Minority Language. *International Journal of the Sociology of Language* 248: 49–71. [CrossRef]
- Mac Giolla Chríost, Diarmait, Partick Carlin, Sioned Davies, Tess Fitzpatrick, Anys Prys Jones, Rachel Heath-Davies, and Jennifer Marshall. 2012. *Welsh for Adults Teaching and Learning Approaches, Methodologies and Resources: A Comprehensive Research Study and Critical Review of the Way Forward*. Cardiff: Welsh Government.
- Mayr, Robert, and Hannah Davies. 2011. A Cross-Dialectal Acoustic Study of the Monophthongs and Diphthongs of Welsh. *Journal of the International Phonetic Association* 41: 1–25. [CrossRef]
- Mayr, Robert, Jonathan Morris, Ineke Mennen, and Daniel Williams. 2017. Disentangling the effects of long-term language contact and individual bilingualism: The case of monophthongs in Welsh and English. *International Journal of Bilingualism* 21: 245–67. [CrossRef]
- McEwan-Fujita, Emily. 2010. Ideology, Affect, and Socialization in Language Shift and Revitalization: The Experiences of Adults Learning Gaelic in the Western Isles of Scotland. *Language in Society* 39: 27. [CrossRef]
- McLeod, Wilson, and Bernadette O’Rourke. 2015. “New speakers” of Gaelic: Perceptions of linguistic authenticity and appropriateness. *Applied Linguistics Review* 6: 151–72. [CrossRef]
- McLeod, Wilson, Bernadette O’Rourke, and Stuart Dunmore. 2014. *New Speakers of Gaelic in Edinburgh and Glasgow*. Soillse: The National Research Network for the Maintenance and Revitalisation of Gaelic Language and Culture.
- Mennen, Ineke, Niamh Kelly, Robert Mayr, and Jonathan Morris. 2020. The Effects of Home Language and Bilingualism on the Realization of Lexical Stress in Welsh and Welsh English. *Frontiers in Psychology* 10: 3038. [CrossRef]
- Mitterer, Holger, Nikola Anna Eger, and Eva Reinisch. 2020. My English Sounds Better than Yours: Second-Language Learners Perceive Their Own Accent as Better than That of Their Peers. *PLoS ONE* 15: e0227643. [CrossRef]
- Morris, Jonathan. 2013. *Sociolinguistic Variation and Regional Minority Language Bilingualism: An Investigation of Welsh-English Bilinguals in North Wales*. Manchester: University of Manchester.
- Morris-Jones, John, W. J. Gruffydd, T. Gwynn Jones, J. Lloyd-Jones, T. H. Parry-Williams, Ifor Williams, Robert Williams, and Henry Lewis. 1928. *Orgraff yr Iaith Gymraeg*. Caerdydd: Gwasg Prifysgol Cymru.
- Müller, Nicole, and Martin J. Ball. 1999. Examining the Acquisition of Welsh Phonology in L1 English Learners. *Journal of Celtic Language Learning* 4: 16–32.
- Nance, Claire, Wilson McLeod, Bernadette O’Rourke, and Stuart Dunmore. 2016. Identity, Accent Aim, and Motivation in Second Language Users: New Scottish Gaelic Speakers’ Use of Phonetic Variation. *Journal of Sociolinguistics* 20: 164–91. [CrossRef]
- National Centre for Learning Welsh. 2020. Statistics 2018–2019’. Available online: <https://learnwelsh.cymru/about-us/statistics-2018-19/> (accessed on 3 March 2021).
- O’Rourke, Bernadette, and Joan Pujolar. 2013. From Native Speakers to “New Speakers”—Problematising Nativeness in Language Revitalization Contexts. *Histoire Epistemologie Langage* 35: 47–67.
- Robert, Elen. 2009. Accommodating “new” speakers? An attitudinal investigation of L2 speakers of Welsh in south-east Wales. *International Journal of the Sociology of Language* 195: 93–115. [CrossRef]
- Rees, Iwan Wyn, and Jonathan Morris. 2018. Astudiaeth o Ganfyddiadau Tiwtoriaid Cymraeg i Oedolion o Anawsterau Ynganau Ymhlith Dysgwyr Yr Iaith. *Gwerddon* 27: 39–66.
- Selleck, Charlotte. 2018. We’re Not Fully Welsh: Hierarchies of Belonging and ‘New’ Speakers of Welsh. In *New Speakers of Minority Languages: Linguistic Ideologies and Practices*. Edited by Cassie Smith-Christmas, Noel P. Ó. Murchadha, Michael Hornsby and Máiréad Moriarty. London: Palgrave Macmillan, pp. 45–65.
- Trosset, Carol S. 1986. The Social Identity of Welsh Learners. *Language in Society* 15: 165–91. [CrossRef]
- Tyler, Michael D. 2019. PAM-L2 and phonological category acquisition in the foreign language classroom. In *A Sound Approach to Language Matters—In Honor of Ocke-Schwen Bohn*. Edited by Anne Mette Nyvad, Michaela Hejrná, Anders Højen, Anna Bothe Jespersen and Mette Hjortshøj Sørensen. Aarhus: Aarhus University, pp. 607–30.
- Van Leussen, Jan-Willem, and Paola Escudero. 2015. Learning to Perceive and Recognize a Second Language: The L2LP Model Revised. *Frontiers in Psychology* 6: 1000. [CrossRef] [PubMed]
- Watkins, T. Arwyn. 1961. *Ieithyddiaeth*. Caerdydd: Gwasg Prifysgol Cymru.
- Watkins, Colin. 2020. 2020 Duolingo Language Report: United Kingdom. *Duolingo Blog*. December 15. Available online: <https://blog.duolingo.com/uk-language-report-2020/> (accessed on 3 March 2021).
- Welsh Government. 2011. Welsh Speakers by Local Authority, Gender and Detailed Age Groups, 2011 Census. Cardiff: Welsh Government. Available online: <https://stats.wales.gov.wales/Catalogue/Welsh-Language/Census-Welsh-Language/welshspeakers-by-localauthority-gender-detailedagegroups-2011census> (accessed on 1 March 2021).
- Welsh Government. 2017. *Cymraeg 2050: A Million Welsh Speakers*. Cardiff: Welsh Government.

Article

# Social Influences on Phonological Transfer: /r/ Variation in the Repertoire of Welsh-English Bilinguals

Jonathan Morris

School of Welsh, Cardiff University, Cardiff CF10 3EU, UK; morrisj17@caerdydd.ac.uk

**Abstract:** It is well known that cross-linguistic interactions can exist between the two languages in a bilingual speaker's repertoire. At the level of phonetics and phonology, this interaction may result in the transfer of a feature from one language to the other or the 'merging' of phonetic properties between languages. Although there are numerous studies of bilingual speakers which show such interactions, relatively little is known about the nature of transfer in communities of long-term bilingualism. The current study investigates phonological transfer of /r/ in Welsh-English bilinguals' speech in north Wales. Specifically, it compares the influence of speaker gender, home language, and speech context on the production of /r/ in both English and Welsh in two communities which differ in the extent to which Welsh is spoken as a community language. It is commonly assumed that the alveolar trill [r] and alveolar tap [ɾ] are the variants of /r/ in Welsh. In English, the alveolar approximant [ɹ] is typical across Wales, but the trill and tap are reported in areas where a high proportion of the population speaks Welsh. Data in both languages were collected from 32 Welsh-English bilinguals (aged 16–18) via sociolinguistic interview and wordlist tasks. The sample was stratified equally by speaker gender, home language, and area (predominantly Welsh-speaking vs. predominantly English-speaking). The results show areal differences in the production of /r/ in both languages, which, I argue, could be attributed partly to differing social structures in the communities under investigation. Consequently, the results showed evidence of bi-directional phonological transfer, which is community-specific and influenced by a number of social factors.

**Keywords:** language variation; bilingualism; phonological transfer; Welsh; Welsh English



**Citation:** Morris, Jonathan. 2021. Social Influences on Phonological Transfer: /r/ Variation in the Repertoire of Welsh-English Bilinguals. *Languages* 6: 97. <https://doi.org/10.3390/languages6020097>

Academic Editors: Juana M. Licerias and Raquel Fernández Fuertes

Received: 22 March 2021  
Accepted: 19 May 2021  
Published: 25 May 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The interaction between the sound systems of a bilingual's languages can be attributed to (1) cross-linguistic interactions at the level of phonetics, whereby the phonetic implementation of a particular sound might be influenced to varying degrees by the other language(s) in a speaker's repertoire (e.g., Mennen 2004; Simonet 2010; Mayr and Montanari 2015; Simonet 2014; Simonet and Amengual 2020), or (2) the 'wholesale' transfer of features from one language to another (see Odlin 1989). More specifically, the process of phonological transfer in synchronic speech describes the substitution of phonemic segments, phonotactic patterns, and prosodic features from one language in another (Simon 2010, p. 63). This results in the production of features typically associated with one of the bilingual speaker's languages in the other language. The extent to which this process is realised in speech has been well-studied in the field of Second Language Acquisition (SLA) and can depend on a number of factors including age of acquisition, level of formal instruction, use of the two languages, social networks, the interlocutor, and the speech context (see Piske et al. 2001; Grosjean 2001 for overviews).

The current study contributes to a growing body of variationist work examining transfer in situations of long-term bilingualism (e.g., Davidson 2015; Nagy 2015; Mooney 2019; Gafter and Horesh 2020). One particular strand of this research has been the analysis of speech in contexts of language revitalisation, wherein a clear distinction in speech patterns has been found between so-called 'new speakers', who have acquired the language

outside of the home, and ‘traditional speakers’ (e.g., Nance 2013, 2015). In the case of Welsh-English bilingualism, however, previous work on monophthong and lexical stress production in one community has shown few cross-linguistic differences and little influence of extra-linguistic factors such as the home language of the speaker (e.g., Morris 2017; Mayr et al. 2017; Mennen et al. 2020).

The study builds on previous work (1) by examining a consonantal feature that exhibits clear cross-linguistic differences between Welsh and English (the patterning of which has hitherto not been quantified) and (2) by comparing two communities. Ultimately, this advances our knowledge of the ways in which local community social structures might influence transfer in situations of long-term bilingualism. Specifically, I report on an analysis of the influence of linguistic and extra-linguistic factors on the production of /r/ in the Welsh and English of 32 bilingual speakers aged 16–18 in the towns of Mold (Flintshire) and Caernarfon (Gwynedd). The two towns have comparable populations but differ in the extent to which Welsh is spoken.

Section 2 outlines the background of the study and focuses on previous studies of cross-linguistic interactions in the speech of Welsh-English bilinguals, previous studies of the /r/ production in both languages, and research questions. Section 3 provides background on the community, speakers, and methodological approach in the study. Section 4 presents the analysis of the results, and these results are discussed with reference to the research questions in Section 5.

## 2. Background

### 2.1. The Welsh Context

The Welsh language is spoken by approximately 19% ( $n = 562,016$ ) of the population of Wales (Welsh Government 2012). The concentration of Welsh speakers in a given area varies throughout Wales, and in some western areas, Welsh is spoken by the majority of the local population. For instance, 65.4% ( $n = 77,000$ ) of the population of the western county of Gwynedd are bilingual, compared to 7.8% ( $n = 5284$ ) in the south-eastern county of Blaenau Gwent (Welsh Government 2020). In Welsh-dominant areas, the language may still be used as a community language, which increases the amount of Welsh to which children are exposed, whereas, in areas where Welsh is spoken by a minority, exposure may be restricted to caregivers or ‘more narrowly drawn social networks’ (Coupland and Ball 1989, p. 10).

The establishment of Welsh-medium education in the twentieth century has proved popular as both first language education for children from Welsh-speaking homes and as immersion education for children with non-Welsh-speaking parents. In western areas, where Welsh is used as a community language, especially in north-west Wales, the majority of schools teach most subjects in Welsh, and only a minority of children come from non-Welsh-speaking homes. In eastern areas, parents may choose English-medium or Welsh-medium education for their child. Welsh-medium education in eastern areas has proven a popular choice in areas where both English- and Welsh-medium schools exist (see Hodges 2012). Consequently, the majority of children in Welsh-medium schools in eastern areas come from English monolingual homes.

Welsh-English minority language bilingualism is complicated by two sets of intertwined dichotomies, which makes the situation inherently interesting for a study of language variation. Firstly, there is a distinction between the western heartland and other anglicised areas. This results in a group of speakers for whom Welsh is the main community language and a group for whom Welsh is arguably limited to certain domains or interlocutors. Secondly, there is a group of speakers who acquired Welsh in the home and a group who acquired Welsh through Welsh-medium education. Consequently, there are bilingual speakers who have different experiences of acquiring their languages. These groups are intertwined because there are increasing numbers of speakers in the heartland areas who speak English at home, and there has always been a proportion of the population in eastern areas who speak Welsh at home. Both having Welsh as a home language and

living in a community where over 60% of the population speak the language are two factors that have been shown to influence daily use of Welsh (e.g., Jones 2008).

## 2.2. Phonological Transfer and Cross-Linguistic Phonetic Interactions in the Speech of Welsh-English Bilinguals

A number of accounts of Welsh English differentiate between accents based on the perceived influence of Welsh (see Durham and Morris 2016, pp. 14–16 for an overview). These differences are accounted for either by a substratum influence of Welsh (Wells 1982, p. 377) or synchronic transfer from Welsh to English in areas where most of the population speak Welsh (Awbery 1997, p. 88). Until recently, however, there have been few attempts to examine synchronic transfer in the speech of Welsh-English bilinguals.

Using the same dataset as the present study, Morris (2017) examined /l/-darkening in the Welsh and English of bilinguals and the extent to which there were cross-linguistic differences in the production of /l/. The results of the statistical modelling showed that female speakers were more likely to differentiate between English and Welsh when producing /l/ in the onset position. A more cursory glance at the data revealed that it was the female speakers in the majority Welsh-speaking town of Caernarfon who showed the greatest differences between Welsh and English compared to the north-eastern town of Mold, where the majority of the population do not speak Welsh (Morris 2017, p. 200).

More recently, and again using the same dataset, Morris (forthcoming) investigated the extent to which there are cross-linguistic differences as well as linguistic and social influences, on four measures of Fundamental Frequency Range (minimum F0, maximum F0, mean F0, and the difference between the minimum and maximum F0). The results showed no cross-linguistic differences but clear areal differences between Caernarfon and Mold. Whereas gender was the most powerful predictor of variation across all measures, speakers' home language was also significant in Caernarfon. Those from Welsh-speaking families in Caernarfon produced a higher minimum, maximum, and mean F0 compared to those from English-speaking homes in the same area (once gender differences had been accounted for). Home language was the most powerful predictor of pitch span (the difference between minimum and maximum F0), with those from Welsh-speaking homes having a greater pitch span compared to those from English-speaking homes regardless of gender (see also Ordin and Mennen 2017 for a similar study which found cross-linguistic differences between Welsh and English in women's speech).

The home-language differences, which seem to be associated most strongly with Welsh-speaking areas (and where home language seems to be a more salient aspect of peer-group formation, see Morris 2014 and Section 3.2), have not been found in other areas. Mayr et al. (2017) examined the production of monophthongs in a community in south Wales. They compared bilingual data from male Welsh speakers from both Welsh- and English-speaking families and English data from male monolinguals. They found few cross-linguistic differences and no effect of home language in the bilinguals' data. There were also few differences between bilinguals and monolinguals in English. Similarly, Mennen et al. (2020) examined correlates of lexical stress in the same data and found no significant differences across most measures.

## 2.3. /r/ in Pre-Vocalic and Intervocalic Positions in Welsh and Welsh English

Descriptions of /r/ in both Welsh and Welsh English tend to be provided in general summaries of the phonology of both varieties or dialectological surveys (e.g., Parry 1977; Penhallurick 1991). In word-initial prevocalic and word-medial intervocalic positions, the voiced alveolar trill [r] is reported as being the most commonly realised variant of /r/ in Welsh, with partial devoicing occurring when it follows a preceding voiceless consonant. The voiced alveolar tap [ɾ] is often used in word-medial intervocalic position in the north-west. The voiced uvular trill [ʀ] or voiced uvular fricative [ʁ] is a dialectal feature of the Bala area of Gwynedd (the county to which Caernarfon belongs), though there is no mention of the uvular variants in the speech of Caernarfon. The voiced alveolar approximant [ɹ] may appear in the clusters /tr/ and /dr/ and, according to Jones (1984, pp. 49–50), is



an idiosyncratic feature for some speakers. The approximant is noted as being a dialectal feature of east Powys only (an area in Mid-Wales which borders England; Davies 1971).

The trilled and tapped variants of /r/ are cited as being a feature of English for Welsh-English bilinguals, and, in particular, in the speech of those living in the north-west (Wilson 2014). Elsewhere, it is assumed that it is the approximant that tends to occur in Welsh English (Penhallurick 1991, p. 132). In their investigation of the extent to which both Welsh and non-Welsh-speaking people in Wales are able to differentiate between the English of bilingual and monolingual speakers, Mayr et al. (2020) asked participants to listen to extracts of English produced by Welsh-English bilinguals and English monolinguals from the same area. They found that the ‘rolled r’ was the most-cited feature that was associated with bilinguals’ English (Mayr et al. 2020, p. 754). In a subsequent linguistic analysis of the extracts, they found a significant difference between monolingual and bilingual speakers, with bilingual speakers (all of whom came from Welsh-speaking homes) producing significantly more trills and taps than monolinguals (Mayr et al. 2020, p. 758).

#### 2.4. Research Questions

The work outlined thus far suggests that there has been little variationist research which examines phonological transfer in the speech of Welsh-English bilinguals. In English, previous dialectological work indicates that the trill and tap are often produced in the speech of those from areas where Welsh is widely spoken by the local population (e.g., Penhallurick 1991). In Welsh, it appears that the alveolar approximant might appear in certain consonant clusters (Jones 1984, pp. 49–50), and there is limited evidence that the approximant might be a feature of some varieties closer to the border with England (Davies 1971).

The extent to which /r/ varies in the two languages of Welsh speakers remains to be seen. Specifically, there has been hitherto no attempt to consider the role of linguistic and extra-linguistic factors on /r/ variation in speakers’ bilingual repertoires. As outlined previously, the consideration of phonological transfer through a variationist lens seems especially important in the Welsh context given the changes to the demographic profile of Welsh speakers in different areas across Wales. The study, therefore, aims to address the following research questions:

1. Is there evidence for phonological transfer in the production of /r/ in Welsh-English bilinguals’ speech?
2. To what extent do linguistic and extra-linguistic factors influence the production of /r/ in English and Welsh?
3. To what extent are there differences in the patterning of variation between two communities where community dynamics and the use of Welsh differ?

### 3. Materials and Methods

Sociolinguistic interview and wordlist data in both Welsh and English were collected from 32 Welsh-English bilinguals aged 16–18 years old from north Wales. The area defined as north Wales for the purposes of this study included the counties of the Isle of Anglesey, Gwynedd, and Conwy County Borough to the west, and Denbighshire, Flintshire, and Wrexham County Borough to the east (population: 698, 400, see Welsh Government 2020). The remainder of this section outlines the communities, speakers, and methods in more detail.

#### 3.1. Communities

The study included data collected from the areas around Caernarfon (Gwynedd) in north-west Wales and Mold (Flintshire) in north-east Wales as part of a wider project (Morris 2013). Both Caernarfon and Mold have comparable populations (around 10,000 people) and are so-called ‘anchor towns’ for outlying villages. There are, however, clear demographic and accentual differences between the two areas (see Morris 2017, for an overview of the accentual differences).

The most relevant demographic difference between the two areas lies in the percentage of Welsh speakers. In Caernarfon and the surrounding area, 83.9% of the population are reported as being able to speak Welsh compared to 22.18% of the population in Mold (Welsh Government 2012). Consequently, Welsh can be viewed as more of a community language in Caernarfon, used as the main language interaction among most residents, whereas the use of Welsh in Mold is arguably much more restricted to Welsh-speaking social networks and certain community organisations.

### 3.2. Speakers

A total of 32 Welsh-English bilingual speakers were recruited from local education providers, where most of the speakers' subjects were delivered in Welsh. Speakers were aged between 16 and 18 at the time of data collection and had also been born in the wider local area (or moved to the area during infancy). Despite coming from the wider local area, speakers stated that either Caernarfon or Mold was the town with which they most closely identified. All participants self-identified as white and either Welsh, British, or both.

The sample for the study was stratified equally by area (Caernarfon or Mold), home language (Welsh or English), and gender (female or male). All speakers had either acquired Welsh at home and/or had received all of their education through the medium of Welsh. All students indicated either a male or female gender identity. Table 1 shows the sample for the current study.

Table 1. The sample.

	Caernarfon		Mold	
	Male Speakers	Female Speakers	Male Speakers	Female Speakers
Welsh at home	4	4	4	4
English at home	4	4	4	4

The inclusion of a binary distinction between Welsh and English home languages is problematic, especially as no observations were made of language use in situ. This decision was taken, however, in light of the aims of the study to examine broad differences between those who acquired Welsh via parental transmission and those who acquired Welsh through education. The sample in the current study contained speakers who either had two Welsh-speaking parents (or were being raised by a Welsh-speaking single parent) who used Welsh with their children or had non-Welsh-speaking parents with whom they spoke in English. There were no mixed-language households or speakers of other languages in the current sample.

Unlike in some similar linguistic contexts (see Tomé Lourido and Evans 2021), the speakers in this study were not aware of the term 'new speakers', and those from non-Welsh-speaking homes would not identify as such. The results of the written questionnaire, sociolinguistic interview, and my own observations indicated differences between language use and the salience of linguistic background among peers (see Morris 2014). While overt attitudes towards Welsh were generally positive, for instance, certain negative opinions were expressed by some speakers from English-speaking homes in Caernarfon. Such negative attitudes were related to either the speakers' own perceived ability in Welsh or Welsh language policy. Extracts (1) and (2) below show examples from the English sociolinguistic interviews of these overt attitudes from two different speakers from English-speaking homes in Caernarfon.

1. I think I can't [speak Welsh] though 'cause I remember the school was quite bitchy 'cause I was in primary school they used to start on me for it a lot for not being able to say it properly. 'Cause I was like one of the better performing kids in my year for everything apart from reading Welsh . . . and that was really embarrassing and I think that's why I don't like it. I got good [grades] in the GCSEs, like I got Bs and stuff but I just don't like it [CEF2].

2. When I'm working . . . I get a lot of people coming to us saying we need to do more signs in Welsh. It's unnecessary [CEM3].

In Caernarfon, home language tended to be a primary indicator of peer-group membership (although certain individuals floated between friendship groups, see also Musk 2006), and it was common to hear those from English-speaking families speak English to those from Welsh-speaking families and receive a reply in Welsh (and vice versa, see (Gafaranga and Calvo 2001) for a similar finding in Catalonia).

The situation among the bilinguals in Mold was much different. Overt attitudes to Welsh were positive regardless of home-language background<sup>1</sup>, and language was rarely discussed in the sociolinguistic interviews. It became apparent, at least among the wider peer groups, that English was the main language of interaction and Welsh was reserved for educational contexts. This had been long-established, with one speaker from a Welsh-speaking home commenting in the Welsh sociolinguistic interview that she had switched to using English with her friends at the beginning of secondary school (some six years previous) because of comments from her wider peer group (author's translation):

3. I used to speak Welsh with my friends (from primary school) and then I remember there was someone who used to call me swot for speaking Welsh and I felt really upset . . . I still speak Welsh with my best friend but it's English that we speak (with other people) really [MWF1].

### 3.3. Procedure

Data were collected in a quiet room on school premises by the author, who is a Welsh-English bilingual from north Wales. During the first session, speakers completed a sociolinguistic interview and wordlist task in Welsh. Another sociolinguistic interview and wordlist, held in English, took place the following day. Interview modules were devised which aimed to elicit informal speech. The topics included in the Welsh interview were (1) childhood, (2) family, and (3) travelling. In the English interviews, speakers were asked questions regarding (1) their experiences of school, (2) the local area, and (3) their leisure activities. Each interview lasted for around 35 minutes.

### 3.4. Data Coding and Analysis

#### 3.4.1. Data Coding

Up to 50 tokens of /r/ in prevocalic and intervocalic positions were extracted from each speaker's interview data (up to 25 tokens in each language). Only tokens that occurred following the first ten minutes of the interview were extracted, and only three instances of each word were included ( $n = 1577$ ). A further 46 tokens were extracted from the wordlist data, although there were 57 instances of repetitions, which were also included in the analysis ( $n = 1529$ ). Table 2 shows the number of tokens included in the analysis for each language by speaker area, gender, and home language.

**Table 2.** Number of tokens included in the analysis of /r/ in prevocalic and intervocalic positions (EHL = English home language, WHL = Welsh home language).

	Mold				Caernarfon				Total
	Female Speakers		Male Speakers		Female Speakers		Male Speakers		
	EHL	WHL	EHL	WHL	EHL	WHL	EHL	WHL	
<b>English</b>	189	196	197	194	196	182	190	196	1540
<b>Welsh</b>	188	182	197	197	184	230	188	200	1566
<b>Total</b>	377	378	394	391	380	412	378	396	3106

<sup>1</sup> As one reviewer pointed out, attitudes towards the language may differ among the monolingual population of Mold.

Tokens were coded auditorily by the author, though each token was checked acoustically in Praat (Boersma and Weenink 2021) before a final decision was made (e.g., Chand 2010). Table 3 details the acoustic cues used for the identification of /r/ variants:

**Table 3.** Acoustic cues used for the identification of /r/ variants.

Variant	Acoustic Cues
[ɹ]	Decrease in the distance between F2 and F3, caused by lowering F3 and increasing F2 out of the vowel accompanied by a decrease in amplitude (Ladefoged 2003, p. 149).
[r]/[r̥]	Decrease in glottal energy at the end of the vowel, shown on the spectrogram as a decrease in amplitude and loss of formant structure, followed by a single contact for [r̥] and two or more contacts for [r] (Jones 2009).

Each token was categorised as approximant, tap, trill, uvular fricative, or zero realisations. A total of 0.74% of tokens contained no audible production of /r/ (all in syllable onset position) and were omitted from the subsequent statistical analysis ( $n = 23$ ). A ternary distinction was made between voiced, partially devoiced, and devoiced productions, but an analysis of voicing was not included in the subsequent analysis.

A sample of 320 tokens (10.30% of the total number of tokens) was re-checked by the author. Ten interview tokens from each speaker were extracted and re-coded blindly following the same procedure, as outlined above (auditory coding and acoustic checking). The same decision was reached on the second round of coding for 317 of the tokens that yielded an agreement rate of 99.06%. In the case of the three tokens, about which there was doubt, the original coding was retained.

### 3.4.2. Statistical Analysis

The statistical analysis of the /r/ data was conducted using mixed-effects logistic regression analyses. All statistical modelling was carried out using the *lme4* (Bates Douglas et al. 2015) and *lmerTest* (Kuznetsova Alexandra and Christensen 2017) packages for R (R Core Team 2020). Mixed-effects modelling allows the researcher to distinguish between fixed and random effects (Baayen 2009, p. 241). Fixed effects (e.g., Gender) are those factors that are replicable in further studies, whereas random effects (such as speaker and word) are sampled randomly (Baayen 2009, p. 241). By including speaker and word as random effects, the modelling is able to account for inter-speaker or inter-item variation when predicting which factors influence variation (Johnson 2009, p. 365).

In order to make comparisons between areas, separate models were intended to be conducted on the Mold and Caernarfon datasets, which included language as a fixed effect. For Mold, this was not possible as there was no variation in English, and modelling was undertaken on the Welsh data only. For Caernarfon, I provide the final model based on the Welsh and English data in Appendix A and present the models for English and Welsh separately in order to aid the interpretation of the interactions between language and social factors. Table 4 shows the factors included in the models. Interactions were also included in the models and are reported in Section 4.

Model selection proceeded as follows. Firstly, the model included the maximal random effects structure. Secondly, the random effects structure was iteratively simplified until the model converged (Baranowski and Turton 2020; Bates Douglas et al. 2015). The example below shows the R code for the Mold Welsh model containing the maximal random effects structure:

```
VARIANT ~ HOME LANGUAGE (HL) + TASK + STRESS + GENDER + CONTEXT + SYLLABLES + HL:TASK + HL:GENDER + GENDER:TASK + CONTEXT:STRESS + (1 + TASK + CONTEXT + STRESS + CONTEXT:STRESS | SPEAKER) + (1 | WORD)
```

The most complex model tested, shown above, included by-speaker random slopes for task, context, stress, and interaction of context and stress. The final models, reported in Section 4 below, contained speaker and word as random effects with no random slopes.

**Table 4.** Factors included in the statistical analysis of /r/ variation (speaker and word included in all final models as random effects).

Factor	Levels (Bold = Baseline)
<i>Dependent Variable</i>	
Variant	<b>Approximant</b> Other
<i>Independent Variables</i>	
Language	<b>English</b> Welsh
Gender	<b>Female</b> Male
Home Language	<b>English</b> Welsh
Task	<b>Interview</b> Wordlist
Phonological Context	<b>V_V</b> #_V C_V
Syllable Type	<b>Monosyllabic</b> Polysyllabic
Syllable stress	<b>Stressed</b> Unstressed

The individual results tables show the fixed factors (independent variables), which were significant predictors of /r/ variation and coefficients ( $\beta$ ), z-values, and p-values for the levels associated with each factor. Taking into consideration the random factors (see above), these coefficients represent a deviation from the baseline (see Table 4 for the levels of each factor taken as the baseline). A positive significant coefficient suggests that the named factor level was more likely to influence the production of the approximant than the baseline factor level. Similarly, a negative significant coefficient indicates that the named factor level was less likely to result in the production of the approximant than the baseline factor level.

#### 4. Results

The presentation of the results proceeds as follows. Firstly, descriptive statistics of the overall frequency of tokens in each language are provided. This gives an overview of the results for reference and also establishes the extent to which transfer occurs between English and Welsh in these data. Secondly, the results of the mixed-effects modelling are described for (1) Mold and (2) Caernarfon. In both areas, the English-language subsets are presented first and then are followed by the Welsh-language subsets.

##### 4.1. Overall Frequency

Looking at the dataset on the whole, excluding the zero realisations, it is clear that the overwhelming majority of tokens were realised as approximants. In total, 79.43% of tokens were produced as approximants ( $n = 2497$ ). Of the non-approximant tokens, 11.24% were produced as taps ( $n = 349$ ), 6.67% were produced as trills ( $n = 207$ ), and 0.97% were produced as uvular fricatives ( $n = 30$ ). Table 5 shows the percentage and number of tokens by language.

**Table 5.** Percentage and number of /r/ variants in the English and Welsh subsets (*n* = 3083).

Variant	%		<i>n</i>		
	English	Welsh	English	Welsh	Total
<b>Uvular fricative</b>	0	1.93	0	30	30
<b>Trill</b>	4.12	9.27	63	144	207
<b>Tap</b>	5.16	17.39	79	270	349
<b>Approximant</b>	90.72	71.41	1388	1109	2497
<b>Total</b>	-	-	1530	1553	3083

Before proceeding to the statistical analysis of the Mold and Caernarfon datasets, it is worth paying attention to the appearance of the uvular fricatives in the data. The uvular fricatives were found in the Welsh of one speaker from the Mold area. Despite not being expected in the areas under discussion (see Section 2.3), further questioning of the speaker revealed family connections to the Bala area, where uvular variants have been previously reported.

4.2. Mold Dataset

As stated in Section 3.4.2, no statistical modelling was undertaken on the entire Mold dataset as there was no variation in the English-language data (see Section 4.2.1 below). Instead, the data were divided into English-language and Welsh-language subsets.

4.2.1. English-Language Subset

Inspection of the data indicated that no variation was present in the English data and that all tokens were produced as approximants (*n* = 771).

4.2.2. Welsh-Language Subset

In the Mold Welsh-language subset (*n* = 754), 84.35% of tokens were realised as approximants (*n* = 636), compared to 8.75% of tokens realised as taps (*n* = 66) and 2.92% of trills (*n* = 22). One speaker produced instances of the uvular fricative associated with the Welsh of the north-western town of Bala and the surrounding area (see Section 4.1). These tokens accounted for 3.98% of the total Mold Welsh-language subset (*n* = 30).

Table 6 shows the final model for the Mold Welsh-language subset (see Section 3.4.2). The model predicted the likelihood of the production of the approximant and contained speaker gender, speaker home language, task, phonological context, syllable type, and syllable stress as fixed effects. Interactions were included between home language and task, home language and gender, gender and task, and phonological context and syllable stress. Speaker and word were included as random effects<sup>2</sup>.

Firstly, the results for the Mold Welsh-language subset suggested that the realisation of the alveolar approximant in Welsh was less likely to occur in the speech of those from Welsh-speaking homes ( $\beta = -3.735, z = -2.711, p = 0.007$ ). Of the Welsh tokens produced by those from Welsh-speaking homes, 73.85% were approximants (*n* = 274) compared to 94.52% of tokens produced by speakers from English-speaking homes (*n* = 362).

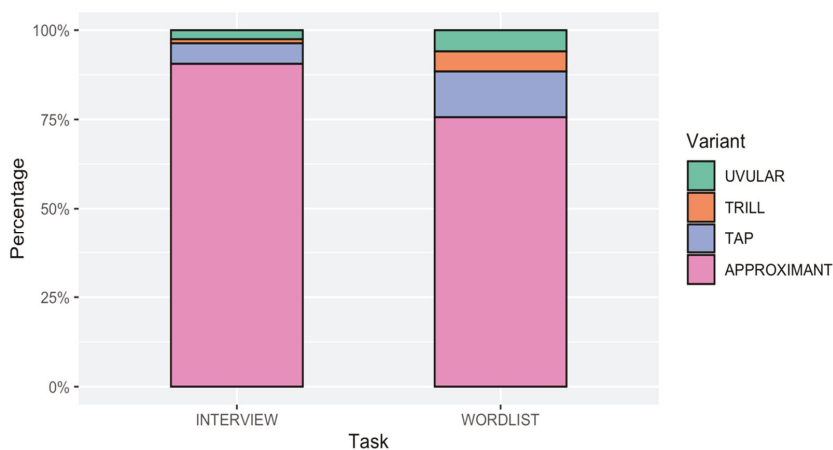
Secondly, the results also suggested that the alveolar approximant was less likely to occur in the wordlist task than in the interview data ( $\beta = -1.385, z = -2.016, p = 0.044$ ). Figure 1 shows the distribution of variants in the interview and wordlist tasks. Of the tokens produced during the sociolinguistic interview, 90.78% were approximants (*n* = 394) compared to 75.63% of tokens in the wordlist data (*n* = 242).

<sup>2</sup> R code: VARIANT ~ HOME LANGUAGE (HL) + TASK + STRESS + GENDER + CONTEXT + SYLLABLES + HL:TASK + GENDER:HL + GENDER:TASK + CONTEXT:STRESS + (1|SPEAKER) + (1|WORD)

**Table 6.** Regression coefficients with *z*- and *p*-values for the final model predicting the production of the alveolar approximant in the Mold Welsh-language subset (*n* = 754). Positive estimates indicate an increased likelihood of the alveolar approximant. AIC = 473.6.

	Estimate	Std. Error	z-Value	Pr(>  z )	Signif.
(Intercept)	3.442	1.199	2.870	0.004	**
Home Language (Baseline: English)					
Welsh	−3.735	1.378	−2.711	0.007	**
Task (Baseline: Interview)					
Wordlist	−1.385	0.687	−2.016	0.044	*
Syllable Stress (Baseline: Stressed)					
Unstressed	−0.582	0.504	−1.155	0.248	
Speaker Gender (Baseline: Female)					
Male	0.029	1.407	0.021	0.983	
Phonological Context (Baseline: V_V)					
#_V	1.184	0.523	2.263	0.024	*
C_V	1.499	0.517	2.903	0.004	**
Syllable Type (Baseline: Monosyllabic)					
Polysyllabic	0.708	0.438	1.615	0.106	
Home Language and Task (Baseline: English and Interview)					
Welsh and Wordlist	0.650	0.690	0.942	0.346	
Home Language and Gender (Baseline: English and Male)					
Welsh and Male	2.402	1.849	1.299	0.194	
Task and Gender (Baseline: Interview and Female)					
Wordlist and Male	−0.635	0.605	−1.049	0.294	
Syllable Stress and Context (Baseline: Stressed and V_V)					
Unstressed and #_V	0.268	0.777	0.345	0.730	
Unstressed and C_V	1.155	1.029	1.122	0.262	

\* *p* ≤ 0.05, \*\* *p* ≤ 0.01.



**Figure 1.** Distribution of /r/ variants in the Mold Welsh-language subset by task (*n* = 754).

Finally, phonological context was shown to be a significant predictor of variation in the Mold Welsh-language subset. The alveolar approximant was more likely to occur in C\_V ( $\beta = 1.499, z = 2.903, p = 0.004$ ) and #\_V contexts ( $\beta = 1.184, z = 2.263, p = 0.024$ ) compared to intervocally (V\_V).

### 4.3. Caernarfon Dataset

As stated in Section 3.4.2, initial statistical modelling was undertaken on the whole Caernarfon dataset and included language as a fixed effect (as well as a number of interactions between language and social factors). The results of this modelling are shown in Table A1 in Appendix A. The remainder of this section presents the modelling on the English- and Welsh-language subsets separately.

#### 4.3.1. English-Language Subset

In the Caernarfon English-language subset ( $n = 759$ ), 81.29% of tokens were produced as approximants ( $n = 617$ ), compared to 10.41% of tokens which were produced as taps ( $n = 79$ ) and 8.3% of tokens produced as trills ( $n = 63$ ).

Table 7 shows the final model for the Caernarfon English-language subset (see Section 3.4.2). The model predicted the likelihood of the production of the approximant and contained speaker gender, speaker home language, task, phonological context, syllable type, and syllable stress as fixed effects. Interactions were included between home language and task, gender and task, and phonological context and syllable stress. The interaction between home language and gender was removed as there was no variation in the data of male speakers from English-speaking homes. Speaker and word were included as random effects<sup>3</sup>.

**Table 7.** Regression coefficients with  $z$ - and  $p$ -values for the final model predicting the production of the alveolar approximant in the Caernarfon English-language subset ( $n = 759$ ). Positive estimates indicate an increased likelihood of the alveolar approximant. AIC = 471.5.

	Estimate	Std. Error	z-Value	Pr(>  z )	Signif.
(Intercept)	6.085	1.557	3.909	<0.001	***
Home Language (Baseline = English)					
Welsh	−4.968	1.669	−2.978	0.003	**
Task (Baseline = Interview)					
Wordlist	−2.373	0.783	−3.030	0.002	**
Syllable Stress (Baseline = Stressed)					
Unstressed	1.247	1.056	1.180	0.238	
Speaker Gender (Baseline = Female)					
Male	1.630	1.462	1.115	0.265	
Phonological Context (Baseline = V_V)					
#_V	0.359	0.468	0.766	0.444	
C_V	0.521	0.592	0.881	0.379	
Syllable Type (Baseline = Monosyllabic)					
Polysyllabic	−0.051	0.386	−0.131	0.896	
Home Language and Task (Baseline = English and Interview)					
Welsh and Wordlist	0.329	0.852	0.386	0.700	
Task and Gender (Baseline = Interview and Female)					
Wordlist and Male	1.896	0.636	2.980	0.003	**
Syllable Stress and Context (Baseline = Stressed and V_V)					
Unstressed and #_V	−1.901	1.248	−1.523	0.128	
Unstressed and C_V	−1.506	1.446	−1.042	0.298	

\*\*  $p \leq 0.01$ , \*\*\*  $p \leq 0.001$ .

<sup>3</sup> R code: VARIANT ~ HOME LANGUAGE (HL) + TASK + STRESS + GENDER + CONTEXT + SYLLABLES + HL:TASK + GENDER:TASK + CONTEXT:STRESS + (1|SPEAKER) + (1|WORD)



The results of the statistical modelling, shown in Table 7, suggested that those from Welsh-speaking homes were less likely to produce the alveolar approximant in the Caernarfon English-language subset compared to those from English-speaking homes ( $\beta = -4.968$ ,  $z = -2.978$ ,  $p = 0.003$ ). Of the English tokens produced by those from Welsh-speaking homes in Caernarfon, 68.53% were approximants ( $n = 257$ ), compared to 93.75% of tokens produced by speakers from English-speaking homes ( $n = 362$ ).

Although there was a significant main effect for task ( $\beta = -2.373$ ,  $z = -3.030$ ,  $p = 0.002$ ), the model also showed a significant interaction between speaker gender and task ( $\beta = 1.896$ ,  $z = 2.980$ ,  $p = 0.003$ ), which requires further examination. Figure 2 shows the percentage of variants produced in the Caernarfon English-language subset by task and speaker gender. The figure shows that 87.04% of the tokens produced by female speakers during the interview task were approximants ( $n = 141$ ) compared to 61.11% of the tokens produced during the wordlist task ( $n = 132$ ). In the male speakers' data, 89.51% of the tokens produced during the interview task were approximants ( $n = 128$ ) compared to 90.38% of the tokens produced during the wordlist task ( $n = 216$ ).

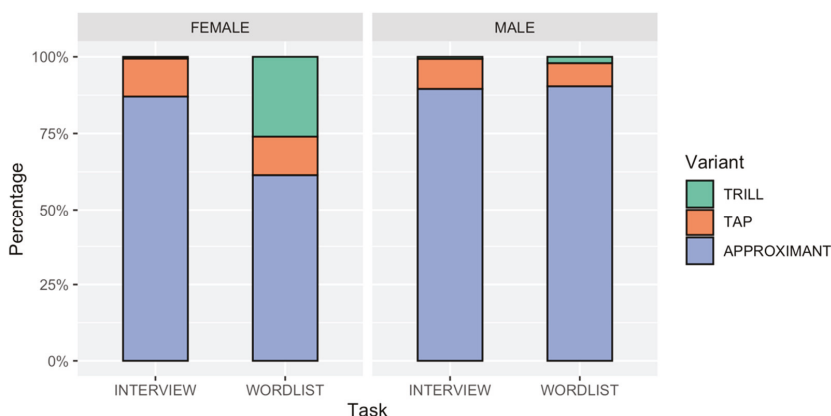


Figure 2. Distribution of /r/ variants in the Caernarfon English-language subset by speaker gender and task ( $n = 759$ ).

To summarise, the alveolar approximant was less likely to occur in the speech of those from Welsh-speaking homes in the Caernarfon English-language subset. A significant interaction between speaker gender and task was also found, with female speakers being more likely to produce fewer approximant tokens in the wordlist task when compared to the interview task. The following section examines the Caernarfon Welsh-language subset.

#### 4.3.2. Welsh-Language Subset

In the Caernarfon Welsh-language subset ( $n = 799$ ), 59.20% of tokens were realised as approximants ( $n = 473$ ), compared to 25.53% of tokens realised as taps ( $n = 204$ ) and 15.27% of tokens realised as trills ( $n = 122$ ).

Table 8 shows the final model for the Caernarfon Welsh-language subset (see Section 3.4.2). The model predicted the likelihood of the production of the approximant and contained speaker gender, speaker home language, task, phonological context, syllable type, and syllable stress as fixed effects. Interactions were included between home language and task, home language and gender, gender and task, and phonological context and syllable stress. Speaker and word were included as random effects<sup>4</sup>.

<sup>4</sup> R code: VARIANT ~ HOME LANGUAGE (HL) + TASK + STRESS + GENDER + CONTEXT + SYLLABLES + HL:TASK + GENDER:HL + GENDER:TASK + CONTEXT:STRESS + (1 | SPEAKER) + (1 | WORD)

**Table 8.** Regression coefficients with *z*- and *p*-values for the final model predicting the production of the alveolar approximant in the Caernarfon Welsh-language subset (*n* = 799). Positive estimates indicate an increased likelihood of the alveolar approximant. AIC = 876.4.

	Estimate	Std. Error	z-Value	Pr(>  z )	Signif.
(Intercept)	0.390	0.632	0.617	0.537	
Home Language (Baseline = English) Welsh	−0.189	0.585	−0.323	0.746	
Task (Baseline = Interview) Wordlist	−1.164	0.397	−2.930	0.003	**
Syllable Stress (Baseline = Stressed) Unstressed	−0.394	0.497	−0.793	0.428	
Speaker Gender (Baseline = Female) Male	1.656	0.628	2.636	0.008	**
Phonological Context (Baseline = V_V) #_V C_V	0.110 0.882	0.431 0.428	0.254 2.059	0.800 0.039	 *
Syllable Type (Baseline = Monosyllabic) Polysyllabic	0.099	0.380	0.259	0.795	
Home Language and Task (Baseline = English and Interview) Welsh and Wordlist	−0.982	0.409	−2.402	0.016	*
Home Language and Gender (Baseline = English and Male) Welsh and Male	−1.517	0.819	−1.852	0.064	
Task and Gender (Baseline = Interview and Female) Wordlist and Male	0.634	0.401	1.581	0.114	
Syllable Stress and Context (Baseline = Stressed and V_V) Unstressed and #_V Unstressed and C_V	0.294 0.427	0.676 0.777	0.435 0.549	0.664 0.583	

\* *p* ≤ 0.05, \*\* *p* ≤ 0.01.

The results for the Caernarfon Welsh-language subset indicated that the alveolar approximant in Welsh was more likely to occur in male speakers’ speech ( $\beta = 1.656$ ,  $z = 2.636$ ,  $p = 0.008$ ). Of the tokens produced by male speakers, 67.62% were approximants ( $n = 261$ ). In the female speakers’ data, 51.33% of tokens were produced as approximants ( $n = 212$ ).

Although home language was not a significant predictor of the realisation of the alveolar approximant in the Caernarfon Welsh-language subset ( $\beta = -0.189$ ,  $z = -0.323$ ,  $p = 0.746$ ), there was a significant interaction between home language and task ( $\beta = -0.982$ ,  $z = -2.402$ ,  $p = 0.016$ ). Those from Welsh-speaking homes were less likely to produce the alveolar approximant in the wordlist task than in the interview task compared to those from English-speaking homes. Figure 3 shows the distribution of variants in the interview and wordlist tasks by speaker home language.

In the Welsh data produced by speakers from Welsh-speaking homes in Caernarfon, 63.94% of tokens produced in the interview task were approximants ( $n = 172$ ) compared to 25.16% of tokens produced during the wordlist task ( $n = 40$ ). In the Welsh data obtained from those from English-speaking homes in Caernarfon, 78.34% of tokens in the interview task were produced as approximants ( $n = 170$ ) compared to 59.09% of tokens in the wordlist task ( $n = 91$ ).

Finally, the phonological context was found to be significant. The alveolar approximant was more likely to occur in C\_V position ( $\beta = 0.882$ ,  $z = 2.059$ ,  $p = 0.039$ ) than intervocally.



Figure 3. Distribution of /r/ variants in the Caernarfon Welsh-language subset by home language and task ( $n = 799$ ).

## 5. Discussion and Conclusions

The study sought to examine (1) the variants of /r/ present in the bilingual repertoire of Welsh-English bilinguals, (2) the linguistic and extra-linguistic factors which influence variation in both Welsh and English in two communities, and (3) the extent to which there were differences between two communities which differ sociolinguistically.

The alveolar approximant was the most common variant in the bilingual repertoire of most of the speakers included in this study. The Welsh variants (the trill, tap, and, in the case of one speaker, the uvular fricative) were wholly absent from the English data in Mold, although they appeared in variation with the approximant in Welsh. Although looking at the raw percentages showed that the vast majority of tokens in the Caernarfon English-language subset were produced as approximants (80.76% compared to 58.98% in the Caernarfon Welsh-language subset), the mixed-effects modelling showed that there were social effects on variation in both languages, and, consequently, the cross-linguistic differences in Caernarfon were less clear.

The high frequency of the alveolar approximant in the Welsh data provides evidence for phonological transfer in the two areas and among both traditional and new speakers. This contradicts previous work on Welsh dialectology, which posits that the alveolar approximant is an idiosyncratic feature of Welsh or restricted to certain border areas (see Section 2.3). Both the descriptive and inferential statistics presented in Section 4 indicate that the appearance of the alveolar approximant in Welsh is not ephemeral ‘interference’ in the speech of Welsh-English bilinguals in both communities and that the alveolar approximant in Welsh is a consistent transfer feature which is subject to social constraints.

It is not possible to comment on the extent to which the production of the alveolar approximant in Welsh constitutes language change in progress (to the extent to which this is possible in certain revitalisation contexts see Nance 2015, p. 573). Further comparisons with older speakers in both areas would be needed to substantiate this claim. However, it is clear that there are linguistic and extra-linguistic factors that influence variation in both English and Welsh and that these factors pattern differently in the two communities under discussion. It is to these two points which I now turn.

The linguistic factors included in the analysis were not significant predictors of /r/ in the English of the two communities. In Welsh, the results in the two areas showed that the alveolar approximant was favoured in C\_V contexts. This supports previous work in Welsh, which claims that the approximant is common in onset clusters (Jones 1984, pp. 59–60).

In work based on the same dataset analysed in the current study, Morris (2017) found no home-language differences for /l/-darkening (a gradient phonetic feature). Home language was found to be significant in Caernarfon, however, for measures related to Fundamental Frequency Range (Morris, forthcoming). In other work, no home language differences were found in the production of monophthongs (Mayr et al. 2017) or lexical

stress (Mennen et al. 2020) in either the Welsh or English of Carmarthenshire or in the production of high back vowels in Cardiff Welsh (Gruffydd, forthcoming). Previous work also showed that certain accentual cues in English helped listeners to differentiate between Welsh- and non-Welsh-speakers (Mayr et al. 2020) and that Welsh variants of /r/, as well as prosodic features, were markers of strongly Welsh-accented speech associated with Welsh-speaking areas (Penhallurick 1991; Wells 1982, p. 390; Wilson 2014; Mayr et al. 2020).

The results of the modelling indicate that those from English-speaking homes were more likely to produce the approximant in both Caernarfon and Mold Welsh. The significance of home language on this feature in Welsh was unsurprising in light of the varying degrees of exposure to Welsh each group had received. Home language tended to correlate with self-reported usage and ability in Welsh, which in turn were highly predictive of the role Welsh is likely to play in the life of an adolescent speaker (Musk 2006; Morris 2014). As was shown in Section 3.2, the importance of home language as a marker of identity among the speakers seemed to be more obvious in Caernarfon, where peer groups often used either Welsh or English. This could explain the more prominent home-language differences. In Mold, the role of home language was much more subtle, and my own observations were that those Welsh-speaking background were much more eager to ‘fit in’ with their peers from English-speaking homes. This could go some way to describe the lack of home-language differences in Mold English, but further ethnographic work would be needed to examine this further.

Task was a significant predictor of variation in Mold Welsh, and the results for Caernarfon showed a significant interaction between home language and task in Welsh. In Caernarfon Welsh, those from Welsh-speaking homes were more likely to style-shift and produce fewer approximant tokens in the wordlist tasks compared to those from English-speaking homes. The results for the Welsh data pointed towards stylistic variation, wherein speakers tended to produce more standard variants in more formal speech. The fact that speakers from English-speaking homes in Caernarfon were less likely to style-shift provided, in my opinion, more evidence that /r/ may have socio-indexical meaning in this area.

Gender differences were also found to operate independently of differences between the two home-language groups in Caernarfon. In the Caernarfon English-language subset, female speakers were more likely to style-shift and produce fewer approximant tokens in the wordlist task. In Welsh, they were also less likely to produce the alveolar approximant regardless of the task. Differences in patterns of style-shifting, and the production of standard variants between male and female speakers, are well-attested in variationist sociolinguistics (Kuznetsova Alexandra and Christensen 2001, p. 274), but generalisations across communities are also problematic (e.g., Eckert and McConnell-Ginet 1992). The fact that female speakers orient away from the alveolar approximant in both Welsh and English cannot be easily reconciled with the idea that women are more likely to orient towards standard norms. Instead, they appear to be more likely to produce the standard variant in their Welsh and style-shift towards a local norm and Welsh-accented speech in their English. Further work on the perception of /r/ in both varieties might shed further light on what social evaluations speakers hold with regard to this feature and whether such evaluations differ between Welsh and English.

The absence of the traditionally Welsh-language features from Mold English contributes to our understanding of the notion of language mode in bilingualism studies (e.g., Grosjean 1989), and, in particular, the notion that socio-psychological factors influence cross-linguistic interactions. I would argue that local accentual norms may be considered one such factor and that the young bilingual speakers in Mold tend to adhere to local norms among English monolinguals in their categorical use of the alveolar approximant. This differs from Caernarfon, where local norms are based on a majority population of bilingual speakers. Similar results were found in perception studies of Welsh English accents, where typically Welsh-influenced features are associated with areas with a high proportion of Welsh-English bilinguals (e.g., Williams et al. 1996).

The results indicate that there are clear differences between communities which can, to a certain extent, be explained by the sociolinguistic differences, which exist between Caernarfon and Mold and, more specifically, the peer groups included in the current study.

**Funding:** This research was partly funded by the Johansson Scholarship, awarded by the School of Arts, Languages and Cultures, University of Manchester.

**Institutional Review Board Statement:** The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Institutional Review Board (or Ethics Committee) of the School of Arts, Languages and Cultures, University of Manchester.

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** The data presented in this study are available on request from the author. The data are not publicly available due to confidentiality restrictions.

**Acknowledgments:** I would like to thank the anonymous reviewers for their insightful comments.

**Conflicts of Interest:** The author declares no conflict of interest.

## Appendix A

**Table A1.** Regression coefficients with *z*- and *p*-values for the final model predicting the production of the alveolar approximant in Caernarfon (*n* = 1558). Positive estimates indicate an increased likelihood of the alveolar approximant. AIC = 1350.7.

	Estimate	Std. Error	z-Value	Pr(>  z )	Signif.
(Intercept)	3.189	0.733	4.348	<0.001	***
Language (Baseline: English) Welsh	−2.418	0.432	−5.593	<0.001	***
Home Language (Baseline: English) Welsh	−1.932	0.821	−2.355	0.019	*
Task (Baseline: Interview) Wordlist	−0.986	0.431	−2.287	0.022	*
Syllable Stress (Baseline: Stressed) Unstressed	−0.077	0.245	−0.312	0.755	
Speaker Gender (Baseline: Female) Male	2.464	0.891	2.764	0.006	**
Phonological Context (Baseline: V_V) #_V C_V	0.152 0.660	0.263 0.308	0.575 2.142	0.565 0.032	 *
Syllable Type (Baseline: Monosyllabic) Polysyllabic	−0.081	0.263	−0.309	0.757	
Home Language and Task (Baseline: English and Interview) Welsh and Wordlist	−0.634	0.355	−1.786	0.074	
Home Language and Gender (Baseline: English and Male) Welsh and Male	−1.620	1.075	−1.507	0.132	
Task and Gender (Baseline: Interview and Female) Wordlist and Male	1.095	0.334	3.278	0.001	**
Language and Home Language (Baseline: English and English HL) Welsh and Welsh HL	1.601	0.425	3.766	<0.001	***
Language and Gender (Baseline: English and Female) Welsh and Male	1.601	0.425	3.766	<0.001	***
Language and Task (Baseline: English and Interview) Welsh and Wordlist	−0.755	0.414	−1.825	0.068	

\* *p* ≤ 0.05, \*\* *p* ≤ 0.01, \*\*\* *p* ≤ 0.001.

## References

- Awbery, Gwenllian. 1997. The English language in Wales. In *The Celtic Englishes*. Edited by Hildegard L. C. Tristram. Heidelberg: Winter, pp. 86–99.
- Baayen, Harald. 2009. *Analyzing Linguistic Data: A Practical Introduction to Statistics Using R*. Cambridge: Cambridge University Press.
- Baranowski, Maciej, and Danielle Turton. 2020. TD-deletion in British English: New evidence for the long-lost morphological effect. *Language Variation and Change* 32: 1–23. [CrossRef]
- Bates Douglas, Martin Mächler, Ben Bolker, and Steve Walker. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67: 1–48.
- Boersma, Paul, and David Weenink. 2021. Praat: Doing Phonetics by Computer [Computer Program]. Version 6.1.40. Available online: <http://www.praat.org/> (accessed on 21 March 2021).
- Chand, Vineeta. 2010. Postvocalic (r) in urban Indian English. *English World-Wide* 31: 1–39. [CrossRef]
- Coupland, Nikolas, and Martin J Ball. 1989. Welsh and English in contemporary Wales: Sociolinguistic issues. *Contemporary Wales* 3: 7–40.
- Davidson, Justin Spencer. 2015. Social Dynamics of Catalan-Spanish Contact in the Evolution of Catalanian Spanish. Ph.D. thesis, University of Illinois at Urbana-Champaign, Champaign, IL, USA.
- Davies, Lyn. 1971. Linguistic interference in East Montgomeryshire. *The Montgomeryshire Collections* 62: 183–94.
- Durham, Mercedes, and Jonathan Morris. 2016. An overview of sociolinguistics in Wales. In *Sociolinguistics in Wales*. Edited by Mercedes Durham and Jonathan Morris. London: Macmillan, pp. 3–28.
- Eckert, Penelope, and Sally McConnell-Ginet. 1992. Think practically and look locally: Language and gender as community-based practice. *Annual Review of Anthropology* 21: 461–88. [CrossRef]
- Gafaranga, Joseph, and Maria-Carme Torras i Calvo. 2001. Language versus medium in the study of bilingual conversation. *International Journal of Bilingualism* 5: 195–219. [CrossRef]
- Gafter, Roey J., and Uri Horesh. 2020. Two languages, one variable? Pharyngeal realizations among Arabic–Hebrew bilinguals. *Journal of Sociolinguistics* 24: 369–87. [CrossRef]
- Grosjean, François. 1989. Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language* 36: 3–15. [CrossRef]
- Grosjean, François. 2001. The bilingual’s language modes. In *One Mind, Two Languages: Bilingual Language Processing*. Edited by Janet Nicol. Oxford: Blackwell, pp. 1–22.
- Gruffydd, Ianto. Forthcoming. Amrywio Ieithyddol Mewn Ysgol Gymraeg yng Nghaerdydd. Ph.D. thesis, Cardiff University, Cardiff, UK.
- Hodges, Rhian Siân. 2012. Welsh-medium education and parental incentives—the case of the Rhymini Valley, Caerffili. *International Journal of Bilingual Education and Bilingualism* 15: 355–73. [CrossRef]
- Johnson, Daniel Ezra. 2009. Getting off the GoldVarb standard: Introducing Rbrul for mixed-effects variable rule analysis. *Language and Linguistics Compass* 3: 359–83. [CrossRef]
- Jones, Glyn E. 1984. The distinctive vowels and consonants of Welsh. In *Welsh Phonology: Selected Readings*. Edited by Martin J. Ball and Glyn E. Jones. Cardiff: University of Wales Press, pp. 40–64.
- Jones, Hywel M. 2008. The changing social context of Welsh: A review of statistical trends. *International Journal of Bilingual Education and Bilingualism* 11: 541–57. [CrossRef]
- Jones, Mark J. 2009. Patterns of variability in apical trills: An acoustic study of data from 19 languages. unpublished manuscript.
- Kuznetsova Alexandra, Per B. Brockhoff, and Rune H. B. Christensen. 2017. lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software* 82: 1–26. [CrossRef]
- Kuznetsova Alexandra, Per B. Brockhoff, and Rune H. B. Christensen. 2001. *Principles of Linguistic Change. Volume 2: Social Factors*. Oxford: Blackwell.
- Ladefoged, Peter. 2003. *Phonetic Data Analysis: An Introduction to Fieldwork and Instrumental Techniques*. Oxford: Blackwell.
- Mayr, Robert, and Simona Montanari. 2015. Cross-linguistic interaction in trilingual phonological development: The role of the input in the acquisition of the voicing contrast. *Journal of Child Language* 42: 1006–35. [CrossRef] [PubMed]
- Mayr, Robert, Jonathan Morris, Ineke Mennen, and Daniel Williams. 2017. Disentangling the effects of long-term language contact and individual bilingualism: The case of monophthongs in Welsh and English. *International journal of bilingualism* 21: 245–67. [CrossRef]
- Mayr, Robert, Llian Roberts, and Jonathan Morris. 2020. Can you tell by their English if they can speak Welsh? Accent perception in a language contact situation. *International Journal of Bilingualism* 24: 740–66. [CrossRef]
- Mennen, Ineke. 2004. Bi-directional interference in the intonation of Dutch speakers of Greek. *Journal of Phonetics* 32: 543–63. [CrossRef]
- Mennen, Ineke, Niamh Kelly, Robert Mayr, and Jonathan Morris. 2020. The effects of home language and bilingualism on the realization of lexical stress in Welsh and Welsh English. *Frontiers in Psychology* 10: 3038. [CrossRef]
- Mooney, Damien. 2019. Phonetic transfer in language contact: Evidence for equivalence classification in the mid-vowels of Occitan–French bilinguals. *International Phonetic Association. Journal of the International Phonetic Association* 49: 53–85. [CrossRef]
- Morris, Jonathan. 2013. Sociolinguistic Variation and Regional Minority Language Bilingualism: An Investigation of Welsh-English Bilinguals in North Wales. Ph.D. thesis, University of Manchester, Manchester, UK.

- Morris, Jonathan. 2014. The influence of social factors on minority language engagement amongst young people: An investigation of Welsh-English bilinguals in North Wales. *International Journal of the Sociology of Language* 2014: 65–89. [CrossRef]
- Morris, Jonathan. 2017. Sociophonetic variation in a long-term language contact situation: /l/-darkening in Welsh-English bilingual speech. *Journal of Sociolinguistics* 21: 183–207. [CrossRef]
- Morris, Jonathan. Forthcoming. Fundamental frequency range in the bilingual repertoire of traditional and new Welsh speakers. *International Journal of Bilingualism*.
- Musk, Nigel. 2006. Performing Bilingualism in Wales with the Spotlight on Welsh: A Study of Language Policy and the Language Practices of Young People in Bilingual Education. Ph.D. thesis, Linköping University, Linköping, Sweden.
- Nagy, Naomi. 2015. A sociolinguistic view of null subjects and VOT in Toronto heritage languages. *Lingua* 164: 309–27. [CrossRef]
- Nance, Claire. 2013. Phonetic Variation, Sound Change, and Identity in Scottish Gaelic. Ph.D. thesis, University of Glasgow, Glasgow, Scotland.
- Nance, Claire. 2015. ‘New’ Scottish Gaelic speakers in Glasgow: A phonetic study of language revitalisation. *Language in Society* 44: 553–79. [CrossRef]
- Odling, Terence. 1989. *Language Transfer: Cross-Linguistic Influence in Language Learning*. Cambridge: Cambridge University Press.
- Ordin, Mikhail, and Ineke Mennen. 2017. Cross-linguistic differences in bilinguals’ fundamental frequency ranges. *Journal of Speech, Language, and Hearing Research* 60: 1493–506. [CrossRef] [PubMed]
- Parry, David. 1977. *The Survey of Anglo-Welsh Dialects. Volume 1: The South-East*. Swansea: University College Swansea.
- Penhallurick, Robert J. 1991. The Anglo-Welsh dialects of north Wales. A survey of conservative rural spoken English in the counties of Gwynedd and Clwyd. *Bamberger Beiträge zur Englischen Sprachwissenschaft* 27.
- Piske, Thorsten, Ian R. A. MacKay, and James E. Flege. 2001. Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics* 29: 191–215. [CrossRef]
- R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. Available online: <https://www.R-project.org/> (accessed on 1 April 2020).
- Simon, Ellen. 2010. Phonological transfer of voicing and devoicing rules: Evidence from L1 Dutch and L2 English conversational speech. *Language Sciences* 32: 63–86. [CrossRef]
- Simonet, Miquel. 2010. Dark and clear laterals in Catalan and Spanish: Interaction of phonetic categories in early bilinguals. *Journal of Phonetics* 38: 663–78. [CrossRef]
- Simonet, Miquel. 2014. Phonetic consequences of dynamic cross-linguistic interference in proficient bilinguals. *Journal of Phonetics* 43: 26–37. [CrossRef]
- Simonet, Miquel, and Mark Amengual. 2020. Increased language co-activation leads to enhanced cross-linguistic phonetic convergence. *International Journal of Bilingualism* 24: 208–21. [CrossRef]
- Tomé Lourido, Gisela, and Bronwen G. Evans. 2021. Sociolinguistic Awareness in Galician Bilinguals: Evidence from an Accent Identification Task. *Languages* 6: 53. [CrossRef]
- Wells, John Christopher. 1982. *Accents of English: The British Isles*. Cambridge: Cambridge Univ. Press.
- Welsh Government. 2012. Welsh Speakers by Local Authority, Gender and Detailed Age Groups, 2011 Census. Available online: <https://stats.wales.gov.wales/Catalogue/Welsh-Language/Census-Welsh-Language/welshspeakers-by-localauthority-gender-detailedagegroups-2011census> (accessed on 21 March 2021).
- Welsh Government. 2020. Summary Statistics for North Wales Region: 2020. Available online: <https://gov.wales/sites/default/files/statistics-and-research/2020-05/summary-statistics-north-wales-region-2020-806.pdf> (accessed on 21 March 2021).
- Williams, Angie, Peter Garrett, and Nikolas Coupland. 1996. Perceptual dialectology, folklinguistics, and regional stereotypes: Teachers’ perceptions of variation in Welsh English. *Multilingua* 15: 171–200. [CrossRef]
- Wilson, Charles. 2014. The English of Welsh-Speakers: Extralinguistic Factors That Influence the Choice of Phonological Variant. Unpublished Master’s dissertation, University of London, London, UK.

Article

# Psycho-Social Constraints on Naturalistic Adult Second Language Acquisition

Azza Al-Kendi <sup>1</sup> and Ghada Khattab <sup>2,\*</sup>

<sup>1</sup> Department of English Language and Literature, Sultan Qaboos University, Al Seeb P.O. Box 42, Oman; azzakindi@sq.u.edu.om

<sup>2</sup> Speech and Language Sciences, School of Education, Communication and Language Sciences, Newcastle University, Newcastle upon Tyne NE1 7RU, UK

\* Correspondence: ghada.khattab@ncl.ac.uk

**Abstract:** The following study investigated a rare case of adult immersion in a second language context without prior exposure to the language. It aimed to investigate whether Length of Residence (LoR) acts as a strong index of L2 speech performance when coupled with daily exposure and interaction with first language speakers. Twenty-two females from Africa and Asia who worked as Foreign Domestic Helpers (FDH) in Omani homes and with varying LoRs performed an AX discrimination and a production task which tapped into Omani consonants and clusters that are absent from their L1s; their accent was also rated by L1 Omani listeners. Results showed a surprising lack of significance of LoR on all the production and perception measures examined. Discrimination results showed a low sensitivity to Arabic consonantal contrasts that are lacking in the L1 across all participants, and a small positive effect of L1 literacy. Production results exhibited low accuracy on all Arabic consonants and a marked foreign accent as judged by L1 listeners, with a small positive effect of L2 literacy. We argue that the nature of the interactions between FDH and employers, along with uneven power relations and social distance, counteract any advantage of LoR and the immersion setting examined here.

**Keywords:** length of residence; foreign domestic helper; foreign accent; naturalistic adult acquisition; L2 speech performance



**Citation:** Al-Kendi, Azza, and Ghada Khattab. 2021. Psycho-Social Constraints on Naturalistic Adult Second Language Acquisition. *Languages* 6: 129. <https://doi.org/10.3390/languages6030129>

Academic Editors: Robert Mayr and Jonathan Morris

Received: 7 June 2021  
Accepted: 22 July 2021  
Published: 28 July 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Over the past decades, a great deal of second language acquisition (SLA) research has focused on sociolinguistic factors that play a role in successful SLA. In naturalistic second language (L2) settings, age of learning (AoL) and length of residence (LoR) are among the most frequently studied predictors that have been found to affect second language speech learning (Piske et al. 2001). AoL has stood the test of time despite disagreements over whether or not there is a critical age or period for learning, but a fair number of studies have pointed out the unequal opportunities to learn the L2 between younger and older speakers (e.g., Klein and Perdue 1997; Craats et al. 2006). Similarly, LoR has proven robust when 'length' goes hand in hand with exposure. A host of factors have, however, been found to attenuate the effects of LoR, including experience with the L2 prior to arrival in the L2 country, the nature of the input, formal instruction, and first language (L1) literacy/education. Here we review some of this work before we turn our attention to an understudied population of L2 learners who fill a gap in terms of enabling us to test what happens when adult L2 learners receive extensive oral input from L1 speakers from the start and for prolonged periods of time.

SLA studies have long highlighted the difficulty adults face when learning L2 speech (Akahane-Yamada 1995; Best and Strange 1992; Flege 1981; Flege et al. 1995; Iverson and Kuhl 1996). The most widely discussed source for this difficulty is AoL. One of the most influential (and controversial) proposals in this area is that of the critical period hypothesis



(CPH), by [Lenneberg \(1967\)](#), who claimed that the ability to learn a new language declines after puberty due to the completion of neural hemispheric lateralization in the brain. A large body of work has both supported and challenged this claim, with disagreements over the existence of an age cut-off and evidence challenging the loss of neural plasticity in older age (e.g., [DeKeyser 2000](#); [Flege 1987, 2018](#); [Moyer 2014](#)). Nevertheless, and perhaps due to age acting as proxy for other optimal opportunities for L2 acquisition, studies involving naturalistic exposure to an L2 support at least a gradual decline in L2 learning outcomes as a function of age (e.g., [Flege 2018](#); [Flege et al. 1995](#); [Oyama 1976](#); [Pfenninger and Singleton 2017](#); [Seliger 1975](#); [Scovel 1988](#)). One linguistic by-product of age which is often attributed to the difficulty to learn the L2 sounds is increased L1 mastery and subsequent influence on the L2 (e.g., [Best et al. 1988](#); [Escudero and Boersma 2004](#); [Escudero 2006](#); [Flege 1995](#); [Kuhl 1991](#)). As L1 categories become more established with age, difficulties in perceiving and producing the L2 seem to arise ([Flege 1995](#)).

Length of residence (LoR) has also been frequently used in SLA research as an index for ultimate attainment in L2 phonology ([Flege 2009](#); [Moyer 2009](#)). [Flege \(2009\)](#) postulates that if the amount of input learners receive matters, then LoR should be correlated with measures of L2 speech attainment. Likewise, [McAllister \(2001\)](#) states that LoR correlates positively with the amount of input an L2 learner has acquired and that the more L2 input one receives the better the opportunities for the L2 learner to master the L2. A range of studies has provided evidence for the positive role of LoR in L2 performance. For instance, [Flege and Fletcher \(1992\)](#) revealed that Spanish adults who had lived in the United States for an average of 14.3 years received significantly better pronunciation ratings of English sentences than individuals with an average LoR of 0.7 years. Similarly, [Flege et al. \(1997\)](#) reported an effect of LoR on the perception of English vowels by L2 adult speakers of English who varied in their LoR between 0.7 and 7.3 years, albeit a modest one. A more significant effect was found on production accuracy, especially for one of the vowels they examined. In more recent work, [Højen \(2019\)](#) found a significant improvement in L1 Danish females' English pronunciation after 7.1 months of short-term immersion in England, with LoR significantly correlating with the participants' pronunciation gain score.

However, other reports on the significance of LoR effects reveal inconsistent results. For instance, [Oyama \(1976\)](#) and [Flege and Fletcher \(1992\)](#) found no effect of LoR on the L2 phonology of Italian and Spanish English speakers in the United States when the effect of their age of arrival (AoA) was controlled for. Similarly, [Flege \(1988\)](#) reported no difference in foreign accents between two groups of Taiwanese adult immigrants to the United States based on their LoR, which varied between 1.1 and 5.1 years. Further investigation revealed that LoR did not influence the degree of foreign accent after a rapid initial stage of L2 learning.

[Piske et al. \(2001\)](#) discuss a number of reasons for the discrepancies in previous studies. First, for studies that found LoR to have an influence on the degree of foreign accent, LoR was a less significant predictor compared to AoL. Second, LoR is more likely to affect the degree of foreign accent if the mean values of LoR of the L2 learner groups differ greatly. Third, additional years of stay in the L2 community are not likely to lead to a decrease in foreign-accented speech in already experienced L2 learners. However, L2 learners who are in the initial phases of learning the L2 when they arrive in the L2 country might benefit from additional years of experience ([Højen 2019](#)). This once again highlights the importance of input from L1 speakers from the start, as well as the importance of the cumulative effect of L2 exposure, as highlighted by [Flege and Bohn \(2021\)](#) in the revised Speech Learning Model (SLM-r); but most studies have investigated adult immigrants who had previously studied the L2 in their countries of origin, therefore being initially exposed to the L2 in a foreign language context. Little is known about the potential influence of total immersion in the L2 from first exposure for adults, bringing them closer to the experience of the children of immigrants.

While a naturalistic setting can be advantageous for the children of immigrants, who also receive formal instruction in the L2 country, adult immigrants may be disadvantaged due to the difficulty in getting access to formal instruction and/or due to coming from

low educational backgrounds (Klein and Perdue 1997; Craats et al. 2006). The research on oral/aural L2 performance of low educated learners is scarce. The vast majority of studies on SLA make use of convenience sampling and thus show an overreliance on a population which is WEIRD (Western, Educated, Industrialized, Rich, Democratic; Henrich et al. 2010). A significant proportion of the research that has been published in journals including Second Language Research, TESOL Quarterly, and Studies in Second Language Acquisition relies on such samples (Bigelow and Tarone 2004; Craats et al. 2006). Furthermore, few studies include L1 formal schooling as a contributing variable (Craats et al. 2006; Haznedar et al. 2018; Young-Scholten 2013). This bias towards recruiting and examining highly educated L2 learners might have skewed our understanding of second language speech and led to an under-explanation of how adults acquire a new language when we isolate factors such as first language literacy.

A body of work on non-literate and low-educated adult immigrant adults to the USA, Europe and Australia has focused on the challenges this poses for L2 literacy and metalinguistic awareness and calls for further research in this area (Craats et al. 2006; Kurvers et al. 2006; Young-Scholten and Strom 2006). However, difficulties do not necessarily arise in all language domains, with work within the area of morphosyntax suggesting that L2 learners can follow a common route in their L2 development of morphosyntax regardless of their age, educational background, L1 or input (e.g., Gass 2013; Ordem and Bada 2017). More work is still needed on the levels of attainment in all domains, including L2 oral production and perception. Little is known about whether L2 learners who are mainly exposed to extensive oral input from L1 speakers in a naturalistic setting achieve L2 speech outcomes that are more similar to those of L1 children.

To summarize, previous research on L2 speech learning has focused on age and LoR as main external factors in adults' ultimate attainment in the L2. However, the strength as well as vulnerability of these factors is due to their interaction with a multitude of other factors which can increase or reduce opportunities for learning. These include the amount of input from the L2 and opportunities for interaction with L1 speakers, the availability of formal instruction in the L2 and the level of L1 literacy prior to arrival in the L2 country, amongst others. The current study reports on a rare situation of naturalistic adult L2 acquisition through total immersion in the L2 context. It focusses on an understudied group of low-educated migrants from East Africa and South Asia who spend long years in the Arab world as domestic helpers, living with their Arab employer and their family and therefore mainly interacting with and receiving input from speakers who are first language users. The aim of the study is to investigate whether the situation provides an optimal context for L2 speech attainment given that conditions that typically strengthen LoR effects are maximized by the context; the learners have little or no previous exposure to foreign-accented Arabic, their LoR varies a great deal due to constant new arrivals, allowing for a comparison of short and long LoR, and their LoR highly correlates with input and interaction. The focus on Arabic as an L2 here is advantageous for two reasons: compared with English, Arabic is relatively understudied in SLA research (but see Ioup et al. 1994; Alhawari 2018); and by focusing on Arabic rather than English as an L2, there is a smaller chance that learners will have had exposure to it prior to arriving in the Arab world (apart from religious practices for some, which will be described later). On the other hand, social factors such as low L1 literacy and uneven power relations between employer and employee may attenuate LoR effects, but these have not been sufficiently considered in the speech-learning literature.

## 2. Materials and Methods

In what follows we present perception and production experiments that were carried out with foreign domestic helpers (FDH) who were living and working in Oman at the time of the study. The main aim was (1) to investigate the extent to which FDH had acquired Arabic consonants and clusters that were expected to pose a challenge in production and perception due to their articulatory and/or phonological complexity and their absence

from the L1s of the FDH; (2) the extent to which successful acquisition correlated with LoR and L1/L2 literacy. Production analyses were also supplemented with foreign-accent ratings that were carried out by L1 Arabic listeners. Here it is important to note that we are not espousing the view that FDH should sound like L1 speakers of Omani Arabic, or that this should be their aspiration. We concur with research that warns against the native speaker ideal (e.g., Holliday 2018), and in fact we avoid using the term ‘native’ wherever possible. The comparison with L1 speakers and the accent rating does, however, allow us to compare how closely L2 speakers approximate the accent patterns of L1 speakers when input is almost exclusively from those speakers. This allows us to address methodological constraints in other studies that have addressed this question, where learners had previously been exposed to accented varieties of the L2 prior to arriving in the L2 country, and/or their residence in the L2 country does not necessarily go hand in hand with increased input.

## 2.1. Participants and Languages under Examination

### 2.1.1. Participants

Twenty-two female FDH who worked for and lived with families in Oman participated in this study. Consent for participation in the study was obtained from the FDH and their employers. The participants completed a questionnaire which elicited information about their demographic and sociolinguistic background. This included information on age, L1(s), age of arrival (AoA) in the Arabic-speaking world, length of residence (LoR) in the Arabic-speaking world, years of formal schooling in the L1 and L2 Literacy (ability in reading or writing in Arabic). The first author read the questionnaire to the FDH and recorded their answers to these questions.

The participants’ background was representative of the very diverse background of FHDs who work in the Arab world. For instance, the FDH came from nine L1 backgrounds (Swahili (5), Indonesian (2), Sinhala (4), Tagalog (5), Bengali (2), Telugu (1), Luganda (1), Yoruba (1) and Oromo (1)). They migrated to the Arabic-speaking world as adults (mean AoA = 27.27), and they had varying Arabic experiences based on their LoR that ranged from 0.7 to 21 years (mean LoR = 6.23). Their mean LoR in Oman was 2.36 years. Nine of them had worked in different Arabic-speaking countries before moving to Oman (e.g., Gulf countries and Lebanon). They all reported that they had been addressed to mainly in Arabic by the family members of the household(s) they had lived in and worked for. Fourteen of them had never been exposed to Arabic before migration, while eight had had access to Arabic via Islam and recitation of the Qur’an. It should be noted that when classifying FDH based on their Arabic literacy, only Muslim FDH who reported being able to read in Arabic via recitation of the Qur’an were considered as literate. Those who reported not to be able to read in Arabic or recite the Qur’an were considered as non-literate in Arabic even if they were exposed to Arabic during other rituals of worship. Other than Arabic, 15 of them reported having some knowledge of English. Ten L1 Omani speaker females were recruited in order to obtain comparative information from L1 patterns for the perception and production tasks. They all had a comparable educational background and were between 19 and 40 years old.

### 2.1.2. L1 Consonant Inventory and Target Sounds and Structures

Table 1 shows the consonant chart of Omani Arabic. In order to control for the variability in the FDH’s L1, we targeted consonants that were absent in all the L1 sound inventories of the FDH participants and that were likely to pose a challenge in perception and/or production due to their complex articulation; these are highlighted in grey. Table 2 details whether the L1 phonology of the FDH participants allows onset and coda consonant clusters in comparison with the target language. It is worth noting that epenthesis in onset clusters is optional in Omani Arabic, but the examination of adult input suggests a prevalence of non-epenthetic realisations (Al-Kendi 2021). L1 phonologies of FDHs that have a restricted onset structure only permit a limited combination of CCs. For

example, Tagalog onset consonant clusters are restricted to a consonant and a glide (e.g., /bwan/ 'month'), while Luganda and Swahili onset structures are restricted to a consonant and a glide or a consonant preceded by a nasal (e.g., /mba/).

**Table 1.** The sound inventory of Omani Arabic (The highlighted consonants are absent from the L1 inventories of the FDH).

	Voicing	Labial	Labio-Dental	Dental	Alveolar	Palatal	Velar	Uvular	Pharyngeal	Glottal
Stop	Voiced	b			d		g			
	Voiceless				t		k	q		ʔ
Emphatic stop	voiceless				tˤ					
Fricative	Voiced			ð	z			ʁ	ħ	
	Voiceless		f	θ	s	ʃ		χ	ħ	h
Emphatic fricative	Voiced			ðˤ						
	Voiceless				sˤ					
Nasal	Voiced	m			n					
Liquid	Voiced				r l					
Approximant	Voiced	w					j			

**Table 2.** The prevalence of CC clusters in the sound systems under investigation. Languages that permit consonant clusters are indicated with the symbol ✓, while those that do not are indicated with the symbol x.

Language	Initial CC	Final CC
Omani Arabic	✓	✓
Bengali	✓	x
Swahili	✓ (restricted)	x
Luganda	✓ (restricted)	x
Tagalog	✓ (restricted)	x
Yoruba	x	x
Oromo	x	x
Indonesian	x	x
Sinhala	x	x
Telugu	x	x

## 2.2. Examining FDH's Discriminability of Arabic Consonants

### 2.2.1. AX Task and Stimuli

An AX forced-choice discrimination paradigm was used to elicit FDH's discriminability of different Arabic consonantal contrasts. In this kind of task, participants are presented with two stimuli in a sequential order whereby the second stimulus is either the same as the first (AA) or different (AB) (Strange and Shafer 2008), and they answer 'same' or 'different'. For the current study, a list of 16 Arabic consonant contrasts was created. The phonemic pairings were created based on their potential confusability for the listeners in terms of perception and/or production, as they only varied in one feature: voicing (e.g., /θ/-/ð/), manner (e.g., /t/-/s/), place (e.g., /χ/-/ħ/; /q/-/k/) or the presence or absence of secondary articulation (e.g., /tˤ /-/t/). The latter refers to Arabic emphatic sounds whose

production is accompanied by a primary articulation at the dental/alveolar area and a secondary articulation that involves a constriction in the upper pharynx.

Two more contrasts were used as control: /r/-/l/, /r/-/w/. Given that the FDH's L1 inventories included these sounds, albeit with potentially different phonetic realizations and phonological patterning, it was likely that all participants would detect these contrasts as different (Aoyama et al. 2004). With the control pairs, a total of 18 consonant contrasts were included in the test. Four test items were created for each of the 18 contrasts (AA, AB, BA, BB), yielding a total of 72 test trials. Each test trial consisted of two monosyllabic pseudo words containing the contrasting sounds in the context Ca:n (where C is a consonant), for instance, /sa:n-s<sup>s</sup>a:n/. Thirty-six trials contained consonants that were acoustically different, while 20 trials contained consonants that were acoustically identical (16 trials were excluded from the list, as they were repetitions of existing trials). This reduced the number of trials to 56. To give an example, /θ/ was paired twice, once with /s/ and another with /t/ because [t] and [s] are likely variants of /θ/ in NNS' productions (Lombardi 2003). When the four test trials were created for each of these pairs, one test item was repeated for both contrasts. So, one of the repeated trials was excluded. When all trials were created, they were submitted to an online randomization software (RANDOM.ORG) to ensure that the four test items for each contrasting pair were not following each other. The stimuli were recorded by the first author in a sound-treated lab using an Edirol digital recorder R-09HR by Roland coupled with a Sennheiser radio microphone, with a sampling rate of 44.100 Hz and WAV-16bit recording mode. Another native Omani speaker trained in linguistics listened to the recorded stimuli for a reliability check. She spoke the same dialect as the NSs in this study. She confirmed that all recorded instances were clearly articulated and checked that time intervals between test items were the same and as specified in the present study.

### 2.2.2. Procedure for the AX Discrimination Task

In the home of their employers, each FDH was presented with the aural stimuli over headphones at a comfortable volume level using the Praat program on a MacBook laptop (Boersma and Weenink 2009). They were instructed, in Arabic, that they needed to decide if the two test items they were about to hear were the same or different. They gave their responses to the first author, who manually entered them on an answer sheet designed specifically for this task. It was not possible to use a full computer version for this task due to the potential difficulty the participants might have faced dealing with technology in case of limited computing education. An inter-stimulus interval (ISI) of 1 s was used between each word in a comparison pair. A longer ISI is shown to facilitate phonemic discrimination rather than phonetic discrimination of contrasts that are absent from the L1 (Werker and Logan 1985). The participants were allowed to listen to the same items again if needed, but they could not change an answer once given (Guion et al. 2000). The trials were presented in two blocks during a twenty-minute session. A three-minute interval separated the two blocks.

To ensure that the participants understood the task procedure, they were presented with a familiarization task prior to the experiment. They were trained to listen to and judge two contrasts (/ʃ vs. /s/, /t vs. /d/) and were given immediate feedback about the accuracy of their responses. The contrasts presented in the familiarization task were different from those used in the real test to avoid providing the participants with help on the target stimuli (Beddor and Gottfried 1995). Adopting a similar approach to Aoyama et al. (2004), the participants had to respond correctly to at least 90% of the stimuli in order to proceed to the actual task. If a participant did not reach this standard, the practice task was repeated up to four times or until they met the standard. Two FDH who performed below 90% in the familiarization task were excluded from this study because they did not display understanding of the task. A similar procedure was used to elicit NSs' responses to the same task. However, the NSs were given an answer sheet to record their own responses.

### 2.3. Examining FDH's Arabic Production

The 22 FDH participants who were recruited for the previous task participated in the picture-naming task.

#### 2.3.1. Stimuli for the Production Task

A picture-naming task was used to elicit single words from the participants. A list of Arabic words that contained the same target Arabic consonants that were used in the AX discrimination was created for this task. Another list that included words with onset and coda consonant clusters was also created. The words selected represented home objects and hence were assumed to be familiar to most FDH. Thirty-six pictures that represented the stimuli words were used to elicit productions from the FDH. The pictures were compiled in a Powerpoint file, one picture per slide.

#### 2.3.2. Procedure for the Production Task

The FDH named the objects in the pictures presented to them using the slideshow function of Powerpoint on a MacBook laptop controlled by the first author. The same high-quality recorder and microphone used in task 1 were used again here to record the participants' productions. When a participant struggled to name an object, a delayed repetition technique was used (Ratner 2000; Guion et al. 2000): if the participant could not name the object in the picture, the first author produced the target word and then asked the participant what the prompt was again. The delay between the prompt and the participant's repetition was approximately 4 s in order to minimize the effect of direct mimicry. The productions were analysed for their target-like accuracy; additionally, clusters were examined for simplification patterns such as epenthesis or deletion of one of the consonants.

### 2.4. Examining FDH's Foreign Accent Rating

#### 2.4.1. Listeners

The listeners for this task were 10 L1 Omani speakers (5 males and 5 females). Their ages ranged between 31 and 40 years at the time they carried out the task. They were all born in Oman and spoke Omani Arabic. None had experience or training in a linguistic-related field. None reported having a history of speech, language or hearing problems.

#### 2.4.2. Material for the Accent Rating Task

The stimuli for this task were taken from production data collected from the picture-naming task, with a focus on words where the target sound of interest was in a word-initial position. Ten words were selected for inclusion in this experiment (/χass/'lettuce', /ħabll/'rope', /qalam/'pen', /ð<sup>ʕ</sup> arf/'envelope', /s<sup>ʕ</sup> aabu:n/'washing liquid', /t<sup>ʕ</sup> a:wleh/'table', /ʕ ar f<sup>ʕ</sup> eh/'bottle', /ʕ alam/'flag', /θo:m/'garlic', /ðurah/'corn'). The stimuli were extracted from the sound files of the picture-naming task in Praat as whole words. The defined stimuli were windowed by a parabolic function and normalized to 50 dB. Two L1 Omani speakers aged 33 and 35 and who spoke Omani Arabic also produced the same words. The total number of stimuli included in the rating experiment was 240. The stimuli were randomized and distributed in three blocks (each including 80 stimuli).

#### 2.4.3. Procedure for the Accent Rating Task

The stimuli were presented to the participants in Praat on a Macbook Pro laptop. The listeners were asked to rate the stimuli they heard on a scale from 1 (not at all native-like) to 9 (completely native-like). They knew that they were going to hear Arabic words produced by L1 and L2 speakers. They could replay the stimuli as many times as they wished before making their choices and moving on to the next stimulus by pressing the 'next' button. They were offered a break after every 80 stimuli. The listeners spent approximately 30–40 min on this experiment.

### 2.5. Statistical Analysis

Data from the AX discrimination task was coded and tabulated manually in Excel, and subsequently imported into R (R Development Core Team 2012). To measure listeners' accuracy and correct for biases, we used some variants from Signal Detection Theory (Macmillan and Douglas 1991). We generated the *d prime* value (sensitivity index) for each individual listener separately. The *d prime* models the difference between the 'true positive' responses and 'false positive' responses in standard units, as in the following formula:  $(d \text{ prime}) = Z(\text{True Positive Rate}) - Z(\text{False Positive Rate})$ . We then used a linear model to statistically measure the difference in the *d prime* mean of NS and FDH groups using the lme4 package in R (Bates et al. 2015). To examine which factors contribute to any differences in *d prime* values among FDH listeners, we used a linear model, with *d prime* as the dependent variable and LoR (continuous), L1 schooling (continuous) and L2 literacy (categorical) as the independent variables (predictors). The model used was diagnosed for the presence of multicollinearity using the variance inflation factor, *vif()*, that measures the influence of collinearity among the predictors in a regression model (Midi et al. 2010). The VIF scores obtained were low (equal to 1), which indicated that it was safe to accurately assess the contribution of the predictors to the model.

To analyse data from the picture-naming task, the productions were first phonetically transcribed in Praat following the labelling of target segments. A Praat script designed by the first author was then used to extract all target word productions and their relevant information from Praat and transfer these to Excel files. Target consonants were assigned a value of 0 or 1 depending on whether the production of the sound was target-like. Descriptive statistics based on the percentages of target-like productions of each consonant averaged across speakers were then provided. To examine which factors play a role in the accuracy of productions, a GLMM was used with LoR, L1 schooling and L2 literacy as predictors and random intercept of speakers as random effects. The model was detected for multicollinearity using VIF. The scores obtained from the VIF were low, indicating the low collinearity of the predictors.

With regard to consonant cluster productions, a GLMM was used to examine the difference in modification between onset and coda consonant clusters. First, consonant clusters were assigned a value of 0 or 1 depending on whether the cluster was in the onset or coda position. Next, the productions of the consonant clusters were assigned a value of 0 or 1 depending on whether the production involved modification<sup>1</sup>. The GLMM had *modification* (yes or no) as a dependent variable and the type of consonant cluster (onset or coda) as predictor. For random effects, speaker was used as random intercept. To examine the effect of psycho-social factors on the pattern of cluster production, onset and coda clusters were analysed differently. For onset clusters, qualitative analysis was carried out because L1 inventories of some of the FDH's contained onset clusters while others did not (see Table 2). This indicates that the L1s of FDH's are incomparable and will affect their production patterns differently. Therefore, using statistical analysis with factors such as L1, LoR, L1 schooling and L2 literacy might not generate meaningful results, and one factor could cancel the significance of other factors (collinearity effects). Therefore, findings from a qualitative analysis will be more reliable and can be indicative of patterns that we can test more thoroughly in future research. As for coda clusters, a GLMM was used to test the effect of LoR, L1 schooling and L2 literacy on the modification of consonant clusters. All FDH's L1 inventories lacked coda clusters, and thus FDH's were considered comparable in relation to their L1s. All three factors were used as predictors, with speaker as random intercept for random effects.

For the foreign accent rating task, raters' responses were first tabulated in an Excel file which was then imported into R. Descriptive statistics were generated in R and included mean, median, SD and variance. In order to determine whether there is a difference in rating scores among the two groups (NS and FDHs), we used cumulative link mixed models (CLMM), using the *ordinal* package in R (Christensen 2019). For the dependent variable, we used an ordered factor, the rating response. For the independent variable, we

used group. For random effects, we used rater and item as intercepts. This was the optimal random effect structure that suited these data.

### 3. Results

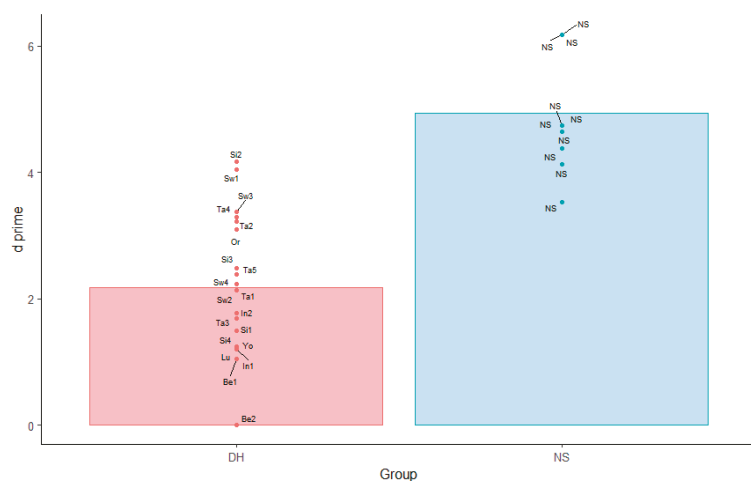
#### 3.1. Discrimination of Arabic Consonant Contrasts

Table 3 shows descriptive statistics that were obtained from a number of operations. The accuracy of the FDH was 0.72, while that of NSs was 0.96. The error rate of FDH (0.27) was higher than that of NSs (0.03). The true positive rate reflects the rate at which listeners responded ‘different’ when the stimulus was ‘different’. On the other hand, the false positive rate reflects the rate at which listeners responded ‘different’ when the stimulus was ‘same’. The sensitivity (which reflects the rate at which listeners responded ‘same’ when the stimulus was ‘same’) of both groups was higher than 50%. Precision reflects the rate at which listeners responded correctly when the stimulus was ‘same’. This was also higher than 70% for both groups. These statistics show that regardless of the difference between FDH and NS groups with regard to accuracy and sensitivity rates, the FDH group has successfully obtained high accuracy and sensitivity results (above 50%).

**Table 3.** Descriptive statistics obtained from the AX discrimination responses.

Group	Accuracy	Error	True Positive Rate (Specificity)	False Positive Rate	True Negative Rate (Sensitivity)	Precision
FDH	0.72	0.27	0.64	0.13	0.86	0.89
NS	0.96	0.03	0.98	0.06	0.93	0.96

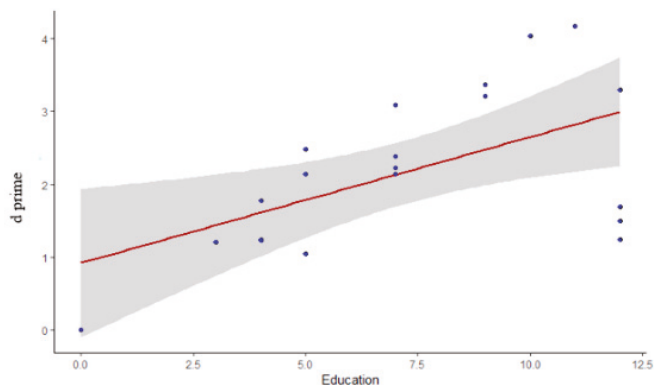
When examining the difference between FDH and NSs’ discriminability of consonant contrasts using *d prime* values, a linear model revealed that the mean *d prime* value of the FDH group (mean = 2.17) was significantly lower than that of the NS group (mean = 4.94,  $\beta = -2.76$ ,  $SD = 0.404$ ,  $p < 0.01$ ). Unsurprisingly, this indicates that the NS group outperformed the FDH group in the AX discrimination task. However, the FDH *d prime* values in Figure 1 reveal a great variation among FDH’s performance (Figure 1), suggesting that some FDH performed as well as some NSs while others performed very poorly.



**Figure 1.** Mean *d prime* scores of the foreign domestic helper listeners (DH) and the native speaker (NS) control group.



Further analyses to examine the factors that affected FDH's variable performance in the AX discrimination task revealed that FDH's L1 schooling played a significant role in their *d prime* scores ( $\beta = 0.18$ ,  $SE = 0.06$ ,  $p = 0.01$ ). The more years the FDH had spent at school in her first language, the greater her discriminability of the contrastive sounds in Arabic was (Figure 2). LoR had no significant effect on FDH's discriminability of consonantal contrasts ( $\beta = 0.02$ ,  $SE = 0.03$ ,  $p > 0.05$ ). Thus, no matter how long the FDH had spent in the Arabic-speaking world, her discriminability of consonant contrasts had not changed or improved. Similarly, L2 literacy did not have a significant effect on FDH's *d prime* scores ( $\beta = 0.25$ ,  $SD = 0.44$ ,  $p > 0.05$ ).



**Figure 2.** The relationship between years of L1 schooling and *d prime* values relevant to the FDH group.

### 3.2. The Production of Arabic Consonants and Consonant Clusters

As expected, complex consonants that are specific to the Arabic inventory posed the greatest challenge for the participants (Figure 3). Common realizations included stopping for fricatives, e.g., [t] for /θ/, [d] for /ð/, [g] or [k] for /ʁ/, [q] for /χ/, [ʔ] for /ʕ/, and [d̪] for /ð̪/; de-emphasis, e.g., [ð] for /ð̪/, [t] for /t̪/ and [s] or [ʃ] for /s̪/; (de-)voicing (along with other processes), e.g., [q] for /q̪/ or /χ/, [g] for /g̪/; fronting/backing (along with other processes), e.g., [k] or [g] for /q/ or /χ/, [t] for /q/, [z] or [s] for /ð/ or [s] for /θ/; and weakening, e.g., [h] for /ʕ/ or /ʁ/ deletion. L2 literacy had a significant effect on FDH's accurate productions of the target sounds ( $\beta = -0.53$ ,  $SE = 0.22$ ,  $p = 0.01$ ). Figure 4 shows that speakers who were literate in Arabic had more target-like productions (50%) of the target consonants than those who were non-literate (37.8%). This suggests that FDH who learned Arabic via recitation of the Qur'an performed better than those who did not. On the other hand, LoR did not play any significant role in the target-like production of Arabic consonants by FDH ( $\beta = 0.02$ ,  $SE = 0.01$ ,  $p > 0.05$ ). In fact, the visual examination of the results showed that speakers with the longest LoRs appeared to be slightly less accurate than those with shorter LoRs. Similarly, L1 schooling did not play a considerable role in FDH's accuracy ( $\beta = 0.03$ ,  $SE = 0.03$ ,  $p > 0.05$ ).

Results showed a high proportion of modified consonant clusters in the production of FDH in the onset and coda position (Table 4). The main strategy used to modify clusters was vowel epenthesis. This indicates that FDH have not acquired complex clusters despite their exposure to them in the target input (Al-Kendi 2021). The tendency to modify onset consonant clusters (89.4%) was more frequent than for coda clusters (48.9%) (Figure 5). A GLMM demonstrated that this difference was significant ( $\beta = -2.51$ ,  $SE = 0.41$ ,  $p < 0.01$ ).

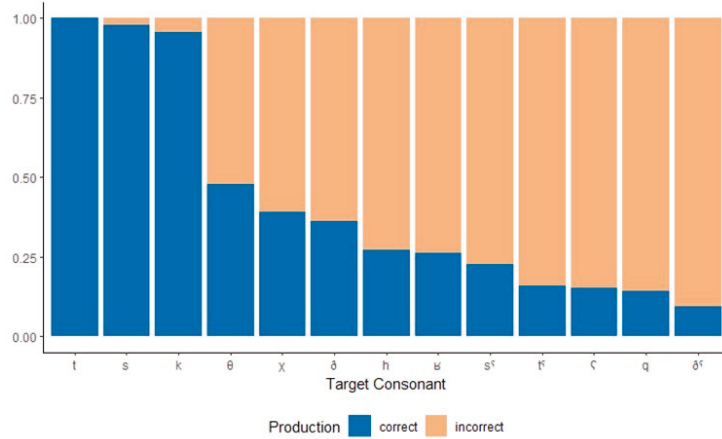


Figure 3. Accuracy in the productions of the target Arabic consonants by FDH.

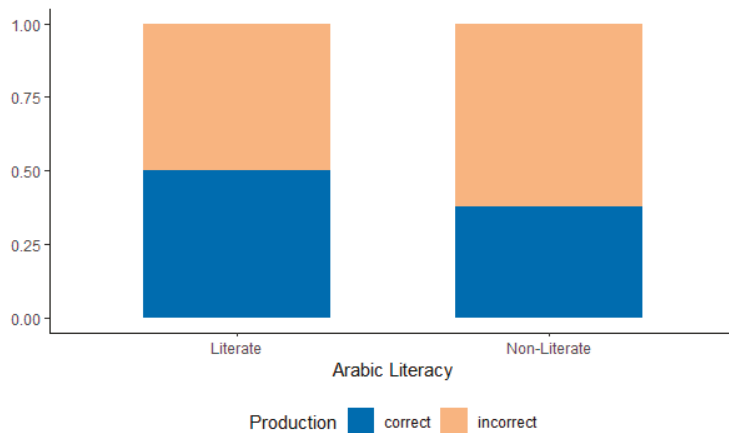


Figure 4. Percent accuracy in the productions of the target Arabic consonants by literate and non-literate FDHs.

Table 4. Percentage of modification in onset and coda consonant clusters (CC) in FDH’s productions and the modification strategy used.

Syllable Type	Syllable Modification		Modification Strategy	
	Yes	No	Consonant Deletion	Vowel Epenthesis
Onset CC	(118) 89.4%	(14) 10.6%	(2) 1.69%	(116) 98.305%
Coda CC	(43) 48.9%	(45) 51.13%	0	(43) 100%

In terms of factors that could have affected FDH production of onset consonant clusters, descriptive statistics and visual inspection showed that, generally, the tendency to produce less marked onset consonant clusters was evident in all FDH’s productions regardless of the L1 (Figure 6). Nevertheless, FDH with Oromo, Sinhala and Telugu L1 backgrounds had the highest rate of onset modification, producing most of the target CCs with epenthetic vowels. FDH with Indonesian, Tagalog, Bengali, Yoruba, Swahili and Luganda L1 backgrounds showed more variation in their production of onsets clusters,

sometimes maintaining the CC realization. To illustrate, these FDH sometimes produced the target word /kta:b/as [kta:b] and others as

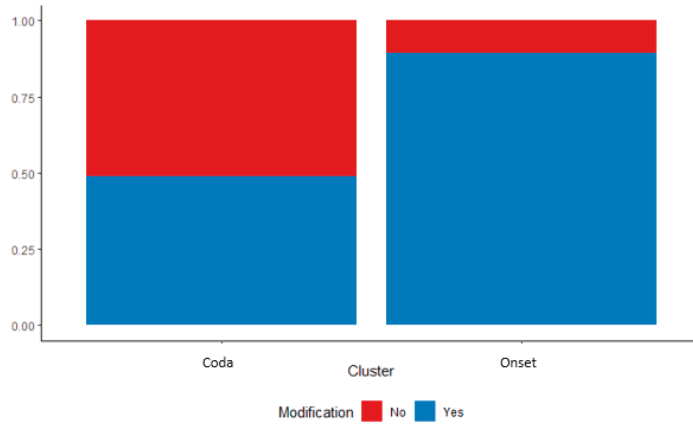


Figure 5. Percentage of onset and coda consonant clusters modification.

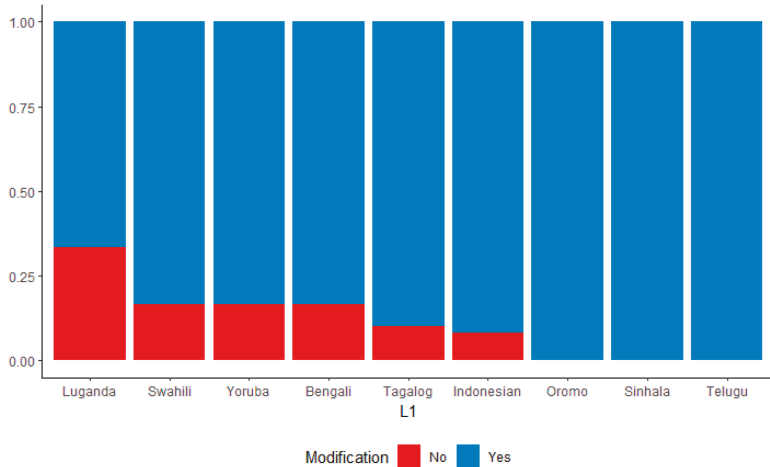
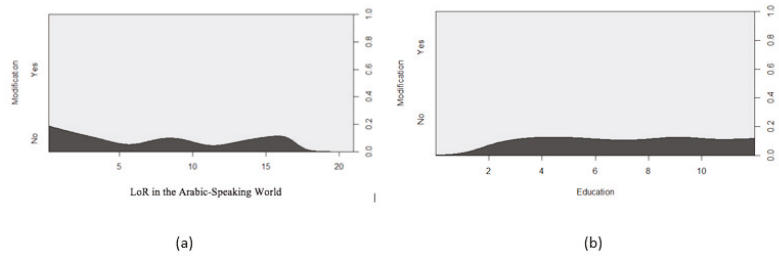


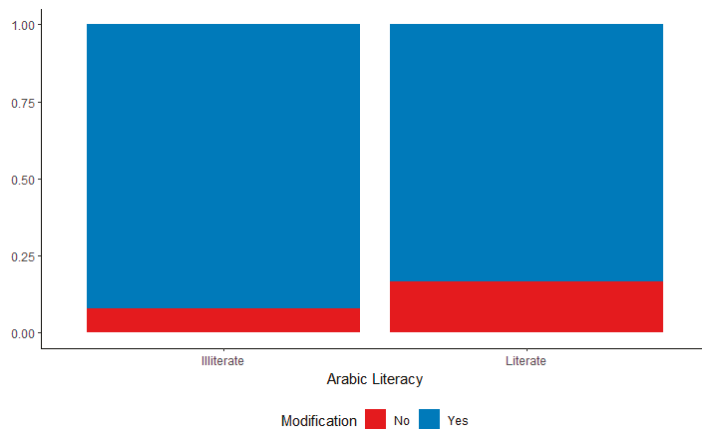
Figure 6. Production accuracy of onset CCs as a function of the L1 of the FDH participants.

Figure 7a shows patterns of onset cluster production by FDH as a function of LoR. Surprisingly, fewer rates of onset simplification are evident in productions of FDH who had the shortest LoR. The trend then shows a stable pattern for FDH with 5 to 15 years of LoR and is then highest for the speakers with the highest LoR. From this, we can conclude that LoR alone does not play a role in the successful production of onset consonant clusters in FDH's productions. Equally, the modification of onset consonant clusters does not appear to change considerably as a function of years of formal schooling. Figure 7b shows a stable trend of onset cluster simplification regardless of the years of schooling.



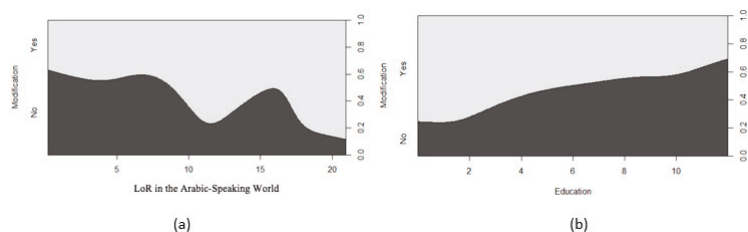
**Figure 7.** The effects of LoR (a) and L1 schooling (b) on FDH’s production accuracy of onset CCs.

Figure 8 illustrates that FDH who were literate in Arabic exhibited fewer modifications of onset consonant clusters (7.77%) than those who were non-literate in Arabic (16.66%).



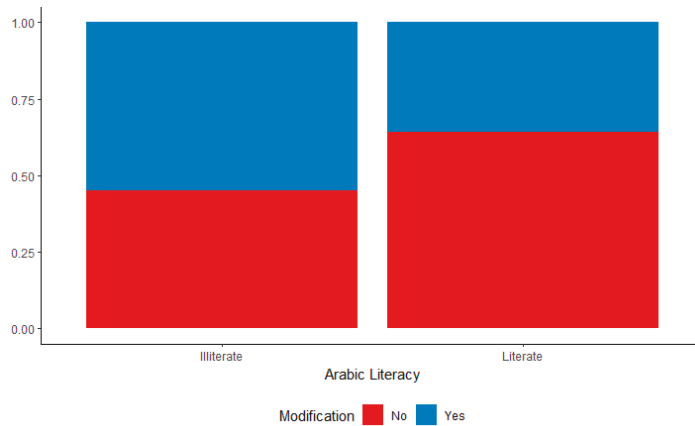
**Figure 8.** The percentages of target-like productions of Arabic onset consonant clusters by literate and non-literate FDHs.

Moving on to factors that may have played a role in FDH’s production of coda consonant clusters, the results once again revealed that LoR had a significantly negative effect on target-like consonant cluster production ( $\beta = -0.102, SE = 0.04, p = 0.03$ ). FDH with the shortest LoR produced clusters in the coda position more frequently than those with a longer LoR (Figure 9a). The trend also shows fluctuation, which implies individual differences in the target-like realization of coda clusters. L1 formal education had a significant effect on the modification of coda consonant clusters ( $\beta = 0.13, SE = 0.06, p = 0.04$ ). The more educated a foreign domestic helper was, the more successful she was at producing a target-like syllable structure (Figure 9b). This is in line with the significant role L1 schooling played in FDH’s performance in the discrimination task.



**Figure 9.** The effects of LoR (a) and L1 (b) schooling on FDH’s production accuracy of coda CCs.

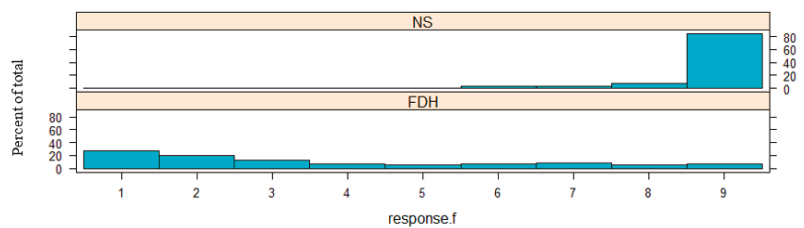
L2 literacy, on the other hand, did not play any significant role in the pattern of coda consonant cluster production ( $\beta = 0.85, SE = 0.55, p > 0.05$ ), though there was a trend for more target-like production by FDH who reported to be literate in Arabic than those who reported to be non-literate (Figure 10). Literate FDH produced coda clusters in 64.28% of the instances, while non-literate FDH produced them in 45% of the instances. Hence, literacy in Arabic—or, more specifically, knowledge of the Arabic script—seems to aid the L2 learners’ acquisition of target-like oral forms, even though this trend was not significant.



**Figure 10.** Production accuracy of Arabic coda consonant clusters by literate and non-literate FDHs.

### 3.3. Foreign Accent Rating of FDHs

Figure 11 illustrates the rating scale of the foreign accent task and the percentage of total scores given to words produced by the two groups (NS and FDHs). It shows that around 80% of the ratings given to NSs fell into the ‘completely native-like’ category, that is number 9 on the scale. However, the highest percentage of rating in the FDH group went to the ‘not at all native-like’ ranking on the scale, that is number 1.

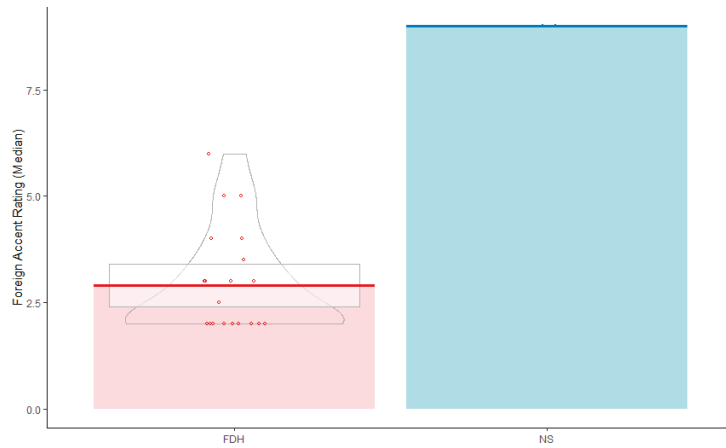


**Figure 11.** Percentages of total responses for each rate in the foreign accent scale for both FDH and NS productions.

Descriptive statistics of the foreign accent rating task indicate that the FDH were rated very low on the foreign accent rating scale (median = 3) compared to the NS group (median = 9), as shown in Table 5. There was, however, considerable variation in the rating scores given to the FDH production (variance = 6.86) compared to those given to the NSs (variance = 1.49). Further analyses of these results revealed that the difference between NSs and FDH’s foreign accent rating was significant ( $\beta = 5.31, SE = 0.22, p < 0.01$ ). Figure 12 illustrates the difference in the foreign accent rating median for the NS and the FDH groups.

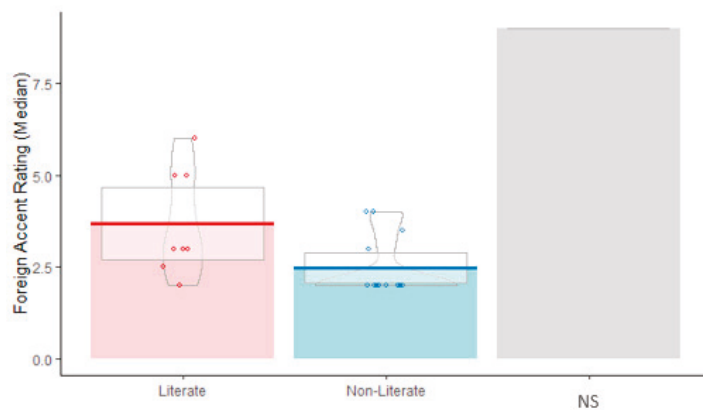
**Table 5.** Descriptive statistics obtained from the foreign accent rating task for both the FDH and the NS groups.

Group	Mean	Median	SD	Variance
FDH	3.64	3	2.62	6.86
NS	8.61	9	1.22	1.49



**Figure 12.** Median and distribution of foreign accent rating scores given to NS and the FDH groups.

When examining the factors that may have affected foreign accent rating, the results revealed that FDH’s L2 literacy played a significant role in their foreign accent rating (Figure 13). Non-literate FDH’s were rated as more foreign-accented than literate FDH ( $\beta = -1.24, SE = 0.49, p = 0.02$ ). However, LoR did not play any significant effect on FDH’s accent rating ( $\beta = -0.03, SE = 0.04, p > 0.05$ ). Likewise, L1 schooling was not found to significantly affect FDH’s foreign accent rating ( $\beta = -0.03, SE = 0.06, p > 0.05$ ). Interestingly, these results are similar to those obtained from the examination of FDH’s production of Arabic consonants.



**Figure 13.** Median and distribution of foreign accent rating scores given to NS and the FDH groups, the latter split according to L1 literacy.

#### 4. Discussion

Despite a constant exposure to Arabic from L1 speakers, the length of residence that FDH spent in the Arab world appeared to play no role in their L2 perception of consonant contrasts or in their production of Arabic consonants or consonant clusters. In some cases, FDH's scores correlated negatively with LoR, as in the production of coda consonant clusters. These findings do not support the assumption that the more L2 input one receives the better the opportunities to master the L2 (McAllister 2001). There have been other studies over the years that have challenged LoR effects on learners' performance (e.g., Flege 1988; Oyama 1976), but their methodologies have rarely included a total immersion in the L2 with input from L1 speakers, as in the current study. Below we reflect on the potential reasons for these surprising results.

It is hard to ignore the different ways in which a lack of effect of LoR has been interpreted in the literature. For example, Flege and Liu (2001) suggest three possible interpretations: (1) the amount of L2 input is not a crucial predictor of L2 performance, (2) L2 performance is constrained by a critical or sensitive period, (3) LoR provides a meaningful index of L2 input for some individuals but not others. The age effect has been dealt with extensively in the literature, but it is the third point that we focus on here. On the one hand, one can interpret these differences in terms of differential access to input from L1 speakers. For instance, Flege (2002) found that LoR can play a role in adults' L2 performance only if they are exposed to a considerable amount of L1 speaker input. In this study, Chinese students with longer LoR in the United States were significantly better in their L2 performance compared to students who had shorter LoR. LoR, however, did not predict the performance of non-students. Flege (2002) concluded that because students had more opportunities for receiving NS input compared to the non-students, their performance improved noticeably over time. However, Moyer (2004, 2009) argues that, in order for late learners to gain sufficient input, they need to engage in the L2 environment in different ways. Situations favourable to such attempts vary across individuals, depending not only on age but also on educational, social and ethnic background. Moyer (2004) further suggests that for LoR to reliably index the L2 experience, an integrated approach that takes into account cognitive, psychological and social factors needs to be carried out. Psychologically, LoR correlates with a sense of overall fluency and satisfaction with L2 attainment as well as motivation to learning the L2. Socially, LoR correlates with the frequency of contact with NSs and the intention for permanent residency in the L2 target community. Cognitively, LoR correlates with L2 instruction and communicative use of the L2 rather than just focusing on form, as well as the amount of feedback on pronunciation and the kind of phonological training.

In light of this integrated model, it is not surprising that LoR was not found to play a role in FDH's phonological performance. Despite the high number of years many FDH spend in the Arab world, their intention is not for permanent residence in the L2 country, but rather to return home once they have saved enough to support their families (Bizri 2014). Their aim of L2 attainment may be restricted to the ability to interact with their employers rather than any motivation to achieve native-like fluency. Furthermore, FDH's language contact with their employers or with other family members is often restricted to conversations around home chores. Due to the task-oriented nature of these interactions, it is unlikely that FDH receive any feedback on their pronunciation. Equally, they do not receive any training on L2 phonology or other linguistic aspects. In addition, the input they receive can be variable within the constricted context in which they work: they attend to children who have not yet fully developed their phonology and hear accented Arabic from other FDH of various nationalities when running errands and during their day off. This is akin to Flege and Liu's (2001) description of immigrants to North America, who are likely to use their L2 English with other NNSs as a lingua franca, though in their study those interactions were happening in the workplace, whereas our participants' main workplace is their NS employers' home.

When the above factors are considered, FDH's experience does not provide an optimal environment for L2 attainment, despite the near total immersion in input from L1 speakers.

Opportunities for meaningful input, contact with other L1 speakers and L2 instruction are very limited in this context, also highlighting the potential role of the lack of variability in the input. This supports the observation that LoR is not a reliable index of L2 experience (Flege and Liu 2001; Moyer 2004, 2009). In order to reliably examine the extent to which input and LoR modulate L2 phonological performance, methodologies always need to take into account cognitive and socio-psychological factors that may shape the L2 speakers' experience.

While LoR was not a significant predictor of FDH's phonological performance, L1 schooling and L2 literacy each played a role in some of the variables examined. L1 schooling correlated significantly with FDH's perceptual sensitivity scores and rate of final consonant cluster productions. However, no effect of L1 schooling was found in the production of Arabic consonants or initial consonant clusters. L2 literacy generally correlated positively with sensitivity scores and production of Arabic consonants as well as initial and final consonant clusters. However, the results were only significant with regard to the production of L2 consonants. The differential roles of L1 formal schooling and L2 literacy in FDH's performance are discussed below.

In terms of the positive role of L1 schooling on the perception task, a first justification for this finding is that the AX discrimination task was an experimental paradigm that required the listeners to understand task instruction and procedure as well as sit for an actual test. This protocol might be more familiar to adults who had attended school and experienced such a situation compared to adults who had little to no experience with carrying out cognitively demanding tasks due to not attending school. Another positive effect of L1 literacy on L2 perception may be due to the higher level of phonological awareness that comes with learning an alphabetic script (e.g., Morais et al. 1979; Adrián et al. 1995; Tarone and Bigelow 2005), and which may have equipped the participants with metalinguistic skills that they could subsequently use in their L2.

An emerging body of work has recently highlighted the role of L1 orthography in L2 production (rather than perception), with results suggesting that L1 orthography leads to a convergence between the L1 and the L2 production patterns (e.g., Bassetti and Atkinson 2015; Escudero et al. 2014; Nimz and Khattab 2020). Our results do not demonstrate strong effects of L1 literacy on L2 production, with the only difference seen in the greater target-like production of final consonant clusters by participants who were literate in the L1. One reason for this may be due to the higher prevalence of CC realizations of consonant clusters in the final than in the initial position in the L2 input that the FDH receive (Al-Kendi 2021), rendering these structures more salient. The production of consonant clusters requires attention to structures in the target input and adjusting L1 phonology accordingly. Schmidt's Noticing Hypothesis (Schmidt 1990) claims that a conscious awareness (i.e., noticing) of the input plays a substantial role in the process of language acquisition. Several researchers have provided support to this hypothesis and confirmed the importance of noticing for language learning (e.g., Jeremy 2002; Lynch 2001; Skehan 1998). Among the factors that Schmidt claims to affect noticing is frequency, and hence the likelihood of FDH's noticing coda more than onset clusters. L1 schooling may have helped FDH develop a language learning aptitude and conscious phonological processing, improving the literate learners' 'noticing' skills that are required for L2 learning (Granena and Long 2013).

It is with L2 literacy that we see a stronger effect on L2 production. There has been a notable increase in attention to the role of orthography in L2 speech learning (e.g., Bassetti et al. 2015; Escudero et al. 2014; Nimz and Khattab 2020), but results are typically mixed, signalling both facilitatory and inhibitory effects in terms of learning L2 phonological categories. Here it is worth focusing on the unusual way in which Arabic literacy is taught for religious purposes to speakers of other languages, like the FDH in this study. While the script is of course key, there is a strong focus on recitation and rote learning in such contexts (Binte Faizal 2019; Supriyadi and Julia 2019), emphasizing the role of production practice in this process. This is likely to have helped FDH who had experience with L2 literacy,



supporting their production of L2 consonants through more advanced motor control along with the establishment of categories for new sounds (e.g., Guenther 1994; Flege 1995).

Note, however, that the L2 literacy effect was mainly seen in single consonant production, with clusters requiring much more motor control and experience with the language before target-like realization. Here it is worth noting that the epenthesis of CC clusters is more common in the onset than in the coda position in Omani Arabic (Al-Kendi 2021), and this is reflected in the patterns found for CC realizations by FDH in those two contexts, albeit with a much higher occurrence of epenthesis in the FDH's production. This supports the expectation that learners will acquire less complex L2 structures (e.g., CV) before more complex ones (Anderson 1987; Eckman 1985; Rice 2007; Zec 2007). This was also reflected in the productions of the speakers whose L2 systems lacked onset consonant contrasts compared with those whose L2 had such contrasts.

## 5. Conclusions

The results from the present study shed light on the perception and production ability of a group of uninstructed low-educated foreigners acquiring the language in a naturalistic setting. The status quo in SLA research has been to investigate the cross-linguistic performance of highly educated adults and to focus on LoR and AoA as primary factors affecting these learners' performance. The results from this study highlight the importance of looking at low-literacy learners and investigating the role of other non-linguistic factors, such as the nature of daily social interactions in an L2 context, the long-term aims of the learners and the power relations between them and their main interlocutors. Despite being totally immersed in Omani Arabic for a number of years, the FDH in this study still struggled with the phonology of Omani Arabic and had a pronounced foreign accent as judged by L1 listeners. Their perception scores were not found to be influenced by LoR or L2 literacy, but rather by the amount of their L1 schooling; this could in itself be a proxy for learning to perform tasks and follow instructions, but L1 literacy may have also increased these learners' metalinguistic awareness. Their production patterns did not show any LoR effect either, and only a modest influence from L2 instruction. The study demonstrates how difficult it is to control for external factors that are beyond the focus of a given study in SLA research. For instance, while the focus of the current study was to investigate what looked like an optimal case of LoR with guaranteed input in order to address previous criticisms of LoR, low literacy and the fact that input and interaction are dominated by the employer and their family may have attenuated any LoR effects, showing how multi-faceted the L2 speech learning experience is.

**Author Contributions:** Conceptualization, A.A.-K. and G.K.; methodology, A.A.-K. and G.K.; formal analysis, A.A.-K.; resources, A.A.-K.; data curation, A.A.-K.; writing—original draft preparation, A.A.-K.; writing—review and editing, G.K.; visualization, A.A.-K.; supervision, G.K.; project administration, A.A.-K.; funding acquisition, A.A.-K. Both authors have read and agreed to the published version of the manuscript.

**Funding:** This research was supported by grants from Sultan Qaboos University.

**Institutional Review Board Statement:** The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Ethics Committee of the School of English Literature, Language and Linguistics at Newcastle University (Faculty of Humanities and Social Sciences Ethics Committee, approved on the 9th of June 2017).

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding authors. The data are not publicly available due to ethical restrictions.

**Acknowledgments:** We thank Martha Young-Scholten for comments on an earlier version.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## Notes

- <sup>1</sup> This refers to whether the consonant cluster was maintained or modified by, for example, by epenthesis of a vowel to break it up or omitting one of the consonants.

## References

- Adrián, José Antonio, Jesus Alegria, and José Morais. 1995. Metaphonological abilities of Spanish illiterate adults. *International Journal of Psychology* 30: 329–51. [CrossRef]
- Akahane-Yamada, Reiko. 1995. Age and acquisition of second language speech sounds perception of American English /r/ and /l/ by native speakers of Japanese. In *Speech Perception and Linguistic Experience*. Baltimore: York Press.
- Alhawari, Mohammed, ed. 2018. *The Routledge Handbook of Second Language Acquisition*. New York: Routledge.
- Al-Kendi, Azza. 2021. Foreigner-Directed Speech and L2 Speech Learning in an Understudied Interactional Setting: The Case of Foreign-Domestic Helpers in Oman. Ph.D. thesis, Newcastle University, Newcastle upon Tyne, UK.
- Anderson, Janet. 1987. The markedness differential hypothesis and syllable structure difficulty. In *Interlanguage Phonology: The Acquisition of a Second Language Sound System*. Edited by Georgette Ioup and Weinberger Steven. Cambridge: Newbury House, pp. 279–91.
- Aoyama, Katsura, James Emil Flege, Susan G. Guion, Reiko Akahane-Yamada, and Tsuneo Yamada. 2004. Perceived phonetic dissimilarity and L2 speech learning: The case of Japanese /r/ and English /l/ and /r/. *Journal of Phonetics* 32: 233–50. [CrossRef]
- Bassetti, Bene, and Nathan Atkinson. 2015. Effects of orthographic forms on pronunciation in experienced instructed second language learners. *Applied Psycholinguistics* 36: 67–91. [CrossRef]
- Bassetti, Bene, Paola Escudero, and Rachel Hayes-Harb. 2015. Second language phonology at the interface between acoustic and orthographic input. *Applied Psycholinguistics* 36: 1–6. [CrossRef]
- Bates, Douglas, Martin Mächler, Ben Bolker, and Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67: 1–48. [CrossRef]
- Beddor, Patrice, and Terry Gottfried. 1995. Methodological issues in adult cross-language research. In *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*. Edited by Winifred Strange. Baltimore: York Press, pp. 207–32.
- Best, Catherine T., and Winifred Strange. 1992. Effects of phonological and phonetic factors on cross-language perception of approximants. *Journal of Phonetics* 20: 305–30. [CrossRef]
- Best, Catherine T., Gerald W. McRoberts, and Nomathemba M. Sithole. 1988. Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance* 14: 345. [CrossRef]
- Bigelow, Martha, and Elaine Tarone. 2004. The role of literacy level in second language acquisition: Doesn't who we study determine what we know? *TESOL Quarterly* 38: 689–700. [CrossRef]
- Binte Faizal, Siti Syuhada. 2019. Visual Word Processing of Non-Arabic-speaking Qur'anic Memorisers. Ph.D. thesis, Newcastle University, Newcastle upon Tyne, UK.
- Bizri, Fida. 2014. Maids' Talk. *Colonization and Domestic Service: Historical and Contemporary Perspectives* 14: 113.
- Boersma, Paul, and David Weenink. 2009. Praat: Doing Phonetics by Computer. Available online: <http://www.fon.hum.uva.nl/praat/> (accessed on 30 September 2016).
- Christensen, Rune Haubo Bojesen. 2019. *Regression models for ordinal data [R package ordinal version 12–10]*.
- Craats, Ineke van de, Jeanne Kurvers, and Martha Young-Scholten. 2006. Research on low-educated second language and literacy acquisition. *LOT Occasional Series* 6: 7–23.
- DeKeyser, Robert M. 2000. The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition* 22: 499–533. [CrossRef]
- Eckman, Fred R. 1985. Some Theoretical and Pedagogical Implications of the Markedness Differential Hypothesis. *Studies in Second Language Acquisition* 7: 289–307. [CrossRef]
- Escudero, Paola, and Paul Boersma. 2004. Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition* 26: 551–85. [CrossRef]
- Escudero, Paola, Ellen Simon, and Karen E. Mulak. 2014. Learning words in a new language: Orthography doesn't always help. *Bilingualism-Language and Cognition* 17: 384–95. [CrossRef]
- Escudero, Paola. 2006. The phonological and phonetic development of new vowel contrasts in Spanish learners of English. *English with a Latin Beat: Studies in Portuguese/Spanish-English Interphonology* 31: 41.
- Flege, James Emil, and Kathryn L. Fletcher. 1992. Talker and listener effects on degree of perceived foreign accent. *The Journal of the Acoustical Society of America* 91: 370–89. [CrossRef]
- Flege, James Emil, and Ocke-Schwen Bohn. 2021. The Revised Speech Learning Model (SLM-r). In *Second Language Speech Learning: Theoretical and Empirical Progress*. Edited by Wayland Ratree. Cambridge: Cambridge University Press, pp. 3–83. [CrossRef]
- Flege, James Emil, and Serena Liu. 2001. The effect of experience on adults' acquisition of a second language. *Studies in Second Language Acquisition* 23: 527–52. [CrossRef]
- Flege, James Emil, Murray J. Munro, and Ian R. A. MacKay. 1995. Factors affecting strength of perceived foreign accent in a second language. *The Journal of the Acoustical Society of America* 97: 3125–34. [CrossRef] [PubMed]

- Flege, James Emil, Ocke-Schwen Bohn, and Sunyoung Jang. 1997. Effects of Experience on Non-native Speakers' Production and Perception of English Vowels. *Journal of Phonetics* 25: 437–70. [CrossRef]
- Flege, James Emil. 1981. The phonological basis of foreign accent: A hypothesis. *TESOL Quarterly* 15: 443–55. [CrossRef]
- Flege, James Emil. 1987. A critical period for L2 learning to pronounce foreign languages? *Applied Linguistics* 8: 162–77. [CrossRef]
- Flege, James Emil. 1988. Factors affecting degree of perceived foreign accent in English sentences. *The Journal of the Acoustical Society of America* 84: 70–79. [CrossRef]
- Flege, James Emil. 1995. Second language speech learning: Theory, findings, and problems. In *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*. Edited by Winifred Strange. Timonium: York Press, pp. 233–77.
- Flege, James Emil. 2002. Interactions between the native and second-language phonetic systems. In *An Integrated View of Language Development: Papers in Honor of Henning Wode*. Edited by Burmeister Petra, Piske Thorsten and Rohde Andreas. Germany: Wissenschaftlicher Verlag Traer, pp. 217–44.
- Flege, James Emil. 2009. Give input a chance. In *Input Matters in SLA*. Edited by T. Piske and M. Young-Scholten. Bristol: Multilingual Matters, pp. 175–90.
- Flege, James Emil. 2018. It's input that matters most, not age. *Bilingualism Language and Cognition* 21: 919–20. [CrossRef]
- Gass, Susan M. 2013. *Second Language Acquisition: An Introductory Course*. New York: Routledge.
- Granena, Gisela, and Michael H. Long. 2013. Age of onset, length of residence, language aptitude, and ultimate L2 attainment in three linguistic domains. *Second Language Research* 29: 311–43. [CrossRef]
- Gunther, Frank H. 1994. A neural network model of speech acquisition and motor equivalent speech production. *Biological Cybernetics* 72: 43–53. [CrossRef] [PubMed]
- Guion, Susan G., James E. Flege, Serena H. Liu, and Grace H. Yeni-Komshian. 2000. Age of learning effects on the duration of sentences produced in a second language. *Applied Psycholinguistics* 21: 205–28. [CrossRef]
- Haznedar, Belma, Joy Kreeft Peyton, and Martha Young-Scholten. 2018. Teaching adult migrants. *Critical Multilingualism Studies* 6: 155–83.
- Henrich, Joseph, Steven J. Heine, and Ara Norenzayan. 2010. Most people are not WEIRD. *Nature* 466: 29. [CrossRef] [PubMed]
- Højen, Anders. 2019. Improvement in Young Adults' Second-language Pronunciation after Short-term Immersion. In *A Sound Approach to Language Matters—In Honor of Ocke-Schwen Bohn*. Edited by Anne Mette, Hejná Michaela, Højen Anders, Jespersen Anna and Sørensen Mette. Aarhus: Department of English, School of Communication & Culture, Aarhus University, pp. 543–59.
- Holliday, Adrian. 2018. Native-Speakerism. In *The TESOL Encyclopedia of English Language Teaching*. Edited by J. I. Lontas. Hoboken: John Wiley & Sons, pp. 1–7. [CrossRef]
- Ioup, Georgette, Elizabeth Boustagui, Manal El Tigi, and Martha Moselle. 1994. Reexamining the critical period hypothesis: A case study of successful adult SLA in a naturalistic environment. *Studies in Second Language Acquisition* 16: 73–98. [CrossRef]
- Iverson, Paul, and Patricia K. Kuhl. 1996. Influences of phonetic identification and category goodness on American listeners' perception of /r/ and /l/. *The Journal of the Acoustical Society of America* 99: 1130–40. [CrossRef]
- Jeremy, Cross. 2002. Noticing in SLA: Is it a valid concept. *TESL-EJ* 6: 1–9.
- Klein, Wolfgang, and Clive Perdue. 1997. The Basic Variety (or: Couldn't natural languages be much simpler?). *Second Language Research* 13: 301–47. [CrossRef]
- Kuhl, Patricia K. 1991. Human adults and human infants show a 'perceptual magnetic effect' for the prototypes of speech categories, monkeys do not. *Perception & Psychophysics* 50: 93–107.
- Kurvers, Jeanne, Ton Vallen, and Roeland van Hout. 2006. Discovering features of language: Metalinguistic awareness of adult illiterates. *LOT Occasional Series* 6: 69–88.
- Lenneberg, Eric H. 1967. The biological foundations of language. *Hospital Practice* 2: 59–67. [CrossRef]
- Lombardi, Linda. 2003. Second language data and constraints on manner: Explaining substitutions for the English interdentals. *Second Language Research* 19: 225–50. [CrossRef]
- Lynch, Tony. 2001. Seeing what they meant: Transcribing as a route to noticing. *ELT Journal* 55: 124–32. [CrossRef]
- Macmillan, Neil A., and Creelman C. Douglas. 1991. *Detection Theory: A User's Guide*. Cambridge: Cambridge University Press.
- McAllister, Robert. 2001. Experience as a factor in L2 phonological acquisition. *Working Papers-Lund University Department of Linguistics* 49: 116–19.
- Midi, Habshah, Saroje Kumar Sarkar, and Sohel Rana. 2010. Collinearity diagnostics of binary logistic regression model. *Journal of Interdisciplinary Mathematics* 13: 253–67. [CrossRef]
- Morais, José, Luz Cary, Jesús Alegria, and Paul Bertelson. 1979. Does awareness of speech as a sequence of phones arise spontaneously? *Cognition* 7: 323–31. [CrossRef]
- Moyer, Alene. 2004. *Age, Accent, and Experience in Second Language Acquisition: An Integrated Approach to Critical Period Inquiry*. Buffalo: Multilingual Matters, vol. 7.
- Moyer, Alene. 2009. Input as a critical means to an end: Quantity and quality of experience in L2 phonological attainment. In *Input Matters in SLA*. Edited by Thorsten Piske and Martha Young-Scholten. Buffalo: Multilingual Matters, pp. 159–74.
- Moyer, Alene. 2014. What's age got to do with it? Accounting for individual factors in second language accent. *Studies in Second Language Learning and Teaching* 4: 443–64. [CrossRef]
- Nimz, Katharina, and Ghada Khattab. 2020. On the role of orthography in L2 vowel production: The case of Polish learners of German. *Second Language Research* 36: 623–52. [CrossRef]

- Ordem, Eser, and Erdogan Bada. 2017. A Variety of Turkish in Mid-South Turkey: Arabic-Turkish Inter-Language. *Journal of Education and Training Studies* 5: 12–19. [CrossRef]
- Oyama, Susan. 1976. A sensitive period for the acquisition of a nonnative phonological system. *Journal of Psycholinguistic Research* 5: 261–83. [CrossRef]
- Pfenninger, Simone E., and David Singleton. 2017. *Beyond Age Effects in Instructional L2 Learning: Revisiting the Age Factor*. Bristol: Multilingual Matters.
- Piske, Thorsten, Ian R. A. MacKay, and James E. Flege. 2001. Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics* 29: 191–215. [CrossRef]
- R Development Core Team. 2012. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing, Available online: <http://www.R-project.org/> (accessed on 18 April 2017).
- Ratner, Nan Bernstein. 2000. Elicited imitation and other methods for the analysis of trade-offs between speech and language skills in children. In *Methods for Studying Language Production*. New York: Psychology Press, pp. 291–312.
- Rice, Keren. 2007. Markedness in phonology. In *The Cambridge Handbook of Phonology*. Edited by Paul de Lacy. Cambridge: Cambridge University Press, pp. 79–97.
- Schmidt, Richard W. 1990. The role of consciousness in second language learning. *Applied Linguistics* 11: 129–58. [CrossRef]
- Scovel, Thomas. 1988. *A Time to Speak: A Psycholinguistic Inquiry into the Critical Period for Human Speech*. Rowley: Newbury House Publishers.
- Seliger, Herbert W. 1975. Maturational constraints in the acquisition of second language accent. *Language Sciences* 36: 20–22.
- Skehan, Peter. 1998. *A Cognitive Approach to Language Learning*. Oxford: Oxford University Press.
- Strange, Winifred, and Valerie L. Shafer. 2008. Speech perception in second language learners. In *Phonology and Second Language Acquisition*. Edited by Jette Hansen and Mary Zampini. Philadelphia: John Benjamins, pp. 153–91.
- Supriyadi, Tedi, and J. Julia. 2019. The problem of students in reading the Quran: A reflective-critical treatment through Action research. *International Journal of Instruction* 12: 311–26. [CrossRef]
- Tarone, Elaine, and Martha Bigelow. 2005. Impact of literacy on oral language processing: Implications for second language acquisition research. *Annual Review of Applied Linguistics* 25: 77–97. [CrossRef]
- Werker, Janet F., and John S. Logan. 1985. Cross-language evidence for three factors in speech perception. *Attention, Perception & Psychophysics* 37: 35–44.
- Young-Scholten, Martha, and Nancy Strom. 2006. First-time L2 readers: Is there a critical period? *LOT Occasional Series* 6: 45–68.
- Young-Scholten, Martha. 2013. Low-educated immigrants and the social relevance of second language acquisition research. *Second Language Research* 29: 441–54. [CrossRef]
- Zec, Draga. 2007. The syllable. In *The Cambridge Handbook of Phonology*. Edited by Paul de Lacy. Cambridge: Cambridge University Press, pp. 161–94.



## Article

# The Effect of Dual Language Activation on L2-Induced Changes in L1 Speech within a Code-Switched Paradigm

Ulrich Reubold <sup>1,\*</sup>, Sanne Ditewig <sup>1</sup>, Robert Mayr <sup>2</sup> and Ineke Mennen <sup>1,\*</sup><sup>1</sup> Department of English, University of Graz, 8010 Graz, Austria; sanne.ditewig@uni-graz.at<sup>2</sup> Centre for Speech and Language Therapy and Hearing Science, Cardiff School of Health Sciences, Cardiff Metropolitan University, Cardiff CF5 2YB, UK; rmayr@cardiffmet.ac.uk

\* Correspondence: johann.reubold@uni-graz.at (U.R.); ineke.mennen@uni-graz.at (I.M.)

**Abstract:** The present study sought to examine the effect of dual language activation on L1 speech in late English–Austrian German sequential bilinguals, and to identify relevant predictor variables. To this end, we compared the English speech patterns of adult migrants to Austria in a code-switched and monolingual condition alongside those of monolingual native speakers in England in a monolingual condition. In the code-switched materials, German words containing target segments known to trigger cross-linguistic interaction in the two languages (i.e., [v–w], [ʃt(ʁ)–st(ɪ)] and [l–l]) were inserted into an English frame; monolingual materials comprised English words with the same segments. To examine whether the position of the German item affects L1 speech, the segments occurred either *before* the switch (“He wants a *Wienerschnitzel*”) or *after* (“I like *Würstel* with mustard”). Critical acoustic measures of these segments revealed no differences between the groups in the monolingual condition, but significant L2-induced shifts in the bilinguals’ L1 speech production in the code-switched condition for some sounds. These were found to occur both before and after a code-switch, and exhibited a fair amount of individual variation. Only the amount of L2 use was found to be a significant predictor variable for shift size in code-switched compared with monolingual utterances, and only for [w]. These results have important implications for the role of dual activation in the speech of late sequential bilinguals.

**Keywords:** L1 attrition; speech; code-switching; English; Austrian German; phonetic drift

**Citation:** Reubold, Ulrich, Sanne Ditewig, Robert Mayr, and Ineke Mennen. 2021. The Effect of Dual Language Activation on L2-Induced Changes in L1 Speech within a Code-Switched Paradigm. *Languages* 6: 114. <https://doi.org/10.3390/languages6030114>

Academic Editor: Elena Babatsouli

Received: 23 April 2021

Accepted: 25 June 2021

Published: 29 June 2021

**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

It is now widely accepted that late sequential bilinguals who are being immersed in a second language (L2) environment long-term may experience a change in linguistic abilities in their native language, a phenomenon commonly referred to as first language (L1) attrition (Köpke and Schmid 2004). While the majority of L1 attrition studies have focused on linguistic levels such as syntax, morphology, and the lexicon (Schmid 2002), recent years have seen a proliferation of research on phonetic and phonological attrition. So far, studies evidenced changes to the L1 in both segmental (Bergmann et al. 2016; de Leeuw et al. 2013; de Leeuw et al. 2018b; de Leeuw 2019a; Guion 2003; Flege 1987; Kornder and Mennen 2021; Major 1992; Mayr et al. 2012; Stoehr et al. 2017; Ulbrich and Ordín 2014) and prosodic (de Leeuw et al. 2012; Mennen and Chousi 2018) areas of L1 production. While the above studies provide evidence for phonetic attrition, changes in L1 phonology have also been reported, resulting in, for instance, a neutralization of phonological contrasts (Cho and Lee 2016; de Leeuw et al. 2018a; Dmitrieva et al. 2010) or a change in L1-specific prominence patterns in anaphora resolution (Gargiulo and Tronnier 2020). Moreover, a few studies have shown that attrition can also affect the perception of segments (Ahn et al. 2017; Celata and Cancila 2010; Dmitrieva 2019), and that native listeners’ global foreign accent ratings may be influenced by long-term immersion in an L2 environment (Major 2010). Evidence of L2-induced changes to L1 pronunciation has not only been observed in long-term residents in an L2 environment, but has also been found

to affect relatively inexperienced L2 learners with little or short-term exposure to the L2 (Chang 2012, 2013, 2019; Dmitrieva et al. 2020; Kartushina et al. 2016) or whose linguistic environment regularly changes by travelling between the L1- and L2-speaking countries (Sancier and Fowler 1997; Tobin et al. 2017). This type of change is, however, typically referred to as “gestural drift” (Sancier and Fowler 1997) or “phonetic drift” (Chang 2012, 2013) rather than L1 attrition, because of the temporary or transient nature of the change.

The changes to L1 pronunciation reported in the literature appear to take the form of either assimilation or dissimilation patterns. Assimilation refers to a shift of the L1 sound in the direction of the L2 sound, resulting in a complete or intermediate merging of L1 and L2 categories (e.g., Alharbi et al. forthcoming; de Leeuw et al. 2013; Flege 1987; Flege and Hillenbrand 1984; Major 1992; Mayr et al. 2012; Ulbrich and Ordin 2014). Dissimilation or polarization refers to a shift of the L1 sound away from both L1 and L2 norms, resulting from “overshooting” the L1 norm (Flege and Eefting 1987; de Leeuw et al. 2012). According to the Speech Learning Model (SLM) (Flege 1995; Flege and Bohn 2021), which has been adopted in a number of previous studies on L1 attrition of speech (Mayr et al. 2012; Bergmann et al. 2016; de Leeuw et al. 2018a; de Leeuw et al. 2018b; Mayr et al. 2020), these assimilation and dissimilation effects arise because L1 and L2 phonetic categories exist in a common phonetic space, and are, therefore, likely to influence one another. Assimilation results from the learners’ inability to discern phonetic differences between L1 and L2 sounds, which “may cause the L1 sound to shift toward (assimilate to) the L2 sound in phonetic space” (Flege and Bohn 2021, p. 42). Dissimilation, on the other hand, is caused by the need “to maintain phonetic contrast between certain pairs of L1 and L2 sounds” (Flege and Bohn 2021, p. 39). While the studies outlined above show that both patterns can occur in L1 attrition of speech, dissimilation patterns are less commonly reported than assimilation patterns (so far, only in Flege and Eefting 1987 and de Leeuw et al. 2012).

The present study aims to contribute to the growing body of work on L2-induced changes to L1 speech production in late sequential bilinguals residing in an L2-speaking environment by acoustically examining a number of segments produced by a group of L1 speakers of Standard Southern British English (SSBE) who have emigrated to Austria in adulthood and have (Austrian) German as their L2.

### 1.1. The Scale of L1 Attrition of Speech

The steadily increasing body of evidence outlined above suggests that the observed changes to L1 pronunciation are not limited to a few isolated cases but are widespread and pervasive (de Leeuw 2019b; Mayr et al. 2020). Nevertheless, the documented changes are unlikely to represent the full scale of L1 attrition in the phonetic/phonological domain, as only limited features have been investigated so far, with a focus on temporal (e.g., Alharbi et al. forthcoming; Flege 1987; Kornder and Mennen 2021; Major 1992; Mayr et al. 2012; Stoehr et al. 2017) and spectral aspects (e.g., de Leeuw et al. 2013; de Leeuw 2019a; Ulbrich and Ordin 2014) of consonant production, and spectral and/or durational aspects of vowel production (e.g., Bergmann et al. 2016; de Leeuw 2019a; Guion 2003; Kornder and Mennen 2021; Mayr et al. 2012). Furthermore, evidence is emerging that not all phonetic aspects are vulnerable to attrition. For instance, in a study of two Dutch monozygotic twin sisters, one of whom moved to the UK in adulthood (Mayr et al. 2012), pervasive changes were found in the L1 (Dutch) voiceless plosives produced by the sister who was immersed in the English-speaking environment, whereas no such effects were found in her productions of the Dutch voiced plosives. Similarly, while extensive attrition was found in her production of Dutch vowels, no evidence of attrition was found in her production of the Dutch /a/. Likewise, Stoehr et al. (2017), in a study of Dutch–German late sequential bilinguals, found evidence of L1 attrition in the Dutch voiceless plosives, whereas the same participants were able to maintain native levels of prevoicing in voiced plosives. Finally, Bergmann et al. (2016) found evidence of L1 attrition in the formant frequencies for /a:/ and /ɪ/, but not for /ɛ/ and /ɔ/ in their study of L1 German immigrants to Canada and the US.

This suggests that not all segments are equally prone to L1 attrition. In order to establish whether some areas of pronunciation are indeed more susceptible to attrition than others, one would need to compare a range of areas of pronunciation within the same group of individuals. Studies examining more than one sound class are, however, rare (but see Bergmann et al. 2016; de Leeuw 2019a; Kornder and Mennen 2021; Mayr et al. 2012).

### 1.2. Individual Variation and the Role of L1 Use and Dual Language Activation

While it is clear that L1 speech may undergo attrition, research has also shown that changes to L1 speech are only evident in some, *but not all*, late sequential bilinguals residing in an L2-speaking environment. For instance, in Major's (1992) study of voice onset time in voiceless plosives, one out of five of the English–Portuguese bilinguals produced the plosives entirely natively in both languages, thus showing no evidence of L1 attrition. Similarly, only three of the ten Albanian–English bilinguals in de Leeuw et al. (2018a) were found to suffer from attrition, by neutralizing or partially neutralizing the contrast between dark and light laterals, while the other seven participants showed no sign of attrition. Likewise, de Leeuw et al. (2013) report that, in a group of ten German–English late bilinguals, nine exhibited variable degrees of changes to their L1 tonal alignment patterns, but one participant produced tonal alignment values that were entirely within the norms of monolingual speakers of either language. These results suggest that not every late sequential bilingual immersed in an L2 environment is affected by phonetic attrition and that there is a fair amount of interpersonal variation in the degree of attrition evidenced in production studies. There is also notable variation in the number of perceived attriters in global foreign accent rating studies, with reported figures ranging from 25% (de Leeuw et al. 2010) to 40% (Bergmann et al. 2016) of the investigated migrant population.

In order to reach a better understanding of the nature of L1 attrition, it is necessary to establish which factors drive the attrition process and lead to apparent changes in the pronunciation of some individuals but not others. While there is a fairly good understanding of the factors that contribute to individual differences in how well L2 learners are able to produce L2 speech, there has been little attempt to establish the variables that predict which individuals are more likely to undergo changes in L1 speech. It is often assumed that some variables—such as an increase in exposure to and use of the L2, as well as a reduction in L1 use associated with L2 immersion (see Dmitrieva et al. 2020)—may be important; however, the role of various predictor variables is rarely systematically investigated. The only studies that have tested predictor variables in L1 speech attrition have investigated their impact on global foreign accent ratings (de Leeuw et al. 2010; Hopp and Schmid 2013) and the limited findings so far remain inconclusive. Hopp and Schmid (2013) tested the influence of age of emigration, length of residence in an L2 environment, amount of L1 and L2 use, and L2 proficiency on global foreign accent ratings in L1 attriters. None of the tested factors were found to predict the degree of perceived foreign accent in the L1. However, in their study on German immigrants to the Netherlands and Anglophone Canada, de Leeuw et al. (2010) found that of the three variables tested (language use, age of arrival in the host country, length of residence in the host country), only language use predicted the perceived degree of foreign accent, while the other two factors were not found to play a role. Bilinguals were more often perceived as non-native in their L1 if they used their L1 in situations where they were likely to code-switch (such as in conversations with family members or friends). The bilinguals who used German in situations where code-switching was unlikely to occur (such as in work settings or conversations with monolingual speakers of the L1) were more likely to be perceived as native speakers of their L1. This suggests that situations where contact with the L1 predominantly takes place in a setting where only the L1 is used may protect against changes in L1 pronunciation; conversely, L1 contact in a situation where bilinguals regularly use both languages may facilitate L1 speech attrition (de Leeuw et al. 2010).

While effects of language use on L1 speech attrition have also been documented in other studies (e.g., Stoehr et al. 2017; Chang 2019), the particular language use situation



referred to in [de Leeuw et al. \(2010\)](#) is one which requires sustained co-activation of the L1 and L2. Although both languages of a bilingual are always activated to some extent, the degree of activation can range from full activation of both languages, referred to as a bilingual language mode, to inhibition of one of the bilingual's languages in a monolingual language mode ([Grosjean 2001](#)). In contexts of dual language activation, and particularly so in code-switching where both languages need to be active, continued cross-linguistic interaction is commonly observed (e.g., [Green 1998](#); [Van Hell and Dijkstra 2002](#)). In [de Leeuw et al.'s \(2010\)](#) study, the dual activation of both languages during the bilinguals' L1 contact, resulting from high levels of code-switching, led to them being perceived as less native in their L1. [Mayr et al. \(2020\)](#), in a study of Spanish language teachers and non-teachers residing in the UK, aimed to differentiate the effects of L1 use and dual language activation in the perceived attrition of L1 speech. It was argued that when teaching their L1 in an L2-speaking environment, language teachers are unlikely to be able to inhibit one of their languages to the extent that non-teachers are able to do. While non-teachers may activate their L2 predominantly at work and their L1 at home, the specific demands on language teachers in foreign language classroom settings will inevitably lead to dual language activation during teachers' professional activities. By investigating two groups of Spanish speakers, teachers and non-teachers, who differed in their need to co-activate their languages in professional settings, but had similar levels of L1 use, [Mayr et al. \(2020\)](#) were able to separate the two possible effects. Their results showed that the group of Spanish language teachers were perceived as less native in their L1 than the group of non-teachers, whereas low L1 use had no effect on the perceived L1 accent. This suggests that bilinguals who are often in situations that lead to cross-linguistic interaction, such as L1 contact where code-switching is common, may be more prone to attrition than bilinguals who "function in alternate monolingual language modes", and reduced L1 use in itself is likely to play less of a role in L1 attrition of speech ([Mayr et al. 2020](#), p. 14).

### 1.3. Code-Switching and Its Effect on L1 Speech Production

Despite the suggestions in the literature that dual language activation and its resulting cross-linguistic interaction may play a role in L1 attrition of speech, few studies have addressed how it influences L1 speech production. Past studies of L1 attrition of speech have investigated the effects on the L1 predominantly within monolingual language modes ([Grosjean 2001](#)), where bilinguals' two languages are carefully separated to avoid dual language activation, and interaction between the two languages is less likely to occur. Comparatively little research has investigated phonetic interaction in situations where both languages are maximally activated—such as in code-switched speech—a situation which is thought to serve "as a catalyst for interaction" ([Olson 2013](#), p. 410). Yet, code-switched speech provides an interesting context in which to examine L1 attrition of speech, as it offers an opportunity to maximize the number and types of possible changes that can occur in L1 speech attrition, whereas speech produced in a non-code-switched context is much less likely to result in cross-language interaction. When the two languages are maximally activated, the full extent of language interaction should become visible, allowing for an examination of the relative susceptibility of different areas of pronunciation to L1 attrition.

So far, however, the majority of studies on the effect of code-switching on speech production have focused on investigations of early bilinguals in languages in contact research. On the one hand, there are some reporting unidirectional transfer with Language A changing towards Language B, but not the reverse, including L1-to-L2 transfer ([Antoniou et al. 2011](#); [Bullock et al. 2006](#); [Goldrick et al. 2014](#); [Muldner et al. 2019](#)) and L2-to-L1 transfer ([Olson 2013](#)). Other studies exhibit bidirectional transfer, with Language A changing towards Language B and Language B changing towards Language A ([Balukas and Koops 2015](#); [Bullock and Toribio 2009](#); [Piccinini and Arvaniti 2015](#)), or no effect of code-switching ([Grosjean and Miller 1994](#); [López 2012](#)). The focus of these studies has been almost exclusively on the production of voice onset time in plosives (but see [Muldner et al. 2019](#); [Olson 2012](#)). There are as yet very few studies that have investigated the effects of

code-switching on the L1 speech of late sequential bilinguals who are being long-term immersed in an L2 environment, and findings have so far been contradictory. Bullock and Toribio (2009) report asymmetrical effects in the L1 production of voice onset times (VOT), with VOT affected in the English–Spanish bilinguals’ L1, but not in the L1 of the Spanish–English bilinguals. Olson (2013), however, found consistent L2-to-L1 effects in VOT production by both Spanish–English and English–Spanish late sequential bilinguals. Neither Bullock and Toribio (2009) nor Olson (2013) reported differences in the relative susceptibility of the segments under investigation, but this is perhaps not surprising, given that they only examined plosives. Moreover, none of the studies investigated which factors may have resulted in L1 attrition occurring in some (groups of) bilinguals but not others. A final point of interest is whether the influence of a switch is more evident and/or more frequent in anticipation of the switch or following the switch. So far, only Bullock and Toribio (2009) examined whether the effect of code-switching on production is influenced by the position of the switch. Their results show that the observed changes in the L1 of the English–Spanish and Spanish–English late sequential bilinguals only occurred in anticipation of a switch, and no changes were observed in L1 VOT productions following a switch. This suggests that the *direction* of L2-induced influences may play a role in the occurrence of L1 speech changes, in that segments occurring before a switch may be more prone to changes than those after a switch. However, more research is needed to confirm or reject this hypothesis.

#### 1.4. The Current Study

The current study investigated L1 speech attrition within a context of dual language activation using a code-switched paradigm in which a bilingual’s two languages are thought to be maximally activated and high levels of language interaction are expected to occur (Green 1998; Olson 2013). The study compares a number of L1 segments, in switched and non-switched conditions, produced by a group of late English–Austrian German (henceforth: English–Austrian) sequential bilinguals, all native speakers of SSBE who emigrated to Austria in adulthood and have (Austrian) German as their L2. This allows us to determine whether some segments are more prone to L2-induced changes in L1 speech than others, and whether this is more evident—as suggested by previous research—in a context of dual language activation than in a monolingual language mode where one of the languages is more likely to be inhibited (de Leeuw et al. 2010; Mayr et al. 2020). The experimental design of the study, in which German words containing segments expected to trigger transfer are inserted into an English frame, and corresponding English segments occur both before and after the German word, also allows us to address whether the L2-induced phonetic influences on L1 production are progressive or regressive. A final objective of this study is to test the influence of a number of predictor variables (i.e., age of emigration, length of residence in an L2 environment, relative amount and quality of L1 and L2 use, and L2 proficiency) on the degree of L1 attrition, with the aim of identifying which predictor variables govern L1 attrition and explaining why L1 attrition occurs in some individuals but not others.

## 2. Materials and Methods

### 2.1. Participants

We recruited a group of late sequential bilinguals (BIL,  $N = 25$ , 11 females, 14 males) who were raised monolingually in the southeast of England as speakers of SSBE and moved to Austria as adults. The speakers in this group were all long-term residents in Austria and lived there continuously for a minimum of two years. The participants were recruited via existing contacts, as well as expatriate communities, language schools, and the British Embassy in Vienna. The age of the participants ranged from 24 to 71 years, with a mean age of 44.8 years. None of the participants reported daily use of foreign languages other than German. In addition, a small group of monolingual speakers of SSBE residing in the UK were recruited for the study (MON,  $N = 10$ , 6 female, 4 male). The monolingual

speakers had never lived outside England and reported no more than high school level knowledge of other languages. Participants in both groups were educated to a tertiary level. The age of the MON participants ranged from 21 to 56 years, with a mean age of 29.6 years. For all speakers in each group, we obtained informed consent before participation, and all participants were offered compensation for their time. None of the participants reported any known speech, language, or hearing impairments.

For the BIL group, we collected some background information about the participants' language use through an online questionnaire. The questionnaire was adapted from Schmid (2011) to the situation of our participant group, and focused on variables that may predict L2-induced changes to L1 speech (e.g., de Leeuw et al. 2010; Hopp and Schmid 2013). These were AoE (age of emigration to Austria), LoR (length of residence in Austria), L2 proficiency, amount of contact with English in a setting in which language mixing was either likely or not likely, amount of L1 use, and amount of L2 use. Participants were specifically targeted such that they varied in age of emigration (AoE) and length of residence (LoR) in Austria, in order to test the effect of these variables on possible L2-induced changes to L1 speech. The speakers' AoE in Austria varied between 21 and 58 years (mean: 33.5), and their LoR in Austria ranged from 2 to 37 years (mean: 11.5).

Following de Leeuw et al. (2010), the type of contact with the L1 was divided into what was termed C+M (L1 contact in settings where language mixing is *likely* to occur) and C−M (L1 contact in settings where language mixing is *unlikely* to occur). The variable C+M describes informal types of L1 contact, i.e., with family and friends residing in Austria,<sup>1</sup> whereas C−M describes contact with monolingual L1 speakers as well as more formal types of contact (i.e., through work). As in de Leeuw et al. (2010), the two contact variables consisted of the means of answers to a number of questions. The answers to the questions were all expressed on a 5-point Likert scale, which was converted to a scale of 0–1, where 0 refers to the minimum and 1 to the maximum amount of contact, in order to compare our results to those in de Leeuw et al. (2010). The variables “amount of L1 use”, “amount of L2 use” and “L2-proficiency” also consisted of answers to a number of questions using 3 to 6-point Likert scales and were also converted on a scale from 0 to 1. The variable “amount of L1 use” refers to the use of and exposure to the L1 with native speakers living in the UK. The variable “amount of L2 use” refers to the use of and exposure to the L2 with native Austrian speakers. The variable “L2 proficiency” refers to self-reported proficiency in pronunciation, fluency, oral comprehension, writing and reading. Scores for each variable are displayed in Table 1. The scores for C−M ranged from 0.31 to 0.94, for C+M from 0.38 to 0.92, for L1 use from 0.17 to 1, for L2 use from 0.18 to 0.97, and for L2 proficiency from 0.20 to 0.96.

**Table 1.** Background details: bilingual participants.

Participant	Gender	AoE (yrs)	LoR (yrs)	C+M	C−M	L1 Use	L2 Use	L2 Proficiency
BIL001	F	28	21	0.75	0.59	0.75	0.71	0.76
BIL002	M	31	14	0.56	0.86	0.92	0.46	0.72
BIL003	M	34	20	0.63	0.61	0.17	0.50	0.40
BIL004	F	58	2	0.88	0.50	0.63	0.39	0.68
BIL005	M	23	16	0.64	0.67	0.50	0.71	0.80
BIL006	M	39	16	0.80	0.72	0.71	0.44	0.44
BIL007	F	22	32	0.50	0.72	0.58	0.83	0.96
BIL008	M	31	30	0.47	0.58	0.25	0.97	0.52
BIL009	M	34	37	0.31	0.38	0.50	0.96	0.84
BIL010	M	29	3	0.63	0.50	0.35	0.60	0.36
BIL012	F	52	6	0.81	0.79	0.67	0.46	0.20
BIL014	M	39	3	0.69	0.84	0.75	0.36	0.44
BIL015	F	23	13	0.44	0.78	0.42	0.49	0.92
BIL016	M	24	11	0.50	0.78	0.92	0.69	0.60

Table 1. Cont.

Participant	Gender	AoE (yrs)	LoR (yrs)	C+M	C–M	L1 Use	L2 Use	L2 Proficiency
BIL017	F	22	2	0.81	0.69	0.92	0.69	0.84
BIL018	M	29	3	0.69	0.75	0.25	0.71	0.60
BIL019	M	21	12	0.75	0.84	0.63	0.68	0.52
BIL021	F	30	8	0.86	0.67	0.67	0.70	0.36
BIL025	M	43	13	0.80	0.59	0.63	0.52	0.44
BIL027	F	34	2	0.80	0.69	0.92	0.18	0.20
BIL028	M	56	11	0.86	0.75	1.00	0.54	0.64
BIL029	F	44	4	0.89	0.75	0.88	0.67	0.40
BIL030	M	29	3	0.63	0.67	0.75	0.69	0.48
BIL031	F	44	2	0.94	0.78	0.54	0.57	0.52
BIL032	F	24	3	0.84	0.92	0.75	0.27	0.24

Note: AoE = age of emigration, LoR = length of residence; C+M = L1 contact in settings where language mixing is likely to occur; C–M = L1 contact in settings where language mixing is unlikely to occur; the scores for C+M, C–M, L1 use, L2 use, and L2 proficiency are on a scale from 0 to 1, where 0 refers to the minimum amount of contact, use, or proficiency and 1 to the maximum amount. As data come from a larger project, the numbering of participants is not continuous.

The participants mostly reported having been taught Standard Austrian German (SAG) during German languages classes whilst residing in Austria (Moosmüller et al. 2015). Moreover, it is likely that they will have also had exposure to a range of different social and regional forms of German as spoken in Austria. However, while this was not formally assessed, in view of their high education levels, they are likely to have oriented towards standard accentual features.

## 2.2. Speech Materials and Recordings

Regulations concerning contact during the COVID-19 pandemic required the implementation of alternative means of data collection. Therefore, despite concerns about technical issues with remote recordings conducted via online tools (cf. Sanker et al. 2021), we decided to use WikiSpeech.<sup>2</sup> This is a content management system for the web-based creation of speech databases (Draxler and Jänsch 2008) and allows for unsupervised online recordings and a project administration workflow, such as the editing of speech content and data download on the part of the project administrator. Participants received step-by-step instructions. They were asked to read aloud the items presented on their screen and to record themselves using either built-in microphones or headsets and to avoid noisy environments. Participants were also asked to adjust recording levels and to listen to the first few test items, in order to avoid poor audio recording quality. All items were recorded in random order and included two repetitions each. There was an automatic pause of 1.5 s between the presentation of individual stimuli. Participants also had the opportunity to pause the recording session at any time and resume the process at a later point, but were advised to finish a recording session within one sitting. The recordings were checked carefully for their recording quality and, on occasion, participants were asked to re-record certain items containing sound quality issues or reading errors.

The recorded items consisted of materials for the study reported here, interspersed with a larger set of materials devised to test for potential segmental and prosodic changes to L1 speech, which will be reported elsewhere. The entire set consisted of 313 items and took each participant approximately 60 min to record. Of those, all 43 items with our target segments ( $\times 2$  repetitions = 86) form part of the current experiment. These comprise two sets: (i) a set of code-switched (CS) materials (recorded by the BIL group only) and (ii) a set of non-CS materials (recorded by both the BIL and MON group). The CS materials consist of 15 sentences, each repeated twice, in an L1 (English) frame in which German items were inserted, containing the following segments, henceforth referred to as sound pairs: <w> [v-w], <st(r)> [ʃt(ɪ)-st(ɪ)];<sup>3</sup> <l> [l-ɪ]. These were carefully selected based on known cross-linguistic differences between English and German (Kufner 1971; Moulton

1962) and evidence from L2 learning contexts (e.g., Hickey 2020). Each sound pair consists of an English consonant or consonant cluster with a phonetically similar “counterpart” in German, which, in turn, was expected to trigger L2-to-L1 transfer. Specifically, we anticipated that English [ɪ] would become lighter under the influence of L2 German [i]. Similarly, the production of English [w] was expected to be realized in approximation of the labiodental fricative [v], a category that is shared by English and German, but only used for the pronunciation of the grapheme <w> in German. Finally, we predicted that English [st] and [stɪ] would approximate the realization of L2 German [ʃt] and [ʃtɪ], respectively.

In the sound pairs <w> [v-w] and <st(r)> [ʃt(ɪ)-st(ɪ)], the segments were varied to occur either *before* or *after* the German item, to determine whether any observed L2-induced influences are progressive or regressive: e.g., <st> [ʃt-st]: “She walked to *Stadteck station*” (progressive), or “They stayed at *Stiftung*” (regressive). For the sound pair <l> [l-ɪ], we only had sentences for the progressive environment. However, after an analysis of the transfer direction for the sound pairs <w> [v-w] and <st(r)> [ʃt(ɪ)]-[st(ɪ)] showed no significant main effect nor an interaction with other factors (see below in Results Sections 3.1.2 and 3.2.2), we decided to add the sound pair <l> [l-ɪ]. Table 2 illustrates the sentences that were used to elicit the sound pairs in this study, along with the number of progressive and regressive items analyzed. A total of 834 items from the set of CS materials were analyzed, consisting of 206 [w] tokens, 461 [s] tokens (adjacent to [t] or [tɪ]), and 167 [ɪ] tokens.

**Table 2.** Code-switching materials analyzed in this study.

Sound Pair	Direction of Expected L2-Induced Change	N	Sentence
<w> [v-w]	Progressive	100	(a) I like <i>Würstel</i> with mustard. ([ˈvʏɐstl̩ wɪd̩])
			(b) There’s a <i>Wienerwald</i> <sup>1</sup> where I live. ([ˈviːnɐvɔltʰ weəʃ])
	Regressive	106	(a) He wants a <i>Wienerschnitzel</i> . ([wɔnts ə ˈviːnɐʃnɪtsl̩])
			(b) He went to <i>Weiz</i> at a fast pace. ([wentʰ tʰə vaɪts])
<st(r)> [ʃtɪ]-[stɪ]	Progressive	243	(a) Real estate prices in <i>Steyr</i> stabilized overnight. ([ˈʃtaɪ.ɐ ˈsteɪbəlɔɪzd])
			(b) She walked to <i>Stadteck</i> station. ([ˈʃtatʰ.ʔekʰ ˈsteɪfɪn])
			(c) He loves <i>Topfenstrudel</i> <sup>2</sup> strongly. ([ˈtɔpfn̩_ʃtʁuːdl̩ ˈstrɔŋli])
			(d) They saw <i>Strache</i> <sup>3</sup> striding along. ([ˈʃtʁaxɐ ˈstrɑɪdɪŋ])
	Regressive	218	(a) He stalked <i>Stephanie</i> . ([stɔkd ˈʃtefaniː])
			(b) They stayed at <i>Stiftung</i> . ([steɪd ət ˈʃtɪftʰɔŋ])
			(c) He stressed <i>Strafe</i> in a weird way. ([strest ˈʃtʁaːfɐ])
			(d) It’s a <b>strange</b> <i>Strafprozess</i> . ([streɪndʒ ˈʃtʁaːf.pʁɔː.tses])
<ɪ> [l-ɪ]	Progressive	167	(a) Was the <i>Meldezettell</i> filled in? ([mɛldɔtsɛtəl fɪt̩d])
			(b) Who in <i>Tulln</i> killed the cat? ([tʰʊln kɪt̩d])
			(c) He knew <i>Knittelfeld</i> well enough. ([knɪtɔlfɛltʰ wɛt̩])

<sup>1</sup> *Wienerwald* refers to the name of a well-known chain of fast-food restaurants. <sup>2</sup> *Topfenstrudel* was chosen instead of *Strudel* in order to avoid confusion with the English word *strudel*. <sup>3</sup> *Strache* is the surname of a well-known Austrian politician.

The set of non-CS materials (which were recorded for both the BIL and MON group) consisted of 28 words repeated twice (see Table 3), starting with the sounds that form the sound pairs in the CS materials, i.e., [w, v, s, ʃ, l, ɫ]. These words had been recorded with two repetitions, one each embedded within the carrier sentences “Say TARGETWORD again” and “We said TARGETWORD together”, respectively. A total of 1867 tokens were extracted from the recordings of the BIL and MON speakers for further analysis. Table 3 shows the number of words that were analyzed for each sound and group.

**Table 3.** Set of non-CS materials.

Speech Sound	Target Words	N (MON)	N (BIL)
[w]	wary, where, west, wet, while, with, wine	259	105
[v]	vary, vest, vet, vile, vine	259	102
[s]	seat, sick, sinking, sit, study, sin	305	121
[ʃ]	sheep, ship	101	40
[l]	lab, leaf, leap, lip	206	83
[ɫ]	feel, pal, peal, pill	205	81

### 2.3. Data Annotation and Acoustic Measures

The recordings of all participants were digitized at 16 kHz. All data were segmented and labelled automatically into individual phonetic segments using WebMaus, a webtool using forced-alignment algorithms (Kisler et al. 2017). They were then converted for further acoustic analysis into the EMU-SDMS format (Winkelmann et al. 2017, 2020), and all pre-calculated boundaries were checked manually and readjusted where needed. All analyses were carried out within R (R Core Team 2020) with Winkelmann et al.’s (2020) R package *emuR*.

We calculated formant frequencies and power spectra for the recordings used for this study. The frequencies of the first five formants (F1–F5) were calculated from the audio signal by means of Praat’s (Boersma and Weenink 2020) built-in standard formant tracker using the Burg method (cf. Childers 1978, pp. 252–55). For female participants, the frequency range was set between 0 and 5500 Hz, whereas for males, it was set between 0 and 5000 Hz. In both cases, we used a frame shift of 6.25 ms, a window length of 25 ms, and pre-emphasis from 50 Hz. Very few but obvious errors in the first two formants of the target sounds for which formants are relevant ([v, w] or [l, ɫ]), such as when the first formant frequency (F1) was mistracked as a second formant frequency (F2), were manually corrected. A power spectrum was calculated for each of the sounds and later analyzed for [s, ʃ], using a Discrete Fourier Transformation with a 40 Hz frequency resolution, a 5 ms Blackman window, and a 5 ms frame shift. Both formant frequencies ([v, w] or [l, ɫ]) and power spectra [s, ʃ] were extracted at the temporal midpoints of the sounds. For the power spectra of [s, ʃ], we calculated the first spectral moment ( $M_1$ ) in the frequency range 2500–6000 Hz using the *moments* function in the R package *emuR* (Winkelmann et al. 2017).

The first spectral moment ( $M_1$ ) represents the mean (sometimes called the centroid or center of gravity) of the spectral slice. We chose this measure for our [s-ʃ] analysis, as it has been shown to be an effective acoustic parameter for distinguishing between these two sounds (Forrest et al. 1988), with  $M_1$  being the most effective of the spectral moments (Haley et al. 2010), and lower values indicating more [ʃ]-like frication. As Kopečková et al. (2019) found that F2 is a useful measure for the distinction between English [v] and [w] (with low values for [w], and high values for [v]), we chose this as our [v-w] measure. For analysis of English “clear” and “dark” laterals [l, ɫ], we chose the distance between F1 and F2, as proposed by Lehiste (1962) and Carter (2002), with lower values showing increasingly “darker” laterals.

We did not apply any extrinsic speaker normalization technique, but instead, used *Speaker* as a random effects variable in the statistical analyses, whenever applicable. We decided against extrinsic speaker normalization, as we expected BIL speakers to possibly assimilate or dissimilate the phonological contrasts under study, and therefore, diminish or

exaggerate the acoustic differences between the sound pairs. Even if sounds other than the ones in the sound pairs were involved in extrinsic speaker normalization, such a technique could potentially factor out these important differences. Note that the acoustic measure used for the lateral approximants [l, ɫ], i.e., the distance between F1 and F2, already involves intrinsic speaker normalization.

#### 2.4. Statistical Analysis

For each of the three sound pairs (<w> [v-w], <st(r)> [[tʰ-st.ɹ], <l> [l-ɫ]), we ran the same series of four linear mixed models (henceforth LMMs) in the R package *lmerTest*, version 3.1-3 (Kuznetsova et al. 2017), which makes use of the techniques in the package *lme4*, version 1.1-26 (Bates et al. 2015). Post hoc tests were carried out using the package *emmeans* (version 1.5.4; Lenth 2021). In both *lmerTest* and *emmeans*, we used Satterthwaite's method to calculate an approximation to the effective degrees of freedom in order to obtain *p*-values. In all these tests, we used the acoustic parameters (i.e., first spectral moment ( $M_1$ ) for [s-ʃ], F2 for [v-w], and the distance between F1 and F2 for [l-ɫ]) as dependent variables. The first LMM (LMM 1) tested whether the BIL speakers already show signs of L1 attrition in their monolingual utterances, by comparing the productions of MON speakers to BIL speakers in the non-CS speech. We took *Phonological Category* (two levels: the two members of a sound pair), *Group* (two levels: BIL vs. MON), and *Gender* (two levels: female vs. male) as fixed factors, and *Speaker* and *Word* as random factors.

The second LMM (LMM 2) compared the BIL speakers' productions of [w], [s] and [ɫ] in CS utterances with the same sounds and the other member of the sound pair ([ʃ], [ʃ] and [l], respectively) in the non-CS utterances, to test whether code-switching had an effect on L2-induced changes, and if so, whether the changes were toward (assimilatory) or away from (dissimilatory) the respective counterpart. Fixed factors were *Phonetic Category* (3 levels: CS vs. non-CS [w], [s] or [ɫ] vs. non-CS ([w], [ʃ] or [l]), *Direction of L2-induced Influence* (two levels: pro- vs. regressive, i.e., whether the tested sound occurs after or before the code-switch, respectively) and *Gender* (female vs. male); *Speaker* and *Word* were random factors. For the sound pair [l-ɫ], we had to leave out the fixed factor *Direction of L2 induced Influence*, as there were no productions in the regressive context.

We then ran a third LMM (LMM 3) in order to quantify individual variation in the effect of code-switching. To this end, we ran LMM 2 (as above) but this time, with *Speaker* as a fixed factor. The BIL speakers who showed an effect of code-switching in LMM 3 were then, as a group, entered into a fourth LMM (LMM 4), in order to establish whether they showed any sign of L1 attrition in their non-CS speech that might have been obscured in the pooled analysis in LMM 1. To this end, their non-CS speech was compared to that of the MON speakers, in what is essentially a reanalysis of LMM 1, but this time, with a reduced pool of BIL speakers.

Finally, in order to determine the influence of the predictor variables *AoE*, *LoR*, *C+M*, *C-M*, *amount of L1 use*, *amount of L2 use* and *L2 proficiency* on L1 changes in the tested speech sound, we ran several multiple linear regression analyses with the BIL speaker group. We calculated speaker-specific estimates taken from LMM 3's Estimated Marginal Means results as a quantification of shift sizes of CS [w, s, ɫ] away from non-CS [w, s, ɫ], and used these measures as dependent variables in the multiple linear regression tests.

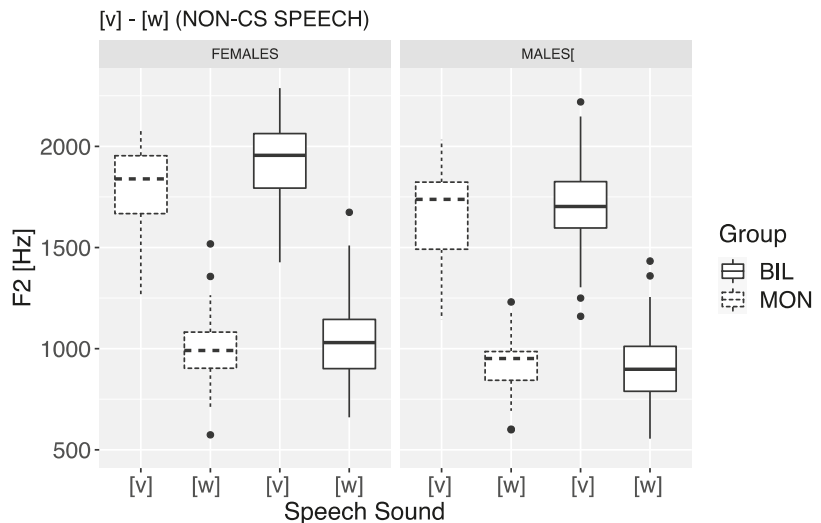
### 3. Results

#### 3.1. [v] vs. [w]

##### 3.1.1. Monolingual vs. Bilingual [v] and [w] in Non-CS Contexts

LMM 1 (see Section 2.4) with F2 as the dependent variable, *Phonological Category* (2 levels: [v] vs. [w]), *Group* (two levels: BIL vs. MON), and *Gender* as fixed factors, and *Speaker* and *Word* as random factors revealed significant main effects for *Phonological Category* ( $F[1, 14.9] = 318.4, p < 0.001$ ), *Group* ( $F[1, 31.1] = 6.6, p < 0.05$ ), and *Gender* ( $F[1, 31.1] = 21.9, p < 0.001$ ) (see Figure 1). We also found significant interactions between *Phonological Category* and *Group* ( $F[1, 30.8] = 6.7, p < 0.05$ ). None of the other two- and

three-way interactions were significant. Post hoc pairwise comparisons with Estimated Marginal Means revealed significant F2 differences between [v] and [w] in all pairwise comparisons ( $p < 0.001$  for each *Gender* and *Group* combination). There was, however, no significant difference between the [w]s produced by the BIL and MON group. Similarly, for the [v]s produced by the males, no significant differences between the BIL and MON were found. However, for the females, F2 in [v]s was found to be significantly higher in BIL as compared to MON ( $p < 0.05$ ).



**Figure 1.** The F2 of intended [v] and [w] sounds in English words, spoken by MON (dashed) and BIL speakers (solid) in non-CS environments. The left panel shows data for females, the right panel for males.

### 3.1.2. Comparison of Bilinguals’ Productions in Non-CS vs. CS Contexts

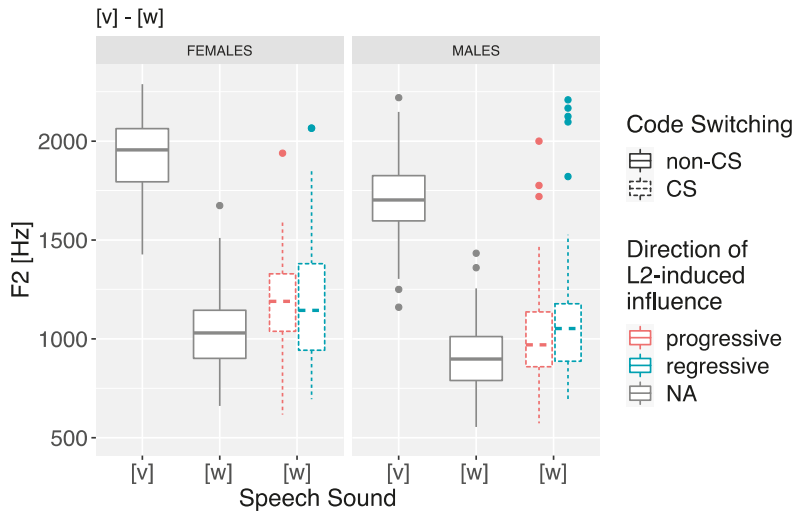
LMM 2 (see Section 2.4) conducted with F2 as the dependent variable, *Phonetic Category* (3 levels: non-CS [v] vs. non-CS [w] vs. CS [w]), *Direction of L2-induced Influence* and *Gender* as fixed factors, and *Speaker* and *Word* as random factors revealed significant main effects for *Phonetic Category* ( $F[2, 16.6] = 194.2, p < 0.001$ ) and *Gender* ( $F[1, 23.0] = 12.4, p < 0.01$ ), but not for *Direction of L2-induced Influence* (as shown in Figure 2). There also were no significant interactions. Estimated Marginal Means showed that non-CS [v] in both males and females is significantly different from non-CS [w], but also CS [w] ( $p < 0.001$  in all cases), showing no evidence of a categorical shift. The shift of CS [w] away from non-CS [w] reaches significance only in males under the regressive influence of German ( $p < 0.05$ ), but not under progressive influence. However, for both females and males, the differences between progressive and regressive German influences on English [w] in CS speech are not significant. Visual inspection of Figure 2 shows that some of the outliers for CS [w] are clearly in the area of the [v] category. This will be discussed further in Section 3.1.3.

### 3.1.3. Individual Variation

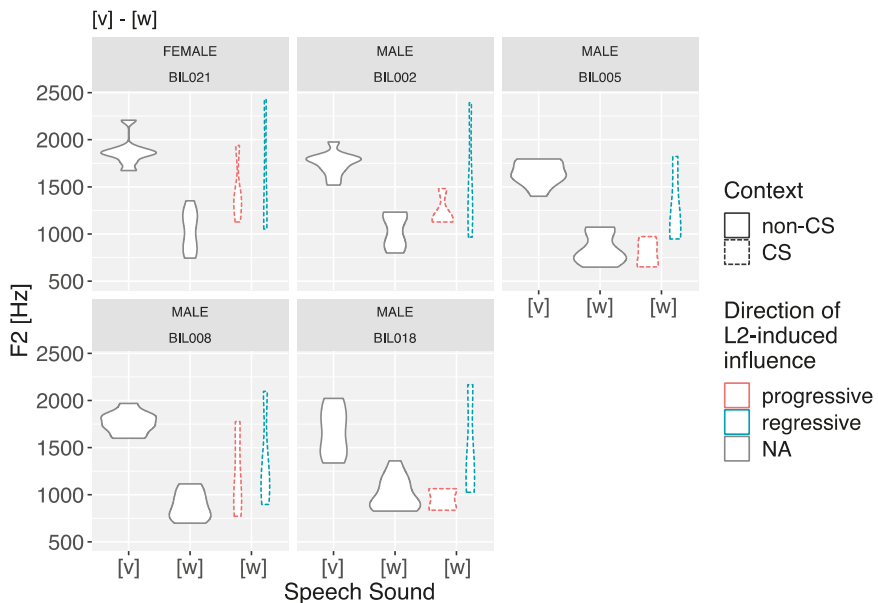
LMM 3 (see Section 2.4) reveals that five speakers (four males and one female) show significant shifts of CS [w] away from non-CS [w], as shown in Figure 3. In the female speaker (BIL021), these shifts are significant for both the progressive ( $p < 0.05$ ) and the regressive ( $p < 0.001$ ) case. In the four males (BIL002, BIL005, BIL008, and BIL018), however, the shift only reaches significance under regressive L2-induced influence ( $p < 0.001$  for each), but not under progressive influence (although there is a tendency for the progressive



case in speaker BIL008, with  $p < 0.1$ ). As can be seen in Figure 3, while some individual shifts of [w] are small acoustic shifts, other shifts are more categorical and overlap with [v].

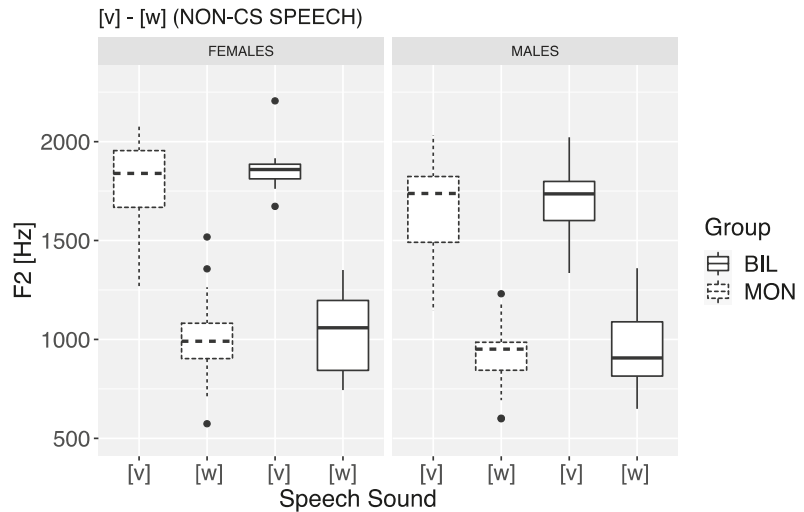


**Figure 2.** F2 of the intended [v] and [w] sounds in English words, spoken by BIL speakers in non-CS (solid) and in CS environments; productions in CS environments are divided into progressive (red) and regressive (blue) L2-induced influences. The left panel shows female data, the right panel male data.



**Figure 3.** F2 of intended [v] and [w] sounds in English words, spoken by 5 (one female, four male) BIL speakers in non-CS (black), and of [w] in CS (colored) environments, with the latter being divided into contexts with possible progressive (red) and regressive (green) L2-induced influences. Only speakers with significant differences of the data, are presented. For the sake of a better presentation of potentially multimodal (and therefore non-normal) distributions of the data, we opted for violin plots instead of boxplots, because a violin plot shows the full distribution of the data as rotated kernel density plots on both sides.

An analysis (LMM 4) of the non-CS [w] and [v] productions of only these five speakers as compared to the MON group with fixed factors *Group* and *Gender* and random factors *Speaker* and *Word* revealed a significant main effect for *Gender* ( $F[1, 11.0] = 9.6, p < 0.05$ ) only (reflecting the sex-related differences in vocal tract sizes and, therefore, formants), but no main effect for group. As an analysis of the Estimated Marginal Means confirmed, there were no significant differences between MON speakers' [v] and BIL speakers' [v] in both males and females; more importantly, the same is true for both speaker groups' [w] sounds (see Figure 4 for descriptive details).



**Figure 4.** F2 of intended [v] and [w] sounds in English words, spoken by the same MON speakers (dashed) that already appeared in Figure 1, and by the five BIL speakers (solid) in non-CS environments, who were shown to have shifted CS [w] away from non-CS [w]. The left panel shows female data, the right panel male data.

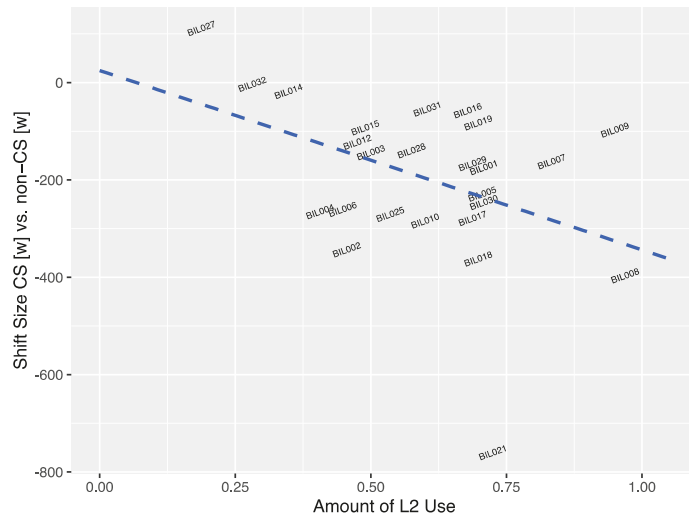
### 3.1.4. Influence of Predictor Variables

As described under Section 2.4, shift sizes of CS [w] away from non-CS [w] were obtained by taking the speaker-specific estimates of the Estimated Marginal Means analysis of LMM 3. The value 0 denotes the position of non-CS [w]; negative numbers denote a shift of CS [w] towards non-CS [v]. Only one of the predictor variables *AoE*, *LoR*, *C+M*, *C-M*, *amount of L1 use*, *amount of L2 use* and *L2 proficiency* turned out to be significantly correlated with shift size: *amount of L2 use* ( $\text{Adjusted } R^2 = 0.13, p < 0.05$ ). As Figure 5 demonstrates, shift size of CS [w] increases with increasing *amount of L2 use*.

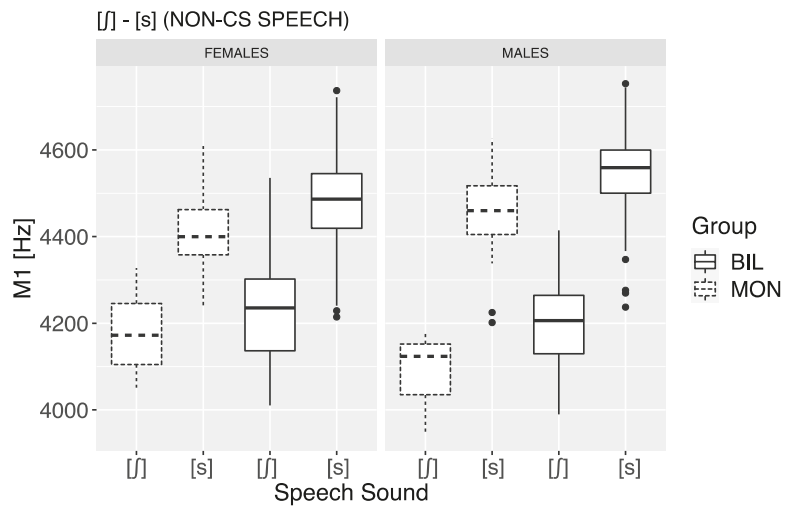
## 3.2. [s] vs. [ʃ]

### 3.2.1. Monolingual vs. Bilingual [s] and [ʃ] in Non-CS Contexts

LMM 1 (see Section 2.4) with *the first spectral moment (M1)* as the dependent variable, *Phonological Category* (2 levels: [s] vs. [ʃ]), *Group*, and *Gender* as fixed factors, and *Speaker* and *Word* as random factors revealed significant main effects for *Phonological Category* ( $F[1, 31.1] = 300.0, p < 0.001$ ) and *Group* ( $F[1, 30.9] = 12.0, p < 0.01$ ), but no effect for *Gender* (see Figure 6). Additionally, there was a significant interaction between *Phonological Category* and *Gender* ( $F[1, 31.1] = 10.9, p < 0.01$ ). Pairwise comparisons by means of the Estimated Marginal Means showed significant differences between [ʃ] and [s] in all possible comparisons ( $p < 0.001$  each). In both male and female speakers, there were no significant differences in either [s] or in [ʃ] between MON and BIL.



**Figure 5.** The shift size of CS [w] vs. non-CS [w] as a function of the predictor variable *amount of L2 use*. Regression line superimposed.

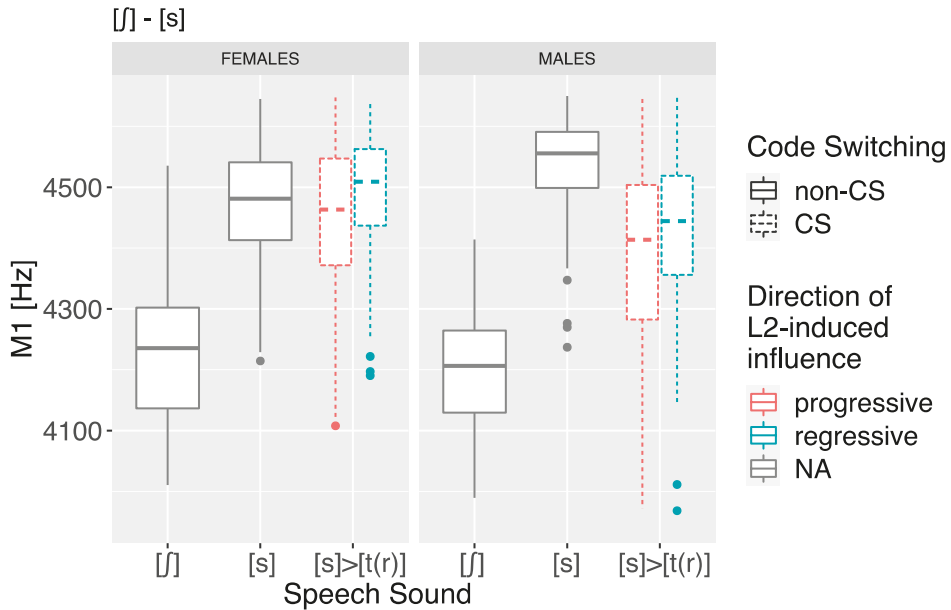


**Figure 6.** The first spectral moment of intended [j] and [s] sounds in English words, spoken by MON speakers (dashed) and BIL speakers (solid) in non-CS speech. The left panel shows female data, the right panel male data.

### 3.2.2. Comparison of Bilinguals’ Productions in Non-CS vs. CS Contexts

LMM 2 (see Section 2.4) with *M1* as the dependent variable, *Phonetic Category* (3 levels: non-CS [j] vs. non-CS [s]<sup>4</sup> vs. CS [s]), *Direction of L2-induced Influence* and *Gender* as fixed factors, and *Speaker* and *Word* as random factors revealed one significant main effect only, namely for *Phonetic Category* ( $F[2, 21.4] = 58.3, p < 0.001$ ), but none for *Direction of L2-induced Influence* or *Gender*. However, there was a significant interaction between *Phonetic Category* and *Gender* ( $F[2, 23.2] = 9.6, p < 0.001$ ). A post hoc analysis with Estimated Marginal Means showed no effects between the CS [s] in progressive vs. regressive environments. As Figure 7 suggests, there are no significant differences between non-CS [s] and CS [s] sounds

in females (under both directions of L2-induced influence). In males, however, CS [s] is significantly shifted away from their non-CS [s] counterparts with  $p < 0.001$  in both cases under both regressive and progressive L2-induced influence. However, the CS [s] does not overlap with non-CS [ʃ] (suggesting a subtle acoustic rather than a categorical shift), showing no evidence of a categorical shift, neither in females nor in males, and irrespective of the direction of the L2 influence.



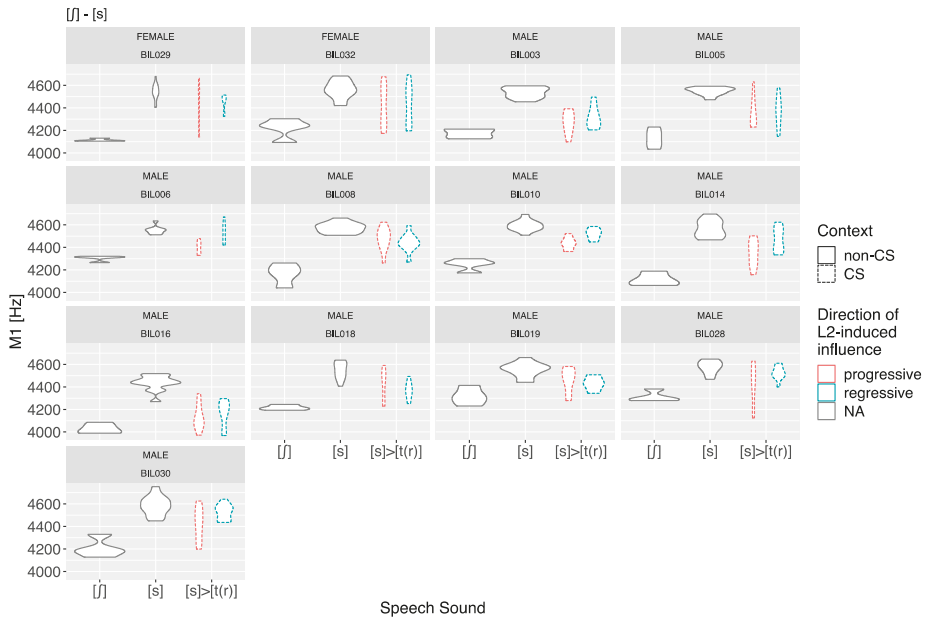
**Figure 7.** The first spectral moment of intended [ʃ] and [s] sounds in English words, spoken by BIL speakers in non-CS (solid) and CS (dashed) speech; productions in CS environments are divided into progressive (red) and regressive (blue) L2-induced influences. The left panel shows female data, the right panel male data.

### 3.2.3. Individual Variation

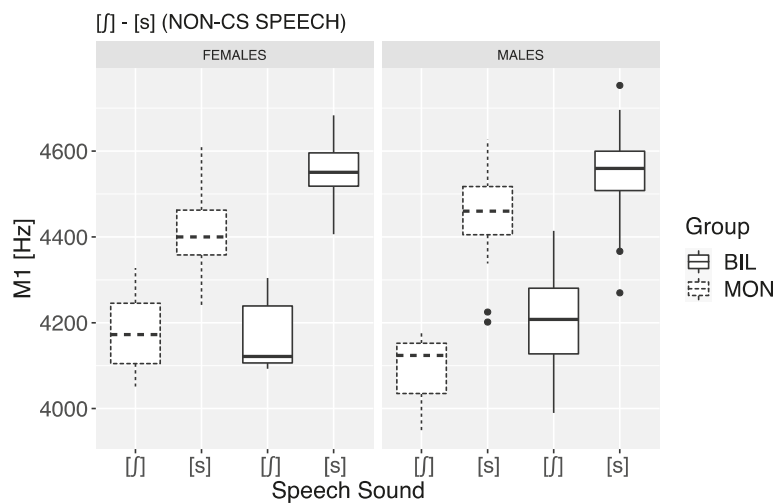
Individual variation was examined by means of LMM 3 (see Section 2.4), which revealed that 2 out of 11 (=18.2%) females showed significant shifts of [s] in CS mode away from their counterparts in non-CS mode. However, out of 14 men, 11 (=78.6%) showed significant shifts in CS [s]. Only these women and men are shown in Figure 8. Of those, six speakers (BIL006, BIL010, BIL014, BIL028, BIL029 and BIL030) only shift CS [s] in progressive, but not regressive, contexts; two speakers (BIL008, BIL019) shift only under regressive, but not progressive, influence away from non-CS [s], towards more [ʃ]-like values. All other speakers (BIL003, BIL005, BIL016, BIL018, BIL032) shift CS [s] significantly away from non-CS [s] under both regressive and progressive German influence; for these, no significant difference was found between the two conditions.

LMM 4 (see Section 2.4), which compared the non-CS [s] and [ʃ] productions of only these 13 speakers as compared to the MON group with fixed factors *Group* and *Gender* and random factors *Speaker* and *Word*, revealed a significant main effect for *Group* ( $F[1, 19.0] = 9.6, p < 0.01$ ), but not for *Gender*, as shown in Figure 9. There are no significant interactions, with the exception of the three-way interaction of *Phonological Category*, *Group*, and *Gender* ( $F[1, 18.9] = 5.2, p < 0.05$ ). The Estimated Marginal Means analysis showed no statistically significant differences in either males or females between the [ʃ] sounds of MON and the BIL speakers. The BIL speakers do, however, differ significantly from the MON speakers for [s] (in both females and males:  $p < 0.01$ ), with the [s] in the BIL shifted away from [s] in MON speakers. Interestingly though, this shift was not towards more

[ʃ]-like values, but rather, in the opposite direction, “overshooting” the [s] values of the MON group.



**Figure 8.** The first spectral moment of intended [ʃ] and [s] sounds in English words, spoken by 13 (2 female, 11 male) of the 25 BIL speakers, in non-CS (black), and of [s] in CS (colored) environments, with the latter being divided into contexts with possible progressive (red) and regressive (green) L2-induced influences.



**Figure 9.** The first spectral moment of intended [ʃ] and [s] sounds in English words in non-CS environments, spoken by the MON speakers (dashed), and by a selection of BIL speakers (solid), who have been shown to have shifted CS [s] away from non-CS [s]. The left panel shows female data, the right panel male data.

### 3.2.4. Influence of Predictor Variables

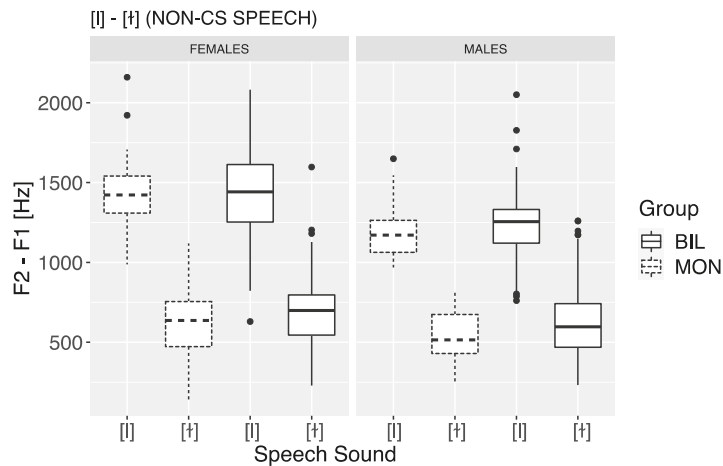
None of the predictor variables *AoE*, *LoR*, *C+M*, *C-M*, *amount of L1 use*, *amount of L2 use* and *L2 proficiency* turned out to be significantly correlated with shift sizes of CS [s].

### 3.3. [l] vs. [ɫ]

As our group analyses for the other two sound pairs showed no effect of the position of the observed L2-induced changes (before or after the German item), we decided to include the sounds [l] and [ɫ], for which we only had sentences for the progressive environment.

#### 3.3.1. Monolingual vs. Bilingual [l] vs. [ɫ] in Non-CS Contexts

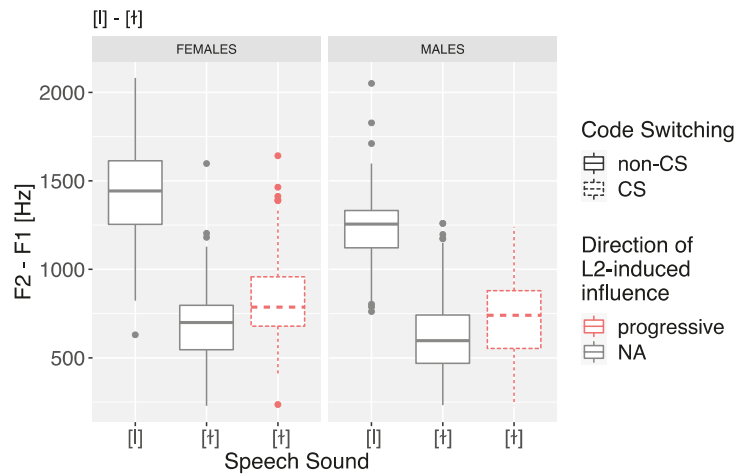
LMM 1 (see Section 2.4) with the distance between F1 and F2 as the dependent variable in a linear mixed model with *Phonological Category* (with 2 levels, [l] vs. [ɫ]), *Group* and *Gender* as fixed factors and *Speaker* and *Word* as random factors revealed a significant main effect each for *Phonological Category* ( $F[1, 9.2] = 106.7, p < 0.001$ ) and for *Gender* ( $F[1, 30.9] = 8.3, p < 0.01$ ), but none for *Group*, as shown in Figure 10. There also was a significant interaction between the factors *Phonological Category* and *Gender* ( $F[1, 30.9] = 4.5, p < 0.05$ ). For both males and females, an Estimated Marginal Means analysis showed that [l] and [ɫ] were well separated in all cases ( $p < 0.001$  in all pairwise comparisons), and that there were no differences between MON and non-CS [l] or between MON and non-CS [ɫ] (cf. Figure 10).



**Figure 10.** The distance between F1 and F2 of intended light [l] and dark [ɫ] sounds in English words, spoken by MON speakers (dashed) and BIL (solid) in non-CS speech. The left panel shows female data, the right panel male data.

#### 3.3.2. Comparison of Bilinguals' Productions in Non-CS vs. CS Contexts

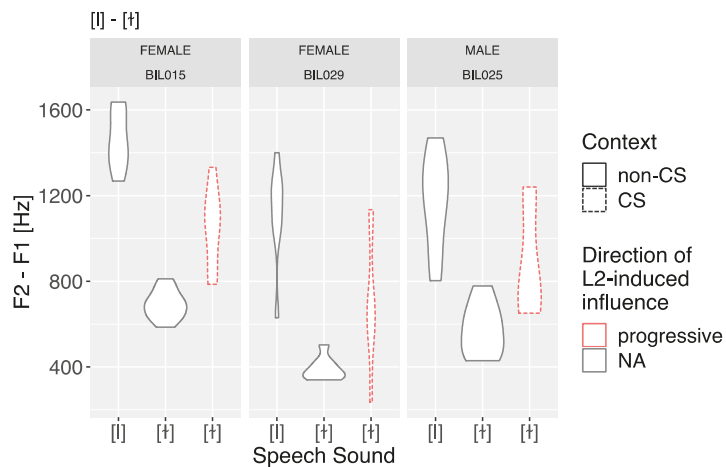
LMM 2 (see Section 2.4) with the *distance between F1 and F2* as the dependent variable, *Phonetic Category* (3 levels: non-CS [l] vs. non-CS [ɫ] vs. CS [ɫ]) and *Gender* as fixed factors, and *Speaker* and *Word* as random factors showed no significant interaction between *Phonetic Category* and *Gender*, but main effects for both factors individually (*Phonetic Category*:  $F[2, 10.8] = 41.0, p < 0.001$ , *Gender*:  $F[1, 22.9] = 5.0, p < 0.05$ ), as shown in Figure 11. An analysis of the Estimated Marginal Means confirmed this: non-CS [l] was always significantly different from both non-CS and CS [ɫ], in both men and women (always  $p < 0.001$ ), but neither showed significant differences between non-CS vs. CS [ɫ]s.



**Figure 11.** The distance between F1 and F2 of intended [l] and [ɫ] sounds in English words, spoken by BIL speakers in non-CS (solid) and CS (dashed) speech; productions in CS environments were only available in progressive (red) L2-induced influence. The left panel shows female data, the right panel male data.

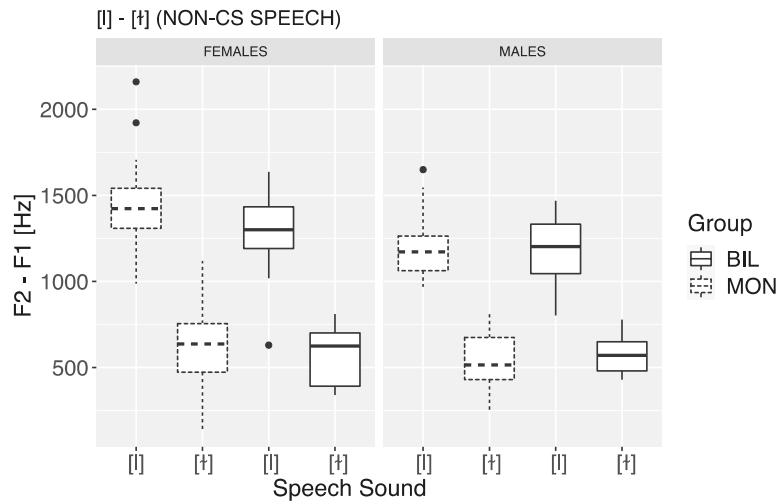
### 3.3.3. Individual Variation

LMM 3 (see Section 2.4) showed that only 3 out of 25 BIL speakers showed significant shifts of CS [ɫ] away from non-CS /l/: 2 females (BIL015:  $p < 0.01$ ; BIL029:  $p < 0.05$ ) and 1 male (BIL025:  $p < 0.05$ ). Figure 12 presents descriptive details of these three speakers.



**Figure 12.** The distance between F1 and F2 of intended [l] and [ɫ] sounds in English words, spoken by 3 (2 female, 1 male) of the 25 BIL speakers, in non-CS (black), and of [ɫ] in CS (colored) environments (only progressive (red) L2-induced influences were available).

As before, we conducted the same analysis as in Section 3.3.1 with these three speakers. The LMM revealed no significant main effects, neither for *Group* nor for *Gender*, and also no significant interactions. This was further confirmed by an Estimated Marginal Means analysis, which showed no significant differences, either between MON [l] and non-CS [l], or between MON-[ɫ] and non-CS [ɫ] (cf. Figure 13), across both males and females.



**Figure 13.** The distance between F1 and F2 of intended [l] and [h] sounds in English words, spoken by the MON speakers (dashed), and by a selection of BIL speakers (solid) in non-CS environments who have been found to have shifted CS [h] away from non-CS [l]. The left panel shows female data, the right panel male data.

### 3.3.4. Influence of Predictor Variables

None of the predictor variables *AoE*, *LoR*, *C+M*, *C−M*, *amount of L1 use*, *amount of L2 use* and *L2 proficiency* were significantly correlated with the shift sizes of CS [h] in our regression analyses.

## 4. Discussion

This paper set out to test the effect of dual language activation on the occurrence of L2-induced changes in the L1 speech of late English–Austrian sequential bilinguals who emigrated to Austria in adulthood. The purpose of the study was threefold. First, it sought to determine whether sounds produced in a code-switched context where both languages are thought to maximally interact (e.g., Green 1998; Olson 2013; Van Hell and Dijkstra 2002) are more prone to L2-induced changes in L1 speech than sounds that are produced in a monolingual (non-code-switched) context. Secondly, where an effect of code-switching was found, we aimed to examine whether it equally affected all the sounds under investigation, and whether the direction of L2-induced influence mattered, i.e., whether influences are more apparent in segments that occur before (regressive influence) or after a switch (progressive influence). Finally, the study sought to determine whether the predictor variables *AoE*, *LoR*, *C+M*, *C−M*, *amount of L1 use*, *amount of L2 use* and *L2 proficiency* could explain any of the observed changes in L1 speech production. In what follows, we will discuss the findings and implications for each of these issues in turn.

### 4.1. The Effect of Dual Language Activation on L2-Induced Changes in L1 Speech

As the BIL participants in our study all immigrated to Austria in adulthood, their L1 speech might have undergone some form of attrition, given that attrition is found to be widespread (de Leeuw 2019b; Mayr et al. 2020) and a considerable number of our participants had lived in Austria for many years. Some changes to their L1 speech may, therefore, have already been present in their productions of monolingual (English) non-code-switched utterances. We therefore first compared the BIL speakers’ production of the three sound pairs under investigation in the monolingual English utterances with the same pairs produced by the MON speakers. Results revealed no significant differences



between the BIL and MON speakers in any of the three sounds that were expected to trigger L2-induced influences (i.e., [w], [s] and [ʃ]). Results also showed that the BIL speakers were able to keep the two members of each sound pair well separated. We then compared the sounds [w], [s] and [ʃ] produced by the BIL speakers in the bilingual (code-switched) utterances with the same sounds and the other member of the sound pair ([v], [ʃ] and [l]), respectively, in order to establish whether sounds produced in a code-switched context are more susceptible to L2-induced changes than those produced in a monolingual (non-code-switched) context, and if so, whether the changes are toward (assimilatory) or away from (dissimilatory) its counterpart. The results showed that the code-switched tokens did differ from the tokens produced in a monolingual context, with [w] and [s] shifted toward (assimilated to) the other member of the sound pair ([v] and [ʃ], respectively), although no difference between the participants' productions was observed across contexts for [ʃ]. This shows that code-switching and its resulting dual language activation leads to an increase in L2-induced changes in L1 speech.<sup>5</sup> Interestingly, although we expected to find more L1 speech changes in the bilingual than in the monolingual context, there were, in fact, no changes at all in the monolingual context in any of the three sounds that were expected to trigger L2-induced transfer. This suggests that the observed L1 speech changes are transient in nature and are reversed when the linguistic environment changes (in this case, from a code-switched to a monolingual context).<sup>6</sup> We therefore prefer not to refer to the observed changes in code-switched speech as L1 attrition but rather as temporary drifts akin to the gestural or phonetic drifts reported in inexperienced L2 learners who are, for a limited time, intensively exposed to the L2, or experience regular changes in their linguistic environment (Chang 2012, 2013; Dmitrieva et al. 2020; Kartushina et al. 2016; Sancier and Fowler 1997; Tobin et al. 2017). These drifts are also reported to be temporary in nature and are fully (Kartushina and Martin 2019) or partially (Chang 2019) reversible. Regularly travelling between their L1- and L2-speaking countries or short-term exposure to the L2 in inexperienced L2 learners who still rely heavily on their L1 will lead to “ad hoc dual language activation” (Mayr et al. 2020, p. 14) where both languages are highly active for a limited period of time, similarly to what happens during code-switching. While dual language activation has led to temporary drifts in L1 speech in a code-switched context, it is certainly possible that they may be a precursor to more persistent changes that may become apparent over time (see also Mayr et al. 2020). After all, the features affected in studies on phonetic drift, such as VOT (e.g., Chang 2012, 2013) and vowel formants (e.g., Kartushina and Martin 2019), are often also the ones that are reported to be affected in studies on long-term phonetic changes to L1 speech (e.g., Bergmann et al. 2016; Kornder and Mennen 2021; Mayr et al. 2012; Stoehr et al. 2017). If this is the case, it may have important implications for research on L1 speech attrition, as studies on code-switched speech may highlight those features that, over time, may lead to more persistent changes in L1 speech, and may thus be able to predict which features are vulnerable to L1 attrition and which may be resistant to change.

#### 4.2. Scale and Direction of L2-Induced Influences

This leads us to the question of whether dual language activation by means of code-switching affected some segments more than others. Our results showed that of the three sounds that were expected to show an influence of the L2, only [w] and [s] were found to have shifted toward the other member of the pair (i.e., in the direction of the inserted German sound) in code-switched utterances, whereas no shift was found for [ʃ]. Just as not all sounds are equally affected by L1 attrition (Bergmann et al. 2016; Mayr et al. 2012; Stoehr et al. 2017), code-switching also does not lead to L2-induced changes in all sounds, with some more affected by code-switching than others. This was also reflected in the individual analyses. None of the individuals who showed an effect of code-switching exhibited shifts in all three sounds, and roughly a quarter (24%) displayed shifts in two of the sounds under investigation (in all cases, [w] and [s]). Most shifts occurred in [s] (52% of participants), followed by [w] (20%) and [ʃ] (8%). This suggests that although all three

sounds may be vulnerable to change, out of the three sounds under investigation, [s] is the most likely to undergo L2-induced changes in a code-switching environment. It should be noted that the tokens for [s] in the code-switched utterances were always adjacent to either /t/ or /tɪ/, whereas due to our restricted set of control materials, [s] in the non-code-switched utterances was only adjacent to /t/ in one word (<study>) and occurred in prevocalic position in the other five words (<seat>, <sick>, <sinking>, <sit>, and <sin>). We therefore cannot entirely exclude the possibility that the reported lowering of the first spectral moment in the code-switched utterances might—in part—be influenced by this imbalance in materials across conditions, particularly given that coarticulatory effects of this kind have been reported (albeit for Australian English varieties) for [s] followed by stop–vowel sequences and stop–rhotic–vowel sequences (Stevens and Harrington 2016). However, as mentioned before (footnote 3), no significant differences were found in MON or BIL speakers between the [s] tokens before /t/ and the prevocalic [s] tokens produced in monolingual utterances, nor in the code-switched tokens with [s] before /t/ and [s] before /tɪ/. Additionally, while in Stevens and Harrington (2016) study, [s] may have retracted in certain contexts, it was never found to overlap with [ʃ]. However, some of the BIL speakers in our study showed an overlap of [s] in their code-switched utterances with [ʃ] in their monolingual utterances (cf. Section 3.2.3). This, and the fact that the other sound pairs (<w> /v-w/ and <l> /l-l/) showed similar patterns of shifts away from one sound towards the other end of the sound pair, makes us confident that we are not reporting mere artifacts.

So how can the differences in observed changes across the three sound pairs be explained? A possible explanation comes from Markedness Theory, according to which sounds that are infrequent in the world's languages, i.e., typologically marked sounds, pose greater articulatory and perceptual difficulties than more frequent, unmarked sounds (Eckman 1977, 1991). Indeed, marked sounds have been shown to be acquired later by children (e.g., Dinnsen et al. 1990; Watts and Rose 2020) and second language learners (Carlisle 1997, 1998). It therefore stands to reason that marked sounds may also be more vulnerable to shifts in code-switched settings than unmarked ones. However, according to the UPSID corpus of 451 languages (Maddieson 1984; Maddieson and Precoda 1990), [ɪ] is much more marked (occurrence: 1.11% of languages) than [i] (occurrence: 38.58% of languages), yet the former shifted the least in the current study. Typological markedness can, therefore, not fully explain the observed patterns.

Alternatively, the frequency with which consonants occur in English may offer an explanation. Thus, one might expect less commonly occurring sounds to be more unstable and hence, more vulnerable to shifts. However, in studies of English consonant phoneme frequency (e.g., Edwards 1992; Wang and Crawford 1960), /s/ is consistently ranked as more frequent than /ʃ/, yet the former was found to shift in the direction of the less frequent [ʃ] in the present study.

Finally, we considered to what extent acoustic distance may be able to account for the observed hierarchy across our three sound pairs. For present purposes, acoustic distance was defined as the mean difference in Hertz in the monolingual control speakers' productions for each sound pair. According to this measure, [s-ʃ] exhibited by far the smallest acoustic difference, i.e., approximately 280 Hz, while that for [v-w] and [l-l] was much greater (approximately 780 Hz and 750 Hz, respectively). Since our findings revealed substantially greater shifts for [s] than [w] and [l], differences in acoustic distance may indeed provide an explanation. In other words, the phonetic proximity of [s] and [ʃ] may have rendered this sound pair more vulnerable to shifts than the two acoustically more distinct sound pairs. As such, the findings of the present study are in line with one of the central tenets of the SLM (Flege 1995; Flege and Bohn 2021): that phonetically similar sounds are more unstable and more likely to be assimilated than dissimilar ones. Future research is needed to explore this issue further and determine the role that the L2 plays in rendering L1 sound contrasts less stable in code-switched settings.

As for the direction of L2-induced influences, based on findings by [Bullock and Toribio \(2009\)](#), it was hypothesized that L2-induced changes during dual language activation may be more common in segments occurring before (regressive) a switch than in those occurring after a switch (progressive L2-influence). This was, however, not confirmed by our data. Our results showed no significant differences in the L2-induced changes occurring before or after the switch. However, a fair amount of individual variation was observed, with individual speakers who had undergone L1 changes, showing shifts either only under regressive influence, or in both regressive and progressive environments. While we are thus unable to draw firm conclusions, overall, our data suggest no privileged direction of L2-induced changes. Future research is needed that explores this issue on a larger scale and using different methodologies to aid our understanding of the cognitive processes that underpin interaction in bilingual sound systems.

#### 4.3. Individual Variation and the Role of Predictor Variables

In general, there was a considerable amount of individual variation, not only in the direction of L2-induced influences, but also between males and females, or in whether or not code-switching resulted in L1 shifts. For instance, while both BIL men and women had a tendency to shift [w] towards [v] in their code-switched tokens, this was only significant in the male participants and only in sounds that occurred before a switch (i.e., under regressive L2 influence). Similarly, while no significant shift was found for [s] in the code-switched speech of BIL females, in male speakers, the [s] had shifted in the direction of [ʃ] under both regressive and progressive influences. For [ɹ], code-switching did not lead to a shift in either males or females. The difference between males and females was also apparent in the individuals who exhibited significant shifts in code-switched speech, where the males clearly outnumbered the females (with 4 males as opposed to 1 female shifting [w] towards [v] and 11 males versus 2 females shifting [s] towards [ʃ]). It may be that adult women (the pattern is less clear in children) are more “experienced” code-switchers, as they are sometimes reported to code-switch more often than men in various social contexts ([Alicea 2001](#); [Hafissatou 2020](#); [Wong 2006](#)). Experienced code-switchers, in turn, have been found to exhibit less short-term cross-language phonetic interaction ([Šimáčková and Podlipský 2015](#)), which may provide a tentative explanation of the observed shifts between males and females in our study. It may be worth exploring the role of gender in future studies.

The variability observed in our study does not come as a surprise, as variability in the extent to which individuals exhibit changes in L1 speech has been widely documented (e.g., [Bergmann et al. 2016](#); [de Leeuw et al. 2012](#); [Major 1992](#); [Mayr et al. 2012](#); [Mennen 2004](#)). While seventeen out of twenty-five participants (68%) in our study showed signs of L1 changes during code-switched speech, the remaining eight (32%) did not. This means that L1 changes in the context of code-switching are not inevitable, and perhaps these speakers were able to suppress the phonetic interaction that typically occurs in code-switched speech where both languages are maximally activated. In an attempt to explain why some speakers in our study exhibited changes while others did not, we examined whether the variables *AoE*, *LoR*, *C+M*, *C−M*, *amount of L1 use*, *amount of L2 use* and *L2 proficiency* could predict the shift size of sounds produced in code-switched compared to monolingual utterances. The only variable that was found to significantly predict this was *amount of L2 use*, and it only did so for the sound [w]. That is, speakers who used their L2 more were found to shift their productions of [w] in code-switched speech toward [v]. While there are no studies on the role of predictor variables in code-switched speech, previous studies on L1 attrition also found an effect of the overall amount of L2 use ([Stoehr et al. 2017](#)). In their study, however, the amount of L2 use was inferred from whether the participants were immersed in an L2-speaking environment (and largely limiting L1 use) or not, rather than factored into a regression analysis. While our study found an effect of the overall amount of L2 use, no effect was found for the overall amount of L1 use. Previous studies on the role of overall amount of L1 use in L2-induced phonetic changes in the L1 have shown

varied effects. While an effect of reduced L1 use was reported in [Stoehr et al. \(2017\)](#); but see our comments above), no effect was found in [Hopp and Schmid \(2013\)](#), and a recent study on phonetic drift also challenges the role of reduced L1 use ([Dmitrieva et al. 2020](#)).

Given that our study tested the shifts in L1 speech during code-switching, where both languages are maximally activated, we expected to find an influence of dual language activation similar to the one reported in [de Leeuw et al. \(2010\)](#) and [Mayr et al. \(2020\)](#). We therefore used predictor variables similar to those in [de Leeuw et al. \(2010\)](#), to test whether the type of contact with the L1 (i.e., in settings where language mixing is likely or unlikely to occur) could predict whether a shift in L1 speech would occur, with the expectation that shifts would be more likely when L1 contact involved language mixing. However, we found no effect of either C+M (L1 contact in settings where language mixing is *likely* to occur) nor C−M (L1 contact in settings where language mixing is *unlikely* to occur). One reason for this difference in findings may be that the questions included in the variables C+M and C−M in our study did not entirely overlap with those included in [de Leeuw et al.'s \(2010\)](#) study. For instance, our C+M variable did not include questions about language used in the church setting, as this was not deemed relevant for our participants. Another reason for this difference in findings may be the fact that [de Leeuw et al. \(2010\)](#) tested the effect of the type of L1 contact on the extent of perceived foreign accent, whereas our study investigated its effect on the produced shifts in L1 sounds. Perceptions of non-nativeness arise from the cumulative effect of a number of characteristics, such as deviations in the realization of vowels, consonants, and prosody ([Jilka 2000](#); [Mennen 2004](#); [Ulbrich and Mennen 2016](#)), and are not based on a single acoustic shift in one sound. It may therefore be more difficult to find a link between acoustic changes and predictor variables, particularly as the shifts in our study were relatively small. In fact, acoustic shifts in L1 sounds are bound to be small given the need for the speaker to maintain sufficient phonetic contrast between sound categories both within and across languages. This therefore restricts the size of shifts that are typically observed, with acoustic values often intermediate between the L1 and L2 norms (e.g., [Flege 1987](#); [de Leeuw et al. 2013](#); [Major 1992](#); [Mayr et al. 2012](#)). Assimilation effects of this kind were also observed in some of the speakers in our study. This was particularly obvious in the productions of [s], which had shifted in the direction of monolingual [ʃ] in the code-switched utterances of thirteen speakers, some of which showed intermediate values, while others had a complete overlap of [s] and [ʃ]. Interestingly, these thirteen speakers also showed a shift in their production of [ʃ] in non-code-switched utterances, overshooting the values of the MON group. This polarization or dissimilation may result from a need to keep phonetic categories distinct: as their [s] has shifted towards [ʃ], the [ʃ] has moved further away—at least in non-code-switched speech—to keep the two categories maximally distinct. Interestingly, we also found evidence of polarization in the [v] produced by the BIL group, which overshot the values of the MON group (with higher F2 values for [v] in the BIL compared to the MON group). While the occurrence of dissimilation in the [s] productions of the individuals discussed above can be explained by the need to keep categories distinct, this explanation cannot account for the polarization of [v], given that the majority of speakers did not show effects of assimilation for [w]. An alternative explanation would be to assume that the dissimilation is instigated by other instances of dissimilation, such that dissimilation of one sound instigates dissimilation at a system-wide level (see [Mayr et al. 2012](#)). However, as the shifts occur in different acoustic parameters and are not observed in all the sounds investigated, this explanation is rather unlikely. In any case, the cross-linguistic interactions between the L1 and L2 system affecting pronunciation are complex and sometimes characterized by unpredictability when the system is reorganizing ([de Bot and Larsen-Freeman 2011](#); [Verspoor et al. 2008](#)), and the ad hoc dual language activation in code-switched speech and its dynamic transient nature might add to this complexity. Further research is needed to fully understand this complexity and how it interacts with predictor variables.

## 5. Conclusions

In conclusion, this study documented dual language activation in the L1 speech of late English–Austrian sequential bilinguals who emigrated to Austria in adulthood. As such, it is one of only few studies to examine the effect of experimentally induced code-switches on the native language of this type of bilingual population, and the first to systematically investigate individual variation and the role of predictor variables in this setting. The results revealed L2-induced shifts in L1 speech production during code-switched contexts, but only for [w] and [s]. An examination of individual variation showed that such shifts are not inevitable though, since nearly a third of participants did not exhibit a difference in their production of the target sounds across contexts. Unlike previous work (Bullock and Toribio 2009), shifts were found to occur both before and after a code-switch, with a range of patterns observed across participants and groups. Finally, our findings indicated that only *amount of L2 use* was a significant predictor, and only in the production of one of the sounds examined, i.e., [w].

Although the study significantly extends our understanding of the role of dual language activation on the L1 speech of late sequential bilinguals, it has a number of limitations that should be considered when planning future research. To begin with, our sample size, whilst in line with much of the existing literature on potential L1 attriters, is modest, limiting the generalizability of our findings. In addition, while we carefully considered the design of our experiments, not all aspects had been fully systematized. Thus, the code-switched sentences for [t] only appeared in a progressive environment. Moreover, the phonetic context for [s] words was not balanced across code-switched and non-code-switched settings, which, in turn, may explain why even MON speakers showed no differences in the first spectral moment across vowel and <st> contexts. Finally, since our results revealed differences in L2-induced changes across the sound pairs included, future work is needed that goes beyond these and systematically investigates how dual language activation affects different segmental and suprasegmental areas of pronunciation, and the extent to which they contribute to listeners' perceptions. Studies of this kind will help us shed new light on the complex, dynamic nature of L1 speech patterns in late sequential bilinguals.

**Author Contributions:** Conceptualization, I.M., R.M., U.R. and S.D.; methodology, I.M., R.M., U.R. and S.D.; formal analysis, U.R.; writing—original draft preparation, I.M. and U.R.; writing—review and editing, I.M., R.M., U.R. and S.D.; visualization, U.R. and S.D.; supervision, I.M.; project administration, I.M.; funding acquisition, I.M. and R.M. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by Austrian Science Fund (FWF), grant number P33007-G.

**Institutional Review Board Statement:** This study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Institutional Ethics Committee of the University of Graz (protocol code GZ. 39/37/63 ex 2019/20; date of approval: 27 February 2020).

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** The data are not publicly available due to ongoing data analyses. Data will be made available upon request from the corresponding authors once all analyses have been completed.

**Acknowledgments:** We would like to thank the Austrian Science Fund (FWF) for their financial support of this research. We would also like to thank Felix Gschier, Kerstin Endes, Sarah Melker, Matthias Wedenig, and Jiaying Li for their help with proof reading, recruitment and/or data collection, and Klaus Jansch for his support with WikiSpeech. Finally, we thank Rebecca Cliff, Kathleen McCarthy, Paul Foulkes, Ghada Khattab and Bettina Beinhoff for their help in accessing participants in the UK.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Notes

- <sup>1</sup> Note that de Leeuw et al. (2010) had additional questions about contact with the native language during church visits. As this did not apply to the majority of our participants, we removed those questions from the survey.
- <sup>2</sup> Sanker et al. (2021) report that acoustic measurements, especially those concerned with spectral properties, are vulnerable to differences in recording equipment and differences in background noise. We dealt with this by using *Speaker* as a random factor in most of our analyses, so that variability in the recording equipment was coupled with the speakers. In group-wise comparisons, the number of speakers in each group will have partially factored out the variability from differences in recording equipment, because any variability observed is likely to be unsystematic. Moreover, in the individual analyses (in Sections 3.1.3, 3.2.3 and 3.3.3), this additional variability is irrelevant given that we compared code-switched versus non-code-switched speech within the same speaker, and thus, in the same recording setting, as both sets were recorded in a single recording session.
- <sup>3</sup> We initially planned to analyze <st> [ʃt-st] and <str> [ʃt(ʁ)-st(ɹ)] as separate sound pairs, because of ongoing sound changes in several English varieties (see, e.g., Rutter (2011), Stevens and Harrington (2016) and Lawrence (2000) for American, Australian, New Zealand and British varieties of English, respectively), in which alveolar fricatives become retracted to more post-alveolar ones especially whenever [s] precedes /t/. This retraction can also be found for British varieties (Cruttenden 2014), especially those spoken in London (Altendorf 2003) and Colchester (Bass 2009). However, as a pairwise comparison of <s(t)> vs. <s(tr)> showed no significant differences, we decided to pool the two sound pairs to increase statistical power.
- <sup>4</sup> Note, however, that CS [s] was always taken from contexts with adjacent /t/ or /tʁ/, whereas non-CS [s] was taken from one word with adjacent /t/ <study>, but otherwise, [s] was adjacent to a vowel (<seat>, <sick>, <sinking>, <sit>, and <sin>). We tested (by means of the same statistics as in Section 3.2.2, but with non-CS [s] subdivided into non-CS /sV/ and non-CS /st/) whether the [s] from a /t/ context in <study> was different from the [s] tokens from vowel contexts, but we found no significant difference, either in females or males. The only difference to the results reported in the main text was that non-CS [s] from <study> in males showed no significant difference to CS [s] in regressive mode, but it differed significantly from CS [s] in progressive mode ( $p < 0.05$ ) (non-CS /sV/ differed from CS [s] in both modes). We therefore decided to keep non-CS /s/ as one group.
- <sup>5</sup> It could be argued that the observed speech changes that occur during code-switched speech may not result from cross-linguistic interaction, but rather from the elicitation of speech production errors, such as those encountered in tongue twisters (e.g., Frisch and Wright 2002; Pouplier and Hardcastle 2005). However, cross-linguistic influences on the phonetic realization of sounds have also been observed in words that are distant from the switch (Piccinini and Arvaniti 2015) and in consecutive interpreting where bilinguals heard a message in their L1 and then, produced the translation in their L2 (Šimáčková and Podlipský 2015). This suggests that dual language activation is the most likely cause of the observed phonetic changes.
- <sup>6</sup> Note though that the assertion that no changes were observed in the non-code-switched target sounds was based on a group comparison of the BIL speakers that were observed to have L2-induced changes in code-switched speech. As we did not run an analysis of individual variation (due to concerns about the effects of extrinsic speaker normalization, cf. Section 2.3), we cannot fully exclude the possibility that some individuals may have shown signs of attrition in their non-code-switched speech.

## References

- Ahn, Sunyoung, Charles B. Chang, Robert DeKeyser, and Sunyoung Lee-Ellis. 2017. Age effects in first language attrition: Speech perception by Korean-English bilinguals. *Language Learning* 67: 694–733. [CrossRef]
- Alharbi, Amira, Anouschka Foltz, and Ineke Mennen. Forthcoming. L2 acquisition and L1 attrition of VOTs of voiceless plosives in highly proficient late bilinguals.
- Alicea, Evelyn. 2001. Gender differences in the use of language switching and the detachment effect in therapy with Latino/Latina, Spanish/English bilingual clients. *Dissertation Abstracts International: Section B: The Sciences and Engineering* 62: 2474.
- Altendorf, Ulrike. 2003. *Estuary English: Levelling at the Interface of RP and South-Eastern British English*. Tübingen: Gunter Narr Verlag.
- Antoniou, Mark, Catherine T. Best, Michael D. Tyler, and Christian Kroos. 2011. Inter-language interference in VOT production by L2-dominant bilinguals: Asymmetries in phonetic code-switching. *Journal of Phonetics* 39: 558–70. [CrossRef]
- Balukas, Colleen, and Christian Koops. 2015. Spanish–English bilingual voice onset time in spontaneous code-switching. *International Journal of Bilingualism* 19: 423–43. [CrossRef]
- Bass, Michael. 2009. ‘Street or shreet?’ Investigating (str-) palatalisation in Colchester English. *Estro: Essex Student Research Online* 1: 10–21.
- Bates, Douglas, Martin Maechler, Ben Bolker, and Steve Walker. 2015. Fitting Linear Mixed Effects Models Using lme4. *Journal of Statistical Software* 67: 1–48. [CrossRef]
- Bergmann, Christopher, Amber Nota, Simone A. Sprenger, and Monika S. Schmid. 2016. L2 immersion causes non-native-like L1 pronunciation in German attriters. *Journal of Phonetics* 58: 71–86. [CrossRef]
- Boersma, Paul, and David Weenink. 2020. Praat: Doing Phonetics by Computer [Computer Program]. Version 6.1.36. Available online: <http://www.praat.org/> (accessed on 12 December 2020).
- Bullock, Barbara E., Almeida J. Toribio, Verónica González, and Amanda Dalola. 2006. Language dominance and performance outcomes in bilingual pronunciation. In *Proceedings of the 8th Generative Approaches to Second Language Acquisition Conference*. Edited by Mary Grantham O’Brien, Christine Shea and John Archibald. Somerville: Cascadilla Press, pp. 9–16.

- Bullock, Barbara E., and Almedia J. Toribio. 2009. Trying to hit a moving target: On the sociophonetics of code-switching. In *Interdisciplinary Approaches to Code-Switching*. Edited by Ludmila Isurin, Donald Winford and Kees de Bot. Amsterdam: John Benjamins, pp. 189–206. [CrossRef]
- Carlisle, Robert S. 1997. The modification of onsets in a markedness relationship: Testing the Interlanguage Structural Conformity Hypothesis. *Language Learning* 47: 327–61. [CrossRef]
- Carlisle, Robert S. 1998. The acquisition of onsets in a markedness relationship: A longitudinal study. *Studies in Second Language Acquisition* 20: 245–60. [CrossRef]
- Carter, Paul G. 2002. Structured Variation in British English Liquids: The Role of Resonance. Ph.D. thesis, University of York, York, UK. Available online: <http://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.274527> (accessed on 7 April 2021).
- Celata, Chiara, and Jessica Cancila. 2010. Phonological attrition and the perception of geminate consonants in the Lucchese community of San Francisco (CA). *International Journal of Bilingualism* 14: 185–209. [CrossRef]
- Chang, Charles B. 2012. Rapid and Multifaceted Effects of Second-Language Learning on First-Language Speech Production. *Journal of Phonetics* 40: 249–68. [CrossRef]
- Chang, Charles B. 2013. A Novelty Effect in Phonetic Drift of the Native Language. *Journal of Phonetics* 41: 520–33. [CrossRef]
- Chang, Charles B. 2019. Language Change and Linguistic Inquiry in a World of Multicompetence: Sustained Phonetic Drift and Its Implications for Behavioral Linguistic Research. *Journal of Phonetics* 74: 96–113. [CrossRef]
- Childers, Donald G., ed. 1978. *Modern Spectrum Analysis*. Piscataway: IEEE Computer Society Press, pp. 252–55.
- Cho, Mi-Hui, and Shinsook Lee. 2016. The impact of different L1 and L2 learning experience in the acquisition of L1 phonological processes. *Language Sciences* 56: 30–44. [CrossRef]
- Cruttenden, Alan. 2014. *Gimson's Pronunciation of English*, 8th ed. Oxford and New York: Routledge. [CrossRef]
- de Bot, Kees, and Diane Larsen-Freeman. 2011. Researching Second Language Development from a Dynamic Systems Theory Perspective. In *A Dynamic Approach to Second Language Development: Methods and Techniques*. Edited by Marjolijn Verspoor, Kees de Bot and Wander Lowie. Amsterdam: John Benjamins, pp. 5–24. [CrossRef]
- de Leeuw, Esther, Aurela Tusha, and Monika S. Schmid. 2018a. Individual phonological attrition in Albanian-English late bilinguals. *Bilingualism: Language and Cognition* 21: 278–95. [CrossRef]
- de Leeuw, Esther, Aurela Tusha, Hui Zhao, Kyle Helke, and Alice Greenfield. 2018b. A case of extreme phonetic attrition in the German rhotic. In *Mind Matters in SLA*. Edited by Clare Wright, Thorsten Piske and Martha Young-Scholten. Bristol: Multilingual Matters, pp. 162–83. [CrossRef]
- de Leeuw, Esther, Ineke Mennen, and James M. Scobbie. 2012. Singing a different tune in your native language: First language attrition of prosody. *International Journal of Bilingualism* 16: 101–16. [CrossRef]
- de Leeuw, Esther, Ineke Mennen, and James M. Scobbie. 2013. Dynamic systems, maturational constraints and L1 phonetic attrition. *International Journal of Bilingualism* 17: 683–700. [CrossRef]
- de Leeuw, Esther, Monika S. Schmid, and Ineke Mennen. 2010. The effects of contact on native language pronunciation in an L2 migrant context. *Bilingualism: Language and Cognition* 13: 33–40. [CrossRef]
- de Leeuw, Esther. 2019a. Native speech plasticity in the German-English late bilingual Stefanie Graf: A longitudinal study over four decades. *Journal of Phonetics* 73: 24–39. [CrossRef]
- de Leeuw, Esther. 2019b. Phonetic Attrition. In *The Oxford Handbook of Language Attrition*. Edited by Monika S. Schmid and Barbara Köpke. Oxford: Oxford University Press, pp. 204–17. [CrossRef]
- Dinnsen, Daniel A., Steven B. Chin, Mary Elbert, and Thomas W. Powell. 1990. Some constraints on functionally disordered phonologies: Phonetic inventories and phonotactics. *Journal of Speech and Hearing Disorders* 33: 28–37. [CrossRef] [PubMed]
- Dmitrieva, Olga, Allard Jongman, and Joan A. Sereno. 2010. Phonological neutralization by native and non-native speakers: The case of Russian final devoicing. *Journal of Phonetics* 38: 483–92. [CrossRef]
- Dmitrieva, Olga, Allard Jongman, and Joan A. Sereno. 2020. The effect of instructed second language learning on the acoustic properties of first language speech. *Languages* 5: 44. [CrossRef]
- Dmitrieva, Olga. 2019. Transferring perceptual cue-weighting from second language into first language: Cues to voicing in Russian speakers of English. *Journal of Phonetics* 73: 128–43. [CrossRef]
- Draxler, Christoph, and Klaus Jänsch. 2008. WikiSpeech—A Content Management System for Speech Databases. Paper presented at Interspeech 2008, the Ninth Annual Conference of the International Speech Communication Association, Brisbane, Australia, September 22–26; pp. 1646–49.
- Eckman, Fred R. 1977. Markedness and the contrastive analysis hypothesis. *Language Learning* 27: 315–30. [CrossRef]
- Eckman, Fred R. 1991. The Structural Conformity Hypothesis and the acquisition of consonant clusters in the interlanguage of ESL learners. *Studies in Second Language Acquisition* 13: 23–41. [CrossRef]
- Edwards, Harold T. 1992. *Applied phonetics: The sounds of American English*. San Diego: Singular.
- Flege, James E. 1995. Second language speech learning: Theory, findings, and problems. In *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues*. Edited by Winifred Strange. Timonium: York Press, pp. 233–77.
- Flege, James Emil, and James Hillenbrand. 1984. Limits on phonetic accuracy in foreign language speech production. *The Journal of the Acoustical Society of America* 76: 708–21. [CrossRef]
- Flege, James Emil, and Ocke-Schwen Bohn. 2021. The Revised Speech Learning Model (SLM-r). In *Second Language Speech Learning: Theoretical and Empirical Progress*. Edited by Ratreay Wayland. Cambridge: Cambridge University Press, pp. 3–83. [CrossRef]

- Flege, James Emil, and Wieke Eefting. 1987. Cross-language switching in stop consonant perception and production by Dutch speakers of English. *Speech Communication* 6: 185–202. [CrossRef]
- Flege, James Emil. 1987. The production of “new” and “similar” phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics* 15: 47–65. [CrossRef]
- Forrest, Karen, Gary Weismer, Paul Milenkovic, and Ronald Dougall. 1988. Statistical analysis of word-initial voiceless obstruents: Preliminary data. *Journal of the Acoustical Society of America* 84: 115–24. [CrossRef] [PubMed]
- Frisch, Stefan A., and Richard Wright. 2002. The phonetics of phonological speech errors: An acoustic analysis of slips of the tongue. *Journal of Phonetics* 30: 139–62. [CrossRef]
- Gargiulo, Chiara, and Mechteld Tronnier. 2020. First language attrition on prosody in a foreign language environment: A speech production study on anaphora resolution. *Journal of Monolingual and Bilingual Speech* 2: 219–44. [CrossRef]
- Goldrick, Matthew, Elin Runnqvist, and Albert Costa. 2014. Language switching makes pronunciation less native-like. *Psychological Science* 25: 1031–36. [CrossRef] [PubMed]
- Green, David W. 1998. Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition* 1: 67–81. [CrossRef]
- Grosjean, François, and Joanne L. Miller. 1994. Going in and out of languages: An example of bilingual flexibility. *Psychological Science* 5: 201–6. [CrossRef]
- Grosjean, François. 2001. The bilingual’s language modes. In *One Mind, Two Languages: Bilingual Language Processing*. Edited by Janet L. Nicol. Oxford: Blackwell, pp. 1–22.
- Guion, Susan G. 2003. The vowel systems of Quichua-Spanish bilinguals: Age of acquisition effects on the mutual influence of the first and second languages. *Phonetica* 60: 98–128. [CrossRef]
- Hafissatou, Kane. 2020. Language Variation: A Case Study of Gender Differences in Wolof-French Codeswitching. *International Journal of Language and Linguistics* 8: 122–27. [CrossRef]
- Haley, Katarina L., Elizabeth Seelinger, Kerry Callahan Mandulak, and David J. Zajac. 2010. Evaluating the spectral distinction between sibilant fricatives through a speaker-centered approach. *Journal of Phonetics* 38: 548–54. [CrossRef] [PubMed]
- Hickey, Raymond. 2020. Persistent features in the English of German speakers. In *English in the German-Speaking World*. Edited by Raymond Hickey. Cambridge: Cambridge University Press, pp. 208–28. [CrossRef]
- Hopp, Holger, and Monika S. Schmid. 2013. Perceived foreign accent in first language attrition and second language acquisition: The impact of age of acquisition and bilingualism. *Applied Psycholinguistics* 34: 361–94. [CrossRef]
- Jilka, Matthias. 2000. The Contribution of Intonation to the Perception of Foreign Accent. Ph.D. thesis, University of Stuttgart, Stuttgart, Germany.
- Kartushina, Natalia, Alexis Hervais-Adelman, Ulrich H. Frauenfelder, and Narly Golestani. 2016. Mutual influences between native and non-native vowels in production: Evidence from short-term visual articulatory feedback training. *Journal of Phonetics* 57: 21–39. [CrossRef]
- Kartushina, Natalia, and Clara D. Martin. 2019. Third-language learning affects bilinguals’ production in both their native languages: A longitudinal study of dynamic changes in L1, L2 and L3 vowel production. *Journal of Phonetics* 77: 100920. [CrossRef]
- Kisler, Thomas, Uwe D. Reichel, and Florian Schiel. 2017. Multilingual processing of speech via web services. *Computer Speech & Language* 45: 326–47. [CrossRef]
- Kopečková, Romana, Ulrike Gut, and Christina Golin. 2019. Acquisition of the /v/-/w/contrast by L1 German children and adults. Paper presented at 19th International Congress of Phonetic Sciences, Melbourne, Australia, August 5–9; pp. 3745–49.
- Köpke, Barbara, and Monika S. Schmid. 2004. First language attrition: The next phase. In *First Language Attrition: Interdisciplinary Perspectives on Methodological Issues*. Edited by Monika S. Schmid, Barbara Köpke, Merel Keijzer and Lina Weilemar. Amsterdam: John Benjamins, pp. 1–43. [CrossRef]
- Kornder, Lisa, and Ineke Mennen. 2021. Longitudinal developments in bilingual second language acquisition and first language attrition of speech: The case of Arnold Schwarzenegger. *Languages* 6: 61. [CrossRef]
- Kufner, Herbert L. 1971. *Kontrastive Phonologie Englisch-Deutsch*. Stuttgart: Klett.
- Kuznetsova, Alexandra, Per Brockhoff, and Rune Christensen. 2017. lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software* 82: 1–26. [CrossRef]
- Lawrence, Wayne. 2000. /str/->/tr/: Assimilation at a distance? *American Speech* 75: 82–87. [CrossRef]
- Lehiste, Ilse. 1962. *Acoustical Characteristics of Selected English Consonants*. The Hague: Mouton.
- Lenth, Russell V. 2021. emmeans: Estimated Marginal Means, aka Least-Squares Means. R Package Version 1.5.4. Available online: <https://CRAN.R-project.org/package=emmeans> (accessed on 23 November 2020).
- López, Verónica González. 2012. Spanish and English word-initial voiceless stop production in code-switched vs. monolingual structures. *Second Language Research* 28: 243–63. [CrossRef]
- Maddieson, Ian, and Kristin Precoda. 1990. Updating UPSID. *UCLA Working Papers in Phonetics* 74: 104–14. [CrossRef]
- Maddieson, Ian. 1984. *Patterns of Sounds*. Cambridge: Cambridge University Press.
- Major, Roy C. 1992. Losing English as a first language. *The Modern Language Journal* 76: 190–208. [CrossRef]
- Major, Roy C. 2010. First language attrition in foreign accent perception. *International Journal of Bilingualism* 14: 163–83. [CrossRef]
- Mayr, Robert, David Sánchez, and Ineke Mennen. 2020. Does Teaching Your Native Language Abroad Increase L1 Attrition of Speech? The Case of Spaniards in the United Kingdom. *Languages* 5: 41. [CrossRef]



- Mayr, Robert, Sasha Price, and Ineke Mennen. 2012. First language attrition in the speech of Dutch-English bilinguals: The case of monozygotic twin sisters. *Bilingualism: Language and Cognition* 15: 687–700. [CrossRef]
- Mennen, Ineke, and Denise Chousi. 2018. Prosody in first-generation adult immigrants and second-generation heritage-language users: The timing of prenuclear rising accents. In *Proceedings of the 9th Speech Prosody Conference*. Edited by Katarzyna Klessa, Jolanta Bachan, Agnieszka Wagner, Maciej Karpinski and Daniel Śledziński. Poznan: University of Poznan, pp. 828–32. [CrossRef]
- Mennen, Ineke. 2004. Bi-directional interference in the intonation of Dutch speakers of Greek. *Journal of Phonetics* 32: 543–63. [CrossRef]
- Moosmüller, Sylvia, Carolin Schmid, and Julia Brandstätter. 2015. Illustrations of the IPA: Standard Austrian German. *Journal of the International Phonetic Association* 45: 339–48. [CrossRef]
- Moulton, William G. 1962. *The Sounds of English and German*. Chicago: University of Chicago Press.
- Muldner, Kasia, Leah Hoiting, Leyna Sanger, Lev Blumenfeld, and Ida Toivonen. 2019. The phonetics of code-switched vowels. *International Journal of Bilingualism* 23: 37–52. [CrossRef]
- Olson, Daniel J. 2012. The phonetics of insertional code-switching: Suprasegmental analysis and a case for hyper-articulation. *Linguistic Approaches to Bilingualism* 2: 443–57. [CrossRef]
- Olson, Daniel J. 2013. Bilingual language switching and selection at the phonetic level: Asymmetrical transfer in VOT production. *Journal of Phonetics* 41: 407–20. [CrossRef]
- Piccinini, Page, and Amalia Arvaniti. 2015. Voice onset time in Spanish-English spontaneous code-switching. *Journal of Phonetics* 52: 121–37. [CrossRef]
- Poupier, Marianne, and William Hardcastle. 2005. A re-evaluation of the nature of speech errors in normal and disordered speakers. *Phonetica* 62: 227–43. [CrossRef]
- R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing, Available online: <https://www.R-project.org/> (accessed on 17 November 2020).
- Rutter, Ben. 2011. Acoustic analysis of a sound change in progress: The consonant cluster /stɪ/ in English. *Journal of the International Phonetic Association* 41: 27–40. [CrossRef]
- Sancier, Michele L., and Carol A. Fowler. 1997. Gestural Drift in a Bilingual Speaker of Brazilian Portuguese and English. *Journal of Phonetics* 25: 421–36. [CrossRef]
- Sanker, Chelsea, Sarah Babinski, Lyn Burns, Marisha Evans, Juhya Kim, Slater Smith, Natalie Weber, and Claire Bowern. 2021. (Don't) Try This at Home! The Effects of Recording Devices and Software on Phonetic Analysis. Lingbuzz: Lingbuzz/005748. Available online: <https://lingbuzz.net/lingbuzz/005748> (accessed on 10 February 2020).
- Schmid, Monika S. 2002. *First Language Attrition, Use and Maintenance: The Case of German Jews in Anglophone Countries*. Amsterdam: John Benjamin. [CrossRef]
- Schmid, Monika S. 2011. *Language Attrition*. New York: Cambridge University Press. [CrossRef]
- Šimáčková, Šárka, and Václav J. Podlipský. 2015. Immediate phonetic interference in code-switching and interpreting. Paper presented at 18th International Congress of Phonetic Sciences, Glasgow, UK, August 10–14.
- Stevens, Mary, and Jonathan Harrington. 2016. The phonetic origins of /s/-retraction: Acoustic and perceptual evidence from Australian English. *Journal of Phonetics* 58: 118–34. [CrossRef]
- Stoehr, Antje, Titia Benders, Janet G. van Hell, and Paula Fikkert. 2017. Second language attainment and first language attrition: The case of VOT in immersed Dutch–German late bilinguals. *Second Language Research* 33: 483–518. [CrossRef]
- Tobin, Stephen J., Hosung Nam, and Carol A. Fowler. 2017. Phonetic Drift in Spanish-English Bilinguals: Experiment and a Self-Organizing Model. *Journal of Phonetics* 65: 45–59. [CrossRef]
- Ulbrich, Christiane, and Ineke Mennen. 2016. When prosody kicks in: The intricate interplay between segments and prosody in perceptions of foreign accent. *International Journal of Bilingualism* 20: 522–49. [CrossRef]
- Ulbrich, Christiane, and Mikhail Ordin. 2014. Can L2-English influence L1-German? The case of post-vocalic /r/. *Journal of Phonetics* 45: 26–42. [CrossRef]
- Van Hell, Janet G., and Ton Dijkstra. 2002. Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic Bulletin and Review* 9: 780–89. [CrossRef]
- Verspoor, Marjolijn, Wander Lowie, and Marijn van Dijk. 2008. Variability in Second Language Development from a Dynamic Systems Perspective. *The Modern Language Journal* 92: 214–31. [CrossRef]
- Wang, William S.-Y., and John Crawford. 1960. Frequency studies of English consonants. *Language and Speech* 3: 131–39. [CrossRef]
- Watts, Erica, and Yvan Rose. 2020. Markedness and implicational relationships in phonological development: A cross-linguistic investigation. *International Journal of Speech-Language Pathology* 22: 669–82. [CrossRef] [PubMed]
- Winkelmann, Raphael, Jonathan Harrington, and Klaus Jänsch. 2017. EMU-SDMS: Advanced speech database management and analysis in R. *Computer Speech & Language* 45: 392–410. [CrossRef]
- Winkelmann, Raphael, Klaus Jaensch, Steve Cassidy, and Jonathan Harrington. 2020. emuR: Main Package of the EMU Speech Database Management System. R Package Version 2.1.1. Available online: <https://cran.r-project.org/web/packages/emuR/index.html> (accessed on 2 October 2020).
- Wong, Kwok-Lan Jamie. 2006. Gender and Codemixing in Hong Kong. Honours thesis, The University of Sydney, Sydney, Australia. Available online: <https://ses.library.usyd.edu.au/bitstream/handle/2123/1726/wong-jamie-thesis.pdf?sequence=2&isAllowed=y> (accessed on 19 June 2021).

Article

# Foreign-Language Phonetic Development Leads to First-Language Phonetic Drift: Plosive Consonants in Native Portuguese Speakers Learning English as a Foreign Language in Brazil

Denise M. Osborne<sup>1</sup> and Miquel Simonet<sup>2,\*</sup>

<sup>1</sup> Department of Languages, Literatures, & Cultures, State University of New York at Albany, Albany, NY 12222, USA; dmosborne@albany.edu

<sup>2</sup> Program in Second Language Acquisition and Teaching, University of Arizona, Tucson, AZ 85721, USA

\* Correspondence: simonet@arizona.edu



**Citation:** Osborne, Denise M., and Miquel Simonet. 2021.

Foreign-Language Phonetic Development Leads to First-Language Phonetic Drift: Plosive Consonants in Native Portuguese Speakers Learning English as a Foreign Language in Brazil. *Languages* 6: 112. <https://doi.org/10.3390/languages6030112>

Academic Editors: Robert Mayr and Jonathan Morris

Received: 23 April 2021

Accepted: 22 June 2021

Published: 25 June 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Abstract:** Fifty-six Portuguese speakers born and raised in Brazil produced Portuguese words beginning in one of four plosives, /p b k g/. Twenty-eight of them were monolinguals (controls), and the rest were learners of English as a foreign language (EFL). The learners were also asked to produce English words beginning with one of four plosives, /p b k g/. We measured the plosives' *voice onset times* (VOT) to address the following research questions: Do foreign-language learners, whose exposure to native English oral input is necessarily limited, form new sound categories specific to their additional language? Does engaging in the learning of a foreign language affect the phonetics of one's native language? The EFL learners were found to differ from the controls in their production of Portuguese voiced (but not voiceless) plosives—prevoicing was longer in learner speech. The learners displayed different VOT targets for voiced (but not voiceless) consonants as a function of the language they were speaking—prevoicing was longer in Portuguese. In EFL learners' productions, English sounds appear to be fundamentally modeled on phonologically similar native sounds, but some phonetic development (or reorganization) is found. Phonetic development induced by foreign-language learning may lead to a minor reconfiguration of the phonetics of native language sounds. EFL learners may find it challenging to learn the pronunciation patterns of English, likely due to the reduced access to native oral input.

**Keywords:** second language acquisition; phonetics; VOT; Portuguese; English

## 1. Introduction

When acquiring a second language (L2), most of us find it difficult to learn its pronunciation. Most people who have learned a L2 have “an accent” in their L2—the pronunciation patterns of L2 learners typically differ from those of monolingual speakers of the L2 (Colantoni et al. 2015; Piske et al. 2001; Simonet 2016; Chang 2019; Wayland 2021). This seems to be true also of many bilinguals who learned their L2 as children (early sequential bilinguals), people who live in bilingual societies (language-contact communities), and people who migrated to a foreign land where the L2 is spoken. Whereas, learning a L2 early and using it often diminishes the saliency of one's accent, it does not always eliminate it (Piske et al. 2001). Perhaps surprisingly, some bilinguals develop “an accent” in their first or native language (L1)—they have pronunciation patterns in their L1 that differ from those of monolingual speakers of the L1 (Kartushina et al. 2016b). In extreme cases, the effects of the L2 on the L1 lead to a phenomenon known as L1 attrition, which we, following others, define as the reduction or decrease of one's fluency and proficiency in a language, the loss of skill (Schmid 2011, pp. 11–17; Köpke and Schmid 2004, p. 5). Some scholars have concluded that the two languages of bilinguals coexist in a common representational

network, thus influencing each other (Grosjean 1989). This has been hypothesized for phonological and phonetic knowledge, as well (Flege and Bohn 2021; Flege 1995).

Much research on L2 pronunciation development (and its limits) focuses on bilinguals immersed in their L2 (Flege 2007, 2018; Flege and Bohn 2021), and research on the potential effects of L2 learning in L1 pronunciation (or L1 phonetic drift) is mostly concerned with bilinguals who seem to be dominant in their L2 and who, in some cases, experience a reduced L1 use (Kartushina et al. 2016b; Hopp and Schmid 2013; de Leeuw et al. 2010; Major 1992; de Leeuw et al. 2018). However, many people who study a L2 do so in communities where the L2 is not commonly spoken, perhaps in their own home country—the L2 is thus a *foreign* language for them. In such cases, learners continue to be immersed in their L1 and use it daily, and they seldom use and practice their L2. Often, learners' experience with their L2 is limited to the classroom setting. The population the present study investigates consists of native speakers of Portuguese—born, raised, and residing in Brazil—learning English as a foreign language (EFL). These learners rarely interact with native English speakers, and they began learning English as adults.

Our study addresses the following research questions: Do foreign-language learners develop pronunciation patterns that resemble those of native speakers of the foreign language? Or, more narrowly (and technically), do foreign-language learners, whose exposure to native oral input in their L2 is very limited, form new sound categories specific to the L2 that approximate target sounds as produced by native speakers of the L2? Moreover, does engaging in the learning of a foreign language affect the phonetics of one's native language? Our study is concerned with foreign-language phonetic learning and native language phonetic drift. We further ask whether these two phenomena are connected.

## 1.1. Review of the Literature

### 1.1.1. L2 Phonetic Development

L2 phonetic development may be investigated in a variety of ways. One of them consists of assessing the strength of “nonnative accent” in L2 learners, as judged by a panel of native-speaking listeners of the learners' target language. Research on the phenomenon of “nonnative accent” has identified several factors that appear to modulate the degree to which L2 learners' pronunciation approximates that of native speakers of the L2, and these include L2 age of acquisition, L1/L2 use, length of L2 experience, motivation, formal instruction, and individual language-learning aptitude (Piske et al. 2001). This research suggests that people who began learning their L2 as adults, particularly if they continue to use their L1 often, are unlikely to acquire the pronunciation patterns of the L2 in a way that closely approximates native speaker norms (e.g., Flege et al. 1997). A second way in which L2 phonetic development has been investigated is by comparing speech samples produced by L2 learners and monolingual controls by means of acoustic analysis, and this research also suggests that late L2 learners are rather unlikely to produce speech samples that do not systematically differ from those of monolingual controls (e.g., Flege 1991). Native-likeness may be unlikely for adult L2 learners, but this does not mean that L2 phonetic development is impossible (e.g., Flege et al. 1995). Rather than asking the extent to which L2-learner speech samples differ from those of monolinguals of the L2, one could ask to what extent learners' L2 samples differ from those they produce in their L1. From this perspective, L2 learning has to do with their having formed sound categories specific to the sounds of their L2—i.e., separate from their L1 categories—even when such new categories may not be identical to those of native speakers of the L2 (Casillas and Simonet 2018, p. 63). This is the perspective we take in the present study. Our study asks whether late L2 learners who seldom use their L2 develop phonetic categories specific to their L2, and we address this question by comparing speech samples across L2 learners' two languages.

Much research on L2 speech development is concerned with populations fully immersed in the L2, such as migrants (Tsukada et al. 2004, 2005; Flege et al. 2003, 2006; MacKay et al. 2001). Many adult learners, however, study their L2 in their home country, typically in classroom settings. Such learners tend to remain dominant in their L1, use their

L1 much more often than they use their L2, and their experience with their L2 is limited to the classroom. Relatively few studies have examined L2 phonetic development in such populations, comparing L1 and L2 productions in a within-speaker design (Dmitrieva et al. 2020; Solon 2016, among others). For instance, Solon (2016) analyzed the production of Spanish /l/ by native English speakers studying Spanish as a L2 in the United States (US). English /l/ is darker (i.e., more pharyngealized) than Spanish /l/, particularly in syllable-coda position. In fact, English has two allophones for /l/ in complementary distribution, one used in syllable onset position and one in syllable-coda position. In a cross-sectional study, Solon (2016) found that learners, as they became more proficient in Spanish, tended to approximate native-Spanish phonetic norms and to reduce the acoustic difference between syllable-onset and coda allophones, which they seemed to transfer in the earlier stages of learning. Interestingly, even the most novice learners in Solon's sample seemed to differentiate between their L2 and the L1 /l/, thus displaying evidence of L2 phonetic development. Dmitrieva et al. (2020) investigated the production of obstruents by native English speakers learning Russian as a L2 in the US. The learners in the Dmitrieva et al. sample also produced slightly different phonetic categories for the obstruents in their L1 and L2.

These findings suggest that L2 phonetic development is possible in cases in which one would think it unlikely. L2 learners may form new (i.e., L2-specific) phonetic categories for their L2 sounds. This does not mean that the new categories closely resemble those of native speakers of the target language. How exceptional are the findings discussed above? In the present study, we replicate and extend these findings with a different, but comparable, population of foreign-language learners.

### 1.1.2. L1 Phonetic Drift

A growing body of research has documented the existence of differences between monolingual and bilingual speech attributed to the influence of the L2 phonetic system on the L1 (Stoehr et al. 2017; Chang 2012; Mayr et al. 2020; Kartushina et al. 2016a; Flege and Eefting 1987; Flege 1987; Guion 2003; Major 1992; Sancier and Fowler 1997; Mora and Nadeu 2012; de Leeuw et al. 2010; Bergmann et al. 2016; Ulbrich and Ordin 2014; Fowler et al. 2008). The phonetic influence of the L2 on the L1 has been termed *L1 drift* (Dmitrieva et al. 2020; Sancier and Fowler 1997; Mayr et al. 2020; Chang 2013), and this is the term we use here.

Some bilinguals become dominant in their L2 and seldom use their L1. For instance, migrants may fully immerse themselves in the culture of their L2 after leaving their home country and moving to a L2-speaking country. For some migrants, this may lead to a phenomenon known as L1 attrition, the reduction or decrease of one's fluency and proficiency in their L1. The phonetic consequences of L1 attrition have been documented in a number of studies (Flege 1987; Major 1992; de Leeuw et al. 2010, 2018; Hopp and Schmid 2013), but L1 attrition is not the focus of our investigation. It has been shown that L1 phonetic drift may be observed even in the absence of L1 attrition—that is, in bilinguals who continue to use their L1 often and remain very fluent in it (Dmitrieva et al. 2020; Sancier and Fowler 1997; Chang 2012; Mayr et al. 2020). This type of drift is what we are interested in: An effect of the phonetic patterns of L2 on those of the L1 that does not come about as a result of a reduction or decrease of a learner's fluency and proficiency in their L1. The literature, however, has not consistently distinguished between L1 attrition and L1 drift (in the absence of attrition) (Schmid 2011, pp. 11–17).

In a review of the literature, Kartushina et al. (2016b) identified several factors that appear to modulate the nature and size of L1 drift. The chief among them is the age of L2 acquisition, and a close second is the L2/L1 use. According to this review, late L2 learners are less likely than early learners to experience L1 drift. The later in life a L2 is learned, the less likely it is to influence the L1. In late learners, L1 drift, when found, tends to be the result of full interlingual equivalence classification and assimilatory in nature, to use the terminology of the Speech Learning Model (SLM) (Flege 1995; Flege and Bohn 2021). For

instance, late learners may produce a single phonetic category for L1 and L2 sounds that, in the speech of monolingual speakers of each language, are similar but not identical. This merged category, therefore, differs both from the one used by monolingual speakers of the L2 and the one used by monolingual speakers of the L1. An intergroup difference in this direction—one in which bilinguals produce the sounds of two languages as more similar to each other (or even fully merged, identical) than those produced by monolingual speakers of those languages—is evidence of assimilatory L1 phonetic drift (Fowler et al. 2008; Flege 1987). In early learners, on the other hand, L1 phonetic drift is more likely to be the result of new-category formation and dissimilatory in nature. For instance, early learners may produce two quite different phonetic categories for L1 and L2 sounds that, in monolingual speech, are only slightly different, thus magnifying such interlingual phonetic difference in their pronunciation. This sort of intercategory deflection tends to affect the L2 sound more than the corresponding L1 sound (Flege et al. 2003), but may affect both (Mora and Nadeu 2012). These may be general tendencies, but there is no principled reason to expect that assimilatory drift is specific to late learners and dissimilatory drift to early learners. The general observation, at any rate, is that the age of L2 acquisition is associated with the L1 phonetic drift. The second factor that modulates interlingual phonetic interactions is the L1/L2 use. People who use their L2 much more often than their L1 may, in time, become dominant in their L2, and this may lead to their developing “an accent” in their L1, modifying their L1 pronunciation patterns (Mayr et al. 2020; Hopp and Schmid 2013; Major 1992; de Leeuw et al. 2010). Other factors may include speech register, L2 proficiency and experience, cognate status (Amengual 2012), and the most recent linguistic environment in which bilinguals have been immersed (Sancier and Fowler 1997; Simonet 2014; Simonet and Amengual 2019).

A small body of research suggests that the L1 drift may be found even in late L2 learners who continue to use their L1 often and, in some cases, seldom use their L2 (Sancier and Fowler 1997; Dmitrieva et al. 2020; Kartushina et al. 2016b; Chang 2012). Two studies are particularly relevant here since they focus on populations like ours, classroom learners in a foreign-language setting. Chang (2012) found evidence of L1 drift in L1 English learners of Korean who were taking a 6-week Korean language course in Korea, and Dmitrieva et al. (2020) found it in L1 English learners of Russian who were studying their L2 in North America. Dmitrieva et al. (2020) examined a constellation of acoustic correlates of plosive voicing, both of plosives in word-initial and in word-final position. For word-initial plosives, VOT was analyzed, and, for word-final plosives, the authors measured preceding vowel duration, stop closures, friction, and the duration of the voicing period during closure. Chang (2012) examined the acoustics of both plosives and vowels. Regarding the plosives, both VOT and  $f_0$  at onset were measured; for the study of vowel timbre, both  $F_1$  and  $F_2$  were analyzed. Whereas all the participants in Chang’s study were novice L2 learners, the Dmitrieva et al. speakers varied in proficiency between relatively novice to intermediate. In addition, whereas Chang’s participants were learning their L2 in the country where the L2 is spoken (and thus were likely to be exposed to their L2 outside the classroom), the Dmitrieva et al. speakers were learning their L2 in their home country (and were rarely exposed to their L2 outside the classroom). Given the factors that seem to modulate L1 drift (Kartushina et al. 2016b), including age of acquisition and use, one would not have readily anticipated that the populations investigated in Chang (2012) and Dmitrieva et al. (2020) would show evidence of L1 drift, but they did. How exceptional are these findings? Can we replicate them by investigating comparable populations?

### 1.1.3. The Plosives of Portuguese and English

Regarding plosive consonants, the phonemic inventories of Portuguese and English are identical. First, both languages contrast a set of phonologically voiced stops with one of phonologically voiceless stops. Second, both languages contrast three sets of plosives varying in place of articulation: Velars, coronals, and bilabials. In sum, both Portuguese and English have /p t k b d g/. However, the phonetic substance of these phonemes

differs between the two languages. To explain some of these differences, we focus on a single phonetic feature (or acoustic metric), *voice onset time* (VOT), and on how this feature is manifested in utterance-initial position. VOT is an acoustic feature that measures the asynchrony between two acoustic landmarks that correspond to two articulatory events involved in the production of plosives: Articulatory release and the onset of vocal fold vibration (Lisker and Abramson 1964, 1967; Abramson and Whalen 2017). VOT varies as a function of phonological voicing, such that voiced and voiceless consonants differ in terms of VOT patterns (among other features), but it is also affected by place of articulation (Cho and Ladefoged 1999).

In utterance-initial position, the phonologically voiced plosives of Portuguese present a period of prevoicing (i.e., the onset of voicing precedes articulatory release), whereas phonologically voiceless consonants present a brief voiceless period following articulatory release and no voicing during closure (i.e., the onset of voicing follows articulatory release, but such voicing lag is brief) (Lousada et al. 2010; Major 1987; Sancier and Fowler 1997). That said, at least one study reports to have found aspirated voiceless plosives in Brazilian Portuguese—voiceless plosives whose voicing lag is unusually (variably) long (Alves et al. 2008). Portuguese, therefore, is a “true voicing” language (Kirby and Ladd 2016; Beckman et al. 2013), that is, a language that contrasts plosives that present voicing during articulatory closure with plosives that do not (Lousada et al. 2010; Sancier and Fowler 1997; Major 1987, 1992). English, on the other hand, is an “aspirating” language (Beckman et al. 2013). In utterance-initial position, phonologically voiced English plosives present a brief period of devoicing and lack voicing during closure (i.e., the onset of modal voicing follows articulatory release by a few milliseconds), and phonologically voiceless plosives have a relatively long period of aspiration, also lacking voicing during closure (i.e., the voicing lag period is very long since the devoiced burst is followed by a period of voiceless aspiration) (Lisker and Abramson 1967). In English, the phonological contrast between voiced and voiceless plosives is phonetically implemented with the presence (vs. absence) of aspiration rather than voicing. In order to make the association between the phonetics and phonology more transparent, some scholars have proposed that the “voicing” contrast in English does not actually involve phonological voicing (i.e., a [voice] distinctive feature) but aspiration (i.e., a [spread glottis] distinctive feature) (Beckman et al. 2011, 2013). We would thus say that Portuguese is a [voice] language and English is a [spread glottis] language.

A small body of literature has explored VOT in the speech productions of Portuguese–English bilinguals (Major 1987, 1992; Sancier and Fowler 1997). Major (1987) found that native Portuguese speakers learning English as a foreign language in Brazil produced English /p t k/ with a much shorter voicing-lag period than native English-speaking controls. Major (1992), on the other hand, analyzed the Portuguese and English productions of a group of native English speakers fully immersed in Brazilian culture, having resided in Brazil between one and three decades. Major’s findings showed that some of the bilinguals produced Portuguese /p t k/ with longer voicing-lag periods than native Portuguese controls and English /p t k/ with shorter voicing-lag periods than monolingual English controls. This is an example of assimilatory L1/L2 influence leading to L1 drift. Finally, Sancier and Fowler (1997) investigated both the English and Portuguese productions of a single L1 Portuguese speaker who learned English as a L2, a very proficient L2 learner. The speaker was recorded in both languages in three different occasions: After 4 months in the US, immediately after a 2.5-month stay in Brazil, and once again after 4 months in the US. The speaker produced English /p k/ with a much longer voicing-lag period than Portuguese /p k/, which showed that the speaker had formed L2-specific VOT categories. Moreover, there was some systematic variation between the recording sessions, such that voicing-lag periods were longer (in both languages) after 4 months in the US than after 2.5 months in Brazil. The latter suggests that recent phonetic exposure may serve to recalibrate VOT targets. The present study is concerned with the Portuguese and English productions of a group of L1 Portuguese learners of English as a foreign language who remain immersed in their L1 and have never travelled to an English-speaking country. Our

population is, therefore, the same population investigated in Major (1987), but different from the populations investigated in the other two studies (Sancier and Fowler 1997; Major 1992).

#### 1.1.4. The Learning of English in Brazil

In public and private schools in Brazil, English has been a compulsory subject from the 6th grade to the final year of high school since January of 2020. These changes were implemented according to proposals made by the new *Base Nacional Comum Curricular* (BNCC, [basenacionalcomum.mec.gov.br](http://basenacionalcomum.mec.gov.br), accessed on 22 June 2021), a normative document from the Education Ministry of Brazil that defines the essential disciplines taught in K-12. The BNCC had previously established that schools should include at least one compulsory foreign language in their curriculum (Law n. 9.394 of December 1996), but the guidelines did not establish which foreign language should be taught. In 2005, a change in the BNCC (Law n. 11.161) turned Spanish into the compulsory foreign language to be taught in public schools.

Even though English is currently the most frequently taught foreign language in public and private Brazilian schools, its instruction and learning have faced several challenges. Some of the challenges identified in the literature concern the large classes, a lack of resources, teachers with insufficient training and proficiency in English, and the fact that instruction is primarily restricted to grammar (Santos 2011). The first national-level research on the teaching and learning of English was conducted only recently (British Council 2019). This study was intended to provide baselines for compulsory English instruction in Brazil. The findings of the study suggested that the challenges reported in previous studies (e.g., Santos 2011) were shared across the country. The study highlighted two main obstacles for the teaching of English in public schools: The lack of teachers with specialized training, and the lack of a curriculum that focused on the social use of the language. It revealed that about half of the English teachers who taught in public schools did not hold a degree in English language or the teaching of English; 81% of the English teachers in the study complained about the lack or the unsuitability of textbooks and course materials; and less than one fourth of the classrooms had access to the internet.

Along with the challenges for effective English teaching in public schools, there is a general belief that the learning of English in public schools is ineffective and that, if someone wants to learn English, they have to study in a language school (Silva 2004). Language schools are private schools, they are believed to be better equipped, to have fewer students in the classroom, better trained instructors, and more resources (Polidório 2014). This belief is shared among students and teachers.

The data of the present study were collected at one of the branches of *Cultura Inglesa*, ([www.culturainglesa.com.br](http://www.culturainglesa.com.br) accessed on 22 June 2021) a language school franchise that works in partnership with the British Council. *Cultura Inglesa* opened its first school in 1934 in Rio de Janeiro, the capital of Brazil at the time (Tavares 2018). It is currently one of the most popular language schools in Brazil, with more than 70 branches across the country. The school offers its students exchange programs and a variety of English proficiency exams published by Cambridge English Qualifications (e.g., FCE-First Certificate in English). Their English teachers are consistently engaged in training, courses, and seminars—both in Brazil and abroad. Students attend two 80-min classes per week, each class has a relatively small number of students, and classrooms are equipped with computers and projectors. *Cultura Inglesa* uses the communicative approach with no translation to Portuguese and uses books from international publishers.

#### 1.2. The Current Study

We analyze the production of voiced and voiceless plosives in two languages, Portuguese and English. Native Portuguese speakers were recruited for our production study. Some were monolingual Portuguese speakers, and others were learning English as a foreign language (EFL) in a private language school in Brazil. The monolinguals were recorded

only in their native language and served as controls; the EFL learners were recorded in both Portuguese (L1) and English (L2).

Firstly, we ask whether EFL learners differ from monolingual Portuguese speakers in their production of Portuguese plosives, with a focus on VOT. If the experience of actively learning a L2 leads to modifications in L1 pronunciation patterns (L1 phonetic drift), we would find that learners and monolinguals differ in the way they pronounce Portuguese sounds. If, on the other hand, L2 learning (in this population of Brazilian EFL learners) does not lead to L1 phonetic drift, we would not find any significant differences between the two groups of native Portuguese speakers. Our hypothesis, given the findings in the contextual literature, was that the Brazilian EFL learners in our study would be unlikely to display any effects of L1 phonetic drift, since they seldom use their L2, continue to be immersed in their L1, and rarely interact with native English speakers.

Secondly, we ask whether EFL learners develop new phonetic categories specific to their L2, English. To address this question, we compare the VOTs of Portuguese and English plosives produced by the EFL learners. Do these learners use the same VOT categories for their two languages or different ones? If the learners transferred the phonetic categories of the L1 into their L2 and had failed to develop new VOT targets for their L2, we would find that they produced a single prevoiced VOT category for voiced plosives and a single short-lag VOT category for voiceless plosives in both languages. Note that we do not ask whether the EFL learners pronounce English sounds in the same way native English speakers do (Major 1987), but, rather, whether their English VOT categories differ from their own Portuguese VOT categories. The focus, therefore, is on new-category formation (Flege 1995; Flege and Bohn 2021), not natelikeness. Our hypothesis, given the findings in the literature, was that the Brazilian EFL learners would fundamentally transfer their native categories to their L2 and would be unlikely to have developed phonetic categories specific to their L2. In sum, our working hypotheses were the null hypotheses.

## 2. Method

### 2.1. Sample

A sample of 56 adults participated in an elicited production study. All of the participants were native speakers of Portuguese born and raised in Brazil. The participants were divided into two groups according to whether they were learning English as a foreign language (EFL) or not. Twenty-eight participants were EFL learners and, at the time of the study, were enrolled in English classes in a private language school. This was our experimental group. The remaining 28 participants did not consider themselves learners of English, at least not at (or before) the time of the study. This was our control group. The difference between the two groups is that between a group of emerging bilinguals (i.e., learners actively engaged in the task of studying a foreign language in school) and one of functional monolinguals.

The control group consisted of adults raised as monolingual speakers of Portuguese ( $N = 28$ ). All of the participants in this group were born and raised in the state of Minas Gerais, most of them in Araxá. Other cities of origin included Bambuí, Campos Altos, Frutal, Ibiá, Perdizes, São Gotardo, Três Marias, Uberaba, and Uberlândia. They were recruited in a variety of locations around the city of Araxá, mostly in a university setting. The median age of the control group was 29 years old, with 18 and 55 being the minimum and maximum ages, respectively. Twenty-one of the members of this group were women, and seven were men. Nine of the participants had obtained a postgraduate degree, 12 had graduated from college, and seven had a high school diploma. None of the members of this group reported having had any significant exposure to English—they had never studied English or traveled to an English-speaking country. Some had studied a Romance language—Spanish, mostly—for a few months or up to a year. Immediately after their participation in the production study, they were asked if they were able to produce a full sentence in English (any sentence). None were able to do so.



The target, experimental group consisted of adults raised as monolingual Portuguese speakers ( $N = 28$ ). All of the members of this group were born and raised in the state of Minas Gerais, most of them in Araxá. Other cities of origin included Belo Horizonte, Ibiá, Ponto Nova, São Gotardo, São João del Rei, and Uberaba. The median age was 26, the minimum age was 18, and the maximum was 50. Eighteen of them were women and 10 were men. Two participants in this group had obtained a postgraduate degree, 19 had graduated from college, and seven had a high school diploma. The participants in this group were enrolled as EFL students in the Araxá branch of *Cultura Inglesa* ([culturainglesaaraxa.com.br](http://culturainglesaaraxa.com.br) accessed on 22 June 2021). The EFL learners in our sample were recruited and tested in the language school.

In addition to English, some of the EFL learners reported having studied some Spanish or Italian, but none had been studying any language (besides English) in the months preceding the study. We collected additional relevant data from the participants in the EFL group, most of the data pertained to their experience as EFL learners and their English proficiency. We asked them whether they had ever had a native English speaker as a teacher—fifteen of them (54%) had—and how long had they been taking English classes at *Cultura Inglesa*—this ranged from one to 21 years, with 4 years being the median age (the 25th percentile was two and the 75th percentile was seven). None of the participants had ever visited any English-speaking country.

A survey was administered to all EFL learners. The survey asked participants to rate their English proficiency on a 1 (poor) to 7 (excellent) scale in each of the following skills: Grammar, listening, pronunciation, reading, speaking, and vocabulary. The mean score for all these skills was 4.2 ( $SD = 1.2$ ). The seven skills were highly correlated with each other ( $r = 0.5\text{--}0.8$ ), and a reliability analysis yielded a high score: Cronbach's  $\alpha = 0.93$ . The survey also asked their estimated percentage of English use in the following environments: With friends, at home, on the internet, in the media (including music and television), in online classes, at school, and at work. The overall percentage of use of English, averaged over all settings and learners, was 38% ( $SD = 13.5\%$ ). Percentage scores were generally not correlated with each other, and they varied from a high percentage of use of English at the language school (82%) to a low use with friends (16%). At work, the mean percentage was 41.2%, and this was the setting that induced the largest variance in the sample ( $SD = 38\%$ ). The survey also asked participants to rate their motivation to learn English on a 1 (disagree) to 7 (agree) scale in response to the following prompts: *I am learning English because it will help me find a better job* ( $M = 6.1$ ), *I am learning English because I love English or American culture* ( $M = 5.6$ ); *Learning English makes me feel important* ( $M = 5.3$ ); *I do not ever want to stop learning English* ( $M = 6.4$ ); *I like learning English* ( $M = 6.5$ ); *When I speak English, I do my best to avoid using Portuguese* ( $M = 5.6$ ); *When I speak English I try to imitate the English or American accent* ( $M = 5.6$ ). Unsurprisingly, the motivation questions did not reliably measure the same construct, Cronbach's  $\alpha = 0.78$ . Overall, the survey suggests that the participants in our sample were highly motivated to learn English and that their use of English was mostly restricted to the language school. In terms of their self-assessed proficiency, the participants rated themselves, on average, as intermediate learners, but a variety of proficiency levels is represented in the sample.

Finally, general English proficiency was assessed by means of a brief cloze test focusing on grammar and vocabulary (Brown 2002; Tremblay 2011), the St. George's International English Placement Test ([stgeorges.co.uk/online-english/online-english-test](http://stgeorges.co.uk/online-english/online-english-test) accessed on 22 June 2021). This test is comprised of 40 individual sentences, all of them very brief, out of which one word has been substituted by a blank. Four options are given to test takers to fill in each blank—this is a multiple-choice test. Out of 40 points, our participant sample obtained a median score of 26.5 ( $SD = 7.9$ ). The minimum score was 10 and the maximum was 38; the 25th percentile was 19.75 and the 75th percentile was 31.25. The self-assessed proficiency scores were positively correlated with the results of the cloze test,  $r = 0.5$ , 95% CI [0.15, 0.73],  $p = 0.007$ . This further suggests that we were able to recruit learners from a range of English proficiency levels, from relative novice to advanced learners.

## 2.2. Instrument

The main experiment was an elicited production task with both auditory and visual prompts presented simultaneously. For each target word—that is, in each experimental trial—participants heard an acoustic stimulus (or auditory prompt) and they saw on a computer screen both a written rendering of the target word (or orthographic prompt) and a drawing that represented the word meaning (or figure prompt). In other words, the participants simultaneously received three types of prompts to elicit their production of each target word. For instance, for the Portuguese word *pato* “duck,” the participants heard a recording by a native speaker of the utterance *pato é a palavra* “duck is the word,” played over headphones. Simultaneously, the computer screen showed an orthographic rendering of the target word, <pato>, and a line drawing representing the bird. Line drawings were creative commons figures or are in the public domain, they were all outline drawings in black and white. The simultaneous presentation of prompts in three modes was done to ensure that all EFL learners, including the beginners, had the best chance to recognize the English word they were being asked to produce (this method was probably redundant in the Portuguese production task).

The control group of Portuguese monolinguals were asked to produce only the Portuguese words, and the EFL learners were asked to produce both the Portuguese (L1) and English (L2) words in two separate sessions. The present study focuses on the production of bilabial and velar plosives. Dental (or alveolar) plosives were not included since, in Brazilian Portuguese, they are known to be pronounced as postalveolar affricates when followed by high front vowels (Barbosa and Albano 2004; Albano 2001, pp. 68–86). This study is concerned with the voicing contrast as manifested in VOT. In sum, we examine both voiced and voiceless plosives in two places of articulation, bilabial and velar.

In our materials, the target plosives appeared always in word- and utterance-initial position to be able to reliably measure prevoicing. For each of the four plosives (and each of the two languages), we chose 20 words that began with that plosive. For half of those 20 words, the target plosive was followed by a low vowel; the other half was followed by a high vowel, either front or back. We were not interested in assessing the potential role of contextual vowels (Lousada et al. 2010; Nearey and Rochet 1994; Yavaş and Wildermuth 2006), but we included such variation for the sake of generalizability to all vowel contexts. We obtained a balanced number of observations of the four target consonants, /p b k g/. All in all, we manipulated phoneme, vowel context, and language, and we controlled for utterance and word position (initial). Target words were placed in a constant carrier phrase: *\_\_ é a palavra* (Portuguese), *\_\_ is the word* (English). When possible, we used minimal pairs contrasting in voicing, such as *pond-bond* (English) and *panda-banda* (Portuguese), in both languages. Most English words were monosyllabic and most Portuguese words were disyllabic. In the Portuguese disyllabic words, lexical stress occurred in word-initial position. These design principles resulted in a list of materials comprising 80 words per language: 20 (lexical items) × 4 (phonemes). The list of target words is found in Table 1.

The auditory stimuli were recorded from one male talker of each language. The talkers were asked to read out loud a list of utterances. The utterances were comprised of the target word in a constant carrier phrase: *\_\_ é a palavra* (Portuguese), *\_\_ is the word* (English). The talkers were also asked to record the question *Qual é a palavra?* (Portuguese) or *What is the word?* (English) at the end of the recording session. To record the auditory stimuli, the talkers sat inside a sound-attenuated booth on the campus of the University of [Removed for Review]. The stimuli were recorded with a Fostex DC-R302 digital recorder and a Shure SM10A head-worn dynamic microphone. The signal was digitized at 44.1 kHz and 16-bit quantization. The talkers read the entire list of materials in their native language three times. One rendering of each target item was selected to be used as auditory stimuli. The sound files were normalized for peak intensity at 75 dB. The talker who produced the English materials was a native speaker of English born and raised in [Removed for Review]. When he was recorded, he was 22 years old and did not speak any language other than English. The talker who produced the Portuguese materials was a native speaker of

Portuguese born and raised in the city of São Paulo, Brazil. At the time of the study, he was living in [Removed for Review]. An exchange student at the University of [Removed for Review], the Portuguese talker had been in the US for 8 months when he was recorded. He assessed himself as being an intermediate English learner.

**Table 1.** Word lists produced by participants.

Phoneme	Portuguese	English
/p/	<i>panda, pata, pato, pano, paca, pala, palha, pasta, papa, passo, pia, pica, picho, pilha, pinga, pico, pingo, pulo, puxa, puxo.</i>	<i>pond, pack, par, part, pan, pat, park, pox, pot, path, pea, pin, peep, peach, pig, pit, pull, Pete, put, push.</i>
/b/	<i>banda, bata, bato, banho, baga, bala, balha, basta, baba, baço, Bia, bica, bicho, bilha, binga, bico, bingo, burro, bucha, bucho.</i>	<i>bond, back, bar, barb, ban, bat, bark, box, bot, bath, bee, bin, beep, beach, big, bit, bull, beat, book, bush.</i>
/k/	<i>calo, cato, cala, cama, cata, cana, canso, castro, case, cabo, quinto, quilha, quina, quica, quincha, quita, Quito, cuspa, cudo, cura.</i>	<i>cop, car, cot, cod, card, cap, cat, carry, cash, cab, could, kill, kilt, kick, curl, kit, kiss, cool, keys, coo.</i>
/g/	<i>galo, gato, gala, gama, gata, gana, ganso, gastro, gaze, gabo, guincho, guilha, guina, guiga, guincha, guitar, Guido, gume, Guto, gula.</i>	<i>goth, garb, got, god, guard, gap, gas, Gary, gash, gab, good, gill, guilt, gig, girl, git, gift, goose, geese, goo.</i>

### 2.3. Procedure

Speech productions were elicited, as explained above, by three types of simultaneous prompts: An auditory prompt (a recording by a native speaker of the language), a written prompt (an orthographic rendering of the word), and a figure prompt (a conceptual representation of the word in the form of a line drawing). Each trial consisted of the simultaneous presentation of the three prompts followed by a recording of the question *Qual é a palavra?* (Portuguese) or *What is the word* (English). The question was to be followed by the participant’s production of the target utterance. A new trial began every 6 (Portuguese) or 7 s (English). The timing of the trials was determined arbitrarily. There were 80 trials per language, presented in random order within each session.

The Portuguese controls provided speech samples only in their native language and thus participated in a single experimental session. The participants in this group were recruited by a native speaker of Portuguese born and raised in Araxá, the first author. The researcher asked several background questions, such as city of origin and age, and then administered the production experiment. All conversations between the researcher and the participants took place in Portuguese. The EFL learners provided speech samples in both their L1 (Portuguese) and their L2 (English), and thus they participated in two experimental sessions. They were recruited by a native speaker of English born and raised in the state of New York, a research confederate, who visited the language school and invited students to participate. All conversations between the confederate and the participants took place in English. This presumably encouraged the participants to situate themselves in English mode for the English session. The confederate asked the participants a list of background questions—including the proficiency, use, and motivation questions reported above—and then asked them to take the English cloze task. Finally, the confederate administered the elicited production task.

When the participants were done with the English portion of the study, they were approached by the first author, a native Portuguese speaker. She invited them to participate in an “additional study on Portuguese,” the second experimental session. They were invited to stay in the room or to return after a brief break.

All conversations between the first author of the study and the participants took place in Portuguese to encourage them to switch to their native language mode. In both sessions, the randomized presentation of the prompts was managed by a stimulus presentation software, PsychoPy2 (Peirce 2007; Peirce et al. 2019). The survey and cloze test data were collected in paper format.

#### 2.4. Data and Analyses

The acoustic metric of choice in this study is VOT. Since we obtained samples of both English and Portuguese voiced and voiceless plosives, one would expect to find the full range of values, from prevoicing (negative VOT) to aspiration (long lag VOT). The segmentation of the acoustic material was done by hand by the first author, who utilized both waveform and spectrographic displays to locate the acoustic landmarks of interest. Segmentation was done in Praat (Boersma 2001). For each target consonant, the first author placed a mark at the onset of the burst that corresponded to the release of articulatory closure and another one at the onset of modal voicing. Both acoustic landmarks were adjusted so that they occurred at upwards zero-crossings in the waveform. If the onset of modal voicing precedes the burst, VOT is a negative value, and this indicates the presence of prevoicing. If, on the other hand, the onset of modal voicing follows the burst, VOT is a positive value, and this indicates voicing lag. A very long lag is suggestive of the presence of aspiration.

The data set had a theoretical ceiling of 6720 observations. Each participant in the Portuguese control group produced 80 tokens, all of them in Portuguese:  $80$  (words)  $\times$   $28$  (speakers) = 2240 tokens. Each participant in the EFL learners group produced 80 Portuguese tokens and 80 English tokens:  $80$  (words)  $\times$   $2$  (languages)  $\times$   $28$  (speakers) = 4480 tokens. The actual data set comprised 6714 observations, as six tokens were either discarded due to the presence of noise in the recording or simply not recorded due to experimental error, such as a trial for which the participant did not produce a response.

Inferential statistics were conducted on by-speaker averages. We calculated the average VOT per speaker per condition. We call this metric *mVOT* for mean VOT. This resulted in three data sets, which were then combined into larger data sets to conduct a variety of statistical comparisons across speaker groups and conditions. The first data set comprised the Portuguese control data: It included four average values per participant, one per plosive /p b k g/:  $28$  (speakers)  $\times$   $4$  (phonemes) = 112 observations. The second data set comprised the Portuguese (L1) productions of the EFL learners: It included four average values per participant,  $28$  (speakers)  $\times$   $4$  (phonemes) = 112 observations. The third data set comprised the English (L2) productions of the EFL learners,  $28$  (speakers)  $\times$   $4$  (phonemes) = 112 observations. Each of these values is an average over 20 observations. In sum, the data set comprising 6714 raw VOT measurements was reduced, by means of by-speaker and by-condition averaging, to 336 observations. Data reduction and wrangling were done with an R script (R Core Team 2018), with the functions provided by the package *tidyverse* (Wickham et al. 2019). Data analyses were conducted in *Jamovi* (The Jamovi Project 2020), a free open-source GUI for R. The R packages used in *Jamovi* were *afex* (Singman et al. 2020), *emmeans* (Lenth 2018), and *esci* (version 0.9.1 for Jamovi, written by Robert J. Calin-Jageman). See [jamovi.org/library](http://jamovi.org/library) (accessed on 22 June 2021) for a list of available modules. Synthetic data and code, including the *Jamovi* files, may be made available to readers interested in reproducing our analyses. Readers may contact the corresponding author.

### 3. Results

This section first reports on the results concerning the VOT values of Portuguese plosives as produced by our two groups of Portuguese native speakers, the EFL learners and the controls. This is a between-group comparison that keeps the language constant, Portuguese. Secondly, we report on a statistical comparison of the VOT values of both the Portuguese (L1) and English (L2) plosives produced by the EFL learners. This is a within-group comparison of plosives in two languages.

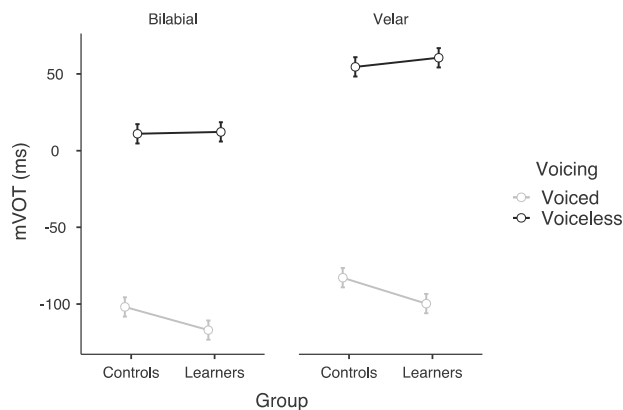
#### 3.1. Portuguese Productions: Between-Subjects Comparison

The Portuguese productions of both groups of native speakers, the controls and the EFL learners, were compared against each other. The dependent variable was *mVOT* (ms), and the factors were *place* (bilabial, velar), *voicing* (voiced, voiceless), and *group* (controls,

EFL learners). The descriptive statistics (mean and standard deviation,  $M$  (SD)) were as follows. For /b/,  $mVOT$  values for the controls were  $-112$  (22) ms, and learners' values were  $-117$  (25) ms. For /g/, the mean for the controls was  $-83$  (22), and the learners' mean was  $-100$  (18). For /p/, the controls' mean was 11 (6) and learners' mean was 12 (6). For /k/, the mean for the controls was 55 (11), and learners' mean was 61 (10).

The data were submitted to a mixed-design ANOVA with *place* and *voicing* as within-subject factors and *group* as a between-subjects factor. This is a  $(2) \times (2) \times 2$  design. The  $\alpha$  criterion was set at 0.05. The ANOVA yielded main effects of both *voicing*,  $F(1,54) = 2373$ ,  $p < 0.0001$ ,  $\eta^2_G = 0.94$ , and *place*,  $F(1,54) = 484$ ,  $p < 0.0001$ ,  $\eta^2_G = 0.48$ . There was a marginally significant effect of *group*,  $F(1,54) = 4.4$ ,  $p < 0.05$  [0.0413],  $\eta^2_G = 0.03$ . There were two significant two-way interactions: A *voicing* and *place* interaction,  $F(1,54) = 120$ ,  $p < 0.0001$ ,  $\eta^2_G = 0.151$ , and an interaction between *voicing* and *group*,  $F(1,54) = 12.5$ ,  $p < 0.001$  [0.0008],  $\eta^2_G = 0.08$ . In general terms, voiced plosives had negative VOT values and voiceless ones had positive values, as one would expect (Lousada et al. 2010). Velar plosive means were “displaced to the right” relative to bilabial means; that is, /k/ ( $M = 60$ , 95% CI [53, 62]) had a longer voicing lag than /p/ ( $M = 12$  [7, 16]) and /g/ had a shorter prevoicing period ( $M = -92$  [-96, -86]) than /b/ ( $M = -109$  [-114, -105]).

The interaction between *voicing* and *place* was due to the fact that the size of the effects of *place* were larger in the voiceless set,  $M^{diff} = 46$ ,  $t(106) = 24$ ,  $p^{tuckey} < 0.0001$ , than in the voiced set,  $M^{diff} = 18$ ,  $t(106) = 9$ ,  $p^{tuckey} < 0.0001$ . On the other hand, the interaction between *voicing* and *group* was due to the fact that there was a significant effect of *group* in the voiced set,  $M^{diff} = 16$ ,  $t(107) = 4$ ,  $p^{tuckey} < 0.001$  [0.0008], but not in the voiceless set,  $M^{diff} = 4$ ,  $t(106) = 0.8$ ,  $p^{tuckey} > 0.05$  [0.815]. The estimated marginal means for this comparison were as follows: Regarding the voiced consonants, the average length of prevoicing of the controls ( $M = -94$ , 95% CI [-98, -87]) was shorter than that of the learners ( $M = -108$  [-114, -103]). Regarding the voiceless consonants, the average voicing lag of the controls ( $M = 33$  [27, 39]) was not reliably different from that of the learners ( $M = 36$  [31, 42]). In sum, the statistical comparisons suggest that there was a difference in the length of prevoicing (in voiced plosives) between controls and EFL learners. However, there was no significant difference in the length of voicing lag in the voiceless set. Prevoicing in utterance-initial voiced plosives seems to be longer in the Portuguese spoken by EFL learners than in that spoken by Portuguese monolinguals, a difference of approximately 16 ms (SE = 4 ms). This appears to be true for both velar and bilabial plosives. Figure 1 plots mean VOT values and 95% confidence intervals as a function of *group*, *place of articulation*, and *voicing*.



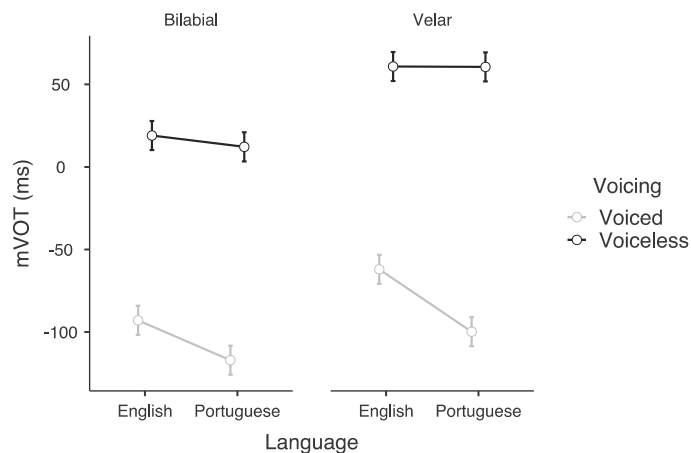
**Figure 1.** Mean (and 95% CI) of Portuguese  $mVOT$  values plotted as a function of *place* (bilabial, velar), *voicing* (voiced, voiceless), and speaker group (EFL learners, controls). Data come from 56 native speakers of Portuguese, 28 of whom are foreign language learners of English.

### 3.2. Learner Productions: Within-Subject Comparison

This section reports on the results of a comparison between the Portuguese (L1) and English (L2) productions of the 28 EFL learners in our sample. The Portuguese monolingual controls are excluded from this analysis. The dependent variable was *mVOT* (ms), and the predictors were *place* (bilabial, velar), *voicing* (voiced, voiceless), and *language* (Portuguese, English). It is important to remember that all factors were within-subject factors—the factor *language* compared the English (L2) and Portuguese (L1) productions of a single group of speakers. The descriptive statistics, mean (and standard deviation), were as follows: For /b/, the mean VOT for L1 productions was  $-117$  (25), and the L2 mean was  $-93$  (36). For /g/, the L1 VOT mean was  $-100$  (18), and the L2 mean was  $-62$  (39). For /p/, the L1 VOT mean was 12 (6), and the L2 mean was 19 (14). For /k/, the mean VOT for L1 tokens was 61 (10), and the L2 mean was 61 (17).

The *mVOT* data were submitted to a repeated measures ANOVA with *place*, *voicing*, and *language* as within-subject predictors. This is a  $(2) \times (2) \times (2)$  design, and the  $\alpha$  criterion was set at 0.05. The ANOVA revealed significant main effects of *voicing*,  $F(1,27) = 674$ ,  $p < 0.0001$ ,  $\eta^2_G = 0.89$ , *place*,  $F(1,27) = 346$ ,  $p < 0.0001$ ,  $\eta^2_G = 0.36$ , and *language*,  $F(1,27) = 25$ ,  $p < 0.0001$ ,  $\eta^2_G = 0.12$ . Voiced plosives had, on average, negative VOT values, and voiceless plosives had, on average, positive ones, a mean difference of approximately 131 ms (SE = 5). Relative to bilabials, velars were generally “displaced to the right,” a mean difference of about 35 ms (SE = 2); that is, /k/ had a longer voicing lag period than /p/, and /b/ had a longer prevoicing period than /g/. Finally, there was a difference between English and Portuguese values such that, in general, English values were “displaced to the right” relative to Portuguese ones, a mean difference of about 17 ms (SE = 3). Since there were several significant interactions between the factors, the main effects may not be interpreted on their own. There were two two-way interactions: *voicing* by *place*,  $F(1,27) = 42$ ,  $p < 0.0001$ ,  $\eta^2_G = 0.05$ , and *voicing* by *language*,  $F(1,27) = 27$ ,  $p < 0.0001$ ,  $\eta^2_G = 0.08$ . The *voicing* by *place* interaction seemed to be due to the fact that the difference between velar and bilabial plosives was larger in the voiceless set,  $M^{\text{diff}} = 45$ ,  $t(53) = 18$ ,  $p^{\text{tuckey}} < 0.0001$ , than in the voiced set,  $M^{\text{diff}} = 24$ ,  $t(53) = 10$ ,  $p^{\text{tuckey}} < 0.0001$ . The *voicing* by *group* interaction was due to the fact that the prevoicing period was significantly longer in the Portuguese voiced plosives than in the English ones,  $M^{\text{diff}} = 31$ ,  $t(50) = 7.2$ ,  $p^{\text{tuckey}} < 0.0001$ , whereas voicing lag was similar for the two languages in the voiceless set,  $M^{\text{diff}} = 3.5$ ,  $t(50) = 0.8$ ,  $p^{\text{tuckey}} > 0.05$  [0.85]. It seems that, in this data set, voiced consonants differ as a function of the language spoken by the learners, whereas voiceless consonants do not.

The omnibus ANOVA also yielded a three-way interaction. There were no significant effects of *language* neither for /p/,  $M^{\text{diff}} = 7$ ,  $t(65) = 1.5$ ,  $p^{\text{tuckey}} > 0.05$  [0.82], nor for /k/,  $M^{\text{diff}} = 0.2$ ,  $t(65) = 0.05$ ,  $p^{\text{tuckey}} > 0.05$  [1]. In addition, whereas there were effects of *language* for both /b/ and /g/, *language* effects were larger for the velars,  $M^{\text{diff}} = 38$ ,  $t(65) = 8.2$ ,  $p^{\text{tuckey}} < 0.0001$ , than for the bilabials,  $M^{\text{diff}} = 24$ ,  $t(65) = 5.2$ ,  $p^{\text{tuckey}} < 0.0001$ . To summarize, voiceless plosives had, on average, a positive VOT in both L1 and L2 productions, and voicing lag was particularly long for /k/. VOT values in L1 and L2 voiceless plosives did not differ from each other. Voiced plosives had, on average, a negative VOT in both L1 and L2 productions, and the prevoicing period was longer for /b/ than for /g/. Most importantly, the prevoicing period was longer in the Portuguese productions than in the English productions. There were, therefore, effects of *language* in the voiced set. It seems that the EFL learners utilized the same VOT targets for the English and Portuguese voiceless plosives. However, they seem to have separate VOT targets for the English and Portuguese voiced plosives. Figure 2 plots average VOT values as a function *language*, *place*, and *voicing*.



**Figure 2.** Mean (and 95% CI) *mVOT* values plotted as a function of *place* (bilabial, velar), *voicing* (voiced, voiceless), and *language* (L2 English, L1 Portuguese). Data come from 28 native speakers of Portuguese learning English as a foreign language in Brazil.

### 3.2.1. Performance Mismatches?

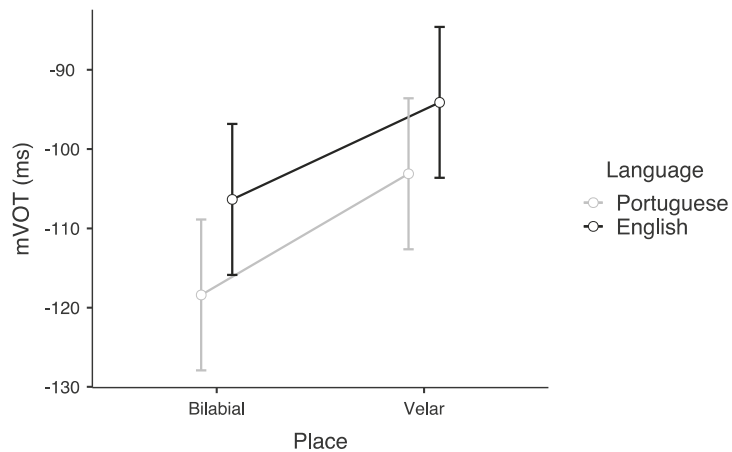
The EFL data presented here suggest that learners develop a “compromise” VOT category for English voiced plosives (Casillas 2021; Flege 1991, p. 395). This compromise category seems to be based on the native Portuguese category, and thus presents significant prevoicing, but it also seems to approximate, to some extent, the native English category, hence the shorter negative VOT. There is, however, an alternative explanation.

Casillas (2021) suggests that the “compromise” VOT values that are sometimes found in the literature on bilingual speech production may result from averaging extreme values and not from the actual development of intermediate or compromise categories in bilingual speech. According to Casillas, bilinguals may be producing “performance mismatches”—when speaking in their less dominant language, bilinguals may fluctuate between L1-like and L2-like tokens. The presence of L1-like tokens in the pool of L2 productions could thus alter the resulting average, displacing it in the direction of L1 categories. Translated to our findings, Casillas’ interpretation would be that our EFL learners present, on average, shorter negative VOT values for their L2 plosives than for their L1 plosives since we have averaged both English productions with Portuguese-like VOT values (i.e., with prevoicing just as long as their native Portuguese plosives) and English productions with English-like VOT values (i.e., with short-lag VOT). Therefore, the averages reported above would not represent any actual production target of the learners but would be the result of averaging extreme values. By hypothesis, our EFL learners could have been producing both prevoiced and short-lag VOT values in their English plosives, and thus the intermediate category we reported as the average would simply indicate the possible presence of performance mismatches. The long prevoicing in the Portuguese (L1) plosives, on the other hand, would simply come from the lack of any short-lag VOT tokens in this pool—there would not be any performance mismatches in the native language.

To address this issue, we focused on the English (L2) productions of the EFL learners only and, particularly, on the voiced plosives, /b g/. There were 1117 voiced plosives in the English data set. Of these, 901 were produced with prevoicing and 216 were produced with short-lag VOT. There were 1120 voiced plosives in the Portuguese data set and, of these, only 19 were produced with short-lag VOT. Performance mismatches appear to occur in both the L1 and L2 but, as Casillas hypothesized, performance mismatches seem to be more common in the L2. This would indeed displace the average VOT of the English plosives in our data set closer to zero, towards the prototypical English category. The difference between the two languages was found to be significant: 98.3% (95% CI [97.4, 98.9]) of the

Portuguese voiced productions were prevoiced, whereas only 80% ([78.2, 82.9]) of the English voiced productions were prevoiced—which is a difference of 17.6% ([15.2, 20.1]) (Cumming 2013, pp. 399–401). A mixed-effects logistic regression model confirmed that the proportion of short-lag VOT tokens was larger in English than in Portuguese:  $\beta = 2.62$ ,  $z = 5.04$ ,  $p < 0.0001$ . The English voiced plosives produced by the EFL learners may not, after all, have a shorter prevoicing period than their Portuguese voiced plosives. It could be that about 20% of their English productions were produced with short-lag VOT.

How long is the prevoicing period of the English voiced plosives that are indeed prevoiced in the speech of the EFL learners? Is it just as long as that of their own Portuguese productions? To answer these questions, we selected all of the English and Portuguese voiced plosives, /b g/, that had been produced with negative VOT, prevoicing. Tokens produced with short-lag VOT were excluded from this analysis. In this sample, there were 1101 Portuguese tokens and 901 English tokens. Then, we calculated by-speaker and by-condition averages. The descriptive statistics for this subset were as follows: The mean VOT for English /b/ was  $-106$  (SD = 31) and that for Portuguese /b/ was  $-118$  (25); the mean VOT for English /g/ was  $-94$  (24) and that for Portuguese /g/ was  $-103$  (18). The dependent variable, *mVOT* (ms), was submitted to a repeated measures ANOVA with *language* (English, Portuguese) and *place* (bilabial, velar). Recall that only voiced plosives with prevoicing were included in this analysis. The ANOVA yielded significant main effects of *place*,  $F(1,27) = 25$ ,  $p < 0.0001$ ,  $\eta^2_G = 0.07$ , and of *language*,  $F(1,27) = 8.1$ ,  $p < 0.001$  [0.0083],  $\eta^2_G = 0.04$ , but there was no significant interaction between the two factors,  $F(1,27) = 1.2$ ,  $p > 0.05$  [0.29],  $\eta^2_G = 0.001$ . Bilabials had, on average, a longer prevoicing period than velars,  $M^{diff} = 14$ ,  $t(27) = 5$ ,  $p^{tuckey} < 0.0001$ . Most importantly, on average, the prevoicing period of the English plosives was shorter than that of the Portuguese plosives,  $M^{diff} = 11$ ,  $t(27) = 3$ ,  $p^{tuckey} < 0.001$  [0.008]. Figure 3 plots average VOT values as a function of *language* and *place*.



**Figure 3.** Mean (and 95% CI) *mVOT* values plotted as a function of *place* (bilabial, velar) and *language* (L2 English, L1 Portuguese). Data come from 28 native speakers of Portuguese learning English as a foreign language in Brazil. The sample includes only voiced plosives with prevoicing.

The difference between English and Portuguese /b/ in this data subset was, on average, 12 ms, 95% CI [3.2, 20.9]. The standardized mean difference, corrected for bias, was  $d^{avg} = 0.42$ , 95% CI [0.13, 0.74]. The correlation between the paired measures was  $r = 0.69$ . As for /g/, the difference between the English and Portuguese VOT values in this subset was, on average, 9 ms, [1.6, 16.3]. Corrected for bias, the standardized mean difference was  $d^{avg} = 0.42$ , 95% CI [0.10, 0.78], and the correlation between the paired measures was  $r = 0.62$ .



To summarize, even when excluding the tokens that have been produced with short-lag VOT, the average prevoicing length was found to be larger in Portuguese than in English in the speech produced by EFL learners. It seems that, in addition to performance mismatches, EFL learners produced intermediate or compromise VOT categories.

### 3.2.2. Effects of Proficiency or Use?

The present subsection explores the possible role of L2 proficiency and amount of usage on the production of English plosives by EFL learners. The analyses we have reported in preceding subsections suggest that only an analysis of voiced plosives is likely to reveal any effects of proficiency or use. Firstly, EFL learners seem to have longer prevoicing production targets for voiced plosives in their native Portuguese than monolingual speakers of Portuguese do. There is no difference, however, between the two groups with regards to voiceless plosives. Secondly, EFL learners seem to have longer prevoicing targets for voiced plosives in their native Portuguese than they do for voiced plosives in English, their L2. There is no difference between L1 and L2 productions for this group with respect to the voiceless plosives. All this suggests that, if we are to find any effects of English proficiency or use on speech production in this sample, we are likely to find them only in the voiced plosives. Therefore, for the analyses reported in this subsection, we focused exclusively on the voiced plosives produced by the EFL learners.

We conducted two sets of analyses. On the one hand, we investigated the potential effects of proficiency and use on the phonetic characteristics of English voiced plosives. For these analyses, we concerned ourselves only with the English voiced plosives produced by the 28 EFL learners in our sample, their L2 plosives. We asked whether English proficiency or use (or both) led to differences in the VOT of the English plosives produced by EFL learners. We hypothesized that increases in English proficiency and use are associated with a shorter length of prevoicing in the English voiced plosives produced by the learners. On the other hand, we analyzed the potential effects of proficiency and use on the size of language mode effects in the voiced plosives produced by the learners. To obtain our dependent variable, we subtracted the mean VOT of a given learner's Portuguese voiced plosive from that of their own corresponding English plosive. We asked whether English proficiency or use (or both) led to differences in the size of the difference between mean L1 and L2 VOT values (for voiced plosives only). We hypothesized that increases in English proficiency and use are associated with larger differences (i.e., larger effects of language mode) between L1 and L2 voiced plosives.

The first set of analyses focused on mean VOT values in the production of English /b/ and /g/. The first analysis was concerned with English /b/. We obtained the mean VOT for each speaker and regressed it against a set of predictors. The chosen predictors were as follows: Measured English proficiency (i.e., the score resulting from the cloze test; range = 0–40), self-assessed English proficiency (i.e., the average of a given learner's various self-assessed proficiency scores; range = 1–7), and self-assessed amount of English use (i.e., the average of a given learner's various estimated usage scores: range = 0–100). These values, four per participant (one metric and three predictors), were submitted to a linear regression model. The overall fit of the model was poor,  $R^2 = 0.16$ , and the results yielded a series of null findings: Measured proficiency,  $\beta = -0.21$ ,  $t = -0.21$ ,  $p > 0.05$  [0.84], self-assessed proficiency,  $\beta = -1.29$ ,  $t = -0.17$ ,  $p > 0.05$  [0.87], and use,  $\beta = -0.95$ ,  $t = -1.49$ ,  $p > 0.05$  [0.15]. The second analysis focused on English /g/. Mean VOT of /g/ was regressed against the proficiency and use predictors: Measured English proficiency, self-assessed English proficiency, and self-assessed amount of English use. Once again, the overall fit of the linear regression model was poor,  $R^2 = 0.19$ , and none of the predictors yielded a significant result: Measured proficiency,  $\beta = -0.75$ ,  $t = -0.71$ ,  $p > 0.05$  [0.48], self-assessed proficiency,  $\beta = -0.53$ ,  $t = -0.07$ ,  $p > 0.05$  [0.95], and use,  $\beta = -0.97$ ,  $t = -1.47$ ,  $p > 0.05$  [0.15]. In sum, there is no evidence that English proficiency (neither measured nor self-assessed) or use affect VOT production in English words in EFL learners whose native language is Portuguese, at least not for /b/ or /g/.

The second set of analyses focused on the size of language mode effects, that is, on the size of the difference in mean VOT between English and Portuguese voiced plosives. For the first analysis in this group, we obtained, for each of the 28 EFL learners, the mean VOT difference between English (L2) and Portuguese (L1) /b/ productions. The mean difference values were regressed against three predictors: Measured proficiency, self-assessed proficiency, and self-assessed amount of English use. The overall fit of the linear regression model was extremely poor,  $R^2 = 0.09$ , and none of the factors were found to account for any significant amount of variance: Measured proficiency,  $\beta = -0.59$ ,  $t = -0.67$ ,  $p > 0.05$  [0.49], self-assessed proficiency,  $\beta = 0.66$ ,  $t = 0.1$ ,  $p > 0.05$  [0.92], and use,  $\beta = -0.51$ ,  $t = -0.09$ ,  $p > 0.05$  [0.37]. The second analysis was concerned with the mean VOT difference between English (L2) and Portuguese (L1) /g/ productions. The same three predictors were used in a linear regression model. The model had a relatively poor fit,  $R^2 = 0.25$ , and none of the predictors were found to be significant: Measured proficiency,  $\beta = -1.22$ ,  $t = -1.37$ ,  $p > 0.05$  [0.18], self-assessed proficiency,  $\beta = 5.16$ ,  $t = 0.75$ ,  $p > 0.05$  [0.45], and use,  $\beta = -1.10$ ,  $t = -1.95$ ,  $p > 0.05$  [0.06].

We found no evidence that either English proficiency or the amount of English use affected speech production in L1 Portuguese EFL learners. There was no evidence that the length of the prevoicing period of English voiced plosives is modulated by any of the experience indicators we used. The size of the difference between the length of the prevoicing period of Portuguese (L1) and English (L2) voiced plosives did not appear to change as a function of proficiency nor amount of L2 use.

## 4. Discussion

### 4.1. Summary of Findings

The present study focused on two main data comparisons and a secondary analysis. In all comparisons, the dependent variable was VOT (ms), and the target sound categories were, also in all cases, velar and bilabial voiced and voiceless stop consonants. The two main data comparisons were as follows. First, two groups of native Portuguese speakers—born, raised, and living in Brazil—produced the target sounds in their native language. One of the two groups comprised active EFL learners and the other, functional monolinguals, with no prior (substantial) exposure to English. The first comparison was concerned with contrasting the productions of the two speaker groups in their native language—a between-speakers comparison. Second, we obtained comparable L2 (English) production data from the EFL learners. Thus, the second main comparison was concerned with contrasting the L1 and L2 speech productions of this particular group—a within-speaker comparison. The secondary analysis, made possible by having recruited EFL learners of various English proficiency levels, explored the potential effects of EFL experience and proficiency on the speech productions of the EFL learners.

Were there phonetic differences between the Portuguese stops produced by L1 Portuguese EFL learners and those of Portuguese monolinguals? In our data, there were significant, albeit small, effects of speaker group. The voiceless plosives did not differ as a function of group, but the voiced ones did. In particular, the EFL learners were found to produce Portuguese /b/ and /g/ with *longer* prevoicing (negative VOT) than the monolinguals. While varying in VOT length, it is important to keep in mind that the average (and mode) Portuguese voiced plosive in both groups was prevoiced. Secondarily, VOT was found to be modulated by place of articulation and voicing, as one would expect (Lousada et al. 2010; Cho and Ladefoged 1999).

Was there evidence of the EFL learners having developed phonetic categories specific to English, that is, separate from those of their Portuguese? In our data, we found significant, albeit small, effects of language. The voiceless consonants were not modulated by language, but the voiced ones were. The English (L2) voiced plosives were found to present significantly shorter negative VOT (prevoicing) than the Portuguese (L1) consonants. Once again, we point out that, while varying in VOT length, the mode (and average) voiced plosive of a Brazilian EFL learner, in both Portuguese and English, is prevoiced. We

conducted a series of statistical analyses as a follow-up of this finding. We found that EFL learners produced a higher proportion of tokens with short-lag VOT in their L2 than in their L1. Prevoiced tokens were still in the majority. The difference in prevoicing length between L1 and L2 tokens remained even after discarding the short-lag VOT tokens—it remained in a data subset that included only truly prevoiced tokens. In other words, a difference in the proportion of tokens with or without prevoicing did not fully explain the language effects found initially with respect to prevoicing length.

The third analysis, a secondary one in the context of our study, was concerned with the possible effects of English proficiency (or experience) on pronunciation. Two analyses were conducted. First, English proficiency was not found to be associated with the variation in VOT measurements in the English data. Second, English proficiency was not found to be associated with the size of the language effect, that is, the size of the difference between English and Portuguese negative VOTs for the voiced plosives was not correlated with English proficiency. In sum, experienced EFL learners did not seem to differ systematically from inexperienced learners in terms of their VOTs. To be clear, EFL learners did differ from each other in their pronunciation patterns—they were not a homogeneous group. However, interlinear variation could not be explained with the meta information we gathered from our EFL learners.

#### 4.2. Interpretation and Implications

##### 4.2.1. L2 Phonetic Development

Do L1 Portuguese EFL learners develop phonetic categories specific to English sounds, that is, separate from those of their Portuguese sounds? Or, more broadly, do foreign-language learners, whose exposure to native L2 oral input is necessarily limited, form new sound categories specific to their L2? The answer is a nuanced “yes”. We found that, for the most part, L1 Portuguese EFL learners used similar, if not identical, phonetic categories for both of their languages. Firstly, there was no evidence of the formation of an aspirated category to be used in English voiceless plosives. EFL learners used ostensibly the same phonetic category, a short- to mid-lag VOT, for all voiceless plosives. Secondly, EFL learners also failed to develop a short-lag phonetic category to be used in English voiced plosives. The learners mostly produced prevoiced tokens both in their L1 and in their L2—that was their mode production pattern. The learners’ productions do not resemble those of native English speakers. The pronunciation of plosives was clearly modeled on the phonetics of their native language. However, and this is important, there were significant, albeit small, differences in the length (and proportion) of prevoicing as a function of language. This suggests that the EFL learners were employing two prevoiced phonetic targets, one for Portuguese voiced plosives and one for English voiced plosives.

Our results are in line with comparable research with foreign-language learners in classroom settings (Dmitrieva et al. 2020). Dmitrieva et al. (2020) found that native English speakers learning Russian in the US produced Russian voiceless plosives mostly with aspirated VOT and Russian voiced plosives mostly with short-lag VOT (Russian resembles Portuguese, and not English, in its use of VOT categories in voicing contrasts: Russian’s voiced plosives are prevoiced and voiceless plosives have short-lag VOT). The learners in the Dmitrieva et al. study modeled the phonetics of their L2 plosives on those of their L1. However, Dmitrieva et al. (2020) also found evidence of L2-specific pronunciation. For instance, one third of Russian voiced plosive tokens were prevoiced and the period of aspiration in the voiceless plosives was shorter in Russian than in English productions. Evidence of phonetic development in foreign-language speech consisted of small, but significant, subcategorical modifications of L1 sounds, like it did in our study.

To make sense of these findings, we make use of some of the basic principles of the SLM (Flege 1995; Flege and Bohn 2021; Flege et al. 2021). Other theoretical models, such as the L2LP (van Leussen and Escudero 2015; Escudero 2005), could be used as well—albeit with some modifications to our explanation. We rely on a single model for explanatory simplicity, as it is not our goal here to compare L2 speech models. Most

L2 speech researchers would agree that (emergent) bilinguals develop “equivalences” or representational connections between the sounds of their L1 and those in the L2 input they are exposed to, which they must mentally store (Simonet 2016). According to the SLM, speakers possess a single representational system for sounds—a common storage space for L1 and L2 phonetic categories (Flege 1995; Flege and Bohn 2021). Emergent bilinguals tend to categorize L2 sounds in the input they receive as a function of the sounds already in their system, thus utilizing mechanisms optimized for the processing of their L1 sounds. This process is typically referred to as “equivalence classification” (Flege 1987), and it arguably results in a warped representational space which, though containing both L1 and L2 sounds, is typically modeled after L1 sounds. The effects of equivalence classification are evident in the voiceless set in our study. We failed to find any difference in VOT between the L1 and L2 voiceless plosives of our EFL learners. It would seem that, in this learner population, English voiceless plosives are classified as instances of Portuguese voiceless plosives, blocking the potential development of an aspirated category specific to English. This is an instance of *full* equivalence classification, a process that results in a single L1/L2 phonetic category for voiceless plosives. What about voiced plosives? Since Portuguese voiced plosives are prototypically prevoiced, it is not surprising that the L1 Portuguese EFL learners in our study also prevoiced the English voiced plosives they produced. This is another instance of equivalence classification. However, in this case, learners did manage to develop two subcategories or two types of prevoiced plosives: One has very long prevoicing and is specific to Portuguese, and the other has shorter prevoicing and is used when speaking English. This seems to be an instance of *partial* equivalence classification, resulting in two subcategories. While EFL learners might use two different phonetic targets for their L1 and L2 voiced plosives, such phonetic targets are likely still associated in their mental representation—they are “equivalent” at some level.

Why is there full equivalence classification in the voiceless set but partial equivalence classification in the voiced set? Or, in other words, why is phonetic development restricted to the voiced set? At this juncture, we can only speculate, since we had not predicted this particular finding. To begin to address these questions, we would like to direct the reader’s attention to a body of findings on the malleability of VOT categories. At least three studies have examined the role of speech rate on VOT in a variety of languages (Kessinger and Blumstein 1997; Magloire and Green 1999; Beckman et al. 2011), and these studies elucidate some facts about VOT categories. It has been found, on the one hand, that aspirated and prevoiced categories are amenable to the effects of speech rate. Short-lag VOT categories, on the other hand, are not. Kessinger and Blumstein (1997) analyzed VOT data from voiced and voiceless plosives in three languages: French, English, and Thai. This research study found that the VOT of French voiced (but not voiceless) plosives is affected by speech rate—in their data, prevoicing was lengthened in slower speech. It was also found that the speech rate altered the VOT of English voiceless (but not voiced) plosives—the aspiration period was longer in slower speech. Finally, in Thai, which has a three-way contrast, speech rate was found to affect both aspirated and prevoiced plosives—slower speech lengthened both the prevoicing and aspiration periods. In all three languages, the short-lag VOT categories were unaffected by manipulations in speech rate. On a different, but related note, Tobin et al. (2017) found, in the speech of proficient Spanish–English bilinguals, an asymmetry between English (aspirated) and Spanish (short-lag VOT) voiceless plosives. In the Tobin et al. study, bilinguals’ productions in both of their languages were recorded in two separate sessions, one in an English-speaking country and one in a Spanish-speaking country. English /p t k/ differed between the two sessions—the aspiration period was shorter when Spanish was the ambient language—while Spanish /p t k/ did not. Interestingly, the authors noted that “the absence of an effect of ambient language on Spanish VOTs in this investigation suggests that the shorter VOTs may some, how be more stable and resistant to accommodation than longer VOTs” (Tobin et al. 2017, p. 52) They also note that this pattern has been documented in various other studies, even if it may not have been explicitly commented on (Antoniou et al. 2011; Chang 2012).

We postulate, based on these observations, that short-lag VOT categories are less malleable (or more resistant to modification) than both prevoiced and aspirated categories, and that this may account for the asymmetry in our findings. Perhaps short-lag VOT categories result from the in-phase coordination of two articulatory events, timed to occur simultaneously, while both prevoiced and aspirated categories result from anti-phase coordination, events timed to occur in a sequence (Browman and Goldstein 2000). In addition, perhaps sequential articulatory coordination is more malleable—more amenable to the development of subcategories in L2 speech—than simultaneous coordination. An alternative explanation makes use of distinctive features in phonology, privative features in particular (Beckman et al. 2011, 2013). Under this account, aspirated categories possess a [spread glottis] featural specification, and prevoiced categories have a [voice] featural specification. Short-lag VOT categories, on the other hand, lack any laryngeal featural specifications. This account has been used to explain the facts regarding speech rate discussed above (Beckman et al. 2011): Categories with a featural specification (but not those lacking one) are affected by speech rate. A possible extension of this account to our findings is this: L2 learners may be more likely to develop phonetic subcategories specific to their L2 for sounds that have an explicit featural specification (aspirated and prevoiced categories, but not short-lag VOT categories), whereas unspecified sounds could be more resistant to the formation of subcategories.

An anonymous reviewer notes that asymmetrical behavior between voiced and voiceless plosives had already been observed in the phonological literature on L2 acquisition and L1 drift situations (Schwartz 2020). Several studies have found cross-linguistic phonetic influence in the voiced (but not the voiceless) set in bilingual speakers whose two languages differ in terms of how they implement the laryngeal contrast in the plosives—that is, people who speak both an “aspirating” language, such as English, and “true voicing” language, such as Portuguese (e.g., Gabriel et al. 2018; Kang et al. 2016; Podlipný et al. 2020; see Schwartz 2020 for a review). It seems that such bilinguals are more likely to assimilate (or cross-linguistically equate) prevoiced and underlyingly voiced short-lag VOT plosives than they are to assimilate underlyingly voiceless short-lag VOT and aspirated plosives. In other words, the plosives in the voiced set seem to be cross-linguistically more similar (in the bilinguals’ behavior) to each other than those in the voiceless set. Schwartz (2020) speculated that this asymmetry revealed a phonological similarity between prevoiced and underlyingly voiced short-lag VOT plosives that does not exist between aspirated and underlyingly voiceless short-lag VOT plosives. According to Schwartz, plosives have three featural levels or nodes of representation: Closure, voice, and vocalic onset. Aspirated plosives have the featural specification [fortis] at all three levels; underlyingly voiceless plain (i.e., short-lag VOT) plosives, on the other hand, receive the featural specification [fortis] only at the voice onset level; and both prevoiced and underlyingly voiced plain plosives lack any featural specification. This phonological account captures the idea that underlyingly voiced and prevoiced plosives are phonologically identical, whereas the two plosives in the voiceless set are phonologically different. In Schwartz’s (2020) account, differences in phonological representation explain the asymmetrical patterns of cross-linguistic influence in the plosives produced by bilinguals. Note, however, that we have found that the Brazilian learners of English in our study produced a single VOT category for both their Portuguese (L1) and their English (L2) voiceless plosives, whereas the voiced consonants were cross-linguistically different (very similar to each other, but still different). Our data suggest that cross-linguistic assimilation or equivalence classification could actually be stronger in the voiceless set than in the voiced set. It is thus not clear whether our results are in line with Schwartz’s (2020) observation or not. This issue requires further research. What seems to be important is that asymmetrical cross-linguistic behavior between phonologically voiced and voiceless plosives seems to be found recurrently in studies of bilingual phonetic behavior (Schwartz 2020), and our data come to corroborate this observation.

The published research study most similar to ours is that of [Dmitrieva et al. \(2020\)](#). Their results were comparable to ours, but there is one aspect of Dmitrieva et al. learners' experience with their L2 that differs from our learners' experience with English. All of the Russian learners in the Dmitrieva et al. sample had exclusively attended language classes taught by native Russian speakers. On the other hand, only about half the EFL learners in our sample had ever had a native English-speaking teacher—never consistently. The role of input in L2 speech has been considered important for some time ([Flege 2008](#)), but input continues to be notoriously difficult to assess, which might have encouraged researchers to focus on other things, such as age of acquisition ([Flege 2018](#)). At this juncture, we must at least mention the possibility that the EFL learners in our sample may have never received enough “normative” English input to be able to develop native-like, L2-specific phonetic categories. The English input some of our learners received was undoubtedly sufficient to develop new grammatical and lexical norms—in fact, some of the learners in our sample were relatively proficient in English. However, if such input was “accented” (in that it was produced by other native Portuguese speakers, EFL learners themselves) it may have never comprised a sizable amount of aspirated voiceless plosives or short-lag voiced plosives. In the absence of large amounts of such input, how could we expect that foreign-language learners would ever develop native-like phonetic targets in their L2? What seems to be worth mentioning, though, is that the EFL learners in our sample *did* develop L2-specific phonetic targets for one of the consonant sets investigated, even if such targets were modeled on L1 sounds and very different from target L2 norms. It is not surprising that the EFL learners did not “sound like” native English speakers, but it is interesting that some phonetic development took place, even if only on the margins. L2 immersion may be needed for natelikeness, but our data suggest that immersion is not crucial at the initial stages of phonetic-category formation ([Dmitrieva et al. 2020](#)).

#### 4.2.2. L1 Phonetic Drift

Are there phonetic differences between the Portuguese plosives produced by L1 Portuguese EFL learners and that of Portuguese monolinguals? Or, less narrowly, does engaging in the learning of a foreign language affect the phonetics of one's native language? The answer, once again, is a nuanced “yes”. We found a phonetic difference between monolingual speakers of Portuguese and EFL learners. The difference concerned the length of the prevoicing period in the voiced plosives. The EFL learners had a longer prevoicing period than the monolinguals. This seems to be evidence of L1 phonetic drift, a modification to one's native pronunciation resulting from foreign-language learning. Evidence of this sort was also found in a comparable study ([Dmitrieva et al. 2020](#)). In our study, the voiceless consonants were not at all affected, and even the voiced consonants were affected only marginally—prevoicing was the norm for both groups of Portuguese speakers. Still, a significant difference between EFL learners and monolinguals was found. Furthermore, there is an apparent connection between our findings concerning L2 development and our findings concerning L1 drift. Since it is the same set of plosives, that is, voiced plosives, that showed evidence of L2 phonetic development (i.e., the formation of a new subcategory) and L1 phonetic drift, it is reasonable to postulate that the two findings are connected.

We speculate that, to “make room” for a new prevoiced plosive specific to English, L1 Portuguese EFL learners shortened the prevoicing period of the English voiced plosive *and* lengthened the prevoicing period of their native voiced plosive. This seems to be an instance of dissimilation, one of the possible epiphenomena, according to the SLM, of new-category formation ([Flege et al. 2003](#); [MacKay et al. 2001](#); [Simonet 2011](#); [Flege and Eefting 1987](#)). Category dissimilation or deflection has been found in other studies, but only in early, proficient bilinguals. For instance, [Flege and Eefting \(1987\)](#) found that Spanish/English bilinguals' Spanish voiceless stops had a shorter VOT period than those of Spanish monolinguals, likely since bilinguals also possessed a long-lag VOT category for their English voiceless stops. Since, in our data, the formation of a new subcategory specific to the L2 results in (or co-occurs with) a modification of the corresponding L1 category, we

conclude that a reorganization, in the form of internal deflection, of the phonetic system of the EFL learners resulted in an instance of L1 phonetic drift. Therefore, it seems that dissimilation is also possible in late L2 learners.

Dmitrieva et al. (2020) found evidence of L1 phonetic drift associated with L2 phonetic development, as we did, but in their study L1 phonetic drift was assimilatory, not dissimilatory—learners modified their L1 phonetic categories in the direction of assimilation to the corresponding L2 category rather than away from it. The Dmitrieva et al. findings are relatively unique in that evidence for both L1 drift and L2 phonetic learning were found for classroom learners in a foreign-language setting (see Chang 2012; Kartushina et al. 2016a). Findings comparable to these had been reported for advanced learners in immersion (or immigrant) settings, as reviewed in the Introduction (de Leeuw et al. 2010, 2018; Baker and Trofimovich 2005; Tobin et al. 2017). Our findings are in line with the Dmitrieva et al. study and add support to their observation that foreign-language learners, who have only very limited exposure to native English oral input, may also undergo L1 phonetic drift. L1 drift may (or may not) be the results of a ‘novelty effect’ (Chang 2012). In our data, L1 drift did not seem to be associated with English experience—it was, therefore, not exclusive to the production of novice learners. As mentioned, what seems to be new about our findings—different from Dmitrieva et al., for instance—is that we documented the existence of dissimilatory (rather than assimilatory) L1 phonetic drift in late L2 learners who seldom use their L2. Kartushina et al. (2016b), in their recent review of the literature, conclude that, in late L2 learners, “the production of L1 categories seems to be unaffected, that is, the L1 categories remain unchanged” (p. 168). They also attribute dissimilation exclusively to early learners. Our findings suggest that the Kartushina et al. statements are a simplification of the facts. L1 categories may be affected by L2 speech learning even in late learners who continue to be dominant in their L1, and dissimilation is also possible in this population. Perhaps a more accurate observation is that L1 drift of a dissimilatory nature depends on the formation of L2-specific sound categories. In other words, dissimilation is a possible result of L2-specific category formation. If a learner, including a late learner, forms a new category specific to their L2, dissimilation is a possible aftereffect.

An anonymous reviewer points out that our EFL learners provided the Portuguese speech data immediately after they had provided the English data. One could postulate that what triggered the lengthening of prevoicing in the Portuguese voiced plosives relative to the English prevoiced tokens were carry-over effects from the speakers’ having participated in the English experimental block immediately before providing the Portuguese data, rather than *bona fide* cross-linguistic influence from the L2 on the L1. Grosjean (2011) distinguishes between two types of cross-linguistic influence, transfer and interference. Transfer is the permanent, static influence of one language’s features on the other—an influence at the level of linguistic competence or long-term representation. Interference, on the other hand, is the ephemeral influence or temporary intrusion of a feature of one language on the other, an influence exclusively at the level of performance or processing. Research has shown that the effects of (dynamic) interference may go beyond those of transfer and are not necessarily affected by bilingual language dominance (Simonet 2014; Simonet and Amengual 2019). Are the L1 drift effects we captured in our current study the result of transfer or of interference? We are afraid we cannot answer this question here, and we must acknowledge that our experimental design results in the presence of a confound. Only future research may resolve this issue. The effects of L2 phonetic development on L1 phonetic drift we have captured in our study could be the result of transfer, interference or both (Grosjean 2011; Simonet 2014). At any rate, we note that both transfer and interference are forms of cross-linguistic influence.

We conclude that, specifically through partial equivalence classification (i.e., the formation of new subcategories specific to the L2), L2 phonetic development leads to L1 phonetic drift, and it may do so even for foreign-language learners in classroom settings, who have only minimal chances to receive native input in their target language. Our findings are in line with Dmitrieva et al. (2020), among others (e.g., Schwartz 2020). Both

sets of findings, Dmitrieva et al. and ours, suggest the existence of a connection between partial equivalence classification and L1 drift. Our findings are novel in that they show that L1 drift may be dissimilatory in nature even in L2 learners who continue to use their L1 daily. L2-specific new-category formation and concomitant dissimilatory L1 phonetic drift are not an exclusive property of life-long bilinguals living in an L2 immersion setting.

## 5. Conclusions

Fifty-six native speakers of Portuguese produced Portuguese words beginning in one of four plosives, /p b k g/. There were two main groups of Portuguese speakers in the sample. One group comprised monolingual speakers and the other, learners of English as a foreign language. The learners (but not the monolinguals) were also asked to produce English words beginning in one of four plosives, /p b k g/. We measured the VOT of all target word-initial stops. We found that the learners produced a single VOT category for voiceless plosives in both Portuguese and English but two subcategories of voiced plosives, one in Portuguese and one in English. While voiced stops were prevoiced in both languages, the English stops had, on average, a shorter period of prevoicing than the Portuguese stops. This was interpreted as evidence of the learners having formed a phonetic subcategory specific to the L2. We also found that the two groups of Portuguese speakers differed from each other precisely in the duration of the prevoicing period in the voiced plosives. The English learners produced voiced Portuguese plosives with a longer period of prevoicing than the monolinguals. Their voiceless plosives did not differ. This was interpreted as evidence of dissimilatory L1 phonetic drift. We postulated that the English learners had been able to develop a L2-specific subcategory for their English voiced consonants by (also) dissimilating their Portuguese voiced consonant from such new subcategory. Such interlingual dissimilation resulted in L1 drift. L2 phonetic development (new-category formation) and L1 drift are possible in foreign language learning, even when exposure to target-language native phonetic norms is severely restricted.

**Author Contributions:** Conceptualization, D.M.O. and M.S.; Data curation, D.M.O.; Formal analysis, M.S.; Funding acquisition, D.M.O.; Investigation, D.M.O. and M.S.; Methodology, D.M.O. and M.S.; Writing—original draft, M.S. All authors have read and agree to the published version of the manuscript.

**Funding:** The authors acknowledge the receipt of funds to purchase equipment (digital recorder and microphone, \$839.98) from the Graduate and Professional Student Council, Research and Project (REaP) Grant, of the University of Arizona. The grant was awarded to D.M.O. in 2018.

**Institutional Review Board Statement:** The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Institutional Review Board (Human Subjects Protection Program) of the University of Arizona (Protocol UAR 1404294632, approved on 18 April 2014).

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** Synthetic data are available from the corresponding author on request. The data are not publicly available because the authors failed to ask participant permission to publicly share their data. The code to reproduce all analyses is also available on request.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Abramson, Arthur, and Douglass Whalen. 2017. Voice Onset Time (VOT) at 50: Theoretical and Practical Issues in Measuring Voicing Distinctions. *Journal of Phonetics* 63: 75–86. [CrossRef]
- Albano, Eleonora. 2001. *O Gesto e Suas Bordas: Esboço de Fonologia Acústico-Articulatória Do Português Brasileiro*. Campinas: Mercado de Letras.
- Alves, Mariane, Izabel Seara, Fernando Pacheco, Simone Klein, and Rui Seara. 2008. On the Voiceless Aspirated Stops in Brazilian Portuguese. In *Computational Processing of the Portuguese Language*. Berlin: Springer, pp. 248–51.
- Amengual, Mark. 2012. Interlingual Influence in Bilingual Speech: Cognate Status Effect in a Continuum of Bilingualism. *Bilingualism: Language and Cognition* 15: 517–30. [CrossRef]



- Antoniou, Mark, Catherine Best, Michael Tyler, and Christian Kroos. 2011. Inter-Language Interference in VOT Production by L2-Dominant Bilinguals: Asymmetries in Phonetic Code-Switching. *Journal of Phonetics* 39: 558–70. [CrossRef] [PubMed]
- Baker, Wendy, and Pavel Trofimovich. 2005. Interaction of Native- and Second-Language Vowel System(s) in Early and Late Bilinguals. *Language and Speech* 48: 1–27. [CrossRef] [PubMed]
- Barbosa, Plinio, and Eleonora Albano. 2004. Brazilian Portuguese. *JIPA* 34: 227–32. [CrossRef]
- Beckman, Jill, Pétur Helgason, Bob McMurray, and Catherine Ringen. 2011. Rate Effects on Swedish VOT: Evidence for Phonological Overspecification. *Journal of Phonetics* 39: 39–49. [CrossRef]
- Beckman, Jill, Michael Jessen, and Catherine Ringen. 2013. Empirical Evidence for Laryngeal Features: Aspirating vs. True Voice Languages. *Journal of Linguistics* 49: 259–84. [CrossRef]
- Bergmann, Christopher, Amber Nota, Simone Sprenger, and Monika Schmid. 2016. L2 Immersion Causes Non-Native-like L1 Pronunciation in German Attriters. *Journal of Phonetics* 58: 71–86. [CrossRef]
- Boersma, Paul. 2001. Praat, a System for Doing Phonetics by Computer. *Glott International* 5: 341–45.
- British Council. 2019. Políticas Públicas Para o Ensino de Inglês: Um Panorama Das Experiências Na Rede Pública. Available online: [https://www.britishcouncil.org.br/sites/default/files/bncc\\_portuguesbx.pdf](https://www.britishcouncil.org.br/sites/default/files/bncc_portuguesbx.pdf) (accessed on 22 June 2021).
- Browman, Catherine, and Louis Goldstein. 2000. Competing Constraints on Intergestural Coordination and Self-Organization of Phonological Structures. *Bulletin de La Communication Parlée* 5: 25–34.
- Brown, James. 2002. Do Cloze Tests Work? Or Is It Just an Illusion? *Second Language Studies* 21: 79–125.
- Casillas, Joseph. 2021. Interlingual Interactions Elicit Performance Mismatches Not ‘Compromise’ Categories in Early Bilinguals: Evidence from Meta-Analysis and Coronal Stops. *Languages* 6: 9. [CrossRef]
- Casillas, Joseph, and Miquel Simonet. 2018. Perceptual Categorization and Bilingual Language Modes: Assessing the Double Phonemic Boundary in Early and Late Bilinguals. *Journal of Phonetics* 71: 51–64. [CrossRef]
- Chang, Charles. 2012. Rapid and Multifaceted Effects of Second-Language Learning on First-Language Speech Production. *Journal of Phonetics* 40: 249–68. [CrossRef]
- Chang, Charles. 2013. A Novelty Effect in Phonetic Drift of the Native Language. *Journal of Phonetics* 41: 520–33. [CrossRef]
- Chang, Charles. 2019. The Phonetics of Second Language Learning and Bilingualism. In *The Routledge Handbook of Phonetics*. Edited by William Katz and Peter Assmann. Abingdon: Routledge, pp. 427–47.
- Cho, Taehong, and Peter Ladefoged. 1999. Variation and Universals in VOT: Evidence from 18 Languages. *Journal of Phonetics* 27: 207–29. [CrossRef]
- Colantoni, Laura, Jeffrey Steele, and Paola Escudero. 2015. *Second Language Speech: Theory and Practice*. Cambridge: Cambridge University Press.
- Cumming, Geoff. 2013. *Understanding the New Statistics: Effect Sizes, Confidence Intervals, and Meta-Analysis*. London: Routledge.
- de Leeuw, Esther, Monika Schmid, and Ineke Mennen. 2010. The Effects of Contact on Native Language Pronunciation in an L2 Migrant Setting. *Bilingualism: Language and Cognition* 13: 33–40. [CrossRef]
- de Leeuw, Esther, Aurela Tusha, and Monika Schmid. 2018. Individual Phonological Attrition in Albanian-English Late Bilinguals. *Bilingualism: Language and Cognition* 21: 278–95. [CrossRef]
- Dmitrieva, Olga, Allard Jongman, and Joan Sereno. 2020. The Effect of Instructed Second Language Learning on the Acoustic Properties of First Language Speech. *Languages* 5: 44. [CrossRef]
- Escudero, Paola. 2005. *Linguistic Perception and Second Language Acquisition: Explaining the Attainment of Optimal Phonological Categorization*. LOT Dissertation Series 113. Utrecht: Utrecht University.
- Flege, James. 1987. The Production of ‘New’ and ‘Similar’ Phones in a Foreign Language: Evidence for the Effect of Equivalence Classification. *Journal of Phonetics* 15: 47–65. [CrossRef]
- Flege, James. 1991. Age of Learning Affects the Authenticity of Voice-Onset Time (VOT) in Stop Consonants Produced in a Second Language. *The Journal of the Acoustical Society of America* 89: 395–411. [CrossRef] [PubMed]
- Flege, James. 1995. Second Language Speech Learning: Theory, Findings, and Problems. In *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*. Edited by Winnifred Strange. Timonium: York Press, pp. 229–73.
- Flege, James. 2007. Language Contact in Bilingualism: Phonetic System Interactions. In *Laboratory Phonology 9*. Edited by Jennifer Cole and José Hualde. Berlin: Mouton de Gruyter, pp. 353–80.
- Flege, James. 2008. Give Input a Chance! In *Input Matters in SLA*. Edited by Thorsten Piske and Martha Young-Scholten. Bristol: Multilingual Matters, pp. 175–90.
- Flege, James. 2018. A Non-Critical Period for Second-Language Learning. In *A Sound Approach to Language Matters—In Honor of Ocke-Schwen Bohn*. Edited by Anne Nyvad, MMichaela Hejná, Anders Højen, Anna Jespersen and Mette Sørensen. Aarhus: Aarhus University, pp. 501–41.
- Flege, James, and Ocke-Schwen Bohn. 2021. The Revised Speech Learning Model (SLM-r). In *Second Language Speech Learning: Theoretical and Empirical Progress*. Edited by Ratreay Wayland. Cambridge: Cambridge University Press, pp. 3–83.
- Flege, James, and Wieke Eefting. 1987. Production and Perception of English Stops by Native Spanish Speakers. *Journal of Phonetics* 15: 67–83. [CrossRef]
- Flege, James, Naoyuki Takagi, and Virginia Mann. 1995. Japanese Adults Can Learn to Produce English /ɹ/ and /l/ Accurately. *Language and Speech* 38: 25–55. [CrossRef]

- Flege, James, Elaina Frieda, and Takeshi Nozawa. 1997. Amount of Native-Language (L1) Use Affects the Pronunciation of an L2. *Journal of Phonetics* 25: 169–86. [CrossRef]
- Flege, James, Carlo Schirru, and Ian MacKay. 2003. Interaction between the Native and Second Language Phonetic Subsystems. *Speech Communication* 40: 467–91. [CrossRef]
- Flege, James, David Birdsong, Ellen Bialystok, Molly Mack, Hyekyung Sung, and Kimiko Tsukada. 2006. Degree of Foreign Accent in English Sentences Produced by Korean Children and Adults. *Journal of Phonetics* 34: 153–75. [CrossRef]
- Flege, James, Katsura Aoyama, and Ocke-Schwen Bohn. 2021. The Revised Speech Learning Model (SLM-r) Applied. In *Second Language Speech Learning: Theoretical and Empirical Progress*. Edited by Ratre Wayland. Cambridge: Cambridge University Press, pp. 84–118.
- Fowler, Carol, Valery Sramko, David Ostry, Sarah Rowland, and Pierre Hallé. 2008. Cross Language Phonetic Influences on the Speech of French-English Bilinguals. *Journal of Phonetics* 36: 649–63. [CrossRef]
- Gabriel, Christoph, Marion Krause, and Tetyana Dittmers. 2018. VOT Production in Multilingual Learners of French as a Foreign Language: Cross-Linguistic Influence from the Heritage Languages Russian and Turkish. *Revue Française de Linguistique Appliquée* 23: 59–72. [CrossRef]
- Grosjean, François. 1989. Neurolinguists, Beware! The Bilingual Is Not Two Monolinguals in One Person. *Brain and Language* 36: 3–15. [CrossRef]
- Grosjean, François. 2011. An Attempt to Isolate, and Then Differentiate, Transfer and Interference. *International Journal of Bilingualism* 16: 11–21. [CrossRef]
- Guion, Susan. 2003. The Vowel Systems of Quichua-Spanish Bilinguals. *Phonetica* 60: 98–128. [CrossRef]
- Hopp, Holger, and Monika Schmid. 2013. Perceived Foreign Accent in First Language Attrition and Second Language Acquisition: The Impact of Age of Acquisition and Bilingualism. *Applied Psycholinguistics* 34: 361–94. [CrossRef]
- Kang, Yoonjung, Sneha George, and Rachel Soo. 2016. Cross-Language Influence in the Stop Voicing Contrast in Heritage Tagalog. *Heritage Language Journal* 13: 184–218. [CrossRef]
- Kartushina, Natalia, Alexis Hervais-Adelman, Ulrich Frauenfelder, and Narly Golestani. 2016a. Mutual Influences between Native and Non-Native Vowels in Production: Evidence from Short-Term Visual Articulatory Feedback Training. *Journal of Phonetics* 57: 21–39. [CrossRef]
- Kartushina, Natalia, Ulrich Frauenfelder, and Narly Golestani. 2016b. How and When Does the Second Language Influence the Production of Native Speech Sounds: A Literature Review. *Language Learning* 66: 155–86. [CrossRef]
- Kessinger, Rachel, and Sheila Blumstein. 1997. Effects of Speaking Rate on Voice-Onset Time in Thai, French, and English. *Journal of Phonetics* 25: 143–68. [CrossRef]
- Kirby, James, and Dwight Robert Ladd. 2016. Effects of Obstruent Voicing on Vowel F0: Evidence from ‘True Voicing’ Languages. *The Journal of the Acoustical Society of America* 140: 2400–2411. [CrossRef]
- Köpke, Barbara, and Monika Schmid. 2004. Language Attrition: The next Phase. In *First Language Attrition: Interdisciplinary Perspectives on Methodological Issues*. Edited by Monika Schmid, Barbara Köpke, Merel Keijzer and Lina Weilemar. Studies in Bilingualism. Amsterdam: John Benjamins, pp. 1–43.
- Lenth, Russell. 2018. Emmeans: Estimated Marginal Means, aka Least-Squares Means (Version 1.4.7). Available online: <https://cran.r-project.org/package=emmeans> (accessed on 22 June 2021).
- Lisker, Leigh, and Arthur Abramson. 1964. A Cross-Language Study of Voicing in Initial Stops: Acoustical Measurements. *Word* 20: 384–422. [CrossRef]
- Lisker, Leigh, and Arthur Abramson. 1967. Some Effects of Context on Voice Onset Time in English Stops. *Language and Speech* 10: 1–28. [CrossRef] [PubMed]
- Lousada, Marisa, Luis Jesus, and Andreia Hall. 2010. Temporal Acoustic Correlates of the Voicing Contrast in European Portuguese Stops. *Journal of the International Phonetic Association* 40: 261–75. [CrossRef]
- MacKay, Ian, James Flege, Thorsten Piske, and Carlo Schirru. 2001. Category Restructuring during Second-Language Speech Acquisition. *The Journal of the Acoustical Society of America* 110: 516–28. [CrossRef]
- Magloire, Joel, and K. P. Green. 1999. A Cross-Language Comparison of Speaking Rate Effects on the Production of Voice Onset Time in English and Spanish. *Phonetica* 56: 158–85. [CrossRef]
- Major, Roy. 1987. English Voiceless Stop Production by Speakers of Brazilian Portuguese. *Journal of Phonetics* 15: 197–202. [CrossRef]
- Major, Roy. 1992. Losing English as a First Language. *The Modern Language Journal* 76: 190–208. [CrossRef]
- Mayr, Robert, David Sánchez, and Ineke Mennen. 2020. Does Teaching Your Native Language Abroad Increase L1 Attrition of Speech? The Case of Spaniards in the United Kingdom. *Languages* 5: 41. [CrossRef]
- Mora, Joan, and Marianna Nadeu. 2012. L2 Effects on the Perception and Production of a Native Vowel Contrast in Early Bilinguals. *International Journal of Bilingualism* 16: 484–500. [CrossRef]
- Nearey, Terrance, and Bernard Rochet. 1994. Effects of Place of Articulation and Vowel Context on VOT Production and Perception for French and English Stops. *Journal of the International Phonetic Association* 24: 1–18. [CrossRef]
- Pearce, Jonathan. 2007. PsychoPy—Psychophysics Software in Python. *Journal of Neuroscience Methods* 162: 8–13. [CrossRef]
- Pearce, Jonathan, Jeremy Gray, Sol Simpson, Michael MacAskill, Richard Höchenberger, Hiroyuki Sogo, Erik Kastman, and Jonas Lindeløv. 2019. PsychoPy2: Experiments in Behavior Made Easy. *Behavior Research Methods* 51: 195–203. [CrossRef]

- Piske, Thorsten, Ian MacKay, and James Flege. 2001. Factors Affecting Degree of Foreign Accent in an L2: A Review. *Journal of Phonetics* 29: 191–215. [CrossRef]
- Podlipský, Václav, Šárka Šimáčková, and Kateřina Chládková. 2020. Phonetic Drift Reveals Interconnected Phonological Representations in Simultaneous Bilinguals: A Case Study of English and Czech Stop Consonants. *International Journal of Bilingualism* 25: 789–99. [CrossRef]
- Polidório, Valdomiro. 2014. O Ensino de Língua Inglesa No Brasil. The English Teaching in Brazil. *Revista Travessias* 8: 340–46.
- R Core Team. 2018. *R: A Language and Environment for Statistical Computing (Version 4.0.1)*. Vienna: R Core Team. Available online: <https://cran.r-project.org> (accessed on 22 June 2021).
- Sancier, Michele, and Carol Fowler. 1997. Gestural Drift in a Bilingual Speaker of Brazilian Portuguese and English. *Journal of Phonetics* 25: 421–36. [CrossRef]
- Santos, Eliana S. S. 2011. O Ensino Da Língua Inglesa No Brasil. *BABEL: Revista Eletrônica de Línguas e Literaturas Estrangeiras* 1: 1–7.
- Schmid, Monika. 2011. *Language Attrition*. Cambridge: Cambridge University Press.
- Schwartz, Geoffrey. 2020. Asymmetrical Cross-Language Phonetic Interaction: Phonological Implications. *Linguistic Approaches to Bilingualism*. [CrossRef]
- Silva, Kleber. 2004. ‘Inglês Só Se Aprende Na Escola de Idiomas’: Importantes Reflexões No Que Concerne Ao Ensino/Aprendizagem de Língua Inglesa. Paper presented at VII Congresso Brasileiro de Linguística Aplicada, São Paulo, Brazil, October 10–14.
- Simonet, Miquel. 2011. Production of a Catalan-Specific Vowel Contrast by Early Spanish-Catalan Bilinguals. *Phonetica* 68: 88–110. [CrossRef]
- Simonet, Miquel. 2014. Phonetic Consequences of Dynamic Cross-Linguistic Interference in Proficient Bilinguals. *Journal of Phonetics* 43: 26–37. [CrossRef]
- Simonet, Miquel. 2016. The Phonetics and Phonology of Bilingualism. In *Oxford Handbooks Online: Scholarly Research Reviews*. Oxford: Oxford University Press, pp. 1–25.
- Simonet, Miquel, and Mark Amengual. 2019. Increased Language Co-Activation Leads to Enhanced Cross-Linguistic Phonetic Convergence. *International Journal of Bilingualism* 24: 208–21. [CrossRef]
- Singman, Henrik, Ben Bolker, Jake Westfall, Frederik Aust, and Mattan Ben-Shachar. 2020. Afex: Analysis of Factorial Experiments (Version 0.27-2). Available online: <https://cran.r-project.org/package=afex> (accessed on 22 June 2021).
- Solon, Megan. 2016. Do Learners Lighten Up? *Studies in Second Language Acquisition* 39: 801–32. [CrossRef]
- Stoehr, Antje, Titia Benders, Janet van Hell, and Paula Fikkert. 2017. Second Language Attainment and First Language Attrition: The Case of VOT in Immersed Dutch–German Late Bilinguals. *Second Language Research* 33: 483–518. [CrossRef] [PubMed]
- Tavares, Ana C. B. 2018. Como Ensinar Cultura: A Abordagem de Aspectos Culturais No Ensino de Língua Inglesa Sob a Ótica Do Professor Brasileiro. Unpublished Ph.D. dissertation, Universidade do Porto, Porto, Portugal.
- The Jamovi Project. 2020. Jamovi (Version 1.2.22). Available online: <https://www.jamovi.org> (accessed on 22 June 2021).
- Tobin, Stephen, Hosung Nam, and Carol Fowler. 2017. Phonetic Drift in Spanish-English Bilinguals: Experiment and a Self-Organizing Model. *Journal of Phonetics* 65: 45–59. [CrossRef]
- Tremblay, Annie. 2011. Proficiency Assessment Standards in Second Language Acquisition Research: ‘Clozing’ the Gap. *Studies in Second Language Acquisition* 33: 339–72. [CrossRef]
- Tsukada, Kimiko, David Birdsong, Molly Mack, Hyekyung Sung, Ellen Bialystok, and James Flege. 2004. Release Bursts in English Word-Final Voiceless Stops Produced by Native English and Korean Adults and Children. *Phonetica* 61: 67–83. [CrossRef] [PubMed]
- Tsukada, Kimiko, David Birdsong, Ellen Bialystok, Molly Mack, Hyekyung Sung, and James Flege. 2005. A Developmental Study of English Vowel Production and Perception by Native Korean Adults and Children. *Journal of Phonetics* 33: 263–90. [CrossRef]
- Ulbrich, Christiane, and Mikhail Ordin. 2014. Can L2-English Influence L1-German? The Case of Post-Vocalic /r/. *Journal of Phonetics* 45: 26–42. [CrossRef]
- van Leussen, Jan-Willem, and Paola Escudero. 2015. Learning to Perceive and Recognize a Second Language: The L2LP Model Revised. *Frontiers in Psychology* 6: 1–12. [CrossRef] [PubMed]
- Wayland, Ratre, ed. 2021. *Second Language Speech Learning: Theoretical and Empirical Progress*. Cambridge: Cambridge University Press.
- Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D’Agostino McGowan, Romain François, Garrett Gromlund, Alex Hayes, Lionel Henry, Jim Hester, and et al. 2019. Welcome to the Tidyverse. *Journal of Open Source Software* 4: 1686. [CrossRef]
- Yavaş, Mehmet, and Renée Wildermuth. 2006. The Effects of Place of Articulation and Vowel Height in the Acquisition of English Aspirated Stops by Spanish Speakers. *International Review of Applied Linguistics in Language Teaching* 44: 251–63. [CrossRef]

Article

# Longitudinal Developments in Bilingual Second Language Acquisition and First Language Attrition of Speech: The Case of Arnold Schwarzenegger

Lisa Kornder \* and Ineke Mennen

Department of English Studies, University of Graz, 8010 Graz, Austria; ineke.mennen@uni-graz.at

\* Correspondence: lisa.kornder@uni-graz.at

**Abstract:** The purpose of this investigation was to trace first (L1) and second language (L2) segmental speech development in the Austrian German–English late bilingual Arnold Schwarzenegger over a period of 40 years, which makes it the first study to examine a bilingual’s speech development over several decades in both their languages. To this end, acoustic measurements of voice onset time (VOT) durations of word-initial plosives (Study 1) and formant frequencies of the first and second formant of Austrian German and English monophthongs (Study 2) were conducted using speech samples collected from broadcast interviews. The results of Study 1 showed a merging of Schwarzenegger’s German and English voiceless plosives in his late productions as manifested in a significant lengthening of VOT duration in his German plosives, and a shortening of VOT duration in his English plosives, closer to L1 production norms. Similar findings were evidenced in Study 2, revealing that some of Schwarzenegger’s L1 and L2 vowel categories had moved closer together in the course of L2 immersion. These findings suggest that both a bilingual’s first and second language accent is likely to develop and reorganize over time due to dynamic interactions between the first and second language system.

**Keywords:** first language attrition; second language acquisition; sequential bilingualism; voice onset time; vowel formants; speech development; English; (Austrian) German; phonetics



**Citation:** Kornder, Lisa, and Ineke Mennen. 2021. Longitudinal Developments in Bilingual Second Language Acquisition and First Language Attrition of Speech: The Case of Arnold Schwarzenegger. *Languages* 6: 61. <https://doi.org/10.3390/languages6020061>

Academic Editors: Robert Mayr, Jonathan Morris, Juana M. Licerias and Raquel Fernández Fuertes

Received: 25 February 2021

Accepted: 19 March 2021

Published: 25 March 2021

**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Learning a second language (L2) late in life often entails that speakers retain a noticeable foreign accent in their L2 resulting from first language (L1) influences on the late-acquired L2 phonetic/phonological system (Flege 1980, 1981; Flege et al. 1996; MacKay et al. 2001; Scovel 1969). Examining L1 influences on pronunciation abilities in the L2 has a long tradition in second language and bilingualism research: Against the background of theories evolving around maturational constraints, such as the critical period hypothesis (Lenneberg 1967; Penfield and Roberts 1959), a prevailing view was that an individual’s native language system—once fully mature—was not likely to be affected and modified by a late-acquired L2 system (Bylund 2009; Flege et al. 2003; Scovel 1969). In the past decades, however, research came to acknowledge the bi-directional nature of interactions between a bilingual’s language systems (e.g., Flege 1987; Mennen 2004; Sancier and Fowler 1997). This change in orientation resulted from the observation that also a mature native system might be affected by L2-induced changes (e.g., Flege 1987). L1 modifications and the resulting decline of linguistic abilities in one’s native language observed in late L2 learners who are being long-term immersed in an L2 environment is commonly referred to as *L1 attrition* (Köpke and Schmid 2004). Interest in gaining a better understanding of whether it is possible for healthy individuals to “unlearn” their L1 as a result of L2-learning experience developed in the early 1980s (see Köpke and Schmid 2004, for an overview). Since then, a considerable amount of research has been conducted to explore attrition effects on native pronunciation abilities, which provides evidence for the permeability of

late bilinguals' L1 pronunciation system with regard to segmental (Bergmann et al. 2016; de Leeuw et al. 2017; de Leeuw 2019; Dmitrieva et al. 2010; Mayr et al. 2012; Stoehr et al. 2017; Ulbrich and Ordin 2014) and prosodic (de Leeuw et al. 2012b; de Leeuw 2019; Mennen and Chousi 2018) features.

The present study aims to further extend the growing body of research on L1 attrition and L2 acquisition and to gain a more profound understanding of the relationship between L2 acquisition and L1 attrition in an immersion context. To this end, long-term changes in the L1 Austrian German and the L2 English of the late consecutive bilingual Arnold Schwarzenegger were examined by conducting acoustic analyses of voice onset time (VOT) and vowel formants based on spontaneous speech samples. To date, this case study is the first which does not exclusively focus on identifying longitudinal pronunciation changes in either the L1 (e.g., de Leeuw 2019) or the L2 (e.g., Saito et al. 2019), or examining L1–L2 interactions over a relatively short period of time (e.g., Chang 2012; Sancier and Fowler 1997) but which traces the trajectory of segmental speech development in both the L1 and a late-acquired L2 over a period of 40 years.

### 1.1. Bi-Directional L1–L2 Influences: An Integrated View of Bilingual Speech Development

Early investigations into second language speech acquisition started from the premise that the “steady state” of a biologically mature L1 pronunciation system is not likely to be disrupted by a late-acquired L2 system (Lado 1957; Lenneberg 1967). This resulted in a rather biased focus on identifying L1 influences on the L2 system and neglected potential interactions between a speaker's languages. Furthermore, a seemingly straightforward definition of *the bilingual* was proposed according to which a bilingual's language systems are representative of two monolingual systems, which resulted in the prevailing view that only those individuals who showed native and balanced proficiency in both languages were to be considered “real” bilinguals (Bloomfield 1933; Thiery 1978). Moving away from this static point of view, Grosjean (1989, 1997) proposed a holistic approach to bilingualism, which reflected bilingual reality more appropriately. He argued that a bilingual's linguistic configuration is per se different from that of a monolingual due to a dynamic interaction between the language systems and, hence, defining bilingual speakers against the background of monolingual competence is rather misleading. With the development of this integrated view, bilingualism and L2 acquisition research shifted perspectives and started to acknowledge bi-directional L1–L2 influences as an inherent characteristic of speech development in second language and bilingual speakers (Flege 1987; Odlin 1989, 2006; Sharwood Smith and Kellerman 1986).

Support for this holistic view and the notion of bi-directionality comes from the Speech Learning Model (SLM) proposed by Flege (1995; see Flege and Bohn 2020, for a revised version of the SLM). According to the SLM, bilinguals' sound systems are not isolated from each other, but exist in a shared phonetic space, which accounts for a mutual interaction between a speaker's language systems (Flege 1995; Flege and Bohn 2020). In this shared space, L1–L2 interactions are determined by different mechanisms which can not only lead to non-native L2 productions, but might also result in a reorganization of L1 categories. A substantial amount of empirical research on bilingual speech development has supported one of the SLM's main predictions that similar L1 and L2 sounds would be difficult to produce authentically due to assimilatory effects (Baker and Trofimovich 2005; Flege 1991; Flege et al. 1996; Flege and Hillenbrand 1984; Thornburgh and Ryalls 1998). That is, late bilinguals may fail to identify fine phonetic differences between acoustically related L1 and L2 sounds and establish a merged L1–L2 category which differs from the respective monolingual categories. Flege and Hillenbrand (1984), for example, found that both L1 French late learners of L2 English, and L1 English late learners of L2 French produced French /t/ with VOT values that considerably exceeded monolingual French short-lag VOT values, but were still too short for long-lag English categories. That is, they had established merged phonetic categories due to L1–L2 assimilation. Dissimilatory effects, by contrast, have been mainly observed among early bilinguals who manage to establish

distinct L1 and L2 categories for closely related sounds, but may overshoot the monolingual targets in both languages in an attempt to maintain contrast (Flege et al. 2003; Flege and Eefting 1988; Mack 1990). A recently revised version of the SLM (Flege and Bohn 2020) stresses the need to focus on individual differences between L2 learners when it comes to mastering an L2 sound system. In fact, inter- as well as intra-subject variability have been widely documented even in rather homogenous speaker groups examined in cross-sectional studies (e.g., Bongaerts et al. 2000; Major 1992; Mennen 2004; Moyer 1999). These studies show that some speakers are well able to attain native proficiency in their L2 while others clearly fall outside the native range, and thus challenge yet again a static view on bilingual speech acquisition.

While research conducted within the framework of the SLM focuses on highly proficient bilinguals with long-term L2 experience (see Flege 1995), bi-directional influences between speakers' linguistic systems and L1 modifications resulting from this interaction have also been observed to occur at a very early stage of L2 acquisition. Chang (2012, 2013), for instance, identified a *phonetic drift* of L1 English plosive and vowel categories towards L2 Korean categories in speakers learning Korean in an instructional setting. As none of the speakers had prior experience with Korean, the L1 changes did not result from long-term L2 exposure and usage, but from recent L2 experience and the novelty of L2 input. Evidence for a drift of L1 categories towards L2 norms in inexperienced individuals also comes from Kartushina et al. (2016), who trained L1 French speakers on the production of Russian and Danish vowels in 1-h training sessions. Similar to the subjects in Chang (2012, 2013), the speakers in this study experienced modifications of their L1 vowel categories in the direction of L2 norms despite lacking previous long-term L2 experience. Kartushina and Martin (2019) showed that the vowel systems of Basque-Spanish bilinguals had moved closer to English norms after completing a two-week study abroad English program. Four months after the program, however, an acoustic re-analysis of the speakers' L1 and L2 vowels revealed that the vowel categories had drifted back to native norms (see Chang 2019), which suggests that L1 modifications occurring in the initial stage of L2 learning are only temporary and are likely to be reversed due to changes in language use and linguistic environment.

Further evidence for the impact of recent L2 experience on L1 speech production and for the reversibility of L1 changes as a function of linguistic environment is provided by Sancier and Fowler (1997; see also Tobin et al. 2017). In an 11-months case study, they examined VOT produced by an L1 Brazilian-Portuguese advanced speaker of L2 English, who travelled between Brazil and the US at monthly intervals. Sancier and Fowler describe a *gestural drift* of VOT in both languages towards the most recently experienced language, i.e., VOT durations were longer and thus more English-like after returning from the US, and shorter and more Portuguese-like after staying in Brazil.

While the studies outlined above explored L2-induced changes in the L1 pronunciation of speakers who were at the onset of L2 learning (Chang 2012, 2013, 2019; Kartushina et al. 2016) or experienced regular changes in their linguistic environment (Sancier and Fowler 1997; Tobin et al. 2017), other studies examined L2-induced changes in the L1 pronunciation of experienced late bilinguals being permanently and long-term immersed in an L2 environment. Research into phonetic and phonological attrition has documented modifications in bilinguals' L1 segmental (de Leeuw 2019; Dmitrieva et al. 2010; Mayr et al. 2012; Stoehr et al. 2017) and prosodic (de Leeuw et al. 2012b; Mennen and Chousi 2018) productions in the direction of L2 norms. In addition, research provides evidence for listeners' perceptual sensitivity to divergences from L1 norms, that is, listeners have been shown to judge bilingual attrited speech as sounding less native compared to monolingual non-attrited speech (Bergmann et al. 2016; de Leeuw et al. 2010; Hopp and Schmid 2013; Mayr et al. 2020; Schmid and Hopp 2014). Research examining potential changes in long-term L2-immersed bilinguals' realizations of plosives and vowels—the two sound classes investigated in the present study—will be discussed in more detail in Section 1.2.

Taken together, these studies show that speakers may experience a restructuring of certain L1 features in the direction of the L2 as a consequence of long-term L2 learning experience and L2 immersion. However, just as bilinguals master L2 pronunciation with varying degrees of success (see [Bongaerts et al. 2000](#); [Major 1992](#); [Mennen 2004](#); [Moyer 1999](#)), the extent to which native abilities in L1 pronunciation decline might also differ among individuals (see [Bergmann et al. 2016](#); [de Leeuw et al. 2017](#); [Major 1992](#); [Mayr et al. 2012](#)). Furthermore, previous research shows that not all features of L1 pronunciation—within the same individual and the same sound category—undergo attrition (e.g., [Bergmann et al. 2016](#); [Mayr et al. 2012](#); [Stoehr et al. 2017](#)).

Notions of an end state of L2 learning (see [Birdsong 2009](#), for a discussion) and of a steady state of the L1 ([Lado 1957](#); [Lenneberg 1967](#)) can hardly be reconciled with the empirical findings outlined above, which document a reorganization of both the L1 and the L2 system as a result of bi-directional interaction processes, and show that L2 acquisition and L1 maintenance are determined by inter- and intra-speaker variability. Such observations offer convincing evidence that bilingual speech development is highly *dynamic* and characterized by a vivid interaction between a speaker's linguistic systems. The notion of dynamic development also lies at the core of dynamic systems theory (DST), a theoretical approach to language development ([de Bot and Larsen-Freeman 2011](#); [Verspoor et al. 2008](#)). According to DST, which essentially supports an integrated view of bilingualism ([Grosjean 1989](#)), language development progresses in a non-linear, often unpredictable manner, and is determined by an intricate interplay of system-internal and external factors ([de Bot et al. 2007](#)). Longitudinal research, following pronunciation changes in both L2 acquisition and L1 attrition and identifying potential interactions between the two, allows us to gain a better understanding of the dynamics of language development over time.

### 1.2. The Present Study

The present investigation aims to contribute to the expanding body of research on bilingual L2 acquisition and L1 attrition of pronunciation by exploring the segmental speech development of Arnold Schwarzenegger (AS), a late consecutive Austrian German–English bilingual who has been immersed in an L2 English environment for the past 52 years. His early (1979–1988) and late (2012–2018) L1 and L2 segmental speech development was examined based on acoustic analyses of VOT of word-initial plosives (Study 1) and formant frequency analyses of the first (F1) and second (F2) formant of stressed monophthongs (Study 2).

The temporal dimension of VOT, which has previously been described as a reliable acoustic–phonetic correlate of the voiced–voiceless distinction in pre-vocalic plosives ([Abramson and Whalen 2017](#); [Lisker and Abramson 1964](#)), refers to the time lapse between the plosive release and the onset of vocal fold vibration ([Lisker and Abramson 1964](#)). English and Austrian German, the two languages examined in the present investigation, exhibit cross-linguistic differences with regard to their implementation of VOT contrast in word-initial pre-vocalic plosives. English is an aspirating language and distinguishes between long-lag VOT for voiceless aspirated plosives and short-lag VOT for phonologically voiced plosives ([Lisker and Abramson 1964](#)), with approximately 35 milliseconds being the threshold value for the long-lag versus short-lag distinction (e.g., [Keating 1984](#)). Some studies report the occurrence of English pre-voiced targets, that is, vocal fold vibration is maintained throughout the stop closure phase ([Flege and Brown 1982](#); [Lisker and Abramson 1964](#)). English pre-voiced plosives have been observed to occur predominantly in controlled speech production contexts ([Roach 2009](#)), particularly in voiced environments ([Docherty 1992](#)), but overall do not appear to be produced systematically across speakers and thus occur in free variation (e.g., [Flege and Brown 1982](#)).

Unlike Standard German, which—similar to English—distinguishes between long-lag and short-lag plosives (e.g., [Braunschweiler 1997](#); [Jessen 1998](#)), Austrian German does typically not maintain this voiced–voiceless distinction in bilabial and alveolar contexts, despite also being an aspirating language ([Moosmüller et al. 2015](#)). That is, speakers of

Austrian German varieties show a tendency to produce both phonologically voiced and voiceless bilabial and alveolar targets with short-lag VOT and, thus, neutralize the contrast between /b/ and /p/, and /d/ and /t/, particularly in spontaneous conversational speech (Hödl 2019; Moosmüller et al. 2015). By contrast, Austrian German velar plosives, similar to English, are produced within two distinct VOT categories, that is, long-lag VOT for voiceless and short-lag VOT for voiced velar plosives (e.g., Hödl 2019). Based on the different VOT patterns observed in English and Austrian German, Study 1 set out to explore if and to what extent AS's realization of L1 and L2 plosives has changed in the past 40 years of L2 immersion, that is, (1) whether his L2 VOT categories have moved closer to L2 norms and (2) whether his L1 categories have shifted away from L1 production norms and, thus, have become less native.

Previous research on L2 acquisition of VOT has examined L1 speakers of voicing languages, such as French, Spanish or Dutch, acquiring an aspirating language, including English or (Standard) German, as their L2 (e.g., Simon 2009; Stoehr et al. 2017; Thornburgh and Ryalls 1998), or vice versa (e.g., Flege and Hillenbrand 1984). It has been observed that late bilinguals often fail to acquire distinct VOT categories in their L2 resulting from assimilation patterns (Flege 1991; Flege and Hillenbrand 1984) or transferring an L1 feature, such as pre-voicing, to the L2 (Mayr et al. 2012; Simon 2009; Stoehr et al. 2017). Even if bilinguals are able to produce distinct L1 and L2 VOT categories, these often do not resemble those of monolingual speakers of the respective language. Mack (1990), for instance, showed that her L1 French L2 English subject was able to maintain phonetic contrast between English and French /p t k/, but realized the target plosives with VOT values that overshot the monolingual targets in both languages.

While a considerable amount of research has explored the acquisition of late-acquired L2 VOT categories (see above), comparatively few studies so far have investigated attrition of VOT in the L1 of late L2 acquirers being long-term residents in an L2 country (Flege and Hillenbrand 1984; Major 1992; Mayr et al. 2012; Stoehr et al. 2017; Sučková 2020). Overall, findings reveal that VOT is indeed sensitive to be affected by modifications due to L2 learning experience; these modifications, however, have been shown to be more prevalent in voiceless plosives while voiced targets seem to be less likely to undergo attrition (Mayr et al. 2012; Stoehr et al. 2017). The L1 French L2 English-immersed bilinguals in Flege and Hillenbrand (1984), for example, experienced a lengthening of their L1 short-lag VOT category for /t/ in the direction of English long-lag /t/. Similarly, the L1 English L2 French-immersed subjects in Flege (1987) had assimilated their L1 long-lag categories for /t/ to French short-lag VOT (see also Major 1992). Mayr et al. (2012) examined VOT in the productions of monozygotic twin sisters who were both L1 Dutch speakers of L2 English, but one sister had moved to an L2-speaking country in adulthood while the other twin had remained in the L1 environment. Results showed that the VOT categories for voiceless plosives produced by the L2-immersed twin had moved closer to L2 norms, that is, she experienced a lengthening of VOT in her Dutch voiceless plosives. At the same time, her L1 voiced categories remained unaffected by L2-induced changes, i.e., she produced Dutch /b d g/ with consistent pre-voicing, and also pre-voiced her English voiced tokens, indicating an L1 influence on the L2.

Unlike previous investigations examining VOT, the present study does not juxtapose a voicing and an aspirating language, but compares two aspirating languages, with Austrian German featuring a neutralization of the voiced–voiceless distinction in bilabial and alveolar plosive targets (Hödl 2019; Moosmüller et al. 2015). Thus, the bilingual subject in the present study is confronted with the task of acquiring an L2 contrast which is essentially absent in his L1, at least in a bilabial and alveolar place of articulation. Successfully implementing this contrast in his L2 presupposes that he has established a distinct long-lag VOT category for all English aspirated plosives. At the same time, the observation that AS implements a VOT contrast in his late L1 bilabial and alveolar plosives would indicate that his L1 plosive categories are affected by modifications in the direction of his L2.



In order to further investigate AS's segmental speech development, an additional sound class was assessed, namely English and Austrian German stressed monophthongs. In Study 2, AS's monophthongal L1 and L2 vowel space was acoustically examined to determine if and to what extent his productions of L1 and L2 vowels have changed in the past 40 years of L2 immersion. The aim was to find out (1) whether his L2 vowel categories have moved closer to L2 production norms, and (2) whether his L1 categories have shifted away from L2 norms and, thus, have become less native. Austrian German comprises eight front vowels, i.e., /i, y, ɪ, ʏ, e, ø, ε, œ/, and five back vowels, i.e., /u, ʊ, o, ɔ, a/. Unlike Standard German, Austrian German varieties lack the mid-central vowel /ə/, which "exists neither phonetically nor phonologically" (Moosmüller 2007, p. 52). Californian English—the variety the subject of the present study is predominantly exposed to<sup>1</sup>—includes five front vowels, i.e., /i, ɪ, e, ε, æ/, four back vowels, i.e., /u, ʊ, o, a/, and the mid-central vowel /ɜ/(Hagiwara 1997; Ladefoged 2005). One characteristic of Californian English is the merging of /ɔ/ and /a/ in the direction of /ɑ/ (e.g., Boberg 2005).

The extent to which late L2 learners are able to successfully acquire L2 vowel categories has been investigated in a remarkable number of studies, of which just a few are listed here (e.g., Baker and Trofimovich 2005; Flege et al. 1997, 2003; Levy and Law 2010; Oh et al. 2011; Piske et al. 2002). Although it has been shown that the amount of L2 experience is positively correlated with L2 vowel production ability, even highly experienced bilinguals may fail to produce L2 vowels in a native manner (Flege et al. 1997; Levy and Law 2010). This failure to establish accurate L2 vowel categories often stems from assimilatory effects, i.e., L2 vowels are assimilated to acoustically related L1 vowels (Baker and Trofimovich 2005; Flege et al. 2003). Flege et al. (2003) have documented such assimilatory effects in a group of late Italian-English bilinguals whose productions of the L2 English vowel /e'/ were characterized by significantly less formant movement compared to monolingual English productions. Their inaccurate L2 productions resulted from assimilating the L2 vowel target to the acoustically related L1 Italian vowel /e/, which is typically produced with less formant movement compared to English /e'/. Similar observations were made by Baker and Trofimovich (2005). While their early Korean-English bilinguals had established distinct L1 and L2 vowel categories, the late bilinguals' L2 vowels were affected by acoustic features of their L1, that is, they did not manage to produce closely related L1 and L2 vowel targets within separate phonetic categories.

L2-induced articulatory changes in L1 vowel production have received little attention so far (Bergmann et al. 2016; de Leeuw 2019; Mayr et al. 2012). Research shows that late bilinguals' L1 vowels might shift towards L2 norms due to L2 learning experience; however, the extent to which a speaker's native vowel categories are affected by such modifications differs. For example, the L1 vowels produced by the Dutch-English bilingual in Mayr et al. (2012) had shifted closer to L2 production norms, which had manifested itself in an overall more open production of L1 target vowels, as typical of L2 English. Resulting from L1–L2 assimilatory effects, the speaker did not manage to maintain contrast between some of her Dutch and English vowels. These assimilatory processes were, however, not observed to affect all L1–L2 vowels as she was able, for instance, to produce distinct vowel targets for English /ɑ/ and Dutch /a/. These and other findings (Bergmann et al. 2016; de Leeuw 2019) suggest that attrition processes are selective and that not all sounds are equally sensitive to undergo modifications in the direction of the L2 system. It remains unclear, however, why some L1 features are more likely to change while others remain largely stable (Mayr et al. 2012).

Against the background of previous research into L1 attrition and L2 acquisition of plosives and vowels, the two studies conducted in the context of the present investigation aim to reveal potential changes in a bilingual's segmental productions over four decades

<sup>1</sup> Note that *Californian English* is commonly used to refer to different regional/local varieties and sociolects spoken in California (Eckert and Mendoza-Denton 2006). Although AS has been living in California for more than 50 years, it is likely that he was and still is exposed to multiple different English varieties and accents, not only as a result of travelling within and outside the US, but also due to contact with different native and non-native speakers of English.

and thus shed light on some of the processes affecting a speaker's accent in both his L1 and a late-acquired L2.

## 2. Materials and Methods

### 2.1. *The Subject: Arnold Schwarzenegger*

The subject of this case study is Arnold Alois Schwarzenegger (AS), born on July 30, 1947 in Thal, a small rural municipality near the Styrian capital city Graz in Austria. He grew up in a monolingual Austrian German<sup>2</sup> environment and started acquiring English as his L2 when he migrated to the US to seek a career in bodybuilding in 1968. Often described as an “embodiment of the American dream” (Allen 2011), AS did not only gain fame as a bodybuilder, but he is also well known for being an action movie actor, a successful businessman, and a politician who held the office of Governor of California for two terms (2003–2011) (Schwarzenegger 2012).

Although AS had received seven years of formal English instruction in school in Austria, his knowledge of English was rather poor when he arrived in the US at the age of 21 (Outland Baker 2006). Before becoming a permanent resident of Los Angeles, California, in the late 1960s, AS spent some time in London, where he lived with Wag and Dianne Bennett, a British couple who supported AS in the initial stages of his bodybuilding career and helped him work on his English skills (Preston 2015). At this point, however, he had not acquired English as a “functional second language” (Flege and Hillenbrand 1984, p. 710). When he arrived in the US, he struggled particularly with acquiring English pronunciation, which he described as one of the most demanding tasks he was confronted with in the first years after migration (Schwarzenegger 2012). Alongside attending English as a Foreign Language classes at Santa Monica Community College in California, AS strived to improve his English conversation skills by engaging with English-speaking friends on a regular basis (Glaister 2006). Still, his heavily accented English turned out to be problematic when he acted in his first movies in the 1970s and early 1980s, which is why he had to take training sessions with professional dialogue coaches (Miller 2012).

Despite never really losing his distinctive foreign accent, AS has grown confident speaking his L2 on a daily basis over the past 52 years living in an L2 immersion setting. In fact, English has become his dominant means of communication which he makes use of even in L1 settings. He prefers, for example, speaking English when being interviewed by German or Austrian broadcast stations and newspapers (von UsLAR 2012; Ziesel 2018). In interviews, he repeatedly stated that it is much easier for him to speak English than German and that the use of his L1 is restricted predominantly to private contexts (Gala 2015; Gersemann 2009; Naumberger Tageblatt 2015). In a 2017-interview, AS reported that he rarely speaks German and, as a result, his L1 proficiency seems to have declined:

[Interviewer:] How often do you still speak German [*sic*]? After all the years in the USA, do you feel more comfortable to speak English?

[Schwarzenegger:] Not much, I am definitely more comfortable in English. Which should tell you how bad my German has gotten. (Muscle & Fitness 2017)

While some argue that AS, after living in the US for several decades, speaks English “with only a slight accent” (Ramos and Krashen 2013, p. 220), others consider him a perfect example of “embracing his Austrian English accent” (Wan 2017). In fact, his accented English pronunciation is vividly debated in online forums (e.g., Quora 2020; Reddit 2021), has led to many comic imitations (e.g., Collins 2019), and is now considered his personal trademark (Daily Mail UK 2015; Gersemann 2009). However, not only his German-accented English pronunciation attracts attention, but also the question of whether AS has “forgotten” how to speak his L1 seems to be intriguing to the general public (Jackson 2020; Quora 2017). This fascination with AS's accent suggests that people are able to perceive changes in both

<sup>2</sup> In this paper, *Austrian German* is used to refer to the Styrian variety of German spoken in Graz and surrounding rural regions (e.g., Wiesinger 2014).

his L1 and L2 pronunciation. The present study examines whether and to what extent such changes can be evidenced in his segmental speech development over four decades.

## 2.2. Speech Materials

The present investigation is based on spontaneous speech samples which were extracted from interviews with AS, collected from the online video platform YouTube. The interviews were conducted in rather informal settings, primarily including TV and radio talk shows. They covered various topics, such as AS's former career as a bodybuilder, his ongoing career in the film business, childhood memories, his political activities as Governor of California, and his commitment to environmental issues. The earliest publicly available interviews with AS were conducted in 1979, i.e., approximately ten years after he had migrated to the US. The reason for this lack of early interviews presumably is that he had not been widely recognized as a celebrity prior to the 1970s. In addition, the recordings varied in quality, with the earliest recordings made in the 1970s and 80s partly being of rather poor quality. For analysis, only those recordings were included which allowed for an exact identification of acoustic landmarks for measuring AS's plosive and vowel productions.

After downloading the online video files, the sampling frequency was set to 44.1 kHz for each file. Audio sequences which were disturbed by background noise, several speakers speaking at the same time, hesitations or disfluencies were excluded. The resulting speech corpus contained 656 audio files with a total duration of approximately 5 h. The individual recordings were then categorized according to two stages after AS's migration to the US, representing his early (1979–1988) and late (2012–2018) L1 and L2 speech. For his L2 English, a mid-stage was additionally included, containing samples from 1994 to 2003. Recordings representing his L1 German pronunciation for this stage were not available.

Each audio file was transcribed orthographically using the web-application OCTRA (Pömp and Draxler 2017). The audio files in combination with the respective orthographic transcripts were automatically segmented and labelled in WebMaus Basic, a web tool for automatic phonetic and phonological transcription of non-prompted speech (Kisler et al. 2017; Schiel 1999). The resulting segmentations were hand-corrected in Praat (Boersma and Weenink 2018).

## 3. Study 1: Schwarzenegger's Plosives

### 3.1. Analysis

From the speech corpus described above, test tokens used to examine AS's realization of VOT contrast in his L1 and L2 were selected manually in Praat (Boersma and Weenink 2018). Tokens selected for analysis contained the pre-vocalic word-initial plosives /p t k b d g/ in stressed position of monosyllabic and disyllabic content words. This resulted in a total of 3459 plosive tokens which were included in the analysis, with  $N = 224$  German and  $N = 3235$  English tokens. The number of tokens representing AS's L2 pronunciation was considerably higher compared to the number of test tokens obtained for his L1 German. The reason for this was that fewer interviews were available in which AS used his L1 German; as previously mentioned, AS shows a preference to speak his L2 also in L1-settings (e.g., von Uslar 2012).

VOT duration was measured between two manually defined boundaries, namely the burst of the plosive, indicated by a sharp peak in the waveform and a corresponding spectral change, and the start of waveform periodicity marking the onset of the following vowel. In some tokens, the release of the plosive could not be reliably identified due to noise or mumbled speech; these tokens ( $N = 42$ ) were discarded from the analysis.

### 3.2. Results

Study 1 aimed to identify long-term changes in AS's L1 and L2 realization of VOT contrast in word-initial stressed plosives across different stages in time. To compare VOT durations obtained for AS's L1 and L2 plosives within and across languages, we ran a

mixed Anova analysis in R (R Core Team 2020, version 3.6.6), using linear mixed-effects models. A linear mixed-effects approach was considered most appropriate for the present investigation because linear mixed models allow controlling for the potential influence of random effects and manage unbalanced data more easily than, for example, repeated measures Anovas (see, e.g., Barr et al. 2013).

The first model was built to examine changes in AS’s VOT in his L1 German plosives across two stages, including VOT *duration* as the dependent variable and *stage* and *phoneme* as independent variables. An interaction between *stage* and *phoneme* was included as an additional fixed effect, and *word* was included as random intercept. The same specifications were applied to the second model which aimed to examine changes in AS’s VOT in his L2 English over time, using English data only. The third model, which contained *language* as an additional independent variable, was built to compare AS’s VOT in his two languages over time.

The lmerTest package (Kuznetsova et al. 2017) was used in R (R Core Team 2020), including the lmerTest function to obtain *p*-values for *t*-statistics. The mixed models were REML-fitted using Satterthwaite’s approximations to estimate degrees of freedom. Throughout analysis, an  $\alpha$ -level of 0.05 was adopted. Pairwise comparisons were conducted using Tukey’s HSD.

### 3.2.1. L1 German Plosives

Figure 1 depicts the VOT durations obtained for AS’s L1 German plosives in the early and the late stage (an overview of the descriptive statistics for AS’s L1 and L2 plosives is provided in Table A1, Appendix A). It can be seen that—as expected—AS realized a VOT contrast in his early and late velar plosives by producing voiceless targets with considerably longer VOT compared to the voiced counterparts. At the same time, both his voiced and voiceless bilabial and alveolar plosives were produced within a short-lag VOT range in the early stage, that is, AS neutralized contrast in his early productions of these targets. In the late stage, however, a substantial lengthening of VOT was identified in his voiceless alveolar productions while maintaining short-lag VOT in his voiced targets, which shows that he produced a voiced–voiceless distinction in the late stage. Note, however, that for his late German productions of /p/ only a small number of test tokens ( $N = 3$ ) could be identified in the speech samples.

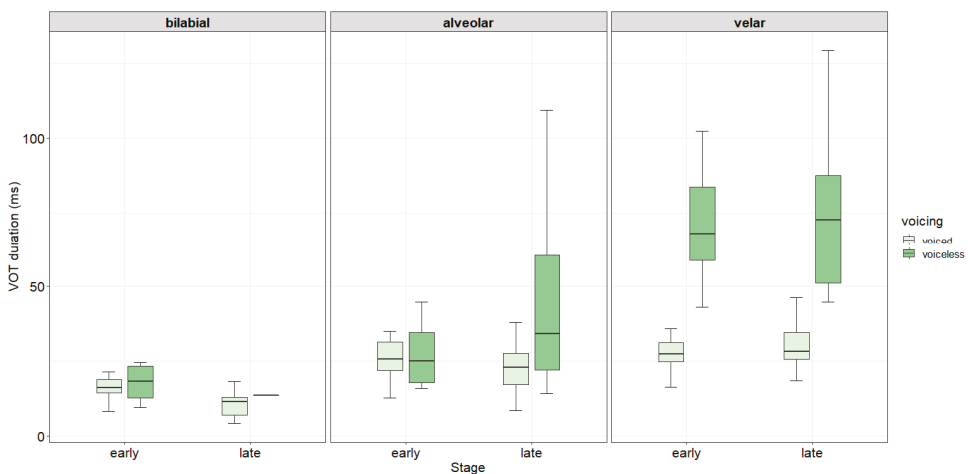


Figure 1. VOT durations (in milliseconds, ms) of AS’s German plosives.

The results of the statistical analysis revealed a main effect for *phoneme* ( $F[5115.4] = 75.6, p < 0.001$ ) and a significant interaction between *stage* and *phoneme* ( $F[5194.6] = 2.4, p < 0.05$ ). Post hoc testing showed a significant difference in VOT duration obtained for AS's voiceless velar plosives in both the early ( $t(139) = -9.75, p < 0.0001$ ) and the late ( $t(153) = -11.48, p < 0.0001$ ) stage. By contrast, no significant differences in VOT duration were observed for his early bilabial and alveolar plosives, which suggests a neutralization of VOT contrast, as depicted in Figure 1. In terms of his late alveolar productions, however, the analysis revealed a significant difference in VOT duration between voiced and voiceless targets ( $t(144) = -4.51, p < 0.001$ ), which indicates a shift in AS's alveolar plosives from neutralizing VOT contrast in the early stage to realizing contrast in the late stage.

### 3.2.2. L2 English Plosives

As depicted in Figure 2, AS showed a tendency to produce a VOT contrast in his L2 for all places of articulation and across all three stages. His voiced plosives covered a short-lag VOT range while his voiceless targets were predominantly produced with long-lag VOT values. However, the broad overlaps between VOT values obtained for his voiced and voiceless plosives suggests that AS did not consistently produce a distinct VOT contrast but was variable in his productions. Furthermore, Figure 2 suggests that his *late* voiceless bilabial and alveolar plosives are characterized by a reduced amount of aspiration—and thus a less native-like production—when comparing his mid and late voiceless bilabials, and his early and late voiceless alveolars.

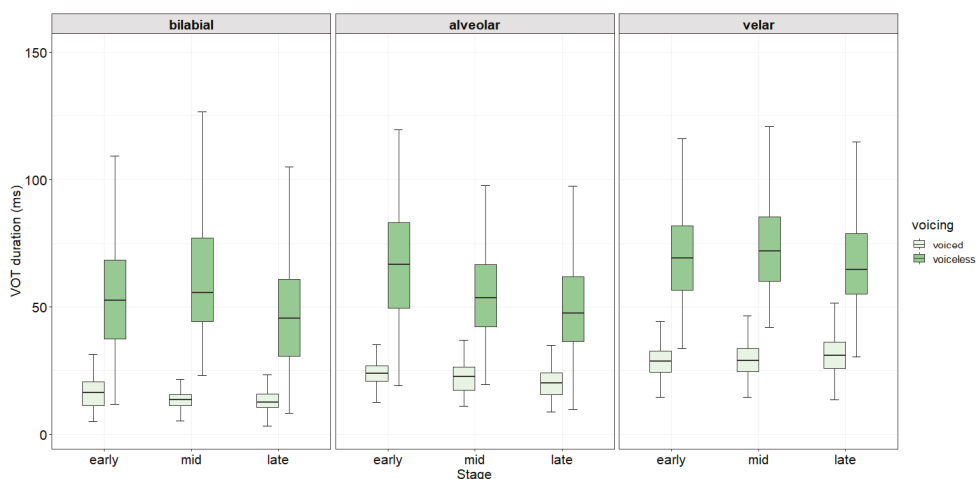


Figure 2. VOT durations (in ms) of AS's English plosives.

The analysis of AS's English VOT showed a main effect for *phoneme* ( $F[5,328] = 296, p < 0.001$ ) and *stage* ( $F[2,97] = 23.2, p < 0.001$ ), as well as significant interaction between *stage* and *phoneme* ( $F[10,97.3] = 8.6, p < 0.001$ ). Post hoc Tukey tests revealed a significant difference in VOT duration for his late and mid voiceless bilabial plosives ( $t(104.8) = -5.21, p < 0.0001$ ). Significant effects were also observed for his voiceless alveolar plosives in the early and the late ( $t(63.5) = 9.14, p < 0.0001$ ) stage, and in the early and mid ( $t(94.9) = 4.86, p < 0.001$ ) stage, which suggests that, as stated above, his voiceless bilabial and alveolar targets were significantly less aspirated in the late stage.

### 3.2.3. Comparison across Languages

Figure 3 compares AS's VOT in his L1 German and L2 English plosives in the early and the late stage. Cross-linguistic differences are most obvious in his early productions

of the bilabial and alveolar targets: While he produced early English /p/ and /t/ with predominantly long-lag VOT, the same plosive targets were realized within a short-lag VOT range in his L1, which resulted in a neutralization of VOT contrast. In the late stage, he was observed to maintain a voiced-voiceless distinction for all L2 targets; at the same time, a tendency to realize a VOT contrast was also identified in his late German productions of the voiceless alveolar plosives, which were characterized by a lengthening of VOT duration in the direction of his L2.

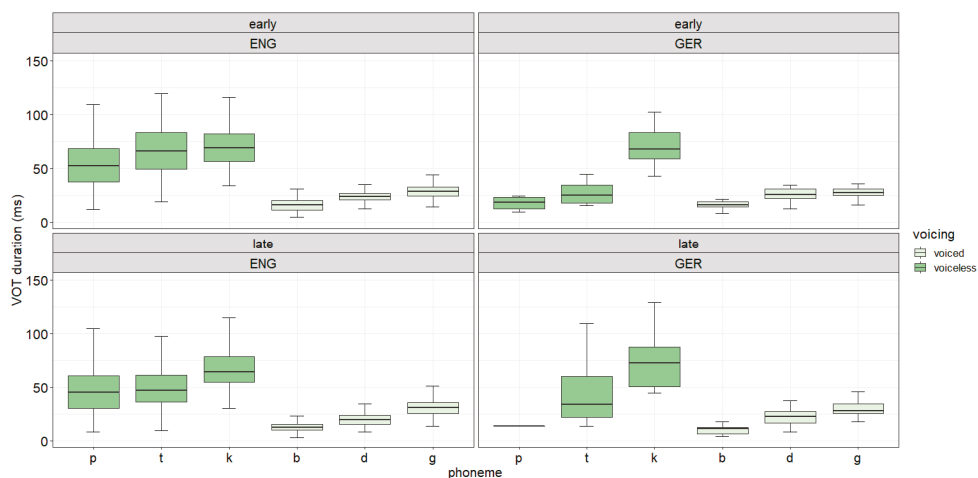


Figure 3. Comparison of VOT durations (in ms) of AS's English (ENG) and German (GER) plosives.

The statistical analysis of AS's L1 and L2 VOT durations showed significant interactions between *stage* and *language* ( $F[1,328.9] = 4.34, p < 0.05$ ), *phoneme* and *language* ( $F[5,255.3] = 12.3, p < 0.001$ ), as well as a three-way interaction between *stage*, *phoneme* and *language* ( $F[5,330.6] = 4.6, p < 0.001$ ). Post hoc results showed a significant difference between AS's English and German early productions of /p/ ( $t(195.1) = 5.96, p < 0.0001$ ) and /t/ ( $t(117.3) = 8.15, p < 0.0001$ ), confirming that, as described above, he produced L1 and L2 voiceless bilabial and alveolar targets within different VOT ranges, respectively. In terms of his late productions, no significant effects were found for /p/ ( $p = 1$ ) and /t/ ( $p = 0.947$ ), which indicates that he realized both English and German targets within a long-lag VOT range.

### 3.3. Discussion

Study 1 aimed to determine if and to what extent AS's productions of L1 Austrian German and L2 English word-initial plosives have changed in the past 40 years of L2 immersion by comparing his L1 and L2 realization of VOT contrast across different stages in time.

The investigation of AS's early L1 German plosives showed a neutralization of VOT contrast in his bilabial and alveolar targets by producing both voiceless and phonologically voiced plosives with short-lag VOT, as commonly observed in Austrian German spontaneous speech (Hödl 2019; Moosmüller et al. 2015). At the same time, he maintained a distinct and native-like VOT contrast in his velar productions, with significantly longer VOT measures obtained for his voiceless velars. In his late German alveolar productions, a significant lengthening of VOT was identified, which indicates a shift of his L1 short-lag VOT categories towards English long-lag VOT. These results confirm the findings of previous studies showing that L2-immersed late bilinguals' voiceless categories are likely to be affected by attrition processes in the direction of L2 norms as a result of L2 learning

experience (Flege and Hillenbrand 1984; Major 1992; Mayr et al. 2012; Stoehr et al. 2017; Sučková 2020).

While neutralizing contrast in his early L1 productions, AS showed a tendency to realize contrast in his L2 productions for all places of articulation and across all three stages. This might have been rather unexpected given that particularly in an early stage of L2 acquisition AS could have experienced L1 influences on the L2 resulting in an inability to produce English voiceless bilabial and alveolar targets with long-lag VOT, as typical of L1 Austrian German. It must be taken into consideration, however, that the recordings representing AS's early English productions were made approximately ten years after migrating to the US, i.e., he had already gained a considerable amount of L2 experience at this point. Moreover, the present investigation showed that his English voiceless plosives were characterized by considerable VOT variability, which suggests that he did not consistently maintain a distinct and native-like contrast between voiced and voiceless L2 targets. Variable L2 VOT productions are frequently observed among late bilinguals (e.g., Flege 1991; Hazan and Boulakia 1993), and can be attributed to different factors, such as diverse L2 input (de Leeuw et al. 2012a) or increased and recent L1 exposure through travelling to an L1-speaking country (de Leeuw 2019; Sancier and Fowler 1997). A further possible explanation for AS's variable L2 productions are potential articulatory constraints particularly affecting the acquisition of aspirated plosives which "require fine temporal coordination to delay the onset of laryngeal vibration relative to oral closure release" (Yu et al. 2015, p. 153). Given that also his late German productions of /t/ were observed to be variable, with some plosives falling in the short-lag VOT range and others in the long-lag range, articulatorily motivated difficulties might have impeded—at least to some extent—the production of consistently aspirated plosives.

Interestingly, AS's late English voiceless plosives were characterized by a significantly reduced amount of aspiration compared to his early and/or mid productions, where no overshooting of monolingual norms was evidenced, which suggests that his L2 productions have moved closer to L1 production norms and have thus become less native in the late stage. Although research shows that mean VOT durations for voiceless plosives typically decrease in elderly speakers as a result of physiological modifications of the vocal tract (e.g., Ryalls et al. 2004; Smith et al. 1987), such biological ageing mechanisms are not likely to have affected AS's late pronunciation of plosives given that biological effects would have resulted in a decrease of VOT in *both* languages. This is, however, not the case since the analysis of AS's late German voiceless targets revealed changes in the opposite direction, as manifested in an *increase* of VOT duration. Instead, the observed shortening of VOT in his late English voiceless plosives and an overall lengthening of VOT in his late German /t/ is indicative of a merging of L1 and L2 categories over time. This assimilation of L1 and L2 VOT categories, as identified in previous acoustic investigations of VOT (e.g., Flege 1987; Major 1992; Mayr et al. 2012), is therefore likely to be the result of cross-linguistic influences affecting pronunciation in both the L1 and the L2.

## 4. Study 2: Schwarzenegger's Vowels

### 4.1. Analysis

To examine AS's vowel space in his L1 and L2, monophthongs occurring in stressed position of monosyllabic and disyllabic content and function words were selected manually from the speech corpus described in Section 2.2 in Praat (Boersma and Weenink 2018). Table 1 depicts the English and German vowels included in the analysis (an overview of the number of tokens included for each vowel target is provided in Table A2, Appendix A). The German vowels /y/, /ʏ/, /ø/ and /œ/ were not included as not enough tokens could be identified in the German audio recordings. In terms of English /ɔ/ and /ɑ/, previous research suggests that speakers of Californian English—the variety AS is predominantly exposed to—tend to neutralize the contrast between these two vowels in the direction of /ɑ/ (e.g., Hagiwara 1997). They were, however, included as separate vowel targets in the present analysis to determine if AS produces distinct vowels. Due to the spontaneous

nature of the speech analyzed in this study, vowel targets included in the analysis occurred in different consonantal contexts. Tokens preceded or followed by the approximants /w/, /l/, /ɹ/, or /j/ were excluded given that their acoustic properties are similar to those of vowels, which may impede the identification of exact measurement points (see Di Paolo et al. 2011).

**Table 1.** German and English vowels included in Study 2.

	German	English		German	English
Front			Back		
/i/	<i>bieten</i> “offer”	<i>bead</i>	/u/	<i>Mus</i> “pulp”	<i>bood</i>
/ɪ/	<i>bitten</i> “request”	<i>bid</i>	/ʊ/	<i>Bus</i> “bus”	<i>book</i>
/ɛ/	<i>Bett</i> “bed”	<i>bed</i>	/ʌ/	–	<i>bud</i>
/e/	<i>Beet</i> “flowerbed”	–	/ɔ/	<i>Motte</i> “moth”	<i>baud</i>
/æ/	–	<i>bad</i>	/ɑ/	<i>satt</i> “sated”	<i>bod</i>
Central					
/ɜ/	–	<i>bird</i>			

The identification of vowel targets resulted in a total of  $N = 262$  German and  $N = 2557$  English vowels which were included in the analysis. In these test tokens, vowel onset and offset were marked manually in Praat (Boersma and Weenink 2018). In plosive, fricative, or affricate contexts, the first glottal striation was determined as the vowel onset, indicating the point where the spectral shape of the formants became visible. Given that the frequency of the first formant in nasal consonants is much lower compared to the F1 frequency of vowels (Ladefoged 2005), the onset and offset of vowels occurring in nasal contexts were marked at the points where the acoustic energy was rising and dropping, respectively.

After determining vowel onset and offset, the frequencies of the first and second formant were measured at the temporal mid-point of the vowel using linear predictive coding in Praat (Boersma and Weenink 2018) with a maximum formant frequency of 5000 Hertz (Hz), a window length of 0.025 s, and a pre-emphasis of 50 Hz. Burg’s algorithm (Childers 1978) was used to extract formant frequencies. While some scholars suggest taking vowel formant measurements at multiple points of the vowel (Di Paolo et al. 2011), the vowel mid-point was chosen as the measurement point to reduce effects of co-articulation (e.g., Reubold et al. 2010).

#### 4.2. Results

Study 2 set out to identify modifications in AS’s L1 and L2 monophthongal vowel space across different stages in time by examining changes in F1 and F2 frequencies over time. Formant frequency measurements of F1 and F2 obtained for his German and English vowels were extracted from Praat (Boersma and Weenink 2018). As in Study 1, linear mixed-effects models were built in R (R Core Team 2020, version 3.6.6), for the same reasons outlined in Section 3.2. The first two models aimed to examine changes in F1 and F2 of AS’s German and English monophthongs, including *F1* and *F2* as dependent variables, respectively, and *stage* and *phoneme* as well as an interaction between the two as fixed effects. *Word* was included as random intercept. The third model explored changes in F1 and F2 comparing AS’s two languages over time and contained *language* as an additional independent variable, otherwise applying the same model specifications outlined above.

In R (R Core Team 2020), the lmerTest package (Kuznetsova et al. 2017) was used, including the lmerTest function to obtain *p*-values for *t*-statistics. The linear mixed models were REML-fitted using Satterthwaite’s approximations to estimate degrees of freedom, adopting an  $\alpha$ -level of 0.05 throughout. Tukey’s HSD was used to conduct pairwise comparisons.



#### 4.2.1. L1 German Vowels

Figure 4 depicts AS’s L1 German vowel space in the early and the late stage (an overview of the descriptive statistics for AS’s L1 and L2 vowels is provided in Table A2, Appendix A). An inspection of the figure shows that his late German vowels /i/, /e/, /u/ and /a/ are characterized by a decrease in F1 compared to his early productions of the same targets, which indicates that these vowel targets have moved to a higher position in the late stage. The back vowel /ʊ/, by contrast, has moved to a more front position in the late stage, as manifested in considerably higher F2 values. A slight decrease in F1 and F2 can also be observed for AS’s production of /ɔ/ in the late stage. Comparatively small changes are evident in AS’s realizations of the front vowels /ɛ/ and /ɪ/.

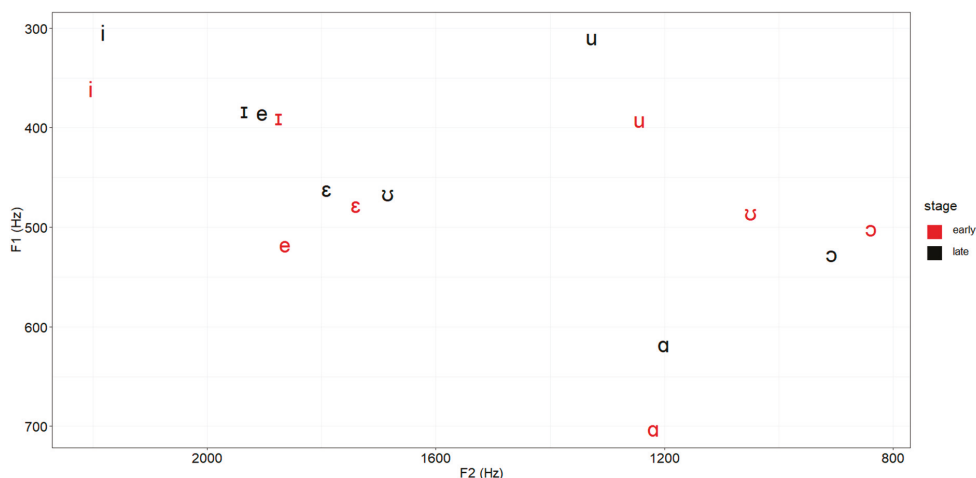


Figure 4. F1–F2 for AS’s German vowels.

The statistical analysis of AS’s German formant frequencies over time revealed significant effects for *stage* ( $F[2,2380] = 8.5, p < 0.001$ ), *phoneme* ( $F[9,523.5] = 250.74, p < 0.0001$ ), and an interaction between *stage* and *phoneme* ( $F[18,2399.2] = 1.76, p < 0.01$ ). Post hoc results showed a significant difference in F1 for /a/ ( $t(246) = 3.77, p < 0.001$ ), /e/ ( $t(231) = 3.47, p < 0.001$ ), and /u/ ( $t(225) = 2.72, p < 0.007$ ) in the early and the late stage, which confirms that AS’s late vowels are produced more close compared to his early vowels. In terms of F2, significant differences were observed for his early and late /ʊ/ ( $t(245) = -8.03, p < 0.0001$ ), with considerably higher F2 values and, thus, a more front production in the late stage.

#### 4.2.2. L2 English Vowels

Figure 5 displays AS’s vowel space in his L2 English across three stages. All of his vowels, with the exception of /u/ and /ʌ/, have moved to a higher position in the late stage, as indicated by lower F1 values in the late compared to the early and/or mid stage. Additionally, a shift in F2 is observable in his productions of /u/ and /ɔ/, with overall lower F2 frequencies and hence a more back production in the late stage. It can also be seen that AS produces distinct vowels for English /a/ and /ɔ/, that is, a merging of these two categories—as typical of Californian English (e.g., Boberg 2005)—is not evident.

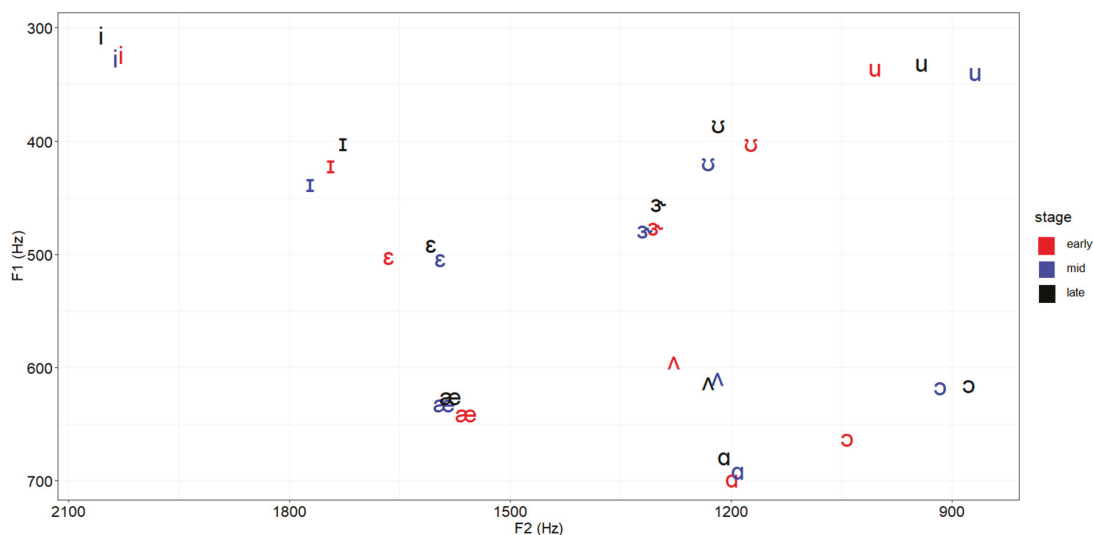


Figure 5. F1~F2 for AS's English vowels.

Results of the statistical analysis showed significant effects for *stage* ( $F[2,2380] = 8.5, p < 0.001$ ), *phoneme* ( $F[9,523.5] = 250.74, p < 0.0001$ ), and an interaction between *stage* and *phoneme* ( $F[18,2399.2] = 1.76, p < 0.01$ ). Post hoc Tukey tests revealed a significant difference in F1 for AS's mid and late productions of the vowel /i/ ( $t(2347) = 3.58, p = 0.001$ ), and for his early and late production of /ɔ/ ( $t(2324) = 2.55, p = 0.03$ ), with overall lower F1 values in the late stage. A significant increase of F2 in the late stage was identified for his early and late /i/ ( $t(102.8) = -2.78, p = 0.022$ ), mid and late /i/ ( $t(102.8) = -3.0, p = 0.009$ ), and early and late /ɔ/ ( $t(85.2) = 4.36, p < 0.001$ ), indicating a shift towards a more front position. Significant effects were also found for F2 in his early and late productions of /ʌ/ ( $t(216.4) = 2.51, p = 0.034$ ), i.e., his late realizations of this target vowel were characterized by a more back production as manifested in a decrease of F2.

#### 4.2.3. Comparison across Languages

Figure 6 compares AS's L1 and L2 vowel space in the early and the late stage. Most notably, a shift of his German front vowel /i/ and German /ɔ/ closer towards the English targets can be identified in the late stage. Similarly, his German production of /ʌ/ has moved closer to his English production of /ʌ/ in the late stage, as manifested in a lowering of F1 for German /a/. By contrast, German and English /ε/ and /ɪ/ have moved further apart in the late stage; the same is true for German /u/ and English /ʊ/, which are nearly identical in the early stage. In the late stage, however, German /u/ is characterized by considerably lower F1 values and has thus shifted away from the English target.

The statistical analysis conducted to compare F1 and F2 across languages and stages revealed a significant effect for *stage* ( $F[1,277.05] = 20.33, p < 0.001$ ) and *phoneme* ( $F[6,420.38] = 260.04, p < 0.001$ ), as well as an interaction between *phoneme* and *language* ( $F[6,422.54] = 7.49, p < 0.001$ ), and a three-way interaction between *stage*, *phoneme* and *language* ( $F[6,260.26] = 2.37, p = 0.03$ ). In the post hoc analysis, significant F1 differences between German and English were identified for AS's late productions of /a/ ( $t(362.7) = 5.1, p < 0.001$ ) and /ɔ/ ( $t(339) = 3.92, p = 0.028$ ), as well as for early /ɔ/ ( $t(258.6) = 5.22, p < 0.001$ ). Significant differences in F2 were identified for AS's late English and German realization of the target vowels /ε/ ( $t(574.8) = -3.91, p = 0.027$ ), /ɪ/ ( $t(585.1) = -4.6, p = 0.0017$ ), /u/ ( $t(574.9) = -5.55, p < 0.001$ ), and /ʊ/ ( $t(686.7) = -6.99, p < 0.001$ ), as manifested in overall lower F2

values identified in his late English productions of these target vowels compared to his German productions.

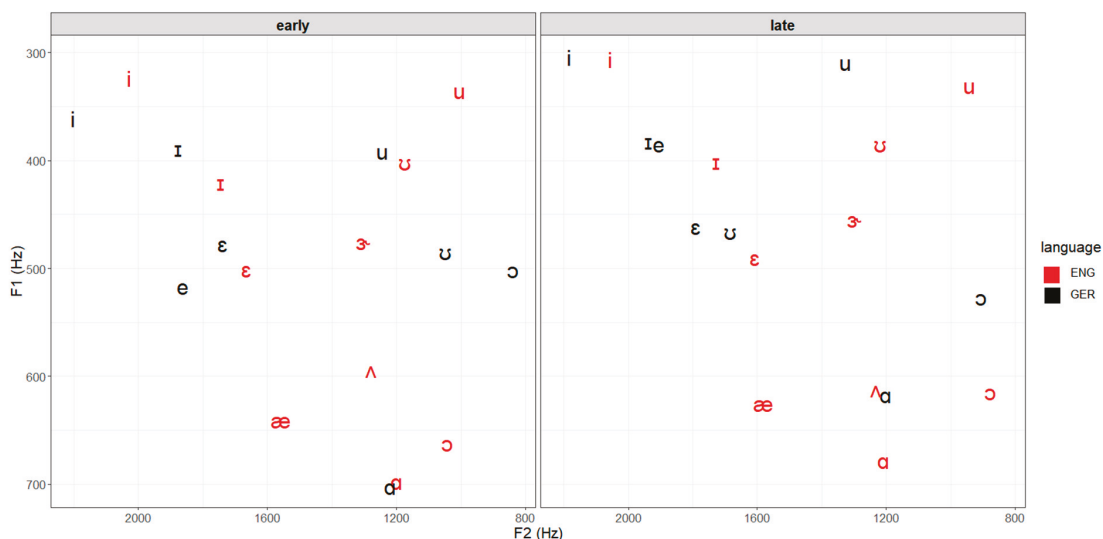


Figure 6. F1~F2 for AS’s German (GER) and English (ENG) vowels.

#### 4.3. Discussion

Study 2 set out to identify potential modifications in AS’s L1 Austrian German and L2 English vowel space across three stages in time. To this end, F1 and F2 of eight German and ten English monophthongs were acoustically examined.

As the analysis revealed, two of AS’s L1 vowels, i.e., /i/ and /a/, were affected by a shift in the direction of the L2, that is, they came to resemble related L2 targets in the late stage, suggesting an influence of the L2 vowel system on the L1. In the case of English /ɔ/, considerable changes in both F1 and F2 were observed in the late stage, with an approximation of the L2 English target closer to the L1 German counterpart. At the same time, some of AS’s L1 vowels, i.e., /u/, /ε/, and /ɪ/, showed changes in the opposite direction, that is, they have moved further away from English targets in the late stage, which might reflect an attempt to enhance contrast between closely related L1 and L2 vowel categories.

Overall, the investigation of AS’s vowels in both of his languages revealed rather diverse modification patterns when comparing his early and late productions, showing that some of his vowels have changed considerably while others exhibited subtle modifications only. In addition, the direction of change was not uniform, with an approximation of L1 and L2 vowel categories, a dispersion of related L1 and L2 categories, and a shift of L2 vowels towards related L1 counterparts. These findings are in line with previous observations concerning the selective nature of L2-induced changes (Bergmann et al. 2016; de Leeuw 2019; Mayr et al. 2012). That is, not all speakers experience modifications in their L1 system to the same extent (Major 1992; Mennen 2004) and even within the same sound class, such modifications are not all-encompassing, as demonstrated in the present investigation. Similarly, in terms of the acquisition of L2 vowel categories, previous research suggests that speakers acquire different L2 vowels with varying degrees of success, often depending on whether an L2 vowel target has a perceptually similar counterpart in the L1 (e.g., Baker and Trofimovich 2005).

Again, one could argue that the changes observed in AS's vowel productions are related to natural ageing mechanisms, which have been shown to lead to a decrease in F1 frequency in elderly speakers (e.g., [Reubold and Harrington 2015, 2017](#)). In fact, an overall decrease in F1 was also identified in some of AS's German and English vowels in the late stage. However, as addressed in Section 3.3, changes resulting from biological ageing processes are not expected to be selective, that is, to have an effect on some vowels only. Hence, the influence of biological mechanisms does not offer a convincing explanation for the modifications observed in AS's German and English vowels. Instead, the changes in AS's vowel productions over time seem to be indicative of complex system-internal mechanisms, which result in various modification patterns. The specific reasons, however, why some vowel categories are affected differently or to a greater extent than others are still to be uncovered.

In order to gain further insight into the extent to which AS's vowels have been affected by bi-directional L1–L2 influences, it would be interesting for future research to also explore the compactness of his vowels over time to identify if the acoustic stability of his vowel targets has changed in the course of L2 immersion (see, e.g., [Kartushina et al. 2016; Kartushina and Martin 2019](#)).

## 5. Overall Discussion

The aim of the present investigation was to trace developments in the L1 and L2 segmental speech production in the late consecutive bilingual Arnold Schwarzenegger over a period of four decades. The findings of Study 1, focusing on AS's realization of VOT contrast in German and English plosives, revealed an assimilation of his late L1 and L2 categories for voiceless targets, which was particularly visible in his alveolar plosives. That is, his short-lag L1 productions of /t/ were characterized by significantly longer and thus more English-like VOT durations in the late stage. At the same time, his late L2 targets were produced with considerably less aspiration, which suggests a drift away from native English production norms closer to L1 short-lag norms. The findings of Study 2, exploring AS's L1 and L2 monophthongal vowel space, to some extent reflect the results of Study 1. In both studies, the changes observed cannot be reliably explained against the background of age-related biological mechanisms, and both investigations provide evidence for a merging of L1 and L2 categories in AS's late productions, which supports one of the main tenets of the Speech Learning Model that closely related L1 and L2 categories may come to resemble each other due to assimilatory effects ([Flege 1995; Flege and Bohn 2020](#)). Study 2 further showed that a bilingual's segmental productions are not necessarily equally affected by L2-induced changes ([Mayr et al. 2012](#)) and that some changes are relatively subtle ([Bergmann et al. 2016; Chang 2012](#)).

Taken together, the present findings confirm that “cross-linguistic transfer is not a one-way street” ([Schmid and Köpcke 2017](#), p. 637) in that the use and development of a late-acquired L2 system can indeed exert influence on a speaker's L1 accent, in the same way as the L1 system affects pronunciation abilities in the L2. In this respect, the findings contradict the notion of an impermeable, invariable L1 system and thus challenge a static view on bilingualism ([Lado 1957; Lenneberg 1967](#)). Instead, they support a dynamic systems-oriented approach to language development, according to which mutual L1–L2 interactions, sensitivity to system-internal and external influences, and sometimes even unpredictability are inherent characteristics of bilingual development (e.g., [de Bot and Larsen-Freeman 2011](#)). This approach considers intra-individual variability in bilingual productions as valuable evidence for the dynamic nature of linguistic development, arguing that “both free and systematic variability will be relatively high when the system is reorganizing” ([Verspoor et al. 2008](#), p. 216). As outlined above, a restructuring and reorganization of a speaker's pronunciation system(s) might be internally motivated, that is, resulting from L1–L2 interactions over time, but also external—social, environmental, and personal—factors can trigger and shape such modification processes ([de Bot et al. 2007; Verspoor et al. 2008](#)).

While our findings are indicative of L1–L2 changes that are most likely internally driven, it might be argued that external factors may have also played a role. For instance, AS’s pronunciation may have been affected by the interview situation leading to stress-induced changes to his pronunciation and a possible adaption to the accent, speaking style and other linguistic features of his interlocutors (e.g., Giles and Ogay 2007). However, we expect this influence to be minimal given that over the years AS has had ample experience conducting interviews and he had been exposed to the speech of many different interviewers, which will have reduced the adaption to an individual’s pronunciation. Another disadvantage of the use of spontaneous speech samples is that it is likely to entail variations in speaking rate, which in turn can influence features of pronunciation, including VOT duration and vowel production (Kessinger and Blumstein 1998). However, controlling for speaking rate has also been found to be problematic in experimental settings given that speakers often have different perceptions of what fast and slow speech is (Harrington 2010). Another possible external factor that might have influenced the results is AS’s recent and enhanced L1 exposure through travelling to his home country Austria, which has previously been shown to have an impact on a speaker’s accent (Sancier and Fowler 1997). Furthermore, potential changes in his private and professional environment might have contributed to the changes and variability observed in his L1 and L2 speech production (see, e.g., Schoonmaker-Gates 2015).

Despite the limitations of using spontaneous speech data outlined above, the present investigation offers a rare insight into the longitudinal development of a late bilingual’s L1 and L2 speech production over a period of 40 years. The investigation thus offers useful insights into the dynamic nature of bilingual speech development and sheds light on the complexity of L1–L2 interaction processes affecting a speaker’s pronunciation. To gradually comprehend this complexity and to fully understand how and to what extent system-internal processes are intertwined with social and environmental factors, further long-term investigations focusing on additional segmental and prosodic variables in both individuals and groups of speakers will be necessary.

**Author Contributions:** Conceptualization, L.K. and I.M.; methodology, L.K. and I.M.; formal analysis, L.K.; writing—original draft preparation, L.K. and I.M.; writing—review and editing, I.M. and L.K.; visualization, L.K.; supervision, I.M. All authors have read and agreed to the published version of the manuscript.

**Funding:** Open Access Funding by the University of Graz.

**Institutional Review Board Statement:** The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Institutional Ethics Committee of the University of Graz (protocol code: GZ. 39/1/63 ex 2019/20; date of approval: 12 November 2019).

**Informed Consent Statement:** Participant consent was waived due to the use and analysis of publicly available speech material. Using the materials for scientific purposes did not entail copyright issues according to the Austrian Copyright Act (UrhG) §42(2).

**Data Availability Statement:** The data are available on request from the corresponding author. The data are not publicly available due to ongoing data analyses. Data will be made available once all analyses have been completed.

**Acknowledgments:** We would like to thank Robert Mayr and the two anonymous reviewers for their helpful comments and feedback. We also thank Ulrich Reubold for his valuable statistical advice, and Kerstin Endes and Matthias Wedenig for their assistance during the data labelling process.

**Conflicts of Interest:** The authors declare no conflict of interest.

### Appendix A

**Table A1.** Means, standard deviations and medians for AS’s VOT durations (reported in milliseconds) obtained for his English and German plosives across different stages in time, including the number of tokens (N) measured for each plosive according to language and stage.

English early				English mid		
<i>Plosive</i>	<i>N</i>	<i>Mean (SD)</i>	<i>Median</i>	<i>N</i>	<i>Mean (SD)</i>	<i>Median</i>
<b>p</b>	69	54.05 (23.33)	52.47	76	61.36 (25.05)	55.34
<b>b</b>	90	15.98 (5.9)	16.4	123	13.81 (3.93)	13.46
<b>t</b>	117	67.58 (22.89)	66.33	103	56.4 (20.13)	53.42
<b>d</b>	97	24.01 (6.02)	23.72	118	22.17 (5.74)	22.67
<b>k</b>	89	69.8 (18.04)	68.96	103	72.75 (17.27)	71.64
<b>g</b>	118	29.35 (7.42)	28.49	100	29.75 (7.3)	28.87
English late				German early		
<i>Plosive</i>	<i>N</i>	<i>Mean (SD)</i>	<i>Median</i>	<i>N</i>	<i>Mean (SD)</i>	<i>Median</i>
<b>p</b>	313	48.22 (24.36)	45.46	4	17.62 (7.27)	18.21
<b>b</b>	403	13.32 (4.7)	12.49	16	16.49 (4.64)	16.07
<b>t</b>	373	50.07 (20.46)	47.42	8	27.46 (11.4)	24.96
<b>d</b>	266	20.83 (7.09)	20.06	15	26.13 (6.29)	25.54
<b>k</b>	287	68.33 (19.37)	64.53	24	70.78 (16.68)	67.97
<b>g</b>	390	31.13 (7.62)	30.85	27	27.70 (4.89)	27.3
German late						
<i>Plosive</i>	<i>N</i>	<i>Mean (SD)</i>	<i>Median</i>			
<b>p</b>	3	16.81 (6.55)	13.6			
<b>b</b>	25	11.16 (4.85)	11.27			
<b>t</b>	12	44.67 (29.54)	34.08			
<b>d</b>	45	22.71 (7.62)	22.76			
<b>k</b>	22	73.06 (22.15)	72.56			
<b>g</b>	25	29.84 (6.28)	28.22			

**Table A2.** Means, standard deviations and medians of F1 and F2 (reported in Hertz) obtained for AS’s English and German monophthongs across different stages in time, including the number of tokens (N) measured for each monophthong according to language and stage.

English early					
<i>Vowel</i>	<i>N</i>	F1		F2	
		<i>Mean (SD)</i>	<i>Median</i>	<i>Mean (SD)</i>	<i>Median</i>
/i/	98	323.6 (60.24)	305.1	2028.1 (135.9)	2018.0
/l/	130	420.2 (69.66)	410.3	1743.2 (118.3)	1752.6
/ɛ/	105	500.6 (77.41)	485.4	1665.2 (188.0)	1667.9
/u/	35	334.9 (45.38)	328.5	1003.6 (267.0)	950.8
/ʊ/	35	400.9 (68.98)	403.5	1172.9 (211.76)	1181.7
/ɔ/	14	661.5 (69.7)	651.1	1042.2 (209.91)	975.1
/a/	84	697.2 (83.84)	707.0	1197.5 (109.31)	1182.6

**Table A2.** *Cont.*

/æ/	98	639.9 (99.28)	650.9	1559.4 (207.47)	1545.2
/ʌ/	86	593.1 (106.16)	601.7	1277.9 (104.57)	1265.8
/ɜ:/	37	474.7 (43.34)	462.7	1302.5 (76.26)	1300.9
<b>English mid</b>					
		F1		F2	
<i>Vowel</i>	<i>N</i>	<i>Mean (SD)</i>	<i>Median</i>	<i>Mean (SD)</i>	<i>Median</i>
/i/	74	326.3 (43.04)	320.0	2035.5 (252.94)	2041.8
/ɪ/	120	436.4 (77.18)	429.1	1771.1 (100.58)	1779.0
/ɛ/	61	502.7 (51.58)	503.3	1595.3 (142.56)	1599.6
/u/	24	338.5 (43.05)	340.3	867.8 (246.98)	778.9
/ʊ/	21	418.0 (50.3)	412.6	1231.2 (209.24)	1236.7
/ɔ/	7	615.4 (71.73)	594.9	915.4 (78.48)	936.1
/ɑ/	69	690.0 (55.31)	696.6	1190.3 (129.17)	1191.7
/æ/	101	630.6 (91.12)	659.9	1589.1 (194.24)	1534.0
/ʌ/	92	607.8 (52.27)	610.3	1218.2 (82.09)	1228.8
/ɜ:/	39	477.1 (42.31)	476.3	1317.4 (74.51)	1323.9
<b>English late</b>					
		F1		F2	
<i>Vowel</i>	<i>N</i>	<i>Mean (SD)</i>	<i>Median</i>	<i>Mean (SD)</i>	<i>Median</i>
/i/	111	306.0 (41.32)	301.9	2054.8 (117.79)	2040.1
/ɪ/	217	400.3 (60.58)	390.5	1727.2 (125.9)	1737.4
/ɛ/	192	489.8 (52.2)	485.9	1607.9 (115.03)	1598.78
/u/	39	330.0 (75.71)	317.9	940.6 (438.66)	799.6
/ʊ/	32	384.6 (39.41)	381.5	1217.6 (257.99)	1272.8
/ɔ/	32	614.1 (78.2)	602.3	876.5 (73.74)	882.4
/ɑ/	108	677.8 (73.99)	681.2	1208.2 (140.06)	1193.6
/æ/	219	624.5 (99.86)	648.2	1580.6 (169.3)	1535.2
/ʌ/	203	611.1 (57.01)	609.0	1231.2 (88.65)	1229.4
/ɜ:/	74	454.0 (36.82)	448.1	1298.6 (85.71)	1292.07
<b>German early</b>					
		F1		F2	
<i>Vowel</i>	<i>N</i>	<i>Mean (SD)</i>	<i>Median</i>	<i>Mean (SD)</i>	<i>Median</i>
/i/	7	360.8 (34.69)	352.7	2203.3 (44.15)	2198.1
/ɪ/	23	388.7 (59.77)	404.0	1875.7 (145.09)	1879.3
/ɛ/	15	476.8 (66.88)	437.4	1739.8 (107.46)	1722.6
/u/	8	391.3 (27.85)	381.4	1242.7 (366.13)	1122.3
/ʊ/	7	484.5 (181.51)	461.2	1048.4 (514.96)	880.4
/ɔ/	32	500.6 (55.08)	490.4	837.6 (52.06)	823.8
/ɑ/	20	701.6 (76.09)	679.7	1218.8 (135.32)	1192.5
/e/	6	516.6 (14.59)	518.4	1863.5 (115.1)	1852.4

Table A2. Cont.

German late		F1		F2	
Vowel	N	Mean (SD)	Median	Mean (SD)	Median
/i/	9	303.8 (16.14)	302.1	2182.1 (58.89)	2199.8
/ɪ/	31	381.4 (35.03)	387.0	1935.5 (110.93)	1964.2
/e/	23	460.7 (30.86)	469.3	1791.9 (122.85)	1799.1
/u/	10	308.7 (27.44)	311.9	1326.7 (227.47)	1398.7
/ʊ/	6	465.2 (91.02)	483.1	1684.3 (129.35)	1645.0
/ɔ/	21	526.1 (52.32)	534.8	906.5 (118.46)	865.2
/a/	39	617.1 (76.84)	598.5	1201.0 (99.85)	1209.4
/ɛ/	5	384.3 (18.81)	390.9	1904.1 (43.14)	1908.8

## References

- Abramson, Arthur S., and D. H. Whalen. 2017. Voice Onset Time (VOT) At 50: Theoretical and Practical Issues in Measuring Voicing Distinctions. *Journal of Phonetics* 63: 75–86. [CrossRef]
- Allen, Nick. 2011. Arnold Schwarzenegger: From Austrian Immigrant to Superstar to Politician. Available online: <https://www.telegraph.co.uk/culture/tvandradio/8419751/Arnold-Schwarzenegger-from-Austrian-immigrant-to-superstar-to-politician.html> (accessed on 30 September 2020).
- Baker, Wendy, and Pavel Trofimovich. 2005. Interaction of Native- and Second-Language Vowel System(S) In Early and Late Bilinguals. *Language and Speech* 48: 1–27. [CrossRef]
- Barr, Dale J., Roger Levy, Christoph Scheepers, and Harry J. Tily. 2013. Random Effects Structure for Confirmatory Hypothesis Testing: Keep It Maximal. *Journal of Memory and Language* 68: 255–78. [CrossRef] [PubMed]
- Bergmann, Christopher, Amber Nota, Simone A. Sprenger, and Monika S. Schmid. 2016. L2 Immersion Causes Non-Native-Like L1 Pronunciation in German Attriters. *Journal of Phonetics* 58: 71–86. [CrossRef]
- Birdsong, David. 2009. Age and the End State of Second Language Acquisition. In *The New Handbook of Second Language Acquisition*. Edited by William C. Ritchie and Tej K. Bhatia. Amsterdam: Elsevier, pp. 401–24.
- Bloomfield, Leonard. 1933. *Language*. New York: Holt.
- Boberg, Charles. 2005. The Canadian Shift in Montreal. *Language Variation and Change* 17: 133–54. [CrossRef]
- Boersma, Paul, and David Weenink. 2018. Praat: Doing Phonetics by Computer (Version 6.0.37). Computer Program. Available online: <http://www.fon.hum.uva.nl/praat> (accessed on 14 March 2018).
- Bongaerts, Theo, Susan Mennen, and Frans van der Slik. 2000. Authenticity of Pronunciation in Naturalistic Second Language Acquisition: The Case of Very Advanced Late Learners of Dutch as a Second Language. *Studia Linguistica* 54: 298–308. [CrossRef]
- Braunschweiler, Norbert. 1997. Integrated Cues of Voicing and Vowel Length in German: A Production Study. *Language and Speech* 40: 353–76. [CrossRef]
- Bylund, Emanuel. 2009. Maturation Constraints and First Language Attrition. *Language Learning* 59: 687–715. [CrossRef]
- Chang, Charles B. 2012. Rapid and Multifaceted Effects of Second-Language Learning on First-Language Speech Production. *Journal of Phonetics* 40: 249–68. [CrossRef]
- Chang, Charles B. 2013. A Novelty Effect in Phonetic Drift of the Native Language. *Journal of Phonetics* 41: 520–33. [CrossRef]
- Chang, Charles B. 2019. Language Change and Linguistic Inquiry in a World of Multicompetence: Sustained Phonetic Drift and Its Implications for Behavioral Linguistic Research. *Journal of Phonetics* 74: 96–113. [CrossRef]
- Childers, Donald G. 1978. *Modern Spectrum Analysis*. New York: IEEE Press.
- Collins, Ben. 2019. This Viral Schwarzenegger Deepfake Isn't Just Entertaining. It's a Warning. Available online: <https://www.nbcnews.com/tech/tech-news/viral-schwarzenegger-deepfake-isn-t-just-entertaining-it-s-warning-n1016851> (accessed on 25 January 2021).
- Daily Mail UK. 2015. Arnold Schwarzenegger Reveals He Can Speak Perfect English—But Keeps Talking with an Accent Because 'Fans Expect It'. Available online: <http://www.dailymail.co.uk/tvshowbiz/article-3141778/Arnold-Schwarzenegger-reveals-speak-perfect-English-keeps-talking-accent-fans-expect-it.html> (accessed on 25 January 2021).
- de Bot, Kees, and Diane Larsen-Freeman. 2011. Researching Second Language Development from a Dynamic Systems Theory Perspective. In *A Dynamic Approach to Second Language Development: Methods and Techniques*. Edited by Marjolijn Verspoor and Kees de Bot and Wander Lowie. Amsterdam: John Benjamins, pp. 5–24.
- de Bot, Kees, Wander Lowie, and Marjolijn Verspoor. 2007. A Dynamic Systems Theory Approach to Second Language Acquisition. *Bilingualism: Language and Cognition* 10: 7–21. [CrossRef]



- de Leeuw, Esther. 2019. Native Speech Plasticity in the German-English Late Bilingual Stefanie Graf: A Longitudinal Case Study over Four Decades. *Journal of Phonetics* 73: 24–39. [CrossRef]
- de Leeuw, Esther, Monika S. Schmid, and Ineke Mennen. 2010. The Effects of Contact on Native Language Pronunciation in an L2 Migrant Setting. *Bilingualism: Language and Cognition* 13: 33–40. [CrossRef]
- de Leeuw, Esther, Ineke Mennen, and James M. Scobbie. 2012a. Dynamic Systems, Maturational Constraints, and L1 Phonetic Attrition. *International Journal of Bilingualism* 17: 683–700. [CrossRef]
- de Leeuw, Esther, Ineke Mennen, and James M. Scobbie. 2012b. Singing a Different Tune in Your Native Language: First Language Attrition of Prosody. *International Journal of Bilingualism* 16: 101–16. [CrossRef]
- de Leeuw, Esther, Aurela Tusha, and Monika S. Schmid. 2017. Individual Phonological Attrition in Albanian-English Late Bilinguals. *Bilingualism: Language and Cognition* 21: 278–95. [CrossRef]
- Di Paolo, Marianna, Malcah Yaeger-Dror, and Alicia Beckford Wassnik. 2011. Analyzing Vowels. In *Sociophonetics: A Student's Guide*. Edited by Marianna Di Paolo and Malcah Yaeger-Dror. New York: Routledge, pp. 87–106.
- Dmitrieva, Olga, Allard Jongman, and Joan Sereno. 2010. Phonological Neutralization by Native and Non-Native Speakers: The Case of Russian Final Devoicing. *Journal of Phonetics* 38: 483–92. [CrossRef]
- Docherty, Gerard J. 1992. *The Timing of Voicing in British English Obstruents*. Berlin and New York: Foris.
- Eckert, Penelope, and Norma Mendoza-Denton. 2006. Getting Real in the Golden State (California). In *American Voices: How Dialects Differ from Coast to Coast*. Edited by Walt Wolfram and Ben Ward. Malden: Blackwell Publishing, pp. 139–43.
- Flege, James E. 1980. Phonetic Approximation in Second Language Acquisition. *Language Learning* 30: 117–34. [CrossRef]
- Flege, James E. 1981. The Phonological Basis of Foreign Accent: A Hypothesis. *TESOL Quarterly* 15: 443–55. [CrossRef]
- Flege, James E. 1987. The Production of 'New' and 'Similar' Phones in a Foreign Language: Evidence for the Effect of Equivalence Classification. *Journal of Phonetics* 15: 47–65. [CrossRef]
- Flege, James E. 1991. Age of Learning Affects the Authenticity of Voice-onset Time (VOT) In Stop Consonants Produced in a Second Language. *The Journal of the Acoustical Society of America* 89: 395–411. [CrossRef] [PubMed]
- Flege, James E. 1995. Second Language Speech Learning: Theory, Findings, and Problems. In *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*. Edited by Winifred Strange. Baltimore: York Press, pp. 233–72.
- Flege, James E., and Ocke-Schwen Bohn. 2020. The Revised Speech Learning Model. Unpublished Preprint. Available online: [https://www.researchgate.net/publication/348265785\\_The\\_revised\\_Speech\\_Learning\\_Model\\_SLM-r](https://www.researchgate.net/publication/348265785_The_revised_Speech_Learning_Model_SLM-r) (accessed on 28 September 2020).
- Flege, James E., and W. S. Brown. 1982. The Voicing Contrast Between English /p/ and /b/ as a Function of Stress and Position-in-Utterance. *Journal of Phonetics* 10: 335–45. [CrossRef]
- Flege, James E., and Wieke Eefting. 1988. Imitation of a VOT Continuum by Native Speakers of English and Spanish: Evidence for Phonetic Category Formation. *The Journal of the Acoustical Society of America* 83: 729–40. [CrossRef] [PubMed]
- Flege, James E., and James Hillenbrand. 1984. Limits on Phonetic Accuracy in Foreign Language Speech Production. *The Journal of the Acoustical Society of America* 76: 708–21. [CrossRef]
- Flege, James E., Murray J. Munro, and Ian R. A. MacKay. 1996. Factors Affecting the Production of Word-Initial Consonants in a Second Language. In *Second Language Acquisition and Linguistic Variation*. Edited by Robert Bayley and Dennis R. Preston. Amsterdam: John Benjamins, pp. 47–74.
- Flege, James E., Ocke-Schwen Bohn, and Sunyoung Jang. 1997. Effects of Experience on Non-Native Speakers' Production and Perception of English Vowels. *Journal of Phonetics* 25: 437–70. [CrossRef]
- Flege, James E., Carlo Schirru, and Ian R. A. MacKay. 2003. Interaction between the Native and Second Language Phonetic Subsystems. *Speech Communication* 40: 467–91. [CrossRef]
- Gala. 2015. Der 'Terminator'-Star Ist Zurück. Available online: <https://www.gala.de/stars/news/interview/arnold-schwarzenegger-der-terminator-star-ist-zurueck-20234970.html> (accessed on 25 January 2021).
- Gersemann, Olaf. 2009. Der Mann, Der Sich Selbst Erfunden Hat. Available online: <http://www.welt.de/politik/ausland/article4895817/Der-Mann-der-sich-selbst-erfunden-hat.html> (accessed on 25 January 2021).
- Giles, Howard, and Tania Ogay. 2007. Communication Accommodation Theory. In *Explaining Communication: Contemporary Theories and Exemplars*. Edited by Bryan B. Whaley and Wendy Samter. Mahwah: Lawrence Erlbaum, pp. 293–310.
- Glaister, Dan. 2006. California Split over Teaching: English Classes for Hispanic Students Branded as 'Return to Segregation'. Available online: <https://www.theguardian.com/education/2006/aug/11/tefl.danglaister> (accessed on 25 January 2021).
- Grosjean, François. 1989. Neurolinguists, Beware! The Bilingual Is Not Two Monolinguals in One Person. *Brain and Language* 36: 3–15. [CrossRef]
- Grosjean, François. 1997. The Bilingual Individual. *Interpreting* 2: 163–87. [CrossRef]
- Hagiwara, Robert. 1997. Dialect Variation and Formant Frequency: The American English Vowels Revisited. *The Journal of the Acoustical Society of America* 102: 655–58. [CrossRef]
- Harrington, Jonathan. 2010. *Phonetic Analysis of Speech Corpora*. Malden: Wiley-Blackwell.
- Hazan, Valerie L., and Georges Boulakia. 1993. Perception and Production of a Voicing Contrast by French-English Bilinguals. *Language and Speech* 36: 17–38. [CrossRef]
- Hödl, Petra. 2019. Production and Perception of Voice Onset Time in Austrian German. Ph.D. thesis, Karl-Franzens-University, Graz, Austria.

- Hopp, Holger, and Monika S. Schmid. 2013. Perceived Foreign Accent in First Language Attrition and Second Language Acquisition: The Impact of Age of Acquisition and Bilingualism. *Applied Psycholinguistics* 34: 361–94. [CrossRef]
- Jackson, Betty. 2020. Does Arnold Schwarzenegger Speak German? Available online: <https://celebanswers.com/does-arnold-schwarzenegger-speak-german/> (accessed on 25 January 2021).
- Jessen, Michael. 1998. *Phonetics and Phonology of Tense and Lax Obstruents in German*. Studies in Functional and Structural Linguistics 44. Amsterdam: John Benjamins.
- Kartushina, Natalia, and Clara D. Martin. 2019. Third-Language Learning Affects Bilinguals' Production in Both Their Native Languages: A Longitudinal Study of Dynamic Changes in L1, L2 and L2 Vowel Production. *Journal of Phonetics* 77: 100920. [CrossRef]
- Kartushina, Natalia, Alexis Hervais-Adelman, Ulrich H. Frauenfelder, and Narly Golestani. 2016. Mutual Influences Between Native and Non-Native Vowels in Production: Evidence from Short-Term Visual Articulatory Feedback Training. *Journal of Phonetics* 57: 21–39. [CrossRef]
- Keating, Patricia A. 1984. Phonetic and Phonological Representation of Stop Consonant Voicing. *Language* 60: 286–319. [CrossRef]
- Kessinger, Rachel H., and Sheila E. Blumstein. 1998. Effects of Speaking Rate on Voice-Onset Time and Vowel Production: Some Implications for Perception Studies. *Journal of Phonetics* 26: 117–28. [CrossRef]
- Kisler, Thomas, Uwe D. Reichel, and Florian Schiel. 2017. Multilingual Processing of Speech via Web Services. *Computer Speech and Language* 45: 326–47. [CrossRef]
- Köpke, Barbara, and Monika S. Schmid. 2004. Language Attrition: The Next Phase. In *First Language Attrition: Interdisciplinary Perspectives on Methodological Issues*. Edited by Monika S. Schmid, Barbara Köpke and Merel Keijzer and Lina Weilemar. Amsterdam: John Benjamins, vol. 28, pp. 1–43.
- Kuznetsova, Alexandra, Per B. Brockhoff, and Rune H. B. Christensen. 2017. lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software* 82: 1–26. [CrossRef]
- Ladefoged, Peter. 2005. *Vowels and Consonants*. Malden: Blackwell Publishing.
- Lado, Robert. 1957. *Linguistics across Cultures. Applied Linguistics for Teachers*. Ann Arbor: University of Michigan Press.
- Lenneberg, Eric H. 1967. *Biological Foundations of Language*. New York: John Wiley and Sons.
- Levy, Erika S., and Franzo F. Law. 2010. Production of French Vowels by American-English Learners of French: Language Experience, Consonantal Context, and the Perception-Production Relationship. *The Journal of the Acoustical Society of America* 128: 1290–1305. [CrossRef]
- Lisker, Leigh, and Arthur S. Abramson. 1964. A Cross-Language Study of Voicing in Initial Stops: Acoustical Measurements. *Word* 20: 384–422. [CrossRef]
- Mack, Molly. 1990. Phonetic Transfer in a French-English Bilingual Child. In *Language Attitudes and Language Conflict. Spracheinstellungen Und Sprachkonflikte*. Edited by Peter H. Nelde. Plurilingua IX. Bonn: Dümmler, pp. 107–24.
- MacKay, Ian R. A., James E. Flege, Thorsten Piske, and Carlo Schirru. 2001. Category Restructuring During Second-Language Speech Acquisition. *The Journal of the Acoustical Society of America* 110: 516–28. [CrossRef]
- Major, Roy C. 1992. Losing English as a First Language. *The Modern Language Journal* 76: 190–208. [CrossRef]
- Mayr, Robert, Sacha Price, and Ineke Mennen. 2012. First Language Attrition in the Speech of Dutch–English Bilinguals: The Case of Monozygotic Twin Sisters. *Bilingualism: Language and Cognition* 15: 687–700. [CrossRef]
- Mayr, Robert, David Sánchez, and Ineke Mennen. 2020. Does Teaching Your Native Language Abroad Increase L1 Attrition of Speech? The Case of Spaniards in the United Kingdom. *Languages* 5: 41. [CrossRef]
- Mennen, Ineke. 2004. Bi-Directional Interference in the Intonation of Dutch Speakers of Greek. *Journal of Phonetics* 32: 543–63. [CrossRef]
- Mennen, Ineke, and Denise Chousi. 2018. Prosody in First-Generation Adult Immigrants and Second-Generation Heritage-Language Users: The Timing of Prenuclear Rising Accents. Paper presented at 9th International Conference on Speech Prosody, Poznań, Poland, June 13–16; pp. 828–32. [CrossRef]
- Miller, Sean. 2012. Acting Tips from Thespian Arnold Schwarzenegger. Available online: <https://www.backstage.com/magazine/article/acting-tips-thespian-arnold-schwarzenegger-51245/> (accessed on 3 April 2020).
- Moosmüller, Sylvia. 2007. Vowels in Standard Austrian German. Habilitation thesis, University of Vienna, Vienna, Austria.
- Moosmüller, Sylvia, Carolin Schmid, and Julia Brandstätter. 2015. Illustrations of the IPA: Standard Austrian German. *Journal of the International Phonetic Association* 45: 339–48. [CrossRef]
- Moyer, Alene. 1999. Ultimate Attainment in L2 Phonology: The Critical Factors of Age, Motivation, and Instruction. *Studies in Second Language Acquisition* 21: 81–108. [CrossRef]
- Muscle & Fitness. 2017. Arnold Schwarzenegger Opens up to Fans on Live Chat. Available online: <https://www.muscleanfitness.com/athletes-celebrities/news/arnold-schwarzenegger-opens-fans-live-chat/> (accessed on 25 January 2021).
- Naumburger Tageblatt. 2015. Arnold Schwarzenegger Im Interview: 'Beim Rechnen Denke Ich Auf Deutsch'. Available online: <https://www.naumburger-tageblatt.de/nachrichten/panorama/arnold-schwarzenegger-im-interview--beim-rechnen-denke-ich-auf-deutsch--22639912-seite3> (accessed on 24 June 2020).
- Odlin, Terence. 1989. *Language Transfer: Cross-Linguistic Influence in Language Learning*. Cambridge: Cambridge University Press.
- Odlin, Terence. 2006. Cross-Linguistic Influence. In *The Handbook of Second Language Acquisition*. Edited by Catherine J. Doughty and Michael H. Long. Malden: Blackwell, pp. 436–86.

- Oh, Grace E., Susan Guion-Anderson, Katsura Aoyama, James E. Flege, Reiko Akahane-Yamada, and Tsuneo Yamada. 2011. A One-Year Longitudinal Study of English and Japanese Vowel Production by Japanese Adults and Children in an English-Speaking Setting. *Journal of Phonetics* 39: 156–67. [CrossRef]
- Outland Baker, Barbara. 2006. *Arnold and Me: In the Shadow of the Austrian Oak*. Bloomington: Author House.
- Penfield, Wilder, and Lamar Roberts. 1959. *Speech and Brain Mechanisms*. Princeton: Princeton University Press.
- Piske, Thorsten, James E. Flege, Ian R. A. MacKay, and Diane Meador. 2002. The Production of English Vowels by Fluent Early and Late Italian-English Bilinguals. *Phonetica* 59: 49–71. [CrossRef]
- Pömp, Julian, and Christoph Draxler. 2017. OCTRA. A Configurable Browser-Based Editor for Orthographic Transcription. Paper presented at 13th Conference of Phonetics and Phonology in the German Language Area (P&P13), Berlin, Germany, September 28–29; pp. 145–48.
- Preston, John. 2015. An Austrian Hick in London: Arnie's Early Years. Available online: <https://www.telegraph.co.uk/film/terminator-genisys/arnold-schwarzenegger-early-life-bodybuilding/> (accessed on 20 May 2020).
- Quora. 2017. Can Arnold Schwarzenegger Speak German? Available online: <https://www.quora.com/Can-Arnold-Schwarzenegger-speak-German> (accessed on 25 January 2021).
- Quora. 2020. Can Arnold Schwarzenegger Speak Without an Accent? Available online: <https://www.quora.com/Can-Arnold-Schwarzenegger-speak-without-an-accent> (accessed on 25 January 2021).
- R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. Computer Program. Vienna: R Foundation for Statistical Computing. Available online: <https://www.r-project.org/> (accessed on 20 September 2020).
- Ramos, Francisco, and Stephen Krashen. 2013. Arnold's Advantages: How Governor Schwarzenegger Acquired English Through De Facto Bilingual Education. *International Multilingual Research Journal* 7: 220–29. [CrossRef]
- Reddit. 2021. Isn't It Kinda Weird That Arnold Schwarzenegger Hasn't Lost His Accent? Available online: [https://www.reddit.com/r/Showthoughts/comments/3hb6bj/isnt\\_it\\_kind\\_a\\_weird\\_that\\_arnold\\_schwarzenegger/](https://www.reddit.com/r/Showthoughts/comments/3hb6bj/isnt_it_kind_a_weird_that_arnold_schwarzenegger/) (accessed on 25 January 2021).
- Reubold, Ulrich, and Jonathan Harrington. 2015. Disassociating the Effects of Age from Phonetic Change: A Longitudinal Study of Formant Frequencies. In *Language Development: The Lifespan Perspective*. Edited by Annette Gerstenberg and Anja Voeste. Amsterdam: Benjamins, pp. 9–37.
- Reubold, Ulrich, and Jonathan Harrington. 2017. The Influence of Age on Estimating Sound Change Acoustically from Longitudinal Data. In *Panel Studies of Language Variation and Change*. Edited by S. E. Wagner and I. Buchstaller. London: Routledge, pp. 129–52.
- Reubold, Ulrich, Jonathan Harrington, and Felicitas Kleber. 2010. Vocal Aging Effects on F0 and the First Formant: A Longitudinal Analysis in Adult Speakers. *Speech Communication* 52: 638–51. [CrossRef]
- Roach, Peter. 2009. *English Phonetics and Phonology: A Practical Course*, 4th ed. Cambridge: Cambridge University Press.
- Ryalls, Jack, Marni Simon, and Jerry Thomason. 2004. Voice Onset Time Production in Older Caucasian- and African-Americans. *Journal of Multilingual Communication Disorders* 2: 61–67. [CrossRef]
- Saito, Kazuya, Yui Suzukida, and Hui Sun. 2019. Aptitude, Experience, and Second Language Pronunciation Proficiency Development in Classroom Settings: A Longitudinal Study. *Studies in Second Language Acquisition* 41: 201–25. [CrossRef]
- Sancier, Michele L., and Carol A. Fowler. 1997. Gestural Drift in a Bilingual Speaker of Brazilian Portuguese and English. *Journal of Phonetics* 25: 421–36. [CrossRef]
- Schiel, Florian. 1999. Automatic Phonetic Transcription of Non-Prompted Speech. Paper presented at XIVth International Congress of Phonetic Sciences, San Francisco, CA, USA, August 1–7; pp. 607–10. [CrossRef]
- Schmid, Monika S., and Holger Hopp. 2014. Comparing Foreign Accent in L1 Attrition and L2 Acquisition: Range and Rater Effects. *Language Testing* 31: 367–88. [CrossRef]
- Schmid, Monika S., and Barbara Köpke. 2017. The Relevance of First Language Attrition to Theories of Bilingual Development. *Linguistic Approaches to Bilingualism* 7: 637–67. [CrossRef]
- Schoonmaker-Gates, Elena. 2015. On Voice-Onset Time as a Cue to Foreign Accent in Spanish: Native and Nonnative Perceptions. *Hispania* 98: 779–91. [CrossRef]
- Schwarzenegger, Arnold. 2012. *A Total Recall. My Unbelievably True Life Story*. London and New York: Simon & Schuster.
- Scovel, Thomas. 1969. Foreign Accents, Language Acquisition, and Cerebral Dominance. *Language Learning* 19: 245–53. [CrossRef]
- Sharwood Smith, M., and E. Kellerman. 1986. *Crosslinguistic Influence in Second Language Acquisition*. New York: Pergamon Press.
- Simon, Ellen. 2009. Acquiring a New Second Language Contrast: An Analysis of the English Laryngeal System of Natives Peakers of Dutch. *Second Language Research* 25: 377–408. [CrossRef]
- Smith, Bruce L., Jan Wasowicz, and Judy Preston. 1987. Temporal Characteristics of the Speech of Normal Elderly Adults. *Journal of Speech and Hearing Research* 30: 522–29. [CrossRef]
- Stoehr, Antje, Titia Benders, Janet G. van Hell, and Paul Fikkert. 2017. Second Language Attainment and First Language Attrition: The Case of VOT in Immersed Dutch-German Late Bilinguals. *Second Language Research* 33: 483–518. [CrossRef]
- Sučková, Magda. 2020. First Language Attrition in Voice Onset Times in Anglophone Expatriates Residing in the Czech Republic. *Brno Studies in English* 46: 47–66. [CrossRef]
- Thiery, Christopher. 1978. True Bilingualism and Second-Language Learning. In *Language Interpretation and Communication*. Edited by David Gerver and H. W. Sinaiko. New York: Plenum, pp. 145–53.
- Thornburgh, Dianne F., and John H. Ryalls. 1998. Voice Onset Time in Spanish-English Bilinguals: Early Versus Late Learners of English. *Journal of Communication Disorders* 31: 215–29. [CrossRef]

- Tobin, Stephen J., Hosung Nam, and Carol A. Fowler. 2017. Phonetic Drift in Spanish-English Bilinguals: Experiment and a Self-Organizing Model. *Journal of Phonetics* 65: 45–59. [CrossRef]
- Ulbrich, Christiane, and Mikhail Ordin. 2014. Can L2-English Influence L1-German? The Case of Post-Vocalic /r/. *Journal of Phonetics* 45: 26–42. [CrossRef]
- Verspoor, Marjolijn, Wander Lowie, and Marijn van Dijk. 2008. Variability in Second Language Development from a Dynamic Systems Perspective. *The Modern Language Journal* 92: 214–31. [CrossRef]
- von Uslar, Moritz. 2012. 'Hallo, Herr Schwarzenegger?'. Available online: <https://www.zeit.de/2012/42/Arnold-Schwarzenegger-Biografie> (accessed on 11 September 2019).
- Wan, Kristopher. 2017. Accents in Hollywood: Arnold Schwarzenegger. Available online: <https://www.lucidaccent.com/post/2017/03/19/accents-in-hollywood-arnold-schwarzenegger> (accessed on 5 December 2019).
- Wiesinger, Peter. 2014. *Das Österreichische Deutsch in Gegenwart Und Geschichte*, 3rd ed. Wien and Berlin: LIT Verlag.
- Yu, Vickie Y., Luc F. de Nil, and Elizabeth W. Pang. 2015. Effects of Age, Sex and Syllable Number on Voice Onset Time: Evidence from Children's Voiceless Aspirated Stops. *Language and Speech* 58: 152–67. [CrossRef] [PubMed]
- Ziesel, Günther. 2018. Arnold Schwarzenegger Im Gespräch mit Günther Ziesel. *Video File*. Available online: <https://www.youtube.com/watch?v=fW3YU3HFO5M> (accessed on 21 May 2020).



MDPI  
St. Alban-Anlage 66  
4052 Basel  
Switzerland  
Tel. +41 61 683 77 34  
Fax +41 61 302 89 18  
[www.mdpi.com](http://www.mdpi.com)

*Languages* Editorial Office  
E-mail: [languages@mdpi.com](mailto:languages@mdpi.com)  
[www.mdpi.com/journal/languages](http://www.mdpi.com/journal/languages)





MDPI  
St. Alban-Anlage 66  
4052 Basel  
Switzerland

Tel: +41 61 683 77 34  
Fax: +41 61 302 89 18

[www.mdpi.com](http://www.mdpi.com)



ISBN 978-3-0365-2278-4