

## The Effects of Language Contact on Non-Native Vowel Sequences in Lexical Borrowings: The Case of Media Lengua

Sky Onosson 

University of Manitoba, Canada

Jesse Stewart 

University of Saskatchewan, Canada

### Abstract

Media Lengua (ML), a mixed language derived from Quichua and Spanish, exhibits a phonological system that largely conforms to that of Quichua acoustically. Yet, it incorporates a large number of vowel sequences from Spanish which do not occur in the Quichua system. This includes the use of mid-vowels, which are phonetically realized in ML as largely overlapping with the high-vowels in acoustic space. We analyze and compare production of vowel sequences by speakers of ML, Quichua, and Spanish through the use of generalized additive mixed models to determine statistically significant differences between vowel formant trajectories. Our results indicate that Spanish-derived ML vowel sequences frequently differ significantly from their Spanish counterparts, largely occupying a more central region of the vowel space and frequently exhibiting markedly reduced trajectories over time. In contrast, we find only one case where an ML vowel sequence differs significantly from its Quichua counterpart—and even in this case the difference from Spanish is substantially greater. Our findings show how the vowel system of ML successfully integrates novel vowel sequence patterns from Spanish into what is essentially Quichua phonology by markedly adapting their production, while still maintaining contrasts which are not expressed in Quichua.

### Keywords

Media Lengua, Quichua, mixed language, generalized additive mixed models, vowels

## Introduction

Mixed languages are considered a special case of contact language, unique in that instead of arising out of communicative necessity, much like a pidgin, creole, or trade language, they are created for expressive purposes (Stewart & Meakins, accepted). This is the result of the originators already being proficient bilinguals in the source languages; leaving no communicative void for the new language to fill. Because of this bilingual knowledge, mixed languages show little in the way of lexical and morphosyntactic simplification as is typical in most creole-based or pidgin-based languages.

At the same time, mixed languages tend to show highly systematic splits between the elements originating from each source language; for example, lexical–functional splits, noun phrase and verb phrase splits, or lexical–structural splits (Meakins, 2013; Stewart & Meakins, accepted). Therefore, the study of mixed languages provides linguists with a rare platform to test everything from language genesis to structural compatibility. Each new study also provides further insights to improve our understanding of cognitive mechanisms that allow humans to take two unrelated, fully functional languages, split them apart and create a new fully functional language.

Prior to the 2010s, the majority of mixed language research focused on the morphosyntactic divisions among the source languages due to their unique arrangements. However, the past decade has benefited from a surge of research in mixed language phonology that centers on the question: “What is the result of competing source language phonologies in a new mixed language?” To answer this question, phonemic conflict sites (i.e., conflicting areas of phonological convergence in the source languages’ phonologies) are compared acoustically and quantitative analyses provide important details into how sounds are treated in the mixed language. Interestingly, unlike the clear splits observed in the morphosyntax, results from acoustic studies (see e.g., Buchan, 2012; Bundgaard-Nielsen & O’Shannessy, 2019; Hendy, 2019; Jones & Meakins, 2013; Jones et al., 2011, 2012; Meakins & Stewart, accepted; Rosen, 2006, 2007; Rosen et al., 2020; Stewart, 2014, 2015a, 2015b, 2020; Stewart & Meakins, accepted; Stewart et al., 2018, 2020b) show that there is a propensity for phonological material to assimilate to the phonology of the ancestral language. In other words, “the language, which was acquired<sup>1</sup> originally as an L2 [second language] . . . essentially conforms to the L1 [first language] phonological system of the ancestral language” (Stewart & Meakins, accepted). However, evidence also shows that the introduced language does not completely relinquish its influence on the mixed language’s phonology and instead supplies material that appears to contribute to the new sound system.

While the resulting phonological system may appear straightforward at first glance, the arrangements of the introduced language’s phonology in the mixed language are anything but. For instance, the resulting system does not always reflect predictions made by models of adaptive dispersion, which hypothesize that newly established categories disperse in acoustic space to avoid crowding neighboring sounds to maintain contrasts (see e.g., Flege, 2007; Johnson, 2000; Liljencrants & Lindblom, 1972; Lindblom, 1986, 1990; Livijn, 2000). Instead, the phonologies of mixed languages are fraught with “near-mergers, overlapping categories, categorical assimilation, categorical maintenance, and overshoot of target categories at the segmental level, in addition to prosodic assimilation, possible preservations of archaic patterns, and innovation at the suprasegmental level” (Stewart & Meakins, accepted).

## 2

## Media Lengua (ML)

One of the most studied mixed languages, from a phonetic standpoint, is ML, a mixed language spoken in the northern Ecuadorian province of Imbabura. ML is a lexical–grammar mixed



**Figure 1.** Communities and cities of Imbabura, Ecuador where data for this study were collected: Pijal for Media Lengua; Cashaloma and Chirihuasi for Quichua; and Ibarra for Spanish. This map is freely licensed under the Open Data Commons Open Database License (ODbL) by the OpenStreetMap Foundation (OSMF).

language (Meakins, 2013; Meakins & Stewart, accepted; Muysken, 1997), which combines Quichua's<sup>2</sup> agglutinating morphology and word order with the Rural Andean Spanish lexicon that has replaced over 90% of Quichua content words primarily through the process of relexification (Deibel, 2017, 2019, 2020; Gómez-Rendón, 2005; Lipski, 2016; Muysken, 1981, 1997; Stewart, 2011, 2015b). ML appears to have formed in the early 20th century and is still currently spoken today by approximately 2000 people in several communities near Lago San Pablo. The data from this study come from the community of Pijal, where the language appears to have originated in Imbabura (see map in Figure 1).

Example (1)<sup>3</sup> provides a sample of ML; the elements in boldface type are of Spanish origin. The first line contains the ML orthography, the second a broad International Phonetic Alphabet transcription with morpheme boundaries, the third the interlinear glosses, and the fourth and fifth lines provide translations in Imbabura Quichua and Spanish, respectively.

- (1) **Mio**    **ciudad puenteca**    **derrumbarircami.**  
mio    siudad puente-ka    dezumba-ri-rka-mi  
1.POSS city bridge-TOP collapse-REF-PST-VAL  
*Ñuka llakta chacaka tunirirkami.* (Imbabura Quichua)

*El puente de mi ciudad se derrumbó.* (Spanish)

“My city’s bridge collapsed.”

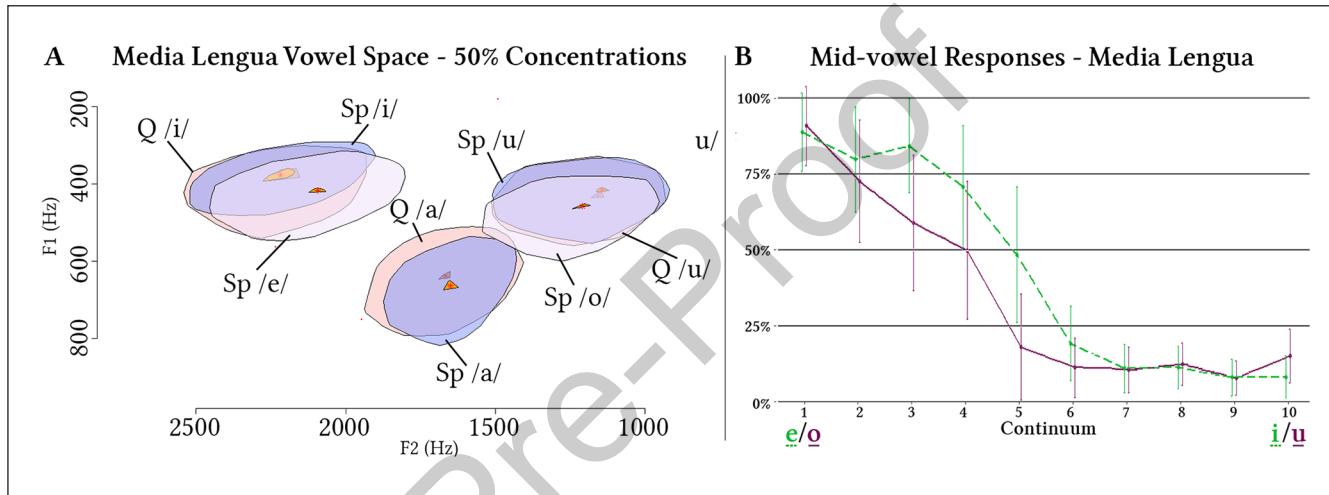
This paper adds to this growing literature with a quantitative acoustic analysis of Spanish origin and Quichua origin vowel sequences in ML. Moreover, our study provides important insights into the functionality of an exceptional vowel system that opposes predicted arrangements in acoustic space.

## 2.1 The ML vowel system

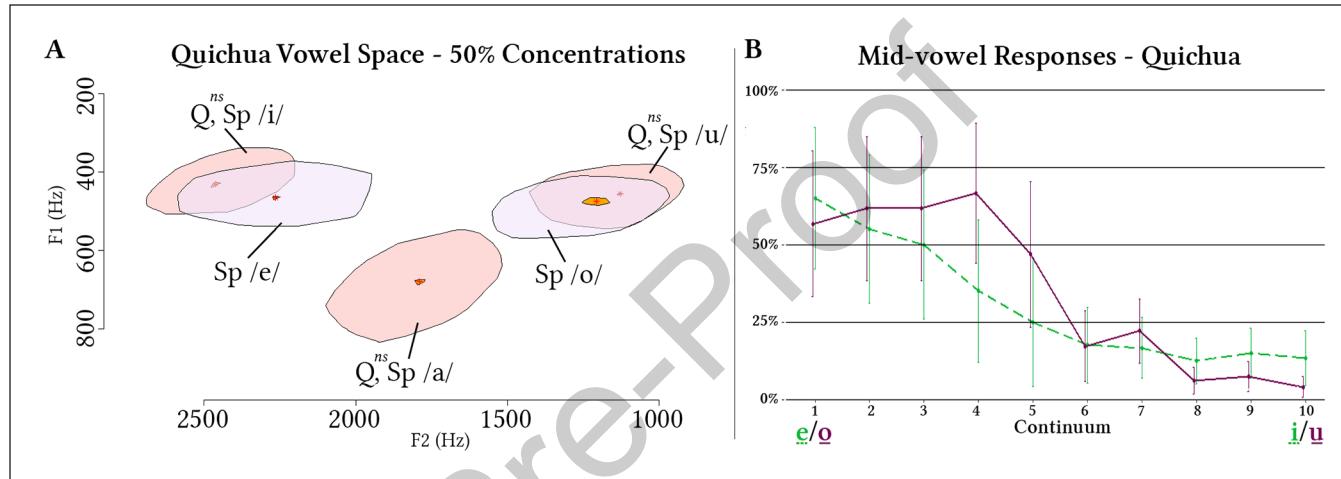
**2.1.1 Monophthongs.** The ML’s vowel system is a combination of Quichua’s three-vowel system, consisting of /i, u, a/, and Spanish’s five-vowel system, consisting of /i, u, e, o, a/. Stewart (2014) describes this system as a complex case of stratification based on the language of origin of the lexeme where a given vowel is found. His analysis of formant 1 (F1) and formant 2 (F2) measurements involved 2515 tokens extracted from ML elicited phrases based on the speech of 10 participants. Results revealed small, albeit significant differences, between Spanish origin and Quichua origin corner vowels (i.e., /i, u, a/) where Spanish origin /i, u, a/ were on average just 13 Hz less centralized in acoustic space compared to Quichua origin /i, u, a/. While at first glance, this difference may appear trivial, the dispersion patterns of the F1 in all the Spanish origin corner vowels fall in line with the directions predicted by models of adaptive dispersion (i.e., the F1 values of the Spanish origin vowels extend further into acoustic space compared to those of Quichua origin), but not with the degree of dispersion expected for contrastive purposes (see e.g., Flege, 2007; Johnson, 2000; Liljencrants & Lindblom, 1972; Lindblom, 1986, 1990; Livijn, 2000, for details on adaptive dispersion theories). This means that at a “micro” level, ML speakers produce consistent differences in lexically non-contrastive vowels of the same quality (similar to near-mergers or covert contrasts) based solely on the language of origin.

Additionally, Stewart (2014) also shows that ML speakers produce Spanish origin mid-vowels in a significantly higher F1 range (i.e., with a lower tongue body position) when compared to Spanish origin and Quichua origin high-vowels (Figure 2A). However, what makes this part of the system interesting is that, like the corner vowels described above, the mid-vowel and high-vowel categories also exist with substantial overlap; yet with just enough distance (0.36 bark, 41 Hz) to meet the threshold of 0.3 bark, established by Kewley-Port (2001), for possible formant discrimination for values between 200 and 3000 Hz. In fact, it was later confirmed by Stewart (2018) that ML speakers are able to consistently identify differences between Spanish origin mid- and high-vowels in a two-alternative forced choice identification task experiment using minimal pairs (Figure 2B).

The results of these studies paint a picture of a system that may operate up to eight vowels,<sup>4</sup> yet it is arranged in an acoustic space that reflects that of a three-vowel system such as Quichua (see Figure 3). Stewart and Meakins (accepted) hypothesize that this unconventional system could have arisen during the genesis of the language given that the originators of ML most likely acquired Spanish as late bilinguals and thus spoke a Quichua-accented Spanish. However, instead of having the Spanish mid-vowels fully assimilate to Quichua high-vowels, enough distance was allotted to maintain some degree of contrast to benefit the newly relexified vocabulary, which brought in contrasting mid- and high-vowels with high phonological function loads. For instance, the maintenance of some Spanish phonemes avoids the realization of possible ambiguous words if complete assimilation to Quichua phonology were instead the case, for example, we might observe the word *pisu* or *pizu*, which could mean *beso* “kiss,” *peso* “weight,” *piso* “floor,” *bezo* “part of the lip,” etc. This overlapping vowel system would have then become nativized, and thus normalized, in subsequent generations of speakers.



**Figure 2.** Image A—Media Lengua (ML) vowel space based on 50% concentrations of F1 and F2 (Hz) measurements of statistically different vowel clusters based on Stewart (2014); and Image B—ML results from a two-alternative forced choice identification task experiment with minimal pair stimuli modified along a 10-step continuum between prototypical ML mid-vowels to prototypical ML high-vowels based on Stewart (2018).



**Figure 3.** Image A—Quichua vowel space based on 50% concentrations of F1 and F2 (Hz) measurements of statistically different vowel clusters based on Stewart, 2014; and Image B—Quichua results from a two-alternative forced choice identification task experiment with minimal pair stimuli modified along a 10-step continuum between prototypical Media Lengua (ML) mid-vowels to prototypical ML high-vowels (nearly identical to Quichua prototypical values) based on Stewart (2018).

Evidence to support this claim also appears in the Quichua vowel system (Stewart, 2014) which has adopted approximately 40% of the Spanish lexicon (see Bakker & Muysken, 1994) over the last 500 years of intense contact. However, unlike ML, Spanish origin corner vowels are shown to be non-significant from native Quichua corner values (implying assimilation, and thus merged in Figure 3A). Moreover, while Spanish origin mid-vowels are significantly different from native Quichua high-vowels and Spanish origin high-vowels, like in ML (see Figure 3A), in Quichua they do not reach the threshold of 0.3 bark, established by Kewley-Port (2001), for possible formant discrimination for values between 200 and 3000 Hz. Indeed, with a difference of just 0.23 bark (26 Hz), Stewart (2018) shows that Quichua speakers do not consistently identify Spanish origin mid-vowels but, as would be expected, consistently identify high-vowels of Spanish origin, which are non-significantly different from native Quichua high-vowels (see Figure 3B).

**2.1.2 Vowel sequences.** This study focuses on what we term “vowel sequences,” by which we mean vowel-like articulations containing two target positions/qualities. The term “vowel sequences” is deliberately imprecise, and intended to be taken agnostically, such that its referents may include what are in fact two different sets, from a phonological point of view: (a) sequences where both members are distinct phonological vowels, sometimes referred to as a hiatus; and (b) sequences where one member is a vowel and one member is a glide—that is, a canonical diphthong. This distinction is, however, less clear than it might initially appear. Similarities between glides and high-vowels in particular have led some phonologists to describe segmental pairs such as /i, j/ as being featurally identical aside from consonantal status (Padgett, 2008). Sánchez Miret (1998) argues that monophthongs, diphthongs, and hiatuses form a cline where the distinction between adjacent categories is one of degree rather than kind. Therefore, “vowel sequences” as units in and of themselves are of primary interest as changes in vowel quality (i.e., a “true” diphthong), component vowels affecting each other’s realization, or even the presence of adjacent vowels in a syllable nucleus (i.e., hiatus), can all be considered useful contrastive structures in the phonology of a language.

However, intense language contact can call into question the functionality of previously contrastive units in lexical borrowings through processes such as assimilation, elision, or reinterpretation. Therefore, acoustic analyses and cross-language comparisons of phonetic correlates such as formant trajectories of borrowed vowel sequences in ML and Quichua will lend to a better understanding of how such units are borrowed, interpreted, and produced. This not only allows us to identify phonological processes such as assimilation, but also gradient modifications which reveal how borrowed sounds conform to the phonological constraints of a new language. Through the phonetic analysis of a large number of vowel sequences we can also identify consistent and variable articulatory gestures (in F1 × F2 space), which can correlate to the importance of a given vowel sequence. For example, variable articulatory gestures might suggest that the sequence carries a low functional load (i.e., contrastive importance is minor) whereas consistent articulatory gestures might suggest that the sequence is more important for discerning meaningful units. Understanding the precise phonological status of each vowel sequence, however, remains an area of future research, and is not something which we intend to explore within the confines of this study.

For ML and Quichua, Spanish origin vowel sequences were organized based on their original Spanish pronunciation. For example, in ML, the word *tierra* “earth, ground, dirt” might be pronounced as [‘tieɾa] or [‘tizɑ], yet the vowel sequence would be considered as /ie/, since its pre-prefixed production was that of /ie/ as in Andean Spanish *tierra* [tiera] “earth, ground, dirt.” Where such individual pronunciation differences between ML versus Spanish occur, our analysis method is designed to determine whether they reach the level of statistical significance within the aggregated dataset.

**Table 1.** Potential Media Lengua vowel sequences. Sequences in boldface type are uniquely of Spanish origin.

a-	e-	i-	o-	u-
<b>ae</b>	<b>ea</b>	ia	<b>oa</b>	ua
ai	<b>ei</b>	<b>ie</b>	<b>oe</b>	<b>ue</b>
<b>ao</b>	<b>eo</b>	<b>io</b>	<b>oi</b>	ui
au	<b>eu</b>	iu	<b>ou</b>	<b>uo</b>

Based on its five-vowel system, ML has the potential to contain 20 distinct non-identical vowel sequences, which are listed in Table 1. Considering ML’s lexical source languages, six of these vowel sequences are shared between both Spanish and Quichua, while 14 occur only in Spanish (boldface type in Table 1).

All 20 identified sequences were found throughout ML’s documented lexicon (see Stewart et al., 2020a), with the exception of /ou/. The lack of this particular sequence is unsurprising given that /ou/ is exceptionally rare and only present in low frequency words in Spanish, often of foreign origin. Moreover, when found in Spanish orthography, the <ou> sequence is produced as one of /u, o, au/ depending on the word, for example, *boutique* [bu'tik] “boutique.” Similarly, the frequency of some vowel sequences is less common than others. Because of this, our study excludes three existing but low-frequency ML vowel sequences: /iu/ with only two lexemes in our data (*viuda* “widow” and *ciudad* “city”); /oi/ which only appeared in four words (*hoy* “today,” *proibido* “prohibited,” *oina* “listen,” and in the name *Zoila*); and, /oe/ which was only found in one word (*oeste* “west”).

### 3

## Method

The aim of this study is to investigate the hypothesized integration and acoustic arrangement of Quichua and Spanish origin vowel sequences in ML. To do so, we conduct an instrumental study of formant trajectories in and across all three languages.

### 3.1 Materials

The number of vowel sequence tokens under analysis in this study totals 1096 across all three languages. ML vowel sequence data were gathered from three sources, totaling 406 tokens. The first source (corpus 1) comes via a 2010–2012 corpus of elicited phrases and contains 150 tokens (37%). The second source (corpus 2) comes via a 2010–2014 corpus of wordlist/sentence list data and contains 195 (48%) tokens. In this corpus collection, participants were informed that they would be asked to read several short sentences (for both the Quichua and ML groups) or words (for the Spanish groups) on a computer. The ML and Quichua data were presented in sentences to avoid possible “switches” in language mode as individual lexical items in isolation may be ambiguous as to their source (e.g., *escuela* “school” is the same word in all three languages). Spelling for all three languages was based on Spanish orthography as the unified Quichua system was still unfamiliar to many Quichua and ML speakers. If a participant could not read (a rare occurrence), the author or the assistant read the sentence and asked the participant to repeat it twice. The second utterance was used for analysis. If a participant struggled with reading, they would be asked to repeat the sentence from memory allowing for a more naturalistic speech sample.

**Table 2.** Speech corpora token quantities.

Corpus	Elicitation type	Media Lengua	Quichua	Spanish
Corpus 1	Phrases	150	—	—
Corpus 2	Word/sentence lists	195	166	524
Corpus 3	Natural speech	61	—	—

The last source (corpus 3) comes via a 2015–2019 corpus of natural speech data and contains 61 tokens (15%). The Spanish and Quichua data come solely via corpus 2, containing wordlist/sentence list data: 166 (100%) tokens for Quichua; and 524 (100%) tokens for Spanish. Quichua has a lower token count as it contains roughly one-third of the number of vowel sequence combinations as Spanish and ML. Token quantities in our data are summarized by language and corpus in Table 2.

During the elicitation sessions for corpus 1, participants were asked to give their best oral translation of a sentence into ML. The participants' oral translations were recorded on a TASCAM DR-1 portable digital recorder using TASCAM's compatible TM-ST1 MS stereo microphone set to 90° stereo width. Elicitations were recorded in 16-bit Waveform Audio File (WAV) format with a sample rate of 44.1 kHz. During the wordlist and sentence list sessions for corpus 2, participants from each language were asked to read words or sentences in their native language from a computer screen. For the Spanish participants there was a mix of sentence list data and wordlist data. However, ML and Quichua required sentences to prime the target language; for example, seeing the word *caña* “cane/stick” instead of *cañahuanmi* “cane-INST-VAL” might elicit a more Spanish-like pronunciation since ML and Quichua speakers are bilingual. The participants were recorded on a TASCAM DR-1 portable digital recorder using a NEXXTECH unidirectional dynamic microphone (50–13,000 Hz response). The wordlist was recorded in 16-bit WAV format with a sample rate of 44.1 kHz. Finally, corpus 3 was created by two assistants from Pijal who recorded the data. Participants were asked to gather in pairs and converse about a topic of their choosing. These data were recorded on a Zoom H6 Handy Recorder with its internal microphone (unidirectional XYH-6 capsule). See Stewart (2015b) for more details on the data collection procedures for corpora 1 and 2.

### 3.2 Participants

For ML, 23 trilingual speakers (Quichua, ML and L2 Spanish) participated in this study. This group consisted of fourteen women and nine men. All participants were from the community of Pijal Bajo and acquired Quichua and ML simultaneously from birth and began learning Spanish upon entering primary school, typically at 6–7 years of age.<sup>5</sup>

Ten Quichua-speaking participants also participated in this study, all of whom were bilinguals (Quichua and L2 Spanish) ranging from low to high proficiency in Spanish.<sup>6</sup> This group consisted of six women and four men. Four women had a rudimentary level of Spanish, one man and one woman were simultaneous bilinguals, and one man acquired Spanish at the age of 18, while the rest acquired Spanish upon entering primary school, typically at 6–7 years of age. Participants were born, raised, and lived in the neighboring communities of Chirihuasi and Cashaloma at the time of recording. These slightly more distant communities were chosen to gather Quichua data over Pijal and neighboring communities to avoid any influence from ML on the Quichua speech of these participants.

From the Spanish group 14 participants took part in this study and all were monolinguals except two with late English L2 acquisition. Eight participants were female and six were male. All

participants had graduated from college and were born, raised, and lived in urban centers (11 from Quito and three from Ibarra) at the time of recording. More demographic information pertaining to all participants is listed in the Appendix.

### 3.3 Procedures

The procedure followed in this study has been successfully implemented in several contexts, including: the comparison of diphthong productions across different phonetic contexts (Onosson, 2018); the cross-dialect comparison of diphthong productions (Onosson, 2018); and, the investigation of vowel-glide sequence production differences between L1 and L2 speakers (Onosson & Bird, 2019). This study potentially represents the first application of this particular methodology to production differences across different languages, and to lexical items deriving from different source languages, and is therefore informative not only for the specific results pertaining to ML, but also with regards to the methodology itself and its potential for application to a diversity of linguistic contexts.

**3.3.1 Vowel extraction.** Each target vowel sequence token was manually segmented in Praat (v6.1.04; Boersma & Weenink, 2019) and acoustic data were extracted using a Praat script (Xu & Gao, 2018) which takes discrete formant measurements at specified intervals. For this study, we retrieved F1, F2, and F3 measurements at 5% intervals across vowel duration, yielding 20 discrete measurements per formant per token. This method permits high-fidelity reconstruction of the original formant trajectories versus more coarse methods which utilize fewer (e.g., 2, 3, or 5) such measurements.

**3.3.2 Generalized additive mixed models (GAMMs).** Following formant extraction, the data were analyzed in R (v3.6.1; R Core Team, 2019) using RStudio (v1.2.5001; RStudio Team, 2019). Our primary method of analysis uses generalized additive models (GAMs; Hastie & Tibshirani, 1990; Wood, 2017). GAMs are designed to analyze and compare dynamic non-linear data, such as the formant trajectories of diphthongs or vowel sequences. A GAM computes a “smooth” across the dynamic component of the data under investigation, in this case being the temporal dimension, which is to say, vowel duration. A pair of smooths can then be directly compared, allowing the computation of a “difference smooth” which indicates temporal regions where the two smooths differ at a given confidence level, typically made at 95%.

The addition of mixed-modeling to GAMs is referred to as generalized additive *mixed* models or GAMMs and allows the inclusion of random effects to cover such factors as unconstrained inter-speaker variance. Our GAMMs analysis was implemented with the bam function of the mgcv package (Wood, 2011) using the formula in Example (2), modeled after Sóskuthy (2017):

- (2) GAMMs formula code:

```
Formant ~ Language +  
s(Time, bs = "cr") +  
s(Time, by = Language, bs = "cr") +  
s(Time, Speaker, bs = "fs", m = 1) +  
s(Time, Word, bs = "fs", m = 1)
```

The main effect tested in each GAMMs analysis is the relationship between *Formant* as the dependent variable and *Language* as the independent variable. This correlation was conducted

across both F1 and F2 for all available languages; wherever possible, all three languages were compared against each other, but as a number of vowel sequences do not occur in Quichua (see subsection 2.1.2), only the relevant ML and Spanish data could be compared for those sequences. Following the main effect, the GAMMs formula includes a series of smooths, each prefaced by the letter s in the formula in Example (2). The first smooth considers only the *Time* value at each formant measurement, taken as a percentage of overall sequence token duration. This method normalizes across individual token durations and thereby accommodates durational variance between different vowel + vowel (V+V) sequences. The second smooth additionally computes separate *Time* smooths per language, which is intended to further accommodate systematic differences that may exist between languages. Finally, two random smooths are also computed: one for *Speaker*; and one for *Word*. The specific coding here, which includes the use of bs = “cr” to specify the use of a *cubic regression spline* as the basis function for fixed effects smooths and *factor smooth interactions* (bs = “fs”) for random effects smooths, follows recommendations by Sóskuthy (2017). The addition of a random smooth for *Speaker* is commonplace in mixed methods, as it allows the existence of individual speaker variation to be incorporated into the model. The random smooth for *Word* is included as a proxy for differing phonetic contexts between V+V tokens. The dataset for our study derives from elicited forms which were selected specifically for the presence of V+V sequences, and do not systematically cover all the possible phonetic or phonological contexts in which a V+V sequence could potentially occur in for every language. In fact, because the phonotactic restrictions on V+V sequences differ between the three languages, and because of differences in the range of possible phonetic environments, which are far greater in Spanish versus Quichua, it would be impossible to build a comprehensive set of elicited forms that covered all possible phonetic contexts, and which were equivalent between all three languages. Because of these inherent limitations, we elected to provide *Word* as a random effect as a means of allowing the specific phonetic context of each elicited form to be included in the model, albeit without necessarily being able to comprehensively compare like contexts between different languages in a more robust and systematic way. In any case, we do not expect phonetic context to exhibit a strong effect on V+V formant production except at the onset and offset points —the central phase of the sequence’s formant trajectories would not be expected to be overly impacted by adjacent segment context, in the majority of cases. We do, however, acknowledge this as a limitation of our methodology, one which could be addressed in a future study with more comprehensive coverage of phonetic context within the elicitation materials.

Our GAMMs comparison method involves the following. For each formant of each V+V sequence, two models are computed: a model which incorporates the main effect of *Language* (M1); and a null model that excludes it (M0). M1 uses precisely the formula listed above, while the formula for M0 removes the sections containing the main effect of *Language* and its associated smooth. The compareML function of the itsadug R package (van Rij et al., 2020) is then used to conduct a comparison between M0 and M1. This function compares each model’s maximum likelihood and degrees of freedom (*df*) results using a Chi-squared ( $\chi^2$ ) test and reports a *p*-value for the overall comparison, which indicates whether the  $\chi^2$  result is deemed significant. With respect to the difference between the two models, which centers around the inclusion versus exclusion of *Language* as a main effect, we interpret the *p*-value from this comparison as an “indicator of surprise” rather than in its traditional use as an indicator of statistical significance, reflective of the exploratory nature of this study (see Baayen et al., 2017, on the evaluation of the GAMMs comparison methods and Roettger et al., 2019, on significance-testing in confirmatory versus exploratory studies).<sup>7</sup> Our model comparison is fundamentally focused on one question: *Do the set of formant trajectories derived from tokens of a particular V+V sequence differ according to language?* As such, we have endeavored to keep our model structure relatively simple, such

that success or failure of the model should be clearly interpretable as pertaining to the inclusion/exclusion of the effect of *Language*. Following the indication by the  $\chi^2$  test result that *Language* is a significant effect, we then analyze the GAMMs difference smooth, or smooths in the case of multiple language-to-language comparisons. This allows identification of specific regions across the vowel sequence duration where the formant trajectories of each of the two (or three) languages differ robustly from each other. An illustration of the application of this method follows below.

To illustrate, we present a comparison of the F1 trajectories of /au/ across the three languages in our study using the procedure outlined above. The M0 versus M1 model comparison for /au/ F1 indicates that M1, which includes the main effect of *Language*, is a better fit than M0,  $\chi^2 = 10.142$ ,  $df = 6$ ,  $p = 0.002$ . Within M1, the difference between ML versus Quichua is non-significant,  $p = 0.0167$ , while the difference between ML versus Spanish is significant,  $F = 4.002$ ,  $p = 0.00133$ . Figure 4 shows the difference smooth comparisons from M1. In Figures 4A and 4B, the ML formant trajectory is represented by the thin, solid black curved line which indicates the computed GAMMs smooth for that trajectory. The horizontal zero-line on the y-axis represents the reference level for each comparison—in Figure 4A this is the Quichua vowel sequence formant trajectory, and in Figure 4B it is the Spanish trajectory. The plotted smooth for ML thus indicates where (i.e., *when*) and in which direction the ML trajectory diverges from the comparison reference trajectory; the shaded gray regions in Figures 4A and 4B represent 95% confidence intervals for the difference smooth. In Figure 4A, the ML confidence interval fully overlaps the reference level zero-line, which indicates that the ML F1 trajectory does not differ substantially from Quichua at any point across the duration of the vowel sequence, a reflection of the non-significant difference between the two languages reported for M1 under the  $\chi^2$  test. The regions where the confidence intervals depart from the zero-line indicate regions of particularly robust difference between the two trajectories under comparison. This region is highlighted in the difference smooth by vertical dashed lines at the edges of the region of difference, accompanied by a red bar along the x-axis. Such a region appears in Figure 4B from the beginning of the sequence until approximately 70% of duration, indicating that the ML F1 trajectory differs most substantially from that of Spanish during this portion of production of the sequence.

Finally, a visual comparison across multiple smooths can be achieved via a comparison plot which overlays each smooth and its associated confidence intervals. Figure 4C shows the comparison of the /au/ F1 smooths across the three languages, where it can be readily observed that the ML and Quichua productions follow a very similar trajectory, while Spanish diverges quite substantially, occurring largely during the initial and medial portions of the trajectory.

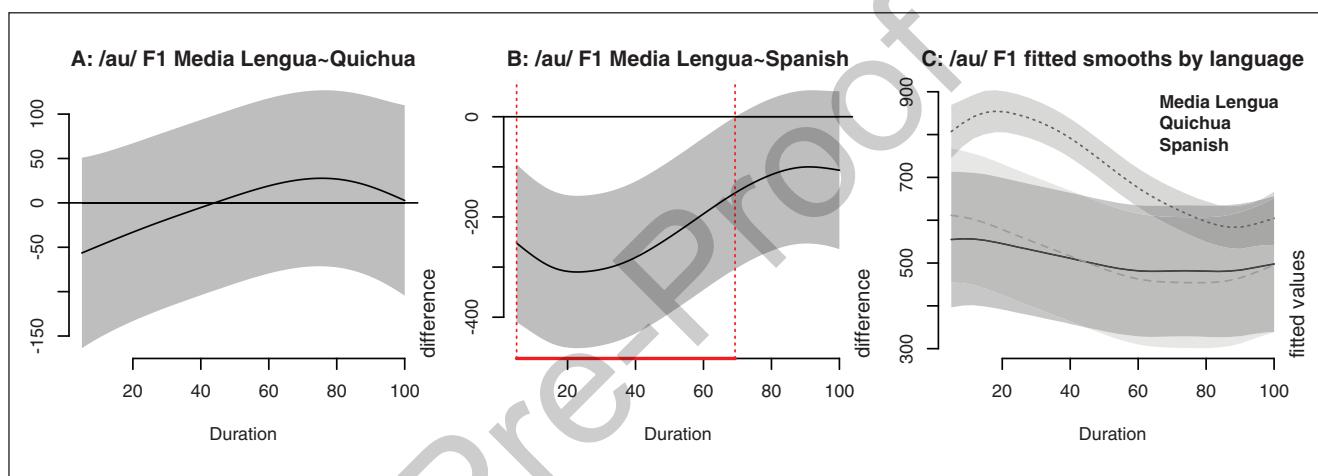
## 4

## Results

### 4.1 Language-internal GAMMs comparisons

The initial set of GAMMs comparisons were run on vowel sequences occurring within the ML data to investigate potential differences between Quichua origin versus Spanish origin lexical items containing the same vowel sequence. Only six distinct vowel sequences occur in the native Quichua lexicon (/ai, au, ia, iu, ua, ui/), and the Quichua origin words in ML do not fully represent this set. As a result, we were able to identify only three sequences occurring across both Quichua and Spanish origin words in ML: /ai, au, ui/.

The GAMMs comparisons were made between F1 and F2 trajectories across the constituent different-source-language lexical items for each sequence, such that, for example, tokens of /au/ in *tauca* ['tau ka] “a lot, much, many” (Quichua origin) were compared to tokens of *bautismo* [bau.'tis.mo] “baptism” (Spanish origin).<sup>8</sup> Main-effect models (M1) were built for each



**Figure 4.** /au/ F1 (Hz) generalized additive mixed models' three-language comparison. In 4C, the solid line indicates the smooth for Media Lengua; the dashed line the smooth for Quichua; and the dotted line the smooth for Spanish.

sequence-formant with *Source Language* as main effect, and an analysis of variance (ANOVA) compared that model's output with a null model (M0) for the same vowel sequence formant. None of these comparisons indicated any significant differences according to *Source Language*. In other words, ML vowel sequence productions do not differ significantly between lexical items from different source languages, but instead represent a single phonological system with similar productions across all similar vowel sequences, regardless of language of origin.

Because ML represents the only dataset in our study derived from all three corpora listed in Table 2, and thus containing a variety of recording types, we ran an additional set of GAMMs models (M1) with *Recording Type* as main effect. ANOVA comparisons of the per-sequence M1 models versus parallel M0 models did not find any significant differences in this case, either, indicating that recording type (and presumably speech register or style) is not correlated with production differences for these vowel sequences.

## 4.2 Cross-language GAMMs comparisons

In Table 1 we identified 20 vowel sequences which exist across ML and Spanish (six of which are also present in Quichua). Of these, /iu, oe, oi, ou/ were excluded from further analysis due to non-occurrence or low token counts in our data (see subsection 2.1.2), leaving the set of sequences listed in Example (3):

1. (3) /ae, ai, ao, au, ea, ei, eo, eu, ia, ie, io, oa, ua, ue, ui, uo/

The GAMMs models including *Language* as main effect were computed for each formant, F1 and F2, of each sequence in Example (3), along with a parallel null model which excluded the effect of *Language*. Comparisons of each null versus non-null model pair (see subsection 3.3.2 for discussion of this methodology) indicated that the *Language* model was superior to the null model (where the comparison *p*-value was below 0.05), for either one or both formants, in eight of the 16 sequences investigated; these are listed in Table 3. Where the ANOVA indicated a significant difference between models, thus confirming that the addition of *Language* as a main effect improved the GAMMs model, we investigated the GAMMs output to identify which languages differ from each other, where more than two languages are involved in the comparison, and critically which specific portions of the formant trajectories differ.

The upper half of Table 3 lists sequences for which Quichua data were available, and thus for which GAMMs comparisons were carried out across all three languages.<sup>9</sup> The lower half of Table 3 lists sequences for which only ML and Spanish data were available for comparison. These two groupings are discussed separately in subsection 4.2.2 and subsection 4.2.3. Before that, we present a descriptive account of the formant trajectories for these eight sequences, in subsection 4.2.1.

**4.2.1 Vowel sequence trajectories.** In subsection 4.2 we established via GAMMs analysis significant differences between ML and ML's two lexical source languages, Spanish and Quichua, in the formant trajectories of eight distinct vowel sequences (see Table 3). Before examining the GAMMs results in more detail, in this subsection we first take a closer look at this set of sequences by plotting their trajectories in F1 × F2 space.

In Figure 5, the trajectories of the eight vowel sequences listed in Table 3 are plotted in F1 × F2 space along with the five ML monophthongs in the background to provide context. For the monophthong data, we use formant measurements extracted at 50% of duration; the positions of the monophthong vowel labels indicate the median F1 and F2 values at this point, and the ellipses represent one standard deviation. For the vowel sequence trajectories, the text label (e.g., “ia”)

**Table 3.** Cross-language vowel sequence generalized additive models' comparisons where the model including *Language* as the main effect is superior to the null model at  $p < 0.05$ .

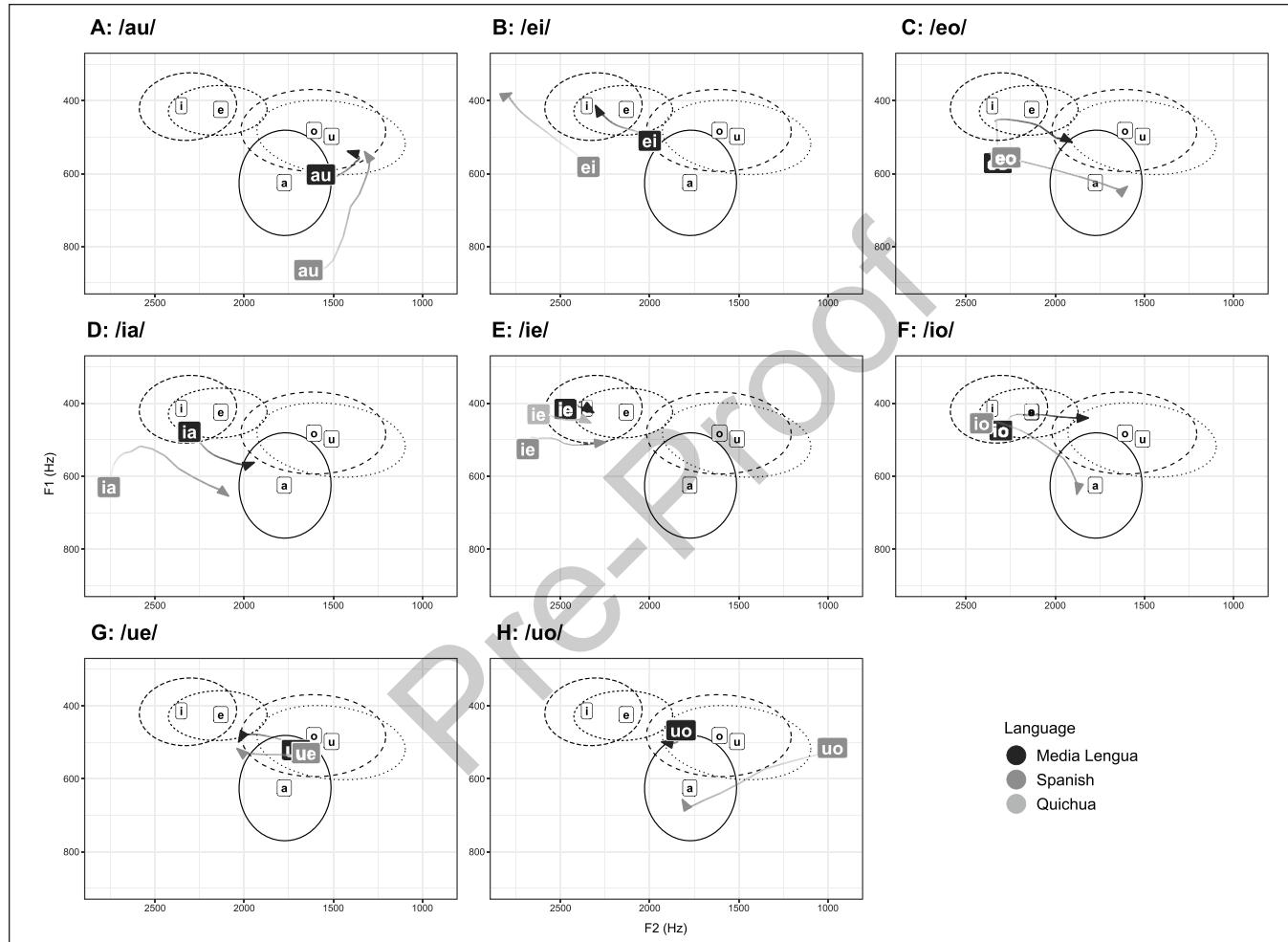
Languages compared	Sequence	F1	F2
Media Lengua (ML), Quichua, Spanish	/au/	$\chi^2 = 10.142, df = 6, p = 0.002$	—
	/eo/	$\chi^2 = 6.337, df = 6, p = 0.049$	—
	/ia/	$\chi^2 = 11.417, df = 6, p < 0.001$	$\chi^2 = 12.376, df = 6, p < 0.001$
	/ie/	—	$\chi^2 = 18.153, df = 6, p < 0.001$
	/ue/	—	$\chi^2 = 14.024, df = 6, p < 0.001$
ML, Spanish	/ei/	$\chi^2 = 5.401, df = 3, p = 0.013$	$\chi^2 = 8.735, df = 3, p < 0.001$
	/io/	$\chi^2 = 13.795, df = 3, p < 0.001$	$\chi^2 = 4.771, df = 3, p = 0.023$
	/uo/	$\chi^2 = 7.574, df = 3, p = 0.002$	—

marks 20% of sequence duration, and the arrowhead marks 80% of duration; we excluded the initial and final 20% portions of the sequences, as coarticulatory effects are more apparent in these regions. Both ML and Spanish vowel sequence trajectories are included in each plot in Figure 5 and distinguished by grayscale-shading as shown in the figure, with ML appearing darker than Spanish; the Quichua /ie/ sequence from Spanish origin words is also included in Figure 5E (in the lightest shading), as this was the only sequence to exhibit a significant difference (restricted to F2, equivalent to articulatory front–back position) between Quichua and ML.

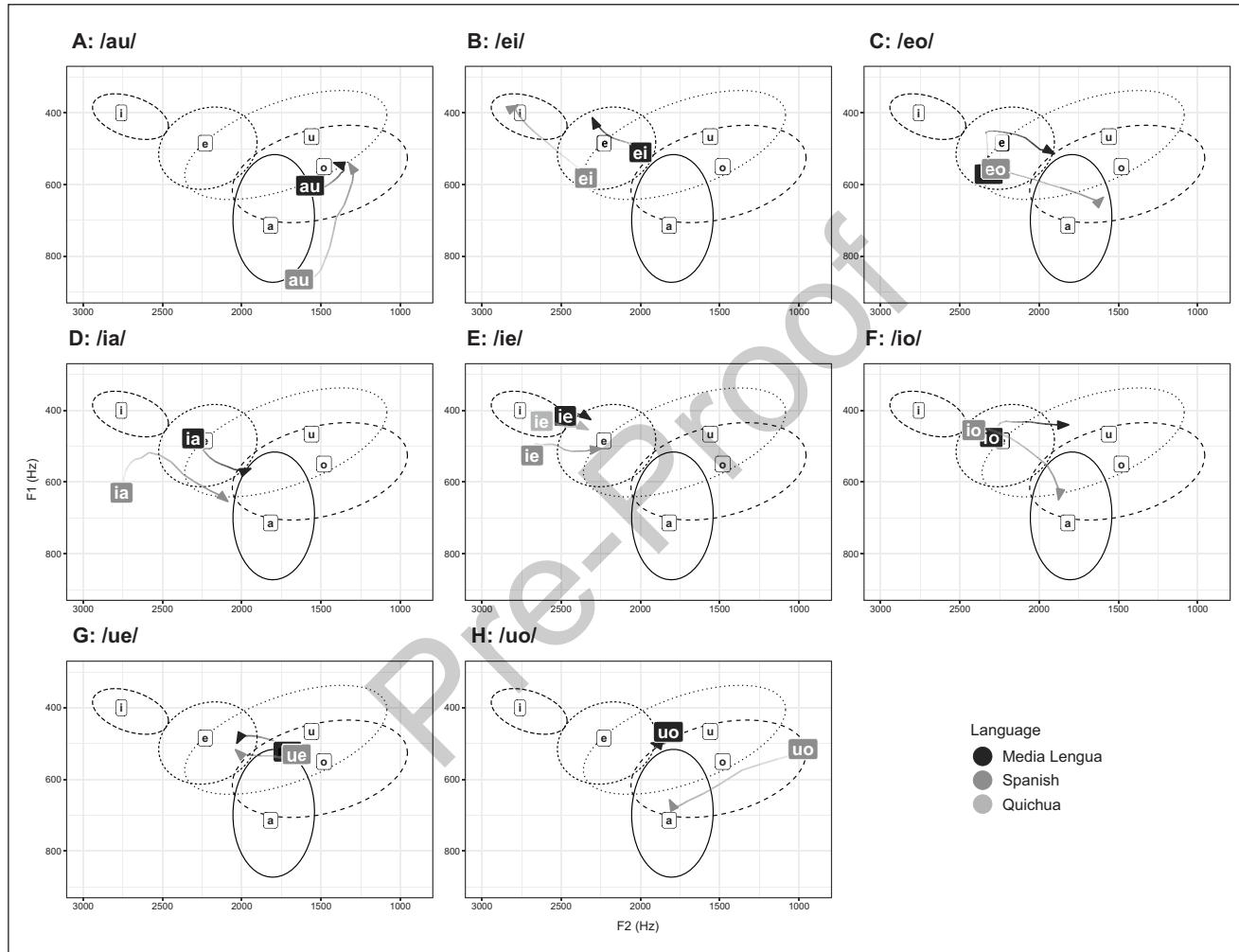
Several differences in acoustic target positions between ML and Spanish can be seen in Figure 5. For /a/, which occurs in /au/ (plot A) and /ia/ (D), the ML target is substantially higher than Spanish, most especially so when it is the initial target in the sequence in /au/. For /e/ the overall pattern is less clear-cut. In /ei/ (B) the ML /e/ target is more retracted than in Spanish, but this is not the case for /eo/ (C) or /ue/ (G) where the targets appear quite close. And in /ie/ (E) the ML /e/ is higher and slightly more advanced than Spanish. ML /i/ is similarly complex in its realizations relative to Spanish. It is more retracted and higher in /ia/ (D) and /ie/ (E), similarly retracted but slightly lower in /ei/ (B), and nearly identical in /io/ (F). ML /o/ exhibits more variation, being substantially higher than Spanish in both /io/ (F) and /uo/ (H), but somewhat retracted and lower in /eo/ (C). Finally, Spanish /u/ exhibits a split pattern. It is fairly similar to ML /u/ in both /au/ (A) and /ue/ (G), but massively more retracted in /uo/ (H).

Regarding intra-language observations of ML, nearly all the vowel sequences in Figure 5 show a noticeable shift in vowel quality or some degree of diphthongization, in the phonological sense (e.g., /au/-A, /ei/-B, /eo/-C, /ia/-D, /io/-F, and /ue/-G). However, there is little in the way of formant movement in /ie/ (E) and /uo/ (H), suggesting that these sequences may be functioning more like monophthongs in ML.

Figure 6 is identical to Figure 5 except that it includes the Spanish monophthongs in the background as context. Cross-comparisons of Figure 5 and Figure 6 suggest that the relationship between Spanish and ML vowel sequences are not necessarily parallel in structure even when considering differences in shape and size of the vowel spaces. For example, there are differences in proportionality with respect to the angular difference (measured in degrees)<sup>10</sup> between monophthong averages and the initial and end point averages of V+V sequences. For example, the ML cline is four times as steep in /ei/ compared to its average cline between monophthongs /i/ and /e/ (a 30° trajectory in /ei/ vs. a 7.4° average difference between monophthongs /e/ and /i/), while the Spanish /ei/ cline is only two times as steep compared to the average cline between Spanish monophthongs /i/ and /e/ (a 22.2° trajectory in /ei/ vs. a 42.2° between /e/ and /i/). Within this sequence, there are also small proportional differences in the average distances



**Figure 5.** Vowel sequence 20%–80% trajectories overlaid on the Media Lengua vowel space; background ellipses indicate one standard deviation.



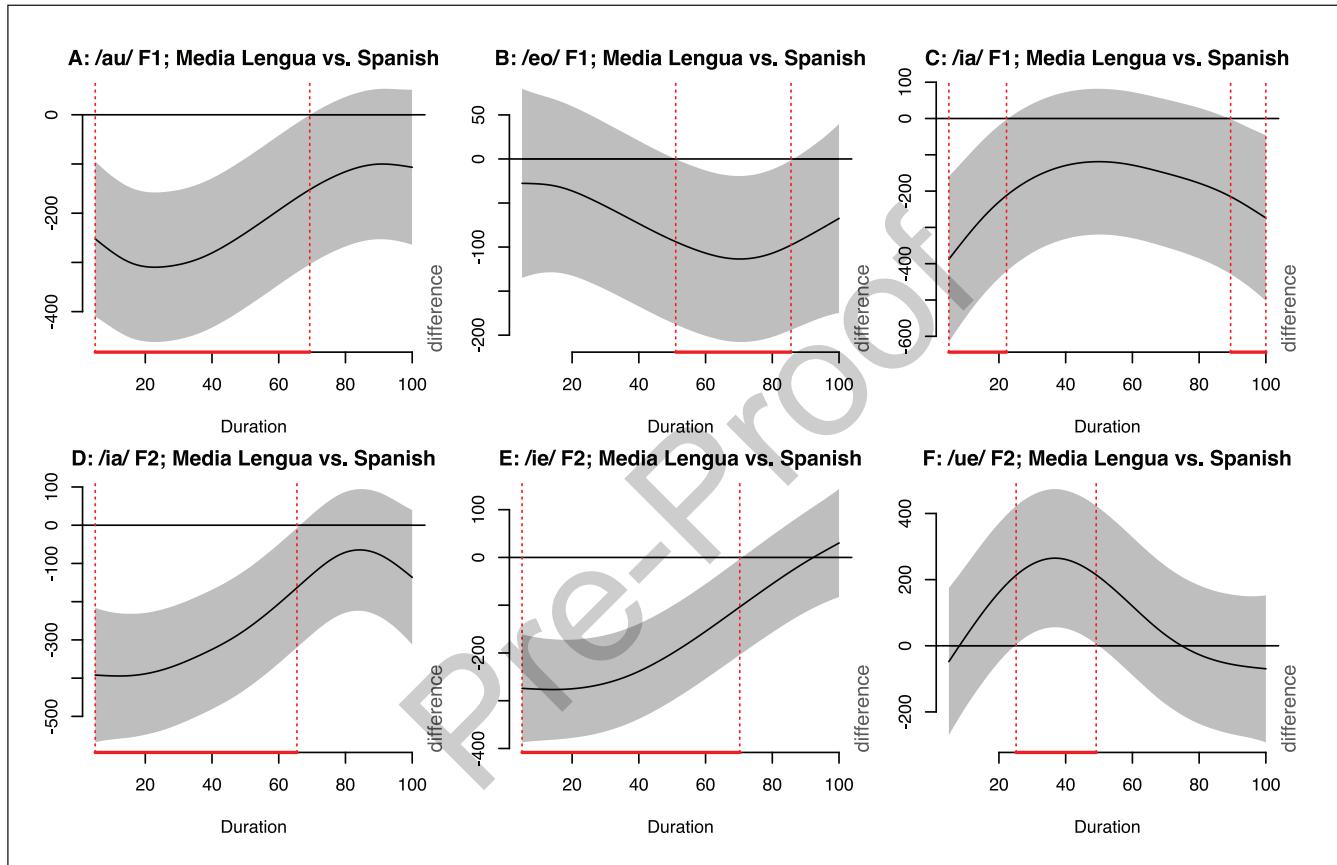
**Figure 6.** Vowel sequence 20%–80% trajectories overlaid on the Spanish vowel space; background ellipses indicate one standard deviation.

between monophthongs and V+V sequences in acoustic space. For instance, the distance between the average /e/ quality and the average /i/ quality in the /ei/ vowel sequence was 1.5 times more distant than the average distance between monophthongs /e/ and /i/, while Spanish showed a difference of just 1.25 times in the same analysis. These indicators may suggest that ML speakers reshaped the /ei/ trajectory to emphasize the change in vowel quality even though, when compared proportionally to Spanish, it is not immediately apparent. Similar proportional differences with respect to trajectory angles can be observed in /eo/ and /eu/ as well. However, two Spanish origin vowel sequences, /ie/ and /uo/, appear to have collapsed into monophthongs in ML (/i/ and either /o/ or /u/, respectively), as there is little in the way of dispersion (angular or distal) from the initial and end qualities in these sequences. This supports the hypothesis that Spanish and ML vowel sequences are not parallel in structure. Rather, ML has reshaped and reanalyzed Spanish origin vowel sequences to fit its phonological demands.

**4.2.2 GAMMs comparisons across three languages.** Five of the vowel sequences which indicated significant differences by *Language* in formant trajectories include data from Quichua as well as ML and Spanish, and are therefore most revealing of the relationship between ML and each of its contributory languages; those sequences are: /au, eo, ia, ie, ue/. Following computation of the GAMMs difference smooths between ML and each of Spanish and Quichua for F1 and F2 of each sequence, there was only a single instance of a significant difference in the M1 GAMM between ML and Quichua, F2 of /ie/; this will be discussed separately at the end of this subsection. The difference smooths between ML and Spanish, for these sequences according to which formants exhibited significant differences under GAMMs analysis, are shown in Figure 7. Note that F1 is negatively correlated with vowel height, such that the *lower* ML F1 productions in Figures 7A, B, and C are indicative of *higher* vowel articulations in terms of tongue position. And, F2 is positively correlated with vowel advancement, such that *lower* ML F2 productions (seen in Figures 7D and E) represent a more *retracted* articulation, while *higher* F2 productions (Figure 7F) represent a more *advanced* or *fronted* articulation. In general, ML vowel sequence productions show greater reduction, more centralization, and less aperture than their Spanish counterparts, which reflects the relatively smaller range of ML's overall vowel space relative to Spanish.

As mentioned above, F2 of /ie/ represents the only example in our data where a significant difference was found between ML and Quichua formant trajectories, illustrated in Figure 8. The difference smooth between ML and Quichua (as reference level) is shown in Figure 8A while the difference between ML and Spanish (as reference level) is repeated in Figure 8B (originally Figure 7E) for comparison. While ML F2 of /ie/ is lower (farther back, i.e., more retracted) than Quichua by between 100 and 200 Hz during the initial 63.5% of duration, this is a notably smaller degree of difference in comparison to that observed between ML and Spanish (Figure 8B), where the difference well exceeds 200 Hz for half of the sequence duration. These differences are further illustrated by layering the fitted per-language smooths into a single plot in Figure 8C, where the closer proximity between the ML and Quichua trajectories is readily apparent. In comparison, Spanish is substantially higher at onset and has a much steeper gradient during the transition towards offset. It is also worth reiterating that /ie/ in Quichua is non-native in origin, occurring only in Spanish origin words.

**4.2.3 GAMMs comparisons across two languages.** The other set of sequences exhibiting significant cross-linguistic differences in formant trajectories are: /ei, eu, io, uo/. As no Quichua data were available for these sequences due to low occurrence in any Quichua lexical items of Spanish origin, all differences reported in this subsection concern only ML versus Spanish. Recall that F1 is



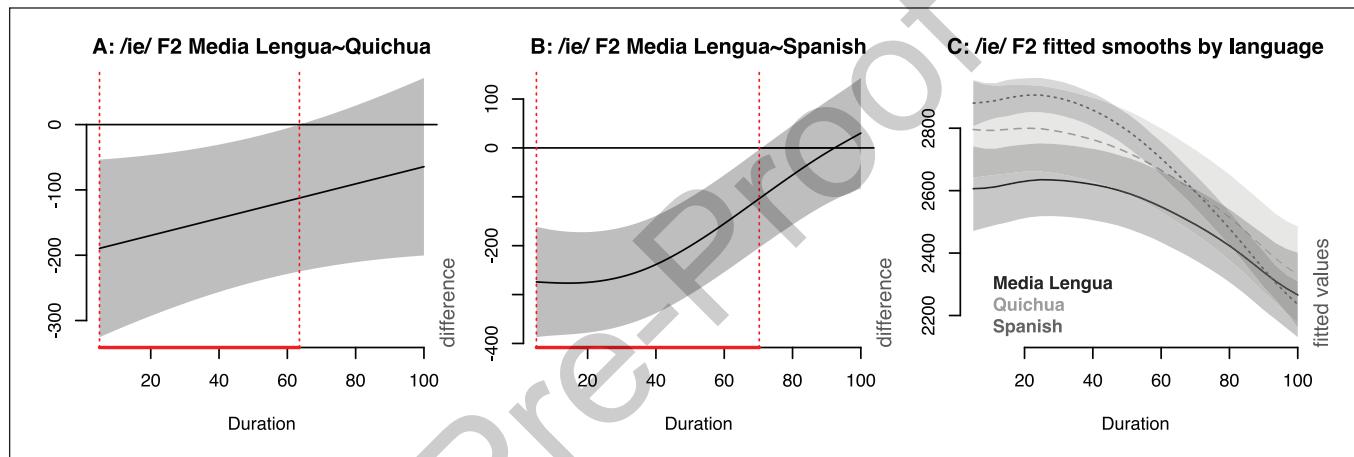
**Figure 7.** Formant (Hz) difference smooths: Media Lengua (ML) versus Spanish (three-language comparisons).

(A) F1 of /au/ (Figure 7A): The generalized additive mixed models (GAMMs) output indicates that ML F1 is lower than Spanish from onset until 69.3% of duration. This indicates that the ML articulation of at least the /a/ portion of /au/ is significantly raised in comparison with Spanish production;

(Continued)

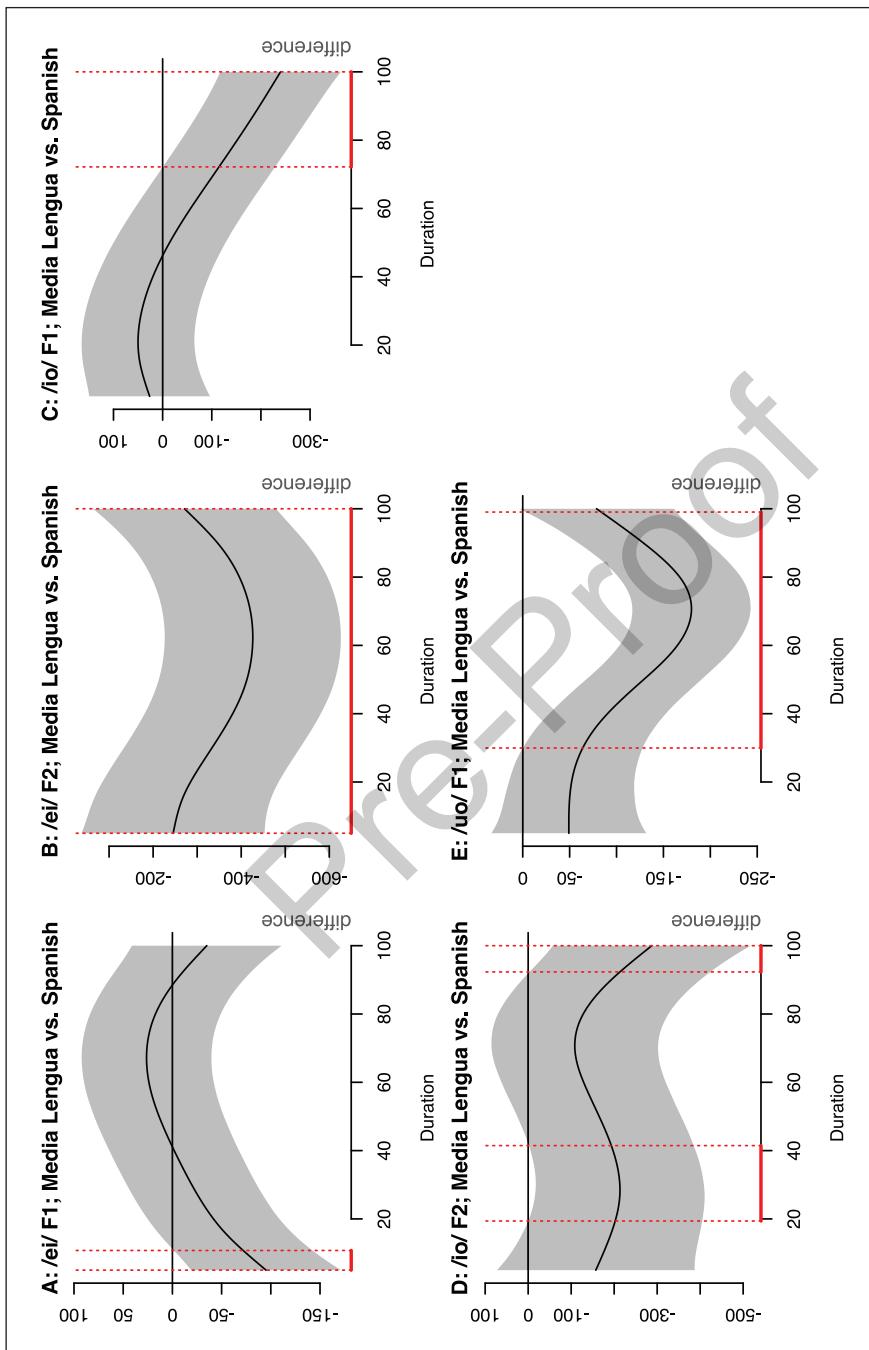
**Figure 7. (Continued)**

- (B) F1 of /eo/ (Figure 7B): The difference between ML F1, which is lower versus Spanish, with the GAMMs indicating the greatest degree of difference occurring between 51% and 85.6% of duration. This suggests that the difference is centered around the /o/ portion of /eo/, with ML having a *raised* articulation here relative to Spanish.
- (C) F1 of /ia/ (Figure 7C): Some differences between ML versus Spanish production are most strongly indicated during onset (prior to 27.3% of duration) and just at offset (after 89.4% duration); due to their limited scope at the edges of the sequence, these differences might be assumed to be due to coarticulatory effects with adjacent consonants (see discussion in subsection 3.3.2 regarding the inclusion of *Word* as a random effect).
- (D) F2 of /ia/ (Figure 7D): For this formant there is a more substantial difference from Spanish, with ML production being lower by several hundred Hz for the initial 65.5% of duration, indicative of a more *retracted* articulation centered around the /i/ portion of /ia/ and extending into the /a/ portion as well.
- (E) F2 of /ie/ (Figure 7E): ML production is substantially lower than Spanish for the initial 70.3% of duration, indicative of a more *retracted* articulation during the /i/ portion of /ie/ and perhaps extending to some degree into the /e/ portion as well.
- (F) F2 of /ue/ (Figure 7F): There is only a small window of robustly higher F2 production in ML versus Spanish for a brief period between 25.2% and 49.1% of duration, indicative of a more *advanced* articulation centered around the /u/ portion of /ue/.



**Figure 8.** /ie/ F2 (Hz) generalized additive mixed models' three-language comparison. In 8C, the solid line indicates the smooth for Media Lengua; the dashed line the smooth for Quichua; and the dotted line the smooth for Spanish.

(Continued)



**Figure 9.** Formant (Hz) difference smooths: Media Lengua (ML) versus Spanish (two-language comparisons).  
(A) F1 of /ei/ (Figure 9A): This difference is very short, occurring only at the onset until 0.8% of duration, and can probably thus be ignored.

**Figure 8.** (Continued)

- (B) F2 of /ei/ (Figure 9B): This difference covers 100% of duration, and is quite substantial, with ML being as much as 400 Hz lower than Spanish close to the mid-point, indicative of a more retracted ML articulation relative to Spanish. Note that, due to this difference extending across the full duration, the reference level indicating the Spanish F2 production is omitted from this plot—it is located above the plotted region, as indicated by the negative values shown on the y-axis. Although this difference occurs for the entire sequence, it reaches its greatest magnitude during the latter half of the trajectory, centered around the /i/ portion of /ei/.
- (C) F1 of /io/ (Figure 9C): The sequence /io/ exhibits some complex cross-linguistic patterning. For F1 the most robust region of difference occurs subsequent to 72.2% of duration, during which ML F1 progressively descends to as much as 200 Hz lower than Spanish, indicative of a more raised ML articulation relative to Spanish.
- (D) F2 of /io/ (Figure 9D): ML is lower than Spanish, with the greatest region of difference occurring during the early portion of the trajectory between 19.4% and 41.5% of duration, and again at the tail end of the sequence subsequent to 92.3% duration, with ML having a lower F2 during both periods, indicating a more retracted articulation in ML.
- (E) F1 of /io/ (Figure 9E): The difference here occurs during the majority of the trajectory, being most robust beginning from 29.9% of duration, ML F1 is lower by a substantial amount over this period, reaching nearly 200 Hz around the 70% mark, indicating a relatively raised ML articulation which is centered around the /o/ portion of /io/.

negatively correlated with vowel height, such that the *lower* ML F1 productions in Figures 9A, C, and E, are indicative of *higher* vowel articulations in terms of tongue position, and that F2 is positively correlated with vowel advancement, such that *lower* ML F2 productions (seen in Figures 9B and D) represent a more *retracted* articulation.

## 5

## Discussion

This is the first study to explore the acoustic production of vowel sequences in a mixed language along with its source languages. The results are novel in that we observe a vowel system that formed essentially like that of the original ancestral source language (Quichua, the original L1) yet has incorporated non-native vowel sequences from the introduced source language (Spanish, the original L2). Interestingly, the trajectories of these “borrowed” vowel sequences are modified to occupy a more central (Quichua-like) region of the vowel space and frequently exhibit markedly reduced trajectories compared to Spanish. At the same time, we find only one case, the sequence /ie/, where a ML vowel sequence differs significantly from its Quichua counterpart—and even in this case the difference from Spanish is substantially greater in comparison. Our findings show how the vowel system of ML successfully integrates nearly all novel (originally L2) vowel sequence patterns from Spanish into what is essentially Quichua phonology by markedly adapting their production while still maintaining contrasts which are not expressed in Quichua. The only two sequences, based on the plots in Figures 5 and 6, that do not appear diphthongal in ML are /ie/ and /uo/ (two phonologically parallel sequences). As an anonymous reviewer pointed out, these pairs look simply like the monophthongs /i/ and /o/ (or possibly /u/) and there appear to be some novel innovations in the behavior of /e/-initial V+V sequences, in that /e/ begins lower than one would expect based on the phonetics of the monophthong /e/. It seems that some interesting changes occurred when ML adapted these sequences from Spanish, and they are no longer easily analyzable as simple sequences of phonologically independent vowels.

The ML vowel space is also interesting in that its contrastive mid- and high-vowels do not conform to models of adaptive dispersion, which predict greater distances between categories in acoustic space to allow for optimal contrastability. Therefore, the shorter trajectories in ML vowel sequences (revealed in this study) and the overlapping mid-vowel and high-vowel categories (revealed in Stewart, 2014) suggest we are looking at a system that was created with unequal influences from its source languages. In fact, ML mid-vowels are only minimally distant enough from their high-vowel counterparts to be contrastive perceptually. This begs the question, why did the mid-vowels not just assimilate to the high-vowels or, opposingly, undergo greater dispersion? As per Stewart and Meakins (accepted), the ML vowel system reflects that of a late bilingual where interference from their L1 (Quichua) impedes native-like production in their L2 (Spanish). Therefore, the arrangement and shape of the vowel categories in acoustic space would suggest that ML was created by late bilinguals instead of early or simultaneous bilinguals. However, they hypothesize that the phonological “stresses” from relexification (e.g., high functional loads of contrastive non-native phonemes) from the mid-vowels may have been driving forces for maintaining/creating contrasts with the high-vowels in the predominately Quichua system. Subsequently, the vowel system could have been nativized with the overlaps “frozen” in place. Evidence from Stewart, et al. (2020a) supports the functional load hypothesis. From a sample of 1415 ML words transcribed by native speakers, approximately 80% of the lexicon contains one of /e, o/ either individually or as part of a vowel sequence. Thus, assimilating every /e/ → /i/ and /o/ → /u/ could have strained the usage of the Spanish lexicon which was optimized for high-vowel versus mid-vowel contrasts, creating unwarranted cognitive processing loads in the process.

Given that ML has indeed adopted contrastive Spanish origin mid-vowels (both productively and perceptually), there is little reason to believe that non-native vowel sequences would not also follow as the categories were already established. However, as the ML vowel categories are more centralized in acoustic space, like that of Quichua, the trajectory of the vowel sequences would also need to adapt to these regions. This can be seen in the ML formant trajectories of the vowel sequences which reflect Quichua vowel sequences in shape, size, and degree of variation. This might suggest why ML vowel sequences act more like those one might expect from Quichua, if such vowel sequences existed in the language, rather than like native Spanish vowel sequences.

One interesting contrast that appears among the vowel sequences which differ between ML versus Spanish (see Figures 5 and 6) concerns the trajectories of /ei/ and /ie/, two sequences which, if they are composed of simple monophthong sequences in hiatus, should be expected to be something like mirror images of each other. In fact, this is true for neither the ML nor the Spanish forms. In both languages, the rising-and-fronting sequence /ei/ shows formant trajectories which reflect this description, moving upwards and forwards in the vowel space. For /ie/, what might be predicted to be a lowering-and-retracting sequence occurrence instead shows in Spanish merely as a retracting movement. While this is also true for ML, this motion is so slight that, as noted above, it could easily be categorized as monophthongal. Ideally, we could compare the behavior of /ie, ei/ with the contrastive pair /uo, ou/ across Spanish and ML. Unfortunately (see subsection 2.1.2) /ou/ does not occur in ML or is not produced as [uo] in Spanish. We do note, however, that ML /uo/ appears, in parallel fashion to /ie/, as a simple monophthong with respect to its formant trajectory, while the same cannot be said of Spanish /uo/ which follows a clear lowering-and-fronting path, marking this development in ML as innovative.

## 6

## Conclusion

This study adds to the mixed language literature with the first acoustic analysis of vowel sequences. It supports findings from other mixed languages studies suggesting that phonological arrangements are not clear-cut cases of stratification or assimilation. Instead, phonological arrangements are more complex and suggest that age of acquisition of the source languages, cognitive factors, and functional load play a much more substantial role in determining the mixed language phonological system than divisions in the language's morphosyntax.

### ORCID iDs

Sky Onosson  <https://orcid.org/0000-0001-9308-7202>

Jesse Stewart  <https://orcid.org/0000-0001-8678-7884>

### Notes

1. The term “ancestral” is used in the chronological sense referring to the original homeland, pre-contact language as opposed to “introduced” which refers to the language introduced to this group, either through trade, colonization, or their own migration (i.e., post-contact).
2. The Quechuan languages spoken in Ecuador are broadly known as Quichua (or Kichwa) with internal variants prefaced by the province where they are spoken; for example, Imbabura Quichua or Cotopaxi Quichua.

3. In this example, Media Lengua (ML) orthography is based on the spelling system developed in Pijal and used in the ML dictionary (Stewart et al., 2020a). The Quichua orthography is based on the unified system adapted to Imbabura.
4. As noted by an anonymous reviewer, the corner vowels are not contrastive vowel categories in the traditional sense. We agree and only mention them to remain true to the original analysis of the Media Lengua vowel system in Stewart (2014). This study does not differentiate between Quichua-origin and Spanish-origin corner vowels due to their negligible differences.
5. Possible variation from each individual participant was considered in the statistical models, which made use of *participant* as a random effect.
6. Since providing a standardized assessment test/written questionnaire would be culturally insensitive, Spanish proficiency for both groups was judged using less intrusive methods: (a) feedback from our Quichua-speaking assistant, who was familiar with the participants; (b) our own interactions in Spanish with the participants after the experiment; and (c) informal oral self-assessment loosely based on a number of questions from the Bilingual Language Profile (BLP)8 (Birdsong et al., 2012).
7. We thank an anonymous reviewer for bringing these papers to our attention.
8. It should be noted that lexical stress in Spanish (e.g., tonic vs. atonic) was not considered as changes in formant trajectories have not been readily identified as a primary correlate in Spanish stress/accent. Instead, most studies focus on pitch, intensity, and duration as acoustic correlates of stress in most Spanish dialects (see e.g., Contreras, 1964; Navarro-Tomás, 1964). In addition, when Ortega-Llebaria and Prieto (2007) analyzed formant frequency as a correlate of stress, they only found a slight tendency towards centralization and only with unstressed [o]; a very similar finding from a forthcoming study on Media Lengua vowel correlates shows that unstressed vowels were only slightly retracted (55 Hz on F2) (Onosson & Stewart, under-review).
9. Although Quichua data were included in the comparisons of /ie/ and /ue/, these sequences are non-native to Quichua and exclusively occur in borrowings from Spanish.
10. Differences in  $F1 \times F2$  averages were measured with a protractor using the initial vowel quality of a sequence as the starting point, for example, [e] in the [ei] vowel sequence.

## References

- Baayen, H., Vasishth, S., Kliegl, R., & Bates, D. (2017). The cave of shadows: Addressing the human factor with generalized additive mixed models. *Journal of Memory and Language*, 94, 206–234. <https://doi.org/10.1016/j.jml.2016.11.006>
- Bakker, P., & Muysken, P. (1994). Mixed languages and language intertwining. In J. Arends, P. Muysken, & N. Smith (Eds.), *Pidgins and creoles: An introduction* (pp. 41–52). John Benjamins.
- Birdsong, D., Gertken, L. M., & Amengual, M. (2012). *Bilingual language profile: An easy-to-use instrument to assess bilingualism*. University of Texas at Austin. <https://sites.la.utexas.edu/bilingual/>
- Boersma, P., & Weenink, D. (2019). Praat: doing phonetics by computer. *Computer application*. <http://www.fon.hum.uva.nl/praat/>
- Buchan, H. (2012). *Phonetic variation in Gurindji Kriol and Northern Australian English: A longitudinal study of fricatives in maternal speech* (Doctoral dissertation, University of Wollongong). <https://ro.uow.edu.au/theses/3789/>
- Bundgaard-Nielsen, R., & O'Shannessy, C. (2019). Voice onset time and constriction duration in Warlpiri stops (Australia). In S. Calhoun, P. Escudero, M. Tabain, & P. Warren (Eds.), *Proceedings of the 19th International Congress of Phonetic Sciences, Melbourne, Australia 2019* (pp. 3612–3616). Australasian Speech Science and Technology Association Inc. [https://assta.org/proceedings/ICPhS2019/papers/ICPhS\\_3661.pdf](https://assta.org/proceedings/ICPhS2019/papers/ICPhS_3661.pdf)
- Contreras, H. (1964). ¿Tiene el español un acento de intensidad? [Does Spanish have an accent of intensity?] *Boletín Del Instituto de Filología de La Universidad de Chile*, 16, 237–239. [In Spanish.]
- Deibel, I. (2017). *VO vs. OV: What conditions word order variation in Media Lengua?* Colloquium on Mixed Languages, Bremen, Germany.
- Deibel, I. (2019). Adpositions in Media Lengua: Quichua or Spanish? – Evidence of a lexical–functional split. *Journal of Language Contact*, 12 (2), 404–439. <https://doi.org/10.1163/19552629-01202006>

- Deibel, I. (2020). *Language representations in the presences of a lexical-functional split: An experimental approach targeting the Quichua–Media Lengua–Spanish interface* (Doctoral dissertation, The Pennsylvania State University).
- Flege, J. (2007). Language contact in bilingualism: Phonetic system interactions. In J. I. Hualde & J. S. Cole (Eds.), *Laboratory phonology 9* (pp. 353–380). Mouton de Gruyter.
- Gómez-Rendón, J. (2005). La Media Lengua de Imbabura [The Media Language of Imbabura.] In H. Olbertz & P. Muysken (Eds.), *Encuentros y conflictos: Bilingüismo y contacto de lenguas en el mundo Andino* [Encounters and conflicts: Bilingualism and language contact in the Andean world] (pp. 39–58). Iberoamericana. [In Spanish.]
- Hastie, T. J., & Tibshirani, R. J. (1990). *Generalized additive models*. Chapman and Hall.
- Hendy, C. R. (2019). *The distribution and acoustic properties of fricatives in Light Warlpiri* (B.A. Honours Thesis, Australian National University). <https://openresearch-repository.anu.edu.au/handle/1885/200483>
- Johnson, K. (2000). Adaptive dispersion in vowel perception. *Phonetica*, 57(2–4), 181–188. <https://doi.org/10.1159/000028471>
- Jones, C., & Meakins, F. (2013). Variation in voice onset time in stops in Gurindji Kriol: Picture naming and conversational speech. *Australian Journal of Linguistics*, 33(2), 196–220. <https://doi.org/10.1080/07268602.2013.814525>
- Jones, C., Meakins, F., & Buchan, H. (2011). Comparing vowels in Gurindji Kriol and Katherine English: Citation speech data. *Australian Journal of Linguistics*, 31(3), 305–326. <https://doi.org/10.1080/0726602.2011.598629>
- Jones, C., Meakins, F., & Mauwiyath, S. (2012). Learning vowel categories from maternal speech in Gurindji Kriol. *Language Learning*, 62(4), 1052–1078. <https://doi.org/1010.1111/j.1467-9922.2012.00725.x>
- Kewley-Port, D. (2001). Vowel formant discrimination II: Effects of stimulus uncertainty, consonantal context, and training. *Journal of the Acoustical Society of America*, 10(4), 2141–2155. <https://doi.org/10.1121/1.1400737>
- Liljencrants, J., & Lindblom, B. (1972). Numerical simulation of vowel quality systems: The role of perceptual contrast. *Language*, 48(4), 839–862. <https://doi.org/1010.2307/411991>
- Lindblom, B. (1986). Phonetic universals in vowel systems. In J. Ohala & J. Jaeger (Eds.), *Experimental phonology* (pp. 13–44). Academic Press.
- Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. In W. Hardcastle & A. Marchal (Eds.), *Speech production and speech modeling* (pp. 403–439). Kluwer Academic.
- Lipski, J. (2016). Language switching constraints: More than syntax? Data from Media Lengua. *Bilingualism: Language and Cognition*, 1(4), 1–25. <https://doi.org/10.1017/S1366728916000468>
- Livijn, P. (2000). Acoustic distribution of vowels in differently sized inventories – hot spots or adaptive dispersion? In *Proceedings of the XIIIth Swedish Phonetics Conference (FONETIK 2000), Skövde, Sweden, May 24–26, 2000* (pp. 93–96). [https://www2.ling.su.se/fon/perilus/2000\\_11.pdf](https://www2.ling.su.se/fon/perilus/2000_11.pdf)
- Meakins, F. (2013). Mixed languages. In Y. Matras & P. Bakker (Eds.), *Contact languages: A comprehensive guide* (pp. 159–228). Mouton de Gruyter.
- Meakins, F., & Stewart, J. (accepted). Mixed languages. In S. Mufwene & A. M. Escobar (Eds.), *Cambridge handbook of language contact*. Cambridge University Press.
- Muysken, P. (1981). Halfway between Quechua and Spanish: The case for relexification. In A. R. Highfield (Ed.), *Historicity and variation in Creole studies* (pp. 57–78). Karoma Publishers.
- Muysken, P. (1997). Media Lengua. In S. Thomason (Ed.), *Contact languages: A wider perspective* (pp. 365–426). John Benjamins.
- Navarro-Tomás, T. (1964). La medida de la intensidad [The measure of intensity]. *Boletín Del Instituto de Filología de La Universidad de Chile*, 16, 231–235. <https://boletinfilologia.uchile.cl/index.php/BDF/article/view/49444/51905> [In Spanish.]
- Onosson, S. (2018). *An acoustic study of Canadian raising in three dialects of North American English*. (Doctoral dissertation, University of Victoria). <https://dspace.library.uvic.ca/handle/1828/9274>
- Onosson, S., & Bird, S. (2019). Differences in vowel-glide production between L1 and L2 speakers of Hul'q'umi'num'. In S. Calhoun, P. Escudero, M. Tabain, & P. Warren (Eds.), *Proceedings of the 19th International Congress of Phonetic Sciences, Melbourne, Australia 2019* (pp. 984–988). Australasian

- Speech Science and Technology Association Inc. [https://assta.org/proceedings/ICPhS2019/papers/ICPhS\\_1033.pdf](https://assta.org/proceedings/ICPhS2019/papers/ICPhS_1033.pdf)
- Onosson, S., & Stewart, J. (under-review). Suppressing biases in correlate identification in acoustic data: The case of Media Lengua. *Laboratory Phonology*.
- Ortega-Llebaria, M., & Prieto, P. (2007). Disentangling stress from accent in Spanish: Production patterns of the stress contrast in deaccented syllables. In P. Prieto, J. Mascaró, & M.-J. Solé (Eds.), *Segmental and prosodic issues in Romance phonology* (pp. 155–176). John Benjamins.
- Padgett, J. (2008). Glides, vowels, and features. *Lingua*, 118(12), 1937–1955. <https://doi.org/10.1016/j.lingua.2007.10.002>
- R Core Team. (2019). R: A language and environment for statistical computing. Programming language. *R Foundation for Statistical Computing, Vienna, Austria*. <https://www.r-project.org/>
- Rosen, N. (2006). Language contact and Michif stress assignment. *Language Typology and Universals*, 59, 170–190. <https://doi.org/10.1524/stuf.2006.59.2.170>
- Rosen, N. (2007). *Domains in Michif phonology*. (Doctoral dissertation, University of Toronto).
- Rosen, N., Stewart, J., & Sammons, O. (2020). How ‘mixed’ is mixed language phonology – An acoustic analysis of the Michif vowel system. *Journal of the Acoustical Society of America*, 147(4), 2989–2999. <https://doi.org/10.1121/10.0001009>
- Roettger, T. B., Winter, B., & Baayen, H. (2019). Emergent data analysis in phonetic sciences: Towards pluralism and reproducibility. *Journal of Phonetics*, 73, 1–7. <https://doi.org/10.1016/j.wocn.2018.12.001>
- RStudio Team. (2019). RStudio: Integrated Development for R. Computer application. RStudio, Inc., Boston, MA. <https://www.rstudio.com/>
- Sánchez Miret, F. (1998). Some reflections on the notion of diphthong. *Papers and Studies in Contrastive Linguistics*, 34, 27–51.
- Sóskuthy, M. (2017). *Generalised additive mixed models for dynamic analysis in linguistics: A practical introduction*. [http://eprints.whiterose.ac.uk/113858/2/1703\\_05339v1.pdf](http://eprints.whiterose.ac.uk/113858/2/1703_05339v1.pdf)
- Stewart, J. (2011). *A brief descriptive grammar of Pijal Media Lengua and an acoustic vowel space analysis of Pijal Media Lengua and Imbabura Quichua*. (Master’s Thesis, University of Manitoba). <http://hdl.handle.net/1993/4882>
- Stewart, J. (2014). A comparative analysis of Media Lengua and Quichua vowel production. *Phonetica*, 71(3), 159–182. <https://doi.org/10.1159/000369629>
- Stewart, J. (2015a). Intonation patterns in Pijal Media Lengua. *Journal of Language Contact*, 8(2), 223–262. <https://doi.org/10.1163/19552629-00802003>
- Stewart, J. (2015b). *Production and perception of stop consonants in Spanish, Quichua, and Media Lengua*. (Doctoral dissertation, University of Manitoba).
- Stewart, J. (2018). Vowel perception by native Media Lengua, Quichua, and Spanish speakers. *Journal of Phonetics*, 71, 177–193. <https://doi.org/10.1016/j.wocn.2018.08.005>
- Stewart, J. (2020). A preliminary, descriptive survey of rhotic and approximant fricativization in Northern Ecuadorian Andean Spanish varieties, Quichua, and Media Lengua. In R. Rao (Ed.), *Spanish phonetics and phonology in contact: Studies from Africa, the Americas, and Spain* (103–140). John Benjamins.
- Stewart, J., & Meakins, F. (accepted). Advances in mixed language phonology: An overview of three case studies. In M. Mazzoli & E. Sippola (Eds.), *New perspectives on mixed languages: From core to fringe. language contact and bilingualism*. De Gruyter Mouton.
- Stewart, J., Ayala, G. P., & Gonza Inlago, L. (2020a). Media Lengua dictionary. *Dictionaria*, 12, 1–3216. <https://dictionaria.clld.org/contributions/media lengua>
- Stewart, J., Meakins, F., Algy, C., Ennever, T., & Joshua, A. (2020b). Fickle fricatives: Obstruent perception in Gurindji Kriol and Roper Kriol. *Journal of the Acoustical Society of America*, 147, 2766. <https://doi.org/10.1121/10.0000991>
- Stewart, J., Meakins, F., Algy, C., & Joshua, A. (2018). The development of phonological stratification: Evidence from stop voicing perception in Gurindji Kriol and Roper Kriol. *Journal of Language Contact*, 11(1), 71–112. <https://doi.org/10.1163/19552629-01101003>
- van Rij, J., Wieling, M., Baayen, R., & van Rijn, H. (2020). itsadug: Interpreting Time Series and Autocorrelated Data Using GAMMs. R package version 2.4. <https://rdrr.io/cran/itsadug/>

- Wood, S. N. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 73(1), 3–36. <https://doi.org/10.1111/j.1467-9868.2010.00749.x>
- Wood, S. N. (2017). *Generalized additive models: An introduction with R* (2nd ed.). Taylor and Francis Group.
- Xu, Y., & Gao, H. (2018). FormantPro as a tool for speech analysis and segmentation. *Revista De Estudios Da Lingagem*, 26(4), 1435–1454. <https://doi.org/10.17851/2237-2083.26.4.1435-1454>

## Appendix: Participants

Tables A1 through A3 provide information pertaining to the participants who took part in the production portion of this study. The data include: speaker code; age at the time of recording; gender; formal education; level of Spanish; Quichua and Media Lengua; frequency of language use; recording type; and place of residence.

**Table A1.** This table provides demographic information about the Media Lengua (ML) participants including: age at the time of recording; gender; level of formal education; level of Spanish; ML; and Quichua; frequency of ML usage; recording type; and place of residency. It also includes the type of recording used to gather the data.

### ML participants

Code	Recording type	Gender	Age	Formal education	Spanish level	ML / Quichua level	ML Usage	Place of residence
41	Conversations	M	65	Primary	High	Native	Intermittently	Pijal
42	Wordlist	F	50	None	High	Native	Daily	Pijal
43	Elicitations, wordlist	F	39	Secondary	High	Native	Daily	Pijal
44	Elicitations, wordlist	F	39	Primary	High	Native	Intermittently	Pijal
48	Wordlist	F	50	Primary	High	Native	Intermittently	Pijal
49	Wordlist	F	50	Primary	High	Native	Intermittently	Pijal
50	Elicitations, wordlist	F	42	Secondary	Mid	Native	Intermittently	Pijal
51	Elicitations, wordlist	F	60	None	Mid	Native	Infrequently	Pijal
52	Wordlist	M	40	Primary	High	Native	Infrequently	Pijal
53	Conversations	M	58	Primary	High	Native	Intermittently	Pijal
56	Conversations, wordlist	F	64	Primary	High	Native	Intermittently	Pijal
58	Wordlist	M	33	University	Native	Native	Infrequently	Pijal
59	Wordlist	M	38	University	Native	Native	Infrequently	Pijal
62	Wordlist	M	54	Primary	High	Native	Intermittently	Pijal
79	Elicitations, wordlist	M	45	Primary	High	Native	Intermittently	Pijal
S011b	Conversations	F	60+	Primary	High	Native	Intermittently	Pijal
S01a	Conversations	F	60+	Primary	High	Native	Intermittently	Pijal
S02a	Conversations	F	45+	Primary	High	Native	Intermittently	Pijal
S03a	Conversations	F	60+	Primary	Mid	Native	Intermittently	Pijal
S03b	Conversations	F	60+	Primary	Mid	Native	Intermittently	Pijal
S04b	Conversations	M	60+	Primary	High	Native	Intermittently	Pijal
S07b	Conversations	M	60+	Primary	High	Native	Intermittently	Pijal
Z013c	Conversations	F	60+	Primary	Mid	Native	Intermittently	Pijal
Z016b	Conversations	F	60+	Primary	High	Native	Intermittently	Pijal

**Table A2.** This table provides demographic information about the Quichua speaking participants including: age at the time of recording; gender; level of formal education; level of Spanish and Quichua; frequency of Quichua usage; recording type; and place of residency.

Quichua participants

Speaker code	Age	Gender	Formal education	Spanish level	Quichua level	Quichua usage	Recording type	Place of residence
63	66/68	M	Primary	Mid/High	Native	Daily	Elicitation/wordlist	Chirihuasi
64	62	F	None	Low	Native	Daily	Elicitation	Chirihuasi
65	45	F	None	Low	Native	Daily	Elicitation	Chirihuasi
69	29	F	None	Low	Native	Daily	Elicitation	Chirihuasi
70	21	F	None	Low	Native	Daily	Elicitation	Chirihuasi
72	42	M	Not applicable (NA)	Mid/High	Native	Daily	Elicitation	Chirihuasi
73	70	M	University	Mid/High	Native	Daily	Elicitation	Chirihuasi
75	52	M	Secondary	Native	Native	Daily	Elicitation	Chirihuasi
76	55	F	NA	Mid/High	Native	Daily	Elicitation	Chirihuasi
77	49	M	Secondary	High	Native	Daily	Elicitation	Chirihuasi

**Table A3.** This table provides demographic information about the Spanish speaking participants including: age at the time of recording; gender; level of formal education; level of Spanish and Quichua; recording type; and place of residency.

Spanish participants

Speaker code	Age	Gender	Formal education	Spanish level	Quichua level	Recording type	Place of residence
68	27	Male (M)	University	Native	None	Sentence	Ibarra
73	34	Female (F)	University	Native	None	Sentence	Ibarra
71	21	M	University	Native	None	Sentence	Ibarra
97	22	F	University	Native	None	Wordlist	Quito
98	35	F	University	Native	None	Wordlist	Quito
100	38	M	University	Native	None	Wordlist	Quito
101	33	M	University	Native	None	Wordlist	Quito
102	35	M	University	Native	None	Wordlist	Quito
103	57	M	University	Native	None	Wordlist	Quito
104	34	F	University	Native	None	Wordlist	Quito
105	30	F	University	Native	None	Wordlist	Quito
106	24	F	University	Native	None	Wordlist	Quito
107	38	F	University	Native	None	Wordlist	Quito
108	24	F	University	Native	None	Wordlist	Quito