Original Paper

Phonetica 2014;71:159–182
DOI: 10.1159/000369629

Received: January 30, 2014
Accepted: November 4, 2014
Published online: February 2, 2015

A Comparative Analysis of Media Lengua and Quichua Vowel Production

Jesse Stewart

University of Manitoba, Winnipeg, Man., Canada

Abstract

This study presents a comparative analysis of F1 and F2 vowel frequencies from Pijal Media Lengua (PML) and Imbabura Quichua. Mixed-effects models are used to test Spanish-derived high and low vowels against their Quichua-derived counterparts for statistical significance. Spanish-derived and Quichua-derived high vowels are also tested against Spanish-derived mid vowels. This analysis suggests that PML may be manipulating as many as eight vowels where Spanish-derived high and low vowels coexist as near-mergers with their Quichua-derived counterparts, while high and mid vowels coexist with partial overlap. Quichua, traditionally viewed as a three-vowel system, shows similar results and may be manipulating as many as six vowels.

1 Introduction

This article investigates the acoustic nature of intra-group (within-group) vowel systems in two languages with varying degrees of lexical interference from Spanish spoken in the Ecuadorian province of Imbabura. The first is the Imbabura dialect of Quichua[1] spoken by 81.9% of the provincial population (Buttner, 1993, p. 48) where an estimated one fifth (21%) of the total lexicon is borrowed from Spanish (Gómez-Rendón, 2007, p. 517). The second language is the Pijal dialect of Media Lengua (ML) spoken by an estimated 300 people in the community of Pijal Bajo. Based on both a 200-word Swadesh list (Swadesh, 1952) and analyses of spontaneous speech, approximately 89–93% of the language's lexicon consists of Spanish-derived borrowings.

Different varieties of ML[2] have emerged throughout the Andean region of Ecuador including several documented cases in the provinces of Cotopaxi (Muysken, 1980, 1981, 1997), Imbabura (Gómez-Rendón, 2007, 2008; Stewart, 2011, 2013) and several

[1] The Ecuadorian variety of Quechua is known as Quichua or Kichwa /ˈki.ʧua/ by both mestizo and indigenous populations.
[2] Media Lengua literally translates to 'half-language' from Spanish. See Muysken (1997), Gómez-Rendón (2005) and Stewart (2011) for a more in-depth description of ML varieties.

KARGER 125

E-Mail karger@karger.com
www.karger.com/pho

0031–8388/15/0713–0159
$39.50/0

Jesse Stewart
Department of Linguistics, University of Manitoba
534 Fletcher Argue Building
Winnipeg, MB R3T 5V5 (Canada)
E-Mail umste247@myumanitoba.ca

lesser studied varieties in the provinces of Cañar and Loja (Muysken, 1997). Several hypotheses exist to its origin and distinct evolutionary path away from Quichua. Muysken (1997) suggests ML arose through ethnic self-identification for indigenous populations who could not identify completely with either rural Quichua or urban Spanish cultures. Gómez-Rendón (2005) posits ML spoken in the community of Angla, near Pijal Bajo, arose due to prolonged contact between the Quichua-speaking with the Spanish-speaking populations. Dikker (2008, p. 121) believes ML was 'created by men who had Quichua as their native language but left to work in Spanish speaking areas. When they returned to their communities, they had been using Quichua on an infrequent basis, while having acquired relatively fluent urban Spanish.' She suggests ML was used as a link between the older monolingual Quichua-speaking generation and the younger monolingual Spanish-speaking generation.

ML is often described as a prototypical bilingual mixed language (Backus, 2003; McConvell and Meakins, 2005) because of its split between roots and suffixes. Relexification, a cognitive process involved in the relabelling of lexical entries from one language to another (Lefebvre, 2005, 2006; Lefebvre and Therrien, 2007; Muysken, 1981), is credited as the primary process for developing ML's lexicon while other processes such as translexification (Muysken, 1980, 1981), lexical freezing (Gómez-Rendón, 2005; Muysken, 1997; Stewart, 2011), adlexification (Shappeck, 2011) and code-switching (Stewart, 2011) also appear to play a more minor role. In ML nearly all the native Quichua lexical roots, including core vocabulary, are replaced by their Spanish counterparts. On the surface, however, the Spanish-derived lexicon in ML appears to conform to Quichua phonology while maintaining Quichua word order and the vast majority of Quichua's agglutinating suffixes. Example (1) illustrates a typical Pijal Media Lengua (PML) sentence. The italicized elements are derived from Spanish:

(1) *si no aseti*-ta *okupa*-kpika *uebo*-ka *sarten*-pi-mi *pega*-ʃpa *keda*-n
if not oil-ACC use-SUB.DS egg-TOP pan-LOC-VAL stick-SUB.SS remain-3
'If you don't use oil, the egg will stick to the pan.' (Consultant 50)

All three languages referenced in this study (Quichua, ML and Spanish) have relatively small vowel inventories. Traditionally, both ML (Muysken, 1997) and Imbabura Quichua (IQ)(Guion, 2003) are considered three-vowel systems made up of /i/, /u/ and /a/ while Spanish contains two additional vowels: /e/ and /o/. Muysken (1997, p. 365) says ML often collapses the mid vowels /e/ and /o/, in the Spanish-derived lexicon, to /i/ and /u/, respectively. Under certain conditions, however, e.g., proper names, interjections, stressed positions and certain lexical items, the mid vowels may be retained.

In the community of Pijal, there is generally a negative language attitude towards both ML and Quichua (more so for ML). Some speakers, however, still value Quichua and would like to see the language maintained in the community. Most people between the ages of ~20 and 35 are passive bilinguals of both ML and Quichua, while those under the age of 20 are often Spanish monolinguals. The morphosyntactic structure of ML is often more conservative for speakers aged ~60 and above while more Spanish influence appears in the grammar of those in the ~35–50 age range. It also appears that women tend to use ML more than men. Similar sociolinguistic trends regarding Quichua are also found in the community of Chirihuasi. People under the age of 20 are often monolinguals or passive bilinguals, while those above the age of 30 are typically L1 Quichua and began learning Spanish upon entering school. People aged ~60 are either late L2 Spanish bilingual or monolingual. The community of Cashaloma is more

DOI: 10.1159/000369629

conservative than Chirihuasi. It is common to find monolingual women in their thirties and most children are early bilinguals (Quichua L1). ML is not found in either of these communities.

In this article, I attempt to answer the question: what is the phonetic nature of vowel production in these contact situations? Do vowels merge into a single system where L2 borrowings undergo complete phonetic assimilation? Do they function in a dual system where separate vowels are used depending on the origin of the morpheme in question? Do they coexist as an intermediate variety with overlapping formant frequencies, i.e., varying degrees of merger?

To date only two studies look at vowel production in mixed languages. Jones et al. (2011) compare vowel production in Gurindji Kriol, a mixed language derived from Gurindji and Kriol (an English-lexifier creole) spoken in the Northern Territory, Australia, to the Australian English spoken in the nearest town of Katherine. Their study reveals significant differences in the relative frequencies of the first and second formants (henceforth F1, F2) resulting in more formant overlap in the front vowels /æ/ and /e/ and back vowels /ʉː/ and /oː/ in Gurindji Kriol compared to those found in Katherine English. They also found that the duration differences between the Gurindji Kriol short and long vowel contrasts were also reduced compared to those in Katherine English.

In the second study, Rosen (2007) shows that in Michif, a mixed language of Plains Cree and Metis French origin, back vowel production is variable. Plains Cree only has one high-mid back vowel (/o/), typically consisting of a higher F1 frequency than French /u/. When the high back vowel is produced in a French-derived lexeme, it may be produced as either [u] or [o], however, when derived from the Plains Cree's 'lower-high' back vowel, it only surfaces as [o].

While other studies are lacking in the mixed language literature, the idea of phonetic duality among bilinguals has been a common topic of linguistic analysis. The Perceptual Assimilation Model proposed by Best et al. (2003) predicts that bilinguals assimilate L2 sounds based on how similar or contrastive a given sound is perceived. This model suggests that bilinguals have only one phonological system where L2 sounds are produced on the basis of L1 patterns. Within this system, categories are allowed to (1) merge into a single category, (2) remain independent, or (3) coexist with varying degrees of overlap. This model would therefore, predict that in contact situations Spanish-derived /e/ and /o/ might emerge as new vowels while /i/, /u/, and /a/ may or may not end up with Quichua and Spanish subsets.

Flege's (2007) Speech Learning Model suggests that when an L2 learner establishes a new category, crowding of the phonetic space occurs, causing dispersion in order to maintain phonetic contrast. The Speech Learning Model proposes that categories operate in the same phonological space and readjust according to external conditions. Adaptive dispersion models (Johnson, 2000; Liljencrants and Lindblom, 1972; Lindblom, 1986, 1990; Livijn, 2000) predict that crowding of phonological categories causes an increase in acoustic range in order to maintain contrast. Therefore high vowels in a five-vowel system like Spanish should be produced with lower and more fronted F1 frequencies than those in a three-vowel system like Quichua. The same model also predicts that Spanish-produced [a] should have a higher F1 frequency (lower vowel range) than that of Quichua [a].

Another important factor regarding phonetic duality in this study is that of merger (assimilation) and near-merger (covert contrast) (Hickey, 2004; Labov et al., 1972,

Phonetica 2014;71:159–182
DOI: 10.1159/000369629

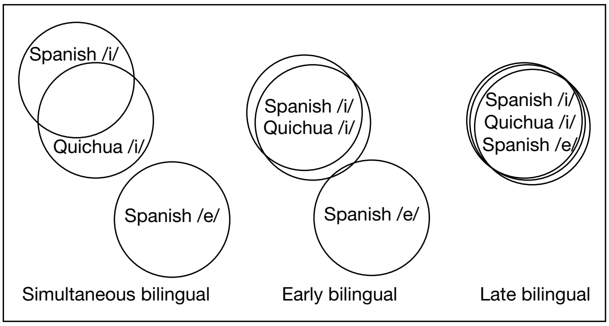


Fig. 1. L1 Quichua speakers' front vowel production of L2 Spanish vowels according to age of acquisition. Based on Guion (2003).

1991). Merger is broadly defined as a unidirectional sound change where two phonemes approximate each other until collapsing into a single sound, at which point no distinction is discernible at the phonetic level. Near-mergers, on the other hand, are produced when a speaker consistently makes small articulatory differences between sounds of two lexical sets but cannot distinguish them auditorily.

Several studies have investigated the phonological nature of vowel systems in monolingual and bilingual groups of Quechuan speakers in Ecuador and Peru. One such study of paramount importance to this investigation is Guion's (2003) article on the phonological systems of Quichua-Spanish bilinguals from Imbabura. In this study, both cross-language and within-language production of vowel data is investigated. Her cross-language results report that simultaneous bilinguals (SBs) maintained three separate front vowels: an [i] with lower F1 frequencies for Spanish production, an [i] with higher F1 frequencies for Quichua production, and an [e] for Spanish production. Early (but not simultaneous) L2 learners, on the other hand, tended to merge Spanish [i] and Quichua [i] into the same vowel space while late L2 learners merged both [i]s and the Spanish [e] into roughly the same Quichua [i] space (fig. 1).

Similar trends were also reported with back vowels [u] and [o] with some variation of Quichua [u] production in early bilinguals which manifested as a rough equivalent to Spanish [o] or [u]. Her findings also suggested that both SBs and early bilinguals maintain separate low vowel categories where Quichua [a] production is lower in F1 frequency than Spanish [a] production. Late bilinguals on the other hand, merged Spanish [a] with Quichua [a]. Her within-language results suggest that SBs and early bilinguals show an upwards shift in vowel space away from monolingual production towards that of Spanish. Guion's (2003) findings also suggest that the distinct organization of vowel categories is linked to the developmental differences related to a speaker's age of L2 acquisition. The earlier a person is exposed to their L2, the greater the chance they will acquire the necessary perceptual information required to produce native-like vowels. According to the significant differences between SBs and early bilingual vowel production in her study, exposure to a speaker's L2 within the first year of life appears to play a key role in acquiring such finely tuned categories for native-like vowel production. Her results concerning SBs and late bilinguals (LBs) provide an important point of comparison for the results of this study that will be discussed in section 4.

In another detailed study regarding phonological variation in a Quechuan language, Pasquale (2001) looked at cross-language changes in high front vowel ([i]) and high back vowel ([u]) height in bilingual and monolingual speakers of Cuzco Quechua as compared to those of monolingual Spanish speakers. His results show that Spanish interference in both Quechua-dominant and Spanish-dominant bilinguals causes the lower Quechua high vowels to shift upwards to Spanish high vowel ranges. Another pertinent finding to come out this study concerns a phonological rule in southern varieties of Quechua which causes high vowels to be lowered and backed in the vicinity of uvular stops ([q, q', q^h]). Evidence suggests that monolingual speakers of Cuzco Quechua are taking advantage of this shift to produce mid vowels in Spanish-derived borrowings at roughly the same range as those constrained in this phonological rule – a range approximately equivalent to that of monolingual-Spanish [e], but shifted further back compared to monolingual-Spanish [o]. Intra-group analysis of the same phonological rule in Pasquale (2001) also suggests that Spanish-dominant bilinguals are not applying this rule when speaking Quechua while Quichua-dominant bilinguals are raising the mid vowels. Since the lowering and backing of high vowels in the vicinity of velars does not exist in northern varieties of Quechua/Quichua, this raises the question: What takes place in Spanish-derived word borrowings in IQ and ML containing mid vowels?

Bilingual preference of Spanish-derived high vowels over Quechua-derived high vowels during Quechua speech in the aforementioned studies prompted further investigation into the motivations behind such contact-induced change. Both Pasquale (2001) and Guion (2003) provide evidence that linguistically induced change is a driving force in the arrangement of Quechua/Quichua vowel spaces. Guion (2003) suggests that bilinguals develop a single phonological system for both languages designed for optimal perceptual contrast. In turn, Pasquale (2001) suggests that a chain shift in the Quechua-dominant bilinguals' vowel space is responsible for dragging up the allophonically conditioned mid vowels in pursuit of the raised high vowels. Pasquale (2009), however, also offers social motivation as a complement to linguistically induced change. As is the case with Quichua in Ecuadorian society, Quechua in Peru is considered stigmatized while Spanish is seen as the prestigious language used in education, government, and urban society. According to Pasquale (2009), the social position of Quechua may be the driving force for setting linguistic change in motion as bilingual speakers desire, albeit subconsciously, to produce Quechua vowels more like those of Spanish.

Another example of social forces as an impetus for linguistic change can be found in Māori, a Polynesian language spoken in New Zealand. Māori, like Quechua, is stigmatized while English is the language of prestige on the island. Both languages have a history of increasingly intense contact over the last 150 years, which can be observed in linguistic changes to the Māori phonological system. Harlow et al. (2009) follow such changes over a 100-year span from the 1880s to the 1980s, and report a variety of English-influenced shifts. Such shifts towards English phonology include the reduction of allophones [ɸ], [ʍ], [h] to /f/, increased aspiration in previously non-aspirated voiceless stops, reduced distinctions in the quality and quantity of phonemically distinct long and short vowels, and apparent reductions of phonemically distinct Māori diphthongs (Harlow et al., 2009). The authors conclude that such shifts are better explained by external social factors rather than internal linguistic change since these sound changes are not innovative but rather assimilative. Socially rooted linguistic changes, like those observed in Quechua and Māori, raise the question: To what degree, if any, have social factors influenced the ML vowel system? Or can any observed changes be better explained by internal motivations?

In the following sections, I present a comparative analysis of intra-language variation regarding F1 and F2 vowel frequencies in both PML and the IQ from the nearby

Phonetica 2014;71:159–182
DOI: 10.1159/000369629

and historically related communities of Chirihuasi and Cashaloma. This section provides acoustic evidence which shows that treating both PML and IQ as either a three- or five-vowel system is an oversimplification and that, depending on one's definition of a vowel category, PML speakers may be manipulating as many as eight vowels while IQ speakers may be manipulating up to six. Here, I provide evidence for the existence of a fourth and fifth vowel, /e/ and /o/, in both PML and IQ and the possibility of three more vowels in PML with Spanish-derived subsets of /i/, /u/, and /a/ which coexist as covert contrasts alongside Quichua-derived /i/, /u/, and /a/ subsets. Similarly, I provide evidence for the possibility of one more vowel subset in IQ with Spanish-derived /a/ which coexists as a covert contrast alongside native Quichua /a/.

2 Method

2.1 Materials

A list containing 100 Spanish sentences was developed for this study. This list was designed to cover all places of articulation in both pre-vowel and post-vowel positions in PML including both voiced and voiceless phonemes and allophones in the bilabial (/p/, /b/, [β], /m/), labiodental (/f/ or [ɸ]), dental/alveolar/postalveolar (/ʧ/, /t/, /d/, /n/, /r/, /ɾ/, /s/, [z], /ʃ/, /ʒ/, /l/), palatal (/ɲ/, /j/), and velar (/k/, /g/ [ɣ], /x/ [h]) positions. Participants were asked to give their best oral interpretation of the sentence in PML. The same sentence list was also used during IQ elicitations in order to maintain the same data gathering conditions.

The participants' oral interpretations 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 Format (WAV) with a sample rate of 44.1 kHz.

2.2 Participants

Ten Quichua/ML/Spanish trilinguals, 6 women and 4 men, and 10 Quichua/Spanish bilinguals, 6 women and 4 men, participated in this study. From the trilingual group, all participants acquired Quichua and ML simultaneously from birth and began learning Spanish upon entering primary school, typically at 6–7 years of age. From the bilingual group, 4 women had a rudimentary level of Spanish, 1 man was a SB and 1 man acquired Spanish at the age of 18, while the rest acquired Spanish upon entering primary school, typically at 6–7 years of age. All ML participants were from the community of Pijal Bajo, while all Quichua participants were from the nearby communities of Chirihuasi and Cashaloma. Participants from both groups reported normal hearing and had lived their entire lives in their respective communities. Table 1 provides details of each speaker's age and gender in this study.

2.3 Procedure

A native Spanish speaker and the author gave all the instructions and read each sentence aloud in Spanish from a printout of the 100-sentence list for the ML participants. The native Spanish speaker elicited the same 100-sentence list in Spanish for the Quichua participants and a native Quichua speaker from Chirihuasi interpreted if confusion arose. The 100 sentences were the same for all the participants and elicitation conditions did not vary.[3] The participants were asked to give their best oral interpretation of each sentence and wait at least 5 s before producing the utterance. We encouraged

[3] It should be noted that since a native speaker of ML or Quichua did not elicit the sentences, there may be an increased chance of accommodation or hypercorrection. To reduce these factors, we held elicitation sessions with 3 or more participants at a time in their homes and asked them to speak in their language when consulting amongst themselves. Even if accommodation was a factor, we would expect an equivalent distribution in both ML and Quichua since the elicitation conditions did not change. Since we are also looking at within-speaker variation within individual words, it is difficult to imagine a scenario where a speaker might only accommodate only one portion of a word and not the rest.

Table 1. Language, age, and gender of the participants in this study

Trilingual group		Bilingual group	
age	gender	age	gender
59	M	66	M
50	F	62	F
39	F	45	F
44	F	29	F
43	F	21	F
42	F	42	F
60	F	28	M
40	M	52	M
58	M	55	F
50	M	49	M
Avg. 49		Avg. 45	

Trilingual group = Quichua/ML/Spanish; bilingual group = Quichua/Spanish.

participants to consult with others if any doubts arose. We also asked participants to speak at a normal conversational speed and to repeat if needed. Consultations with other participants and the 5-second waiting period made it more likely that speakers were accessing their long-term memory and reducing mimicry (Guion, 2003).

F1 and F2 frequencies from 2,515 PML and 2,191 IQ vowel tokens were analyzed for this study. These included 926 tokens from Quichua-derived lexemes/morphemes in PML and 1,589 tokens from Spanish-derived lexemes in PML. From the IQ data, 990 tokens from native Quichua lexemes/morphemes and 1,201 tokens from Spanish-derived lexemes were measured. All vowels were manually measured at their mid-point using Praat 5.2.9 (Boersma and Weenink, 2011). Spanish-derived vowels were organized based on their original Spanish pronunciation, i.e., the /u/ in *k<u>u</u>mina* 'eat', would be considered /o/ and not /u/, since its pre-lexified production was that of /o/ in Spanish *comer* /komeɾ/ 'eat'.

3 Results

The results of this study are presented in six sections (3.1–3.6). The first section (3.1) tests the hypothesis that PML Spanish-derived vowels /i/, /u/, and /a/ differ significantly from their PML Quichua-derived counterparts. For each vowel pair, a separate mixed-effects model was built to test F1 and F2 frequencies between Spanish-derived and Quichua-derived vowel pairs /i/, /u/, and /a/. The second section (3.2) tests the hypothesis that PML Spanish-derived vowels /i/ and /u/ differ significantly from PML Spanish-derived /e/ and /o/, respectively. It was anticipated that vowel formant comparisons from the same language of origin would provide evidence for or against the existence of /e/ and /o/ in PML. The third section (3.3) tests the hypothesis that PML Quichua-derived vowels /i/ and /u/ differ significantly from PML Spanish-derived /e/ and /o/, respectively.

Traditional normalization methods for this data type are often implemented due to out of range differences for within-speaker comparisons since these unequal variances between categories render traditional ANOVAs useless. Because we are only interested in within-speaker comparisons, however, unequal variances are not a problem

Phonetica 2014;71:159–182
DOI: 10.1159/000369629

for mixed-effects models since each speaker receives their own intercept. Therefore, 'it becomes statistically legitimate to include, within the same model, data from speakers who have values that span quite different ranges' (Drager and Hay, 2012, p. 75).

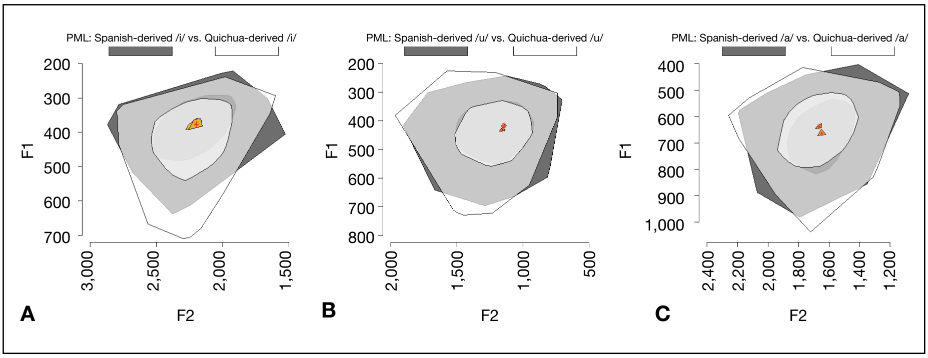
The same hypotheses and statistical analyses are then repeated using IQ data. The fourth section (3.4) tests the hypothesis that IQ Spanish-derived vowels /i/, /u/, and /a/ differ significantly from their native Quichua counterparts. The fifth section (3.5) tests the hypothesis that IQ Spanish-derived vowels /i/ and /u/ differ significantly from IQ Spanish-derived /e/ and /o/. It was anticipated that vowel formant comparisons from Spanish borrowings would provide evidence for or against the existence of /e/ and /o/ in IQ. The sixth and final section (3.6) tests the hypothesis that native Quichua vowels /i/ and /u/ differ significantly from IQ Spanish-derived /e/ and /o/, respectively.

Mixed-effects models were created in R 2.12.2 with the *lmer* function of the *lme4* package (Bates, 2012); p values and 95% confidence intervals (CI_{95}) were estimated by Monte-Carlo Markov chain (P_{MCMC}) sampling using the *pvals.fnc* (Baayen, 2008). All the models included 'speaker' and 'morpheme' as random effects. Optimal models were based on those with the lowest Bayesian Information Criterion where each fixed effect predictor was still significant. Non-significant predictors were removed from the model one by one, based on the closest t value to zero, until only significant predictors remained.

The following possible predictors were considered when building the models: gender (male/female), age (group I, 39–50 years; group II, 51–66 years), position of the syllable relative to the end of the word, features of the pre-vowel environment (including: nasal, stop, fricative, tap, approximant, labial, alveolar, postalveolar, palatal, velar, high front and mid front vowels, high back and mid back vowels, low vowel, word-initial, and word-final) and post-vowel environment (including: nasal, stop, fricative, tap, approximant, labial, alveolar, postalveolar, palatal, velar, high front and mid front vowels, high back and mid back vowels and low vowel, word-initial, and word-final), the part of speech of the word (including: noun, verb, adjective or adverb), if the vowel formed part of a root or suffix, language derivation (is the morpheme in question from Quichua or Spanish?), and if the vowel was found at a language switch, e.g., *kom<u>i</u>-nahun* 'they eat together'.

Each of the following subsections includes a density plot of the residuals from its respective F1 mixed-effects model. They include every possible variable except the contrast being discussed, i.e., the graphs are smoothed histograms summarizing how far away each vowel is from the best prediction of where it 'should' be according to a model that knows everything about the vowel except its language of origin. It is important to note that the models that the graphs are based on contain all the possible predictors, not just those that are statistically significant, therefore there is likely to be a great deal of overfitting to the data.[4]

[4] It is also important to note that some of these predictors are correlated with the contrast being investigated. For example, whether a vowel comes from a root or a suffix is fairly strongly correlated with whether its language of origin was Spanish or Quichua, so it is quite possible that a model is removing some of the variation that is really related to language of origin and incorrectly attributing that variation to the root/suffix distinction. For these reasons, each graph illustrates the worst possible case for the hypothesis that the vowel classes are different. If despite those disadvantages we can still see a difference between the bell curves, for example, Quichua-derived and Spanish-derived vowels in a graph, we can be confident that the difference is real and that it is due to the language of origin.



Color version available online

Fig. 2. Raw PML data of each vowel based on language of origin (Spanish in black and Quichua in white). **A** High front vowels. **B** High back vowels. **C** Low vowels.

F1 and F2 plots of the raw data are also provided to visualize any shifts in the data based on language of origin. It should be noted that due to the high power of the statistical tests, significant differences are not always clear in the raw data plots provided in figures 2 and 6. Depending on the plot, significant differences may be better interpreted by looking at one or more of the following areas: the 95% concentrations (outer hulls), the 50% concentrations (inter bags), or the mean averages (centre concentrations).

The following subsections include the results from the *pvals.fnc* and the model summary of each mixed-effects model. When a result is significant, we are most interested in the coefficient estimate (β), which is a conservative estimate of the average frequency distance in Hertz between the vowel pairs in question. The intercept results of each model are also included. The intercept can be defined as our 'starting point' or the estimated value in Hertz if the predictors were not present, e.g., an F1 intercept of 439 Hz for an /i/ and /e/ comparison where /i/ is significantly different by –13 Hz would mean the intercept for /e/ is 439 Hz and that for /i/ is 426 Hz.

3.1 PML Spanish-Derived versus Quichua-Derived High and Low Vowels

The statistical tests reported in this section were designed to answer the question: is there a statistically significant difference between Spanish-derived high and low vowels, and Quichua-derived high and low vowels in PML?

This section presents the results for PML's Spanish-derived vowel frequencies (/i/, /u/, /a/) when compared to PML's Quichua-derived vowel frequencies of the same shape. Figure 2 shows the raw data plotted according to the F1 and F2 frequency of each vowel. The outer hulls represent 95% concentrations in the data and inter bags represent 50% concentration. White represents PML Quichua-derived vowels and black represents PML Spanish-derived vowels.

The following results are from the F1 frequencies of PML Spanish-derived and Quichua-derived high and low vowels, e.g., the /i/ in the word [*ki̲nse*] 'fifteen', with the /i/ in the word [*abla*-hu-ni̲] 'speak-PROG-1S.PRES'.

- The F1 frequency of Spanish-derived /i/ was significantly lower than that of Quichua-derived /i/ ($t = –2.6$, $p = 0.014$, $\beta = –13$, $CI_{95\%} = –23$ to –2, intercept = 439).
- The F2 frequency of Spanish-derived /i/ was non-significant when compared to Quichua-derived /i/ ($t = –0.6$, $p = 0.99$, $\beta = –10$, $CI_{95\%} = –29$ to 29, intercept = 2,513).

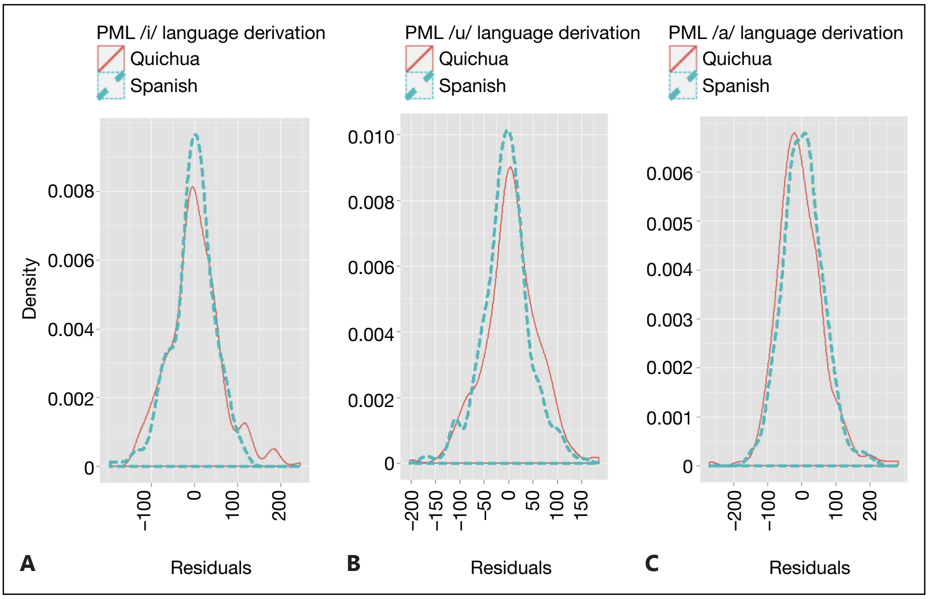


Fig. 3. Residual density plots of F1 frequencies from Spanish-derived (dashed) and Quichua-derived (solid) high and low vowels in PML: /i/ (**A**), /u/ (**B**), and /a/ (**C**).

- The F1 frequency of Spanish-derived /u/ was significantly lower than that of Quichua-derived /u/ (t = –2.5, p = 0.0004, β = –15, $CI_{95\%}$ = –29 to –9, intercept = 505).
- The F2 frequency of Spanish-derived /u/ was non-significant when compared to Quichua-derived /i/ (t = 0.22, p = 0.35, β = –5, $CI_{95\%}$ = –60 to 21, intercept = 1,045).
- The F1 frequency of Spanish-derived /a/ was significantly higher than that of Quichua-derived /a/ (t = 1.98, p = 0.04, β = 11, $CI_{95\%}$ = 0.2 to 21, intercept = 718).
- The F2 frequency of Spanish-derived /a/ was non-significant when compared to Quichua-derived /a/ (t = –0.92, p = 0.37, β = –13, $CI_{95\%}$ = –34 to 12, intercept = 1,699).

The results of these statistical tests report significant differences in tongue body height (F1) in all three Spanish-derived vowels when compared with their Quichua counterparts. The differences in F1 frequency among the Spanish-derived and Quichua-derived high vowels indicate a subtle increase in tongue body position for the Spanish-derived subset. For the low vowel subsets, the differences in F1 frequency indicate a subtle decrease in tongue body position for the Spanish-derived subset.

The results of the same statistical tests regarding the F2 frequencies reported non-significant differences in tongue body frontedness for all three-vowel pairs. Figure 3 provides the residual density plots of the F1 data presented in this section. Residual plots present two bell curves, one for each language of origin. Any deviance between the curves provides further evidence that the 'language of origin' effect is actually causing the significant result in the statistical model rather than another correlated predictor.

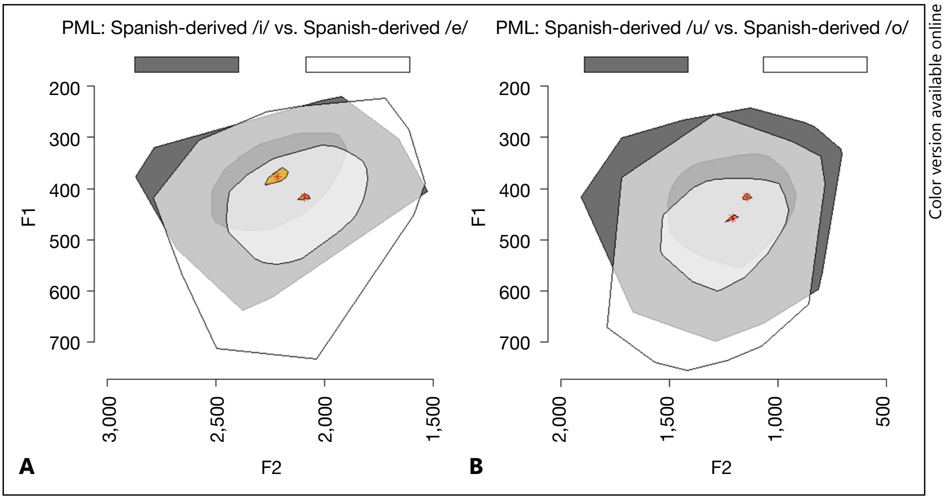


Fig. 4. Raw PML data of each vowel based height (high vowels in black and mid vowels in white). **A** PML front vowels. **B** PML back vowels.

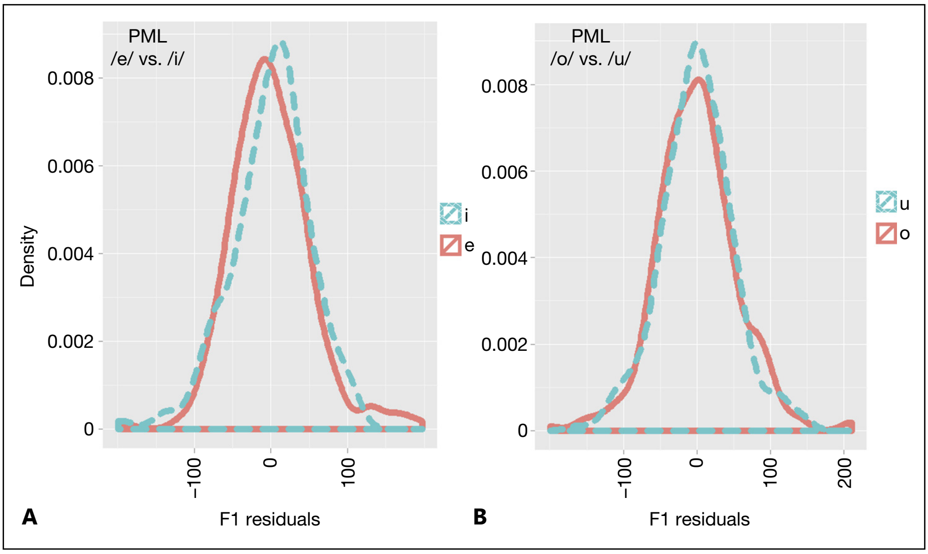
3.2 PML Spanish-Derived High and Mid Vowels

The statistical tests reported in this section were designed to answer the question: is there a statistically significant difference between Spanish-derived high vowels and Spanish-derived mid vowels in PML?

This question is of interest for a number of reasons: (1) No one has yet taken acoustic measurements from ML, and therefore we cannot know to what extent Spanish phonological contrasts have crossed over into ML, i.e., the degree to which PML has incorporated a separate set of mid vowels into its phonology. (2) While data from section 3.1 show Spanish-derived vowels and Quichua-derived vowels have not completely merged, the addition of Spanish-derived mid vowels would provide even more evidence for two coexisting systems. (3) The adoption of the Spanish mid vowels could be a practical strategy for dealing with homophony and ambiguities that might otherwise arise through Quichua vowel assimilation. Figure 4 shows the raw data plotted according to the F1 and F2 frequency of each vowel. The outer hulls represent 95% concentrations in the data and inter bags represent 50% concentration. White represents PML Spanish-derived high vowels and black represent PML Spanish-derived mid vowels.

The following results compare the F1 and F2 frequencies of PML Spanish-derived /i/ and /u/ like those found in the words [*pintur*-kuna-ka] 'painter-PL-TOP' and [*fruta*-ta-ta] 'fruit-ACC-WH.Q' with PML Spanish-derived /e/ and /o/ similar to those found in the words [*eskribi*-ʧun-mi] 'write-SUBJ.DS-VAL' and [*pueblo*-man-mi] 'to the town-DIR-VAL'.

- The F1 frequency of Spanish-derived /i/ was significantly lower than that of Spanish-derived /e/ in PML morphemes ($t = –9.8$, $p < 0.0001$, $\beta = –44$, $CI_{95\%} = –53$ to $–35$, intercept = 473).
- The F2 frequency of Spanish-derived /i/ was significantly higher than that of Spanish-derived /e/ in PML morphemes ($t = 7.3$, $p < 0.0001$, $\beta = 112$, $CI_{95\%} = 85$ to 137, intercept = 2,342).



Color version available online

Fig. 5. Residual density plots of F1 frequencies from Spanish-derived high (dashed) and mid vowels (solid) in PML: /i/ vs. /e/ (**A**) and /u/ vs. /o/ (**B**).

- The F1 frequency of Spanish-derived /u/ was significantly lower than that of Spanish-derived /o/ in PML morphemes ($t = -8.3$, $p < 0.0001$, $\beta = -38$, $CI_{95\%} = -46$ to -28, intercept = 503).
- There was a non-significant difference between the F2 frequencies for Spanish-derived /u/ and Spanish-derived /o/ in PML morphemes ($t = 0$, $p = 0.73$, $\beta = -0.0001$, $CI_{95\%} = -34$ to 23, intercept = 1,290).

The results of these statistical tests reported significant differences in tongue body height between PML Spanish-derived high vowels and mid vowels. Unlike the subtle F1 frequency differences found between Spanish-derived and Quichua-derived high and low vowels in section 3.1, the F1 frequency differences between the PML high vowels and mid vowels are quite apparent. Unlike in section 3.1, the F2 frequencies between PML /i/ and /e/ were also significantly different, though the F2 frequency difference reported between PML /u/ and /o/ was non-significant.

These effects are not being caused by a handful of clear mid-tokens that are dragging the average up and down – rather the entire distribution of /e/ and /o/ has been shifted over relative to /i/ and /u/. Figure 5 provides the residual density plots of the data presented in this section.

3.3 PML Quichua-Derived High Vowels and Spanish-Derived Mid Vowels

We have shown that Spanish-derived /i/ and /u/ are significantly higher and more fronted than Quichua-derived /i/ and /u/ in PML. We have also shown that Spanish-derived /i/ is significantly higher and more fronted than Spanish-derived /e/, while Spanish-derived /u/ is significantly higher than Spanish-derived /o/. It remains unclear, however, whether PML speakers have merged Quichua-derived /i/ and /u/ with Spanish-derived /e/ and /o/, respectively, the way that Guion (2003) found many early Quichua/Spanish bilinguals did, or whether they also maintain the distinction between these two vowels, the way Guion found many simultaneous Quichua/Spanish bilinguals did.

The following results compare the F1 and F2 frequencies of PML Quichua-derived /i/ and /u/ like those found in the words [*komi*-ngiʧi] 'eat-2PL' and [*kasa*-kuna] 'house-PL', with PML Spanish-derived /e/ and /o/ similar to those found in the words [*kafe*-ta] 'cafe-ACC' and [*kaz̲o*-mi] 'car-VAL'.

- The F1 frequency of Quichua-derived /i/ was significantly lower than that of Spanish-derived /e/ in PML morphemes ($t = –6.9$, $p < 0.0001$, $\beta = –39$, $CI_{95\%} = –50$ to –28, intercept = 468).
- The F2 frequency of Quichua-derived /i/ was significantly higher than that of Spanish-derived /e/ in PML morphemes ($t = 7.9$, $p < 0.0001$, $\beta = 139$, $CI_{95\%} = 104$ to 162, intercept = 2,324).
- The F1 frequency of Quichua-derived /u/ was significantly lower than that of Spanish-derived /o/ in PML morphemes ($t = –4.6$, $p < 0.0001$, $\beta = –23$, $CI_{95\%} =$ –34 to –13, intercept = 517).
- There was a non-significant difference between the F2 frequencies for Quichua-derived /u/ and Spanish-derived /o/ in IQ morphemes ($t = –1.5$, $p = 0.21$, $\beta =$ –30, $CI_{95\%} = –57$ to 13, intercept = 1,105). Recall there was also a non-significant difference in F2 between Spanish-derived /i/ and Spanish-derived /e/.

The results from the statistical tests reported significant differences in tongue body height (F1) between Quichua-derived high vowels and Spanish-derived mid vowels in PML. As would be expected, the F1 frequency differences between Quichua-derived high vowels and Spanish-derived mid vowels are not as large as those found between Spanish-derived high vowels and Spanish-derived mid vowels in section 3.2. The combined results from sections 3.1–3.3 suggest that PML may be manipulating as many as eight vowels.

3.4 IQ Spanish-Derived versus Native Quichua High and Low Vowels

The statistical tests reported in this section were designed to answer the question: is there a significant difference between Spanish-derived high and low vowels and native Quichua high and low vowels in IQ? It is also worth noting that Spanish-derived words in IQ are similar to those in PML in that they typically underwent the same processes of lexification, i.e., they are not taken from L1 Quichua speakers speaking Spanish or part of code-switching phrases. Figure 6 shows the raw data plotted according to the F1 and F2 frequency of each vowel. The outer hulls represent 95% concentrations in the data and the inter bags represent 50% concentrations. White represents native Quichua-derived vowels and black represents IQ Spanish-derived vowels.

- There was a non-significant difference between the F1 frequencies of /i/ in Spanish-derived and native IQ /i/ ($t = –0.4$, $p = 0.62$, $\beta = –2$, $CI_{95\%} = –9$ to 6, intercept = 433).
- There was a non-significant difference between the F2 frequencies of /i/ in Spanish-derived and native IQ /i/ ($t = –0.02$, $p = 0.62$, $\beta = –0.4$, $CI_{95\%} = –28$ to 30, intercept = 2,676).
- There was a non-significant difference between the F1 frequencies of /u/ in Spanish-derived and native Quichua /u/ in IQ ($t = 0.8$, $p = 0.28$, $\beta = 4$, $CI_{95\%} = –4$ to 14, intercept = 458).
- The F2 frequency of Spanish-derived /u/ was significantly lower than that of native IQ /u/ ($t = –1.6$, $p = 0.038$, $\beta = –40$, $CI_{95\%} = –91$ to –3, intercept = 1,211).

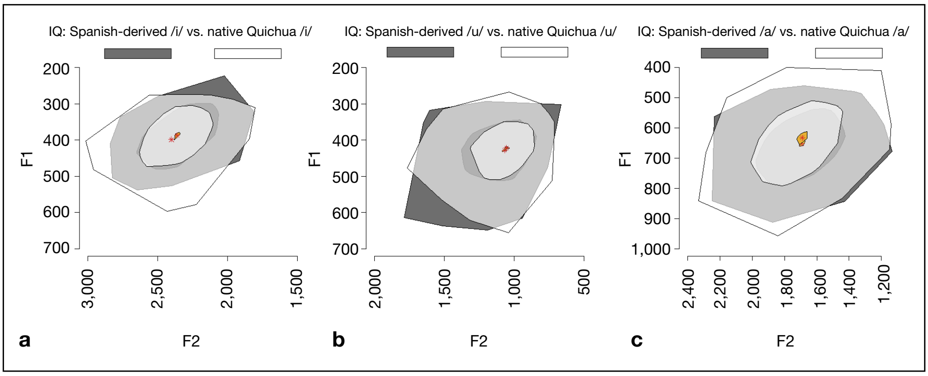


Fig. 6. Raw IQ data of each vowel based on language of origin (Spanish-derived in black and native Quichua in white). **A** High front vowels. **B** High back vowels. **C** Low vowels.

- The F1 frequency for /a/ in Spanish-derived morphemes was significantly higher than that of native Quichua /a/ ($t = 1.7$, $p = 0.045$, $\beta = 11$, $CI_{95\%} = 0.1$ to 23, intercept = 684). I am not fully convinced of this result for two reasons: (1), the t value is suspiciously small (within ±2 is usually non-significant with large datasets) and (2) the P_{MCMC} value is just below 0.05; p value results tend to differ slightly across runs using P_{MCMC} sampling. In order to avoid cherry picking the data of each model was also restricted to only one run of *pvals.fnc*.[5] No corrections were made for multiple comparisons by using methods such as Bonferroni's correction, Scheffé's test or Tukey's Honestly Significant Difference. Therefore, we consider this result not to be strong evidence for a difference between Spanish-derived and native Quichua /a/s in IQ. If this effect is real, it is the biggest F1 difference one will find in IQ.
- There was a non-significant difference between the F2 frequencies of /a/ in Spanish-derived and native IQ /a/ ($t = -0.95$, $p = 0.13$, $\beta = -15$, $CI_{95\%} = -47$ to 7, intercept = 1,824).

The results of these statistical tests reported non-significant differences in tongue body height (F1) with the exception of /a/. Non-significant results in tongue body frontedness (F2) were also found between Spanish-derived vowels and their native Quichua counterparts with the exception /u/ where the Spanish-derived vowel appears to be 40.8 Hz lower than its Quichua-derived counterpart ($t = -1.6$, $p = 0.03$, $\beta = -41$, $CI_{95\%} = -91$ to -3). These non-significant findings regarding the F1 frequencies contrast with the small but significant differences for the same tests in PML. Figure 7 provides the residual density plots of the F1 data presented in this section.

[5] Since exact p-value calculations are not possible in mixed-effects models, pvals.fnc provides an alternative to this calculation using P_{MCMC} sampling. This function runs 10,000 simulations in the model under analysis and varies the coefficients slightly with each run. Because the simulations are never the same for each run, the output of each run also varies ever so slightly. Therefore, if a result is near significance, e.g., 0.051, multiple runs may provide a slightly lower result, e.g., 0.049. In order to avoid cherry picking the results with multiple runs of pvals.fnc I restricted each model to one run.

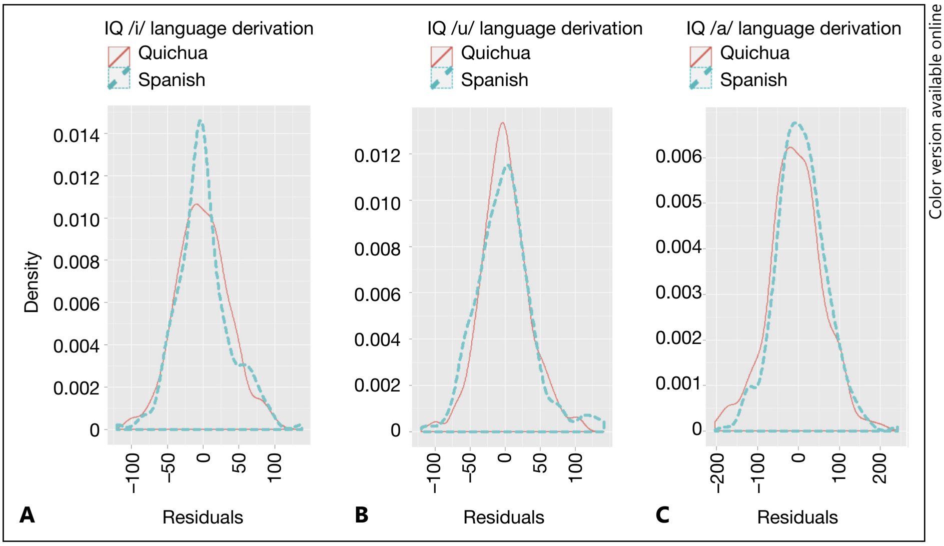


Fig. 7. Residual density plots of F1 frequencies from Spanish-derived (dashed) and native Quichua (solid) high and low vowels in IQ: /i/ (**A**), /u/ (**B**), and /a/ (**C**).

3.5 IQ Spanish-Derived High and Mid Vowels

The statistical tests reported in this section were designed to answer the question: is there a statistically significant difference between Spanish-derived high vowels and Spanish-derived mid vowels in IQ? This question is similar to the one found in section 3.2 and important for Quichua for essentially the same reasons.

The following results compare the F1 and F2 frequencies of IQ Spanish-derived /i/s and /u/s like those found in the words [*amigu*-mi] 'friend-VAL' and [*luna*-ka] 'moon-TOP' with IQ Spanish-derived /e/s and /o/s similar to those found in the word [*kuadernu*-ta] 'notebook-ACC'. Figure 8 shows the raw data plotted according to the F1 and F2 frequency of each vowel. The outer hulls represent 95% concentrations in the data and inter bags represent 50% concentration. White represents IQ Spanish-derived high vowels and black represents IQ Spanish-derived mid vowels.

- The F1 frequency in Spanish-derived /i/ was significantly lower than that of Spanish-derived /e/ in IQ morphemes ($t = -4.9$, $p < 0.0001$, $\beta = -27$, $CI_{95\%} = -38$ to -18, intercept = 481).
- The F2 frequency in Spanish-derived /i/ was significantly higher than that of Spanish /e/ in IQ morphemes ($t = 6.5$, $p < 0.0001$, $\beta = 126$, $CI_{95\%} = 93$ to 163, intercept = 2,543).
- The F1 frequency in Spanish-derived /u/ was significantly lower than that of Spanish-derived /o/ in IQ morphemes ($t = -4.3$, $p < 0.0001$, $\beta = -25$, $CI_{95\%} = -36$ to -14, intercept = 481).
- The F2 frequency for Spanish-derived /u/ was significantly lower than that of Spanish-derived /o/ in IQ morphemes ($t = -2.7$, $p = 0.0056$, $\beta = -61$, $CI_{95\%} = -97$ to -17, intercept = 1,200).

The results of these statistical tests reported significant differences in tongue body height (F1) between Spanish-derived high vowels and mid vowels in IQ. The F2

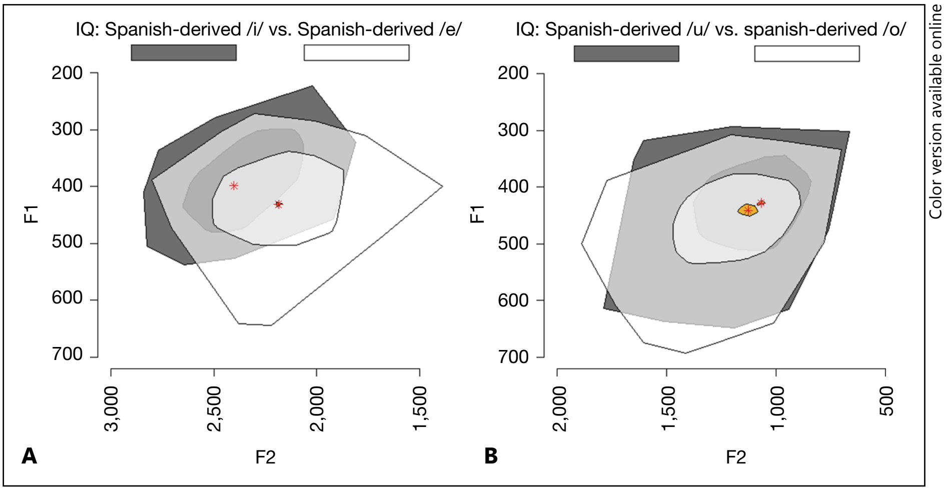


Fig. 8. Raw IQ data of each vowel based height (high vowels in black and mid vowels in white). **A** IQ front vowels. **B** IQ back vowels.

frequencies between Spanish-derived high vowels and mid vowels were significantly different as well.

These effects are not being caused by a handful of clear mid tokens that are dragging the average up and down – rather the entire distribution for /e/ has shifted over relative to /i/. In contrast, a small handful of Spanish-derived /o/ tokens appear to show up as clear /o/ with no appreciable shift in the rest of the distribution. This case of hypercorrection by the Quichua speakers could be causing a significant difference where there may otherwise be a non-significant result. The F1 frequency differences in IQ indicate a noticeable raise in tongue body height but only about half the size of those found in PML, i.e., the Spanish mid vowels are higher (Hz) in PML than in IQ. Figure 9 provides the residual density plots of the data presented in this section.

3.6 IQ Native Quichua High Vowels and Spanish-Derived Mid Vowels

The statistical tests reported in this section were designed to answer essentially the same question proposed in section 3.3 but regarding IQ: is there a statistically significant difference between native Quichua high vowels and Spanish-derived mid vowels in IQ?

The following results compare the F1 and F2 frequencies of native Quichua /i/s and /u/s like those found in the words [ʃi̲mi-ta] 'language-ACC' and [ɾuɾa-nʧi] 'do-1P' with Spanish-derived /e/s and /o/s similar to those found in the word [*kuade̲rnu̲*-ta] 'notebook-ACC'.

- The F1 frequency in native Quichua /i/ was significantly lower than that of Spanish-derived /e/ in IQ morphemes ($t = –6.3$, $p < 0.0001$, $\beta = –29$ $CI_{95\%} = –36$ to $–20$, intercept = 452).
- The F2 frequency in native Quichua /i/ was significantly higher than that of Spanish-derived /e/ in IQ morphemes ($t = 6.5$, $p < 0.0001$, $\beta = 132$, $CI_{95\%} = 85$ to 155, intercept = 2,581).

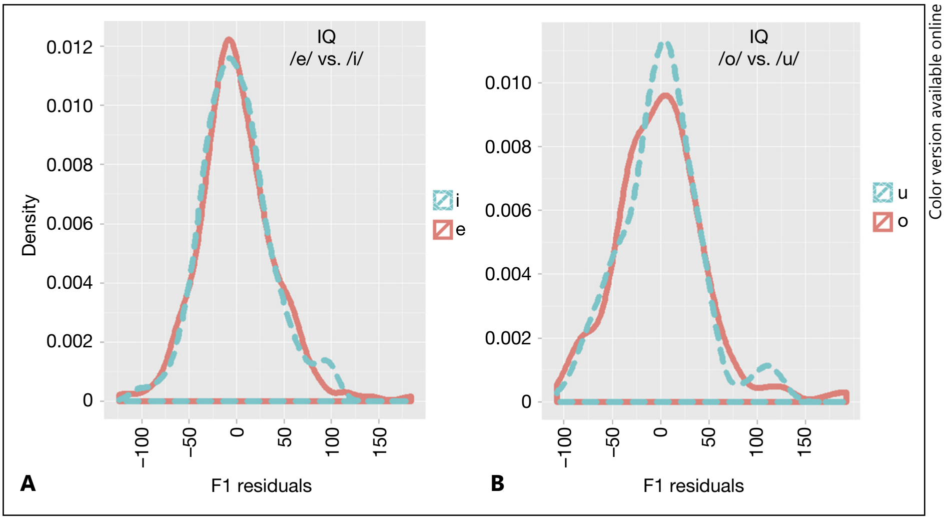


Fig. 9. Residual density plots of F1 frequencies from Spanish-derived high (dashed) and mid (solid) vowels in IQ: /i/ vs. /e/ (**A**) and /u/ vs. /o/ (**B**).

- The F1 frequency in native Quichua /u/ was significantly lower than that of Spanish-derived /o/ in IQ morphemes ($t = -5.2$, $p < 0.0001$, $\beta = -24$, $CI_{95\%} = -32$ to -16, intercept = 489).
- The F2 frequency in native Quichua /u/ was significantly lower than that of Spanish-derived /o/ in PML morphemes ($t = -3.5$, $p = 0.0008$, $\beta = -75$, $CI_{95\%} = -103$ to -30, intercept = 1,178).

The results of these statistical tests reported significant differences in tongue body height (F1) between native Quichua high vowels and Spanish-derived mid vowels in IQ. Similar to the Spanish-derived high vowel and mid vowel tests in section 3.5, all F2 frequencies were significantly different between the native Quichua high vowels and Spanish-derived mid vowels. These results suggest that IQ may manipulate up to six vowels.

4 Discussion and Conclusions

This study had the goal of presenting a comparative analysis of F1 and F2 frequencies from both PML and IQ. Statistical evidence accounts for as many as eight vowels in PML and up to six vowels in IQ. The results show the possibility of a fourth and fifth vowel, /e/ and /o/ in both PML and IQ in what are both traditionally considered three-vowel systems (Guion, 2003; Muysken, 1997). In addition, results from the PML data suggest the possibility of Spanish-derived /i/, /u/, and /a/ subsets which coexist as near-mergers alongside Quichua-derived /i/, /u/, and /a/. Similarly, evidence also exists for an additional vowel subset in IQ Spanish-derived /a/, which may coexist as a near-merger alongside Quichua-derived /a/. Figure 10 provides a side-by-side comparison of the vowel inventories described in this study.

The results of this analysis suggest that PML makes use of two overlapping vowel systems based on the vowels' language of origin. Spanish-derived high vowels (/i/ and

Phonetica 2014;71:159–182
DOI: 10.1159/000369629

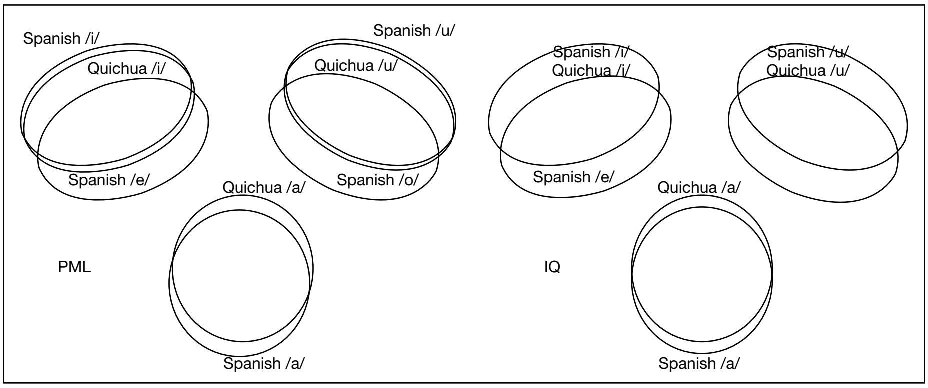


Fig. 10. Side-by-side comparison of PML and IQ vowel inventories.

/u/) have lower F1 frequencies while the Spanish-derived low vowel (/a/) has a higher F1 frequency when compared to their Quichua-derived counterparts. While this is predicted in adaptive dispersion models, since the vowels are being dispersed in the correct direction, they are by no means creating separate categories, i.e., they seem to coexist stably while overlapping each other in an almost identical vowel space. The PML data also contradicts Flege's (2007) Speech Learning Model since this model predicts that two competing systems with stable overlap should be undesirable. The PML data hypothetically fits with Best et al. (2003) Perceptual Assimilation Model, which predicts that bilinguals assimilate L2 sounds based on how similar or contrastive a given sound is perceived to be to the listener's native phonology. Within this system categories are allowed to (1) merge into a single category, (2) stay independent, or (3) coexist with varying degrees of overlap. One possible issue facing this model is the fact that we are not dealing with L2 sounds, instead these coexisting systems appear to have been passed down from generation to generation under conditions of near-merger. Since all ML speakers are trilingual, it may be that knowledge of the source languages helps support these systems since ML speakers are often aware which parts of ML are derived from Spanish and which parts are Quichua.

Regarding the high and low vowel pairs in PML, the significant differences are not large (13 Hz lower for Spanish /i/, 15 Hz lower for Spanish /u/, and 11 Hz higher for Spanish /a/). These frequency differences are, however, perceivable in clinical settings (Kewley-Port and Atal, 1989; Mannell, 1994). Based on this analysis, it appears that these effects are not being caused by a handful of clear tokens dragging up and down the average, rather the entire distribution of Spanish-derived vowels has nearly overlapped the distributions of the Quichua-derived vowels.

On the other hand, regarding IQ high and low vowels, there was a non-significant difference in acoustic vowel space based on the language of origin (with the questionable exception of the F1 frequency in Spanish-derived /a/, as discussed in section 3.4). If Quichua merges Spanish borrowings according to Quichua phonology, why was this process only partial in PML? The answer may lie in the distinctive evolutionary paths of IQ and PML. In IQ, the main influence of Spanish phonology on each lexeme would, hypothetically, have been at its point of borrowing, from a small number of bilinguals before immediately conforming to Quichua phonology when monolinguals adopted

Table 2. Guion's (2003, p. 116) mean /i/ data reproduced in Hz

Group	/i/ mean, Hz	Mean difference compared to monolinguals, Hz	Mean difference compared to SBs, Hz
3-vowel (SBs)	232.6	–15.3 n.s.	–
2-vowel (early bilinguals)	231.8	–16.1 n.s.	–0.8[a]
1-vowel (LBs)	264.8	16.9	32.2[a]
Monolingual Spanish	247.9	–	15.3[a]

n.s. = Non-significant; [a] not tested.

the lexemes. The idea of conforming to Quichua phonology also implies that Spanish-derived vowels underwent complete merger and consecutive generations would have no point of reference to separate the Spanish-derived and Quichua-derived vowels into distinct categories. For PML, however, the influence of Spanish phonology probably came from a large number of bilinguals and lasted for generations.

The complete phonological merger of IQ Spanish-derived high vowels with their native Quichua counterparts is similar to what Guion (2003) found for LBs who speak Spanish without producing significantly different Spanish high vowels from those of Quichua. Within her data, however, there is an untested similarity to the PML data, specifically regarding how SBs maintain separate vowel systems for Quichua and Spanish production. Her data contains the mean results for SBs (comparable to Spanish-like vowel production) and the results for LBs (comparable to Quichua-like vowel production). These results are similar to how PML speakers make use of separate systems for Spanish-derived vowels and Quichua-derived vowels. After converting her data from Bark to Hertz[6], the normalized F1 frequencies of SBs compared to LBs show a mean difference of –32 Hz between the Spanish-like /i/ of SBs and Quichua-like /i/ of LBs (table 2).

Guion (2003) did not test these data for significance, but there appears to be a considerable amount of difference between SBs and LBs. The distance in Hertz between SBs' and LBs' production of Spanish /i/ is roughly double the frequency (32 Hz) as that found between Spanish-derived /i/ and Quichua-derived /i/ in PML (13 Hz). This suggests that PML speakers are maintaining distinct high front vowel categories at roughly half the range of SBs.

The same tendency was also found regarding high back vowels. The normalized F1 frequencies from SBs compared to LBs in Guion's (2003) study revealed a mean difference of –31 Hz between the Spanish-like /u/ of SBs and Quichua-like /u/ of LBs (table 3). Again this data was not tested for significance but a considerable amount of difference between SBs and LBs is apparent. The distance in Hertz between SBs is also nearly double the frequency (31 Hz) as that found between Spanish-derived and Quichua-derived high back vowels in PML (–15 Hz). This suggests that PML speakers are maintaining distinct high back vowel categories at half the range of SBs and LBs.

[6] Guion (2003, p. 107) data was also normalized to male based on F3 values to avoid between-talker variation. My conversions are based on the normalized data. It is also worth noting that my F1 data (>400) typically come from the intercept for women.

Table 3. Guion's (2003, p. 117) mean /u/ data reproduced in Hz

Group	/u/ mean, Hz	Mean difference compared to monolinguals, Hz	Mean difference compared to SBs, Hz
Separate Spanish /u/	302.8	20.5[a]	–
Separate Spanish /o/	280.1	–2.2[a]	–22.7[b]
1-vowel (LBs)	334.2	51.9	31.4[b]
Monolingual Spanish	282.3	–	–20.5[b]

[a] No significant difference compared to the monolingual group. [b] Not tested.

Table 4. Guion's (2003, p. 118) mean /a/ data reproduced in Hz

Group	/a/ mean, Hz	Mean difference compared to monolinguals, Hz	Mean difference compared to SBs, Hz
3-vowel	482.0	–34.4	0
Raised Quichua vowel	469.6	–46.8	–12.4[a]
No Spanish vowel (LBs)	497.4	–19	15.4[a]
Monolingual	516.4	–	34.4[a]

[a] Not tested.

This tendency also is apparent for low vowels. The normalized F1 frequencies from SBs compared to LBs in Guion's data reveal a mean difference of 15 Hz between the Spanish-like /a/ of SBs and Quichua-like /a/ of LBs (table 4). Once more this data was not tested for significance, but it is comparable to the low vowel dispersion seen in PML Spanish-derived /a/ and Quichua-derived /a/. The distance in Hertz for SBs is approximately one third as large as that found between Spanish-derived /a/ and Quichua-derived /a/ in PML (11 Hz). This suggests that speakers of PML are maintaining distinct low vowel categories at roughly one third the distance of SBs and LBs. This data is also comparable to the significant difference between Spanish-derived low vowels in IQ and native Quichua low vowels. Spanish-derived low vowels in IQ are produced, on average, 11 Hz higher than their native IQ counterparts.

Unlike the PML Spanish-derived and Quichua-derived high and low vowels, the significant differences between PML Spanish-derived high and mid vowels are more apparent: the F1 frequency for /i/ was found to be, on average, 44 Hz lower than that of /e/ while the F2 frequency for /i/ was found to be, on average, 112 Hz higher than that of /e/. Regarding /u/ and /o/, the F1 frequency for /u/ was found to be, on average, 38 Hz lower than that of /o/ while there was no significant difference found between F2 values of /u/ and /o/.

Regarding IQ, the significant differences in F1 frequency between Spanish-derived high and mid vowels are roughly half the size when compared with PML: the F1 frequency for /i/ was found to be, on average, 27 Hz lower than that of /e/ while the F2 frequency for /i/ was found to be, on average, 126 Hz higher than that of /e/. Regarding /u/ and /o/, the F1 frequency for /u/ was found to be, on average, 25 Hz

Table 5. Guion's (2003, pp. 116–117) high and mid vowel data reproduced in Hz

Group	/i/ mean, Hz	/e/ mean, Hz	Mean difference, Hz
3-vowel (SBs)	232.6	337.2	104.6
2-vowel (LBs)	231.8	355.5	123.7
1-vowel (LBs)	264.8	287.4	22.6
Monolingual Spanish	247.9	355.5	107.6
Group	**/u/ mean, Hz**	**/o/ mean, Hz**	**Mean difference, Hz**
Separate Spanish /u/	302.8	372.3	69.5
Separate Spanish /o/	280.1	386.2	106.1
1-vowel	334.2	360.6	26.4
Monolingual Spanish	282.3	394.9	112.6

lower than that of /o/ and the F2 frequency for /u/ was found to be, on average, 62 Hz lower than that of /o/.

Guion's (2003) findings for high versus mid vowel production in LBs also have certain untested similarities to the IQ data in this study. After once more converting her data from Bark to Hertz (table 5), it appears that LBs (again, who are comparable to Quichua monolinguals) have a mean F1 difference of 23 Hz between the mean frequencies of /i/ and /e/ and a mean difference of 26 Hz between /u/ and /o/ (section 3.5). These results are nearly identical to the significant differences between IQ Spanish-derived high and mid vowels. The IQ results show an average difference of 27 Hz between Spanish-derived /i/ and /e/ while an average difference of 25 Hz was found between Spanish-derived /u/ and /o/. This suggests IQ speakers are maintaining distinct high vowel and mid vowel categories at roughly the same mean distance as the LBs when producing Spanish high and mid vowels.

PML speakers show similar results between high and mid vowel production but at roughly twice the distance of the IQ speakers. This suggests that PML speakers are performing the impressive task of maintaining distinct high and mid vowel categories at greater acoustic differences than monolinguals, but also at roughly half the distance as SBs. As with the high vowel and low vowel results, this data shows that the current generation of PML speakers has managed to reconstruct a highly overlapping system of categories using only L1 input. This is evident in the fact that the current generation of PML speakers are considered early bilinguals[7]. Their frequency differences, however, are not overshoots like those found in the early bilinguals' group in Guion's (2003) data, instead they are comparable to a lesser degree, to those of SBs *without* being SBs. Figure 11 provides a side-by-side comparison of the front vowels based on my conversions of Guion's (2003) SB data, and the PML and IQ data from this study.

While external social factors are clearly the impetus for the formation of the language itself, it remains unclear if external dynamics played a substantial role in shaping the ML vowel system. One could argue that internal factors such as the insertion of Spanish vocabulary required additional perceptual cues to disambiguate elements such as Spanish-derived minimal pairs, e.g., *piso* 'floor' and *peso* 'weight'. It might be

[7] The typical age of Spanish acquisition for PML speakers was 6–7 years upon entering school.

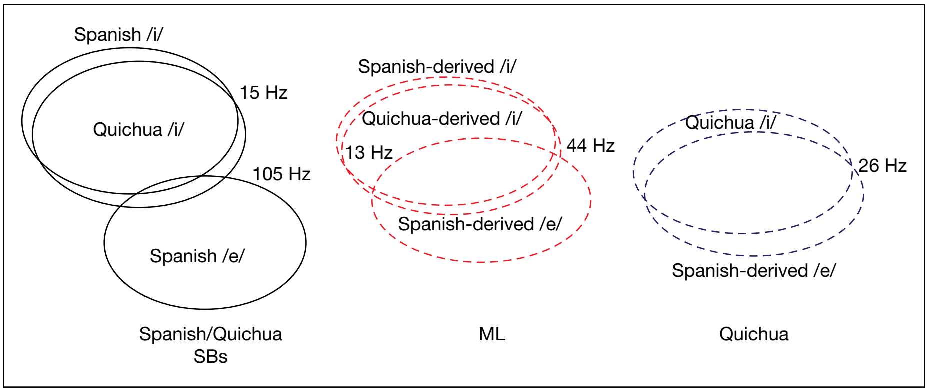


Fig. 11. Scaled representation of high and mid front vowel categories in Spanish/Quichua SBs based on Guion (2003) (left), Spanish- and Quichua-derived categories in ML (mid) and native Quichua and Spanish-derived categories in Quichua (right) from this study.

that the degree of variation in the overlapping systems was enough to distinguish such examples without having to reorganize the entire distribution of vowels – an economic factor for a population of L2 Spanish speakers. The high and low vowel contrasts might suggest that many of the originators and future generations of speakers were in fact early bilinguals since such subtle distinctions do not appear in the speech of mid to late speakers as shown by Guion (2003). The distinctions, however, suggest that the originators or future generations were probably not SBs since the difference in range between Spanish-derived and Quichua-derived high and low vowels is only roughly half that of Guion's (2003) SBs. Because of these factors (the insertion of a vocabulary from a five-vowel system, the degree of overlapping spaces, and the age of acquisition) it appears that internal linguistic dynamics may have a larger role in shaping the ML vowel space (and possibly the IQ vowel space as well) than external social factors.

How does this study compare to the Jones et al. (2011) study of Gurindji Kriol vowels? Similarly, both mixed languages (Gurindji Kriol and ML) show considerable overlap of both source languages' vowel systems. ML, however, appears to have emerging vowel categories (/e/ and /o/) while Gurindji Kriol vowel categories appear to be merging as seen in the extensive overlap of /æ/ and /e/ and /ʉː/ and /oː/ compared to those in Katherine English. Regarding Michif, while it is uncertain the extent of overlap between the Plains Cree high-mid back vowel /o/ and Metis French /u/, Rosen (2007) demonstrates that the high-mid back vowel from French has been added to Michif in French-derived morphemes. It may be that ML and Michif, with relatively small phonological source language vowel inventories (three and five plus a lengthening contrast, respectively), have opted to expand the vowel system as a way to disambiguate homophonous words, e.g., PML *lona* 'tarp' and *luna* 'moon', while Gurindji Kriol speakers felt that word contrasts could still be achieved with a smaller set of vowels.

Future investigations should look at vowel perception in both ML and Quichua in order to determine if the vowel pairs described here are in fact convert contrasts or if speakers are able to parse the overlapped vowel systems in order to disambiguate minimal pairs.

Acknowledgements

This research was funded by the Manitoba Graduate Scholarship (MGS), the University of Manitoba Graduate Fellowship (UMGF), and fieldwork expenses were paid for by the J.G. Fletcher Award for Graduate Research. I would like to thank Kevin Russell, Robert Hagiwara, Enrique Fernandez, and the anonymous reviewers for their comments and suggestions on the drafts of this article. Finally, I would like to thank Gabriela Prado, Antonio Maldonado, Lucía Gonza, Elvis Tuquerres, and the participants for taking part in this investigation.

References

Baayen H (2008): LanguageR: Analyzing Linguistic Data: A Practical Introduction to Statistics Using R. Cambridge, Cambridge University Press.

Backus A (2003): Can a mixed language be conventionalized alternational codeswitching? In Matras Y, Bakker P (eds): The Mixed Language Debate. Theoretical and Empirical Advances. Trends in Linguistics 145. Berlin, Mouton de Gruyter, pp 237–270.

Bates DM (2012): lme4: mixed-effects modeling with R. http://lme4.r-forge.r-project.org/book/.

Best C, Hallé P, Bohn OS, Faber A (2003): Crosslanguage perception of nonnative vowels: phonological and phonetic effects of listeners' native languages. Proc 15th Int Congress of Phonetic Sciences, Barcelona.

Boersma P, Weenink D (2011): Praat: Doing Phonetics by Computer; Version 5.2.9 (computer program). Amsterdam, Institute of Phonetic Sciences, University of Amsterdam.

Buttner T (1993): Uso del Quichua y el castellano en la sierra ecuatoriana. Quito, Abya Yala.

Dikker S (2008): Spanish prepositions in Media Lengua: redefining relexification; in Stolz T, Bakker D, Palomo RS (eds): Hispanisation: The Impact of Spanish on the Lexicon and Grammar of the Indigenous Languages of Austronesia and the Americas. Berlin/New York, Mouton De Gruyter, pp 121–148.

Drager K, Hay J (2012): Exploiting random intercepts: two case studies in sociophonetics. Lang Variation Change 24:59–78.

Flege J (2007): Language contact in bilingualism: phonetic system interactions; in Cole J, Hualde J (eds): Laboratory Phonology 9. Berlin, Mouton de Gruyter, pp 353–380.

Gómez-Rendón J (2005): La Media Lengua de Imbabura; in Olbertz H, Muysken P (eds): Encuentros y conflictos: Bilingüismo y contacto de lenguas en el mundo andino. Madrid, Iberoamericana, pp 39–58.

Gómez-Rendón J (2007): Grammatical borrowings in Imbabura Quichua; in Matras Y (ed): Grammatical Borrowing in Cross-Linguistic Perspective. New York, de Gruyter, vol 1, pp 481–521.

Gómez-Rendón J (2008): Mestizaje linguístico en Los Andes: Génesis y estructura de una lengua mixta. Quito, Abya-Yala.

Guion S (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.

Harlow R, Keegan P, King J, Maclagan M, Watson C (2009): The changing sound of the Māori language; in Stanford J, Preston D (eds): Variation in Indigenous Minority Languages. Amsterdam, Benjamins, vol 25, pp 129–152.

Hickey R (2004): Mergers, near-mergers and phonological interpretation. New Perspectives on English Historical Linguistics: selected papers from 12 ICEHL, Glasgow.

Johnson K (2000): Adaptive dispersion in vowel perception. Phonetica 57:181–188.

Jones C, Meakins F, Buchan H (2011): Comparing vowels in Gurindji Kriol and Katherine English: citation speech data. Aust J Ling 31:305–326.

Kewley-Port D, Atal B (1989): Perceptual differences between vowels located in a limited phonetic space. Acoust Soc Am 85:1726–1740.

Labov W, Karen M, Miller C (1991): Near-mergers and the suspension of phonemic contrast. Lang Variation Change 3:33–74.

Labov W, Yaeger M, Steiner R (1972): A quantitative study of sound change in progress. Philadelphia, US Regional Survey.

Lefebvre C (2005): Relexification: a process available to human cognition. Berkeley Linguistic Society, Berkeley.

Lefebvre C (2006): Creole Genesis and the Acquisition of Grammar: The Case of Haitian Creole. Cambridge, Cambridge University Press, vol 88.

Lefebvre C, Therrien I (2007): On Papiamentu ku; in Siegel J, Lynch J, Eades D (eds): Language Description, History and Development: Linguistic Indulgence in Memory of Terry Crowley. Amsterdam, Benjamins, vol 30, pp 169–182.

Liljencrants J, Lindblom B (1972): Numerical simulation of vowel quality systems: the role of perceptual contrast. Language 48:839–862.

Lindblom B (1986): Phonetic universals in vowel systems; in Ohala J, Jaeger J (eds): Experimental Phonology. Orlando, Academic Press, pp 13–44.

Lindblom B (1990): Explaining phonetic variation: a sketch of the H and H theory; in Hardcastle W, Marchal A (eds): Speech Production and Speech Modeling. Dordrecht, Springer Science & Business Media, pp 403–439.

Livijn P (2000): Acoustic distribution of vowels in differently sized inventories – hot spots or adaptive dispersion? PERILUS 23.
Mannell RH (1994): The Perceptual and Auditory Implications of Parametric Scaling in Synthetic Speech; unpublished PhD diss Macquarie University, North Ryde.
Maye J, Werker J, Gerken A (2002): Infant sensitivity to distributional information can affect phonetic discrimination. Cognition 82:143–178.
McConvell P, Meakins F (2005): Gurindji Kriol: a mixed language emerges from code-switching. Aust J Ling 25:9–30.
Muysken P (1980): Sources for the study of Amerindian contact vernaculars in Ecuador. Amsterdam Creole Stud 3:66–82.
Muysken P (1981): Halfway between Quechua and Spanish: the case for relexification; in Highfield AR (ed): Historicity and Variation in Creole Studies. Ann Arbor, Karoma, vol 57–78.
Muysken P (1997): Media Lengua; in Thomason SG (ed): Contact Languages: A Wider Perspective. Amsterdam, Benjamins, pp 365–426.
Pasquale M (2001): Quechua and Spanish Language Contact: Influence on the Quechua Phonological System; PhD diss University of Michigan, Ann Arbor.
Pasquale M (2009): Phonological variation in a Peruvian Quechua; in Stanford J, Preston D (eds): Variation in Indigenous Minority Languages. Amsterdam, Benjamins, vol 25, pp 245–258.
Rosen N (2007): Domains in Michif Phonology; PhD diss University of Toronto. Toronto Working Papers in Linguistics.
Shappeck M (2011): Quichua-Spanish Language Contact in Salcedo, Ecuador: Revisiting Media Lengua Syncretic Language Practices; PhD diss University of Illinois at Urbana-Champaign.
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, MSpace.
Stewart J (2013): Stories and Traditions from Pijal: Told in Media Lengua. Charleston, CreateSpace.
Swadesh M (1952): Lexicostatistic dating of prehistoric ethnic contacts. American Philosophical Society.
R Core Development Team (2008): R; version 2.12.2. Vienna, R Foundation for Statistical Computing.