Journal of Phonetics 71 (2018) 177–193

Contents lists available at ScienceDirect

Journal of Phonetics

journal homepage: www.elsevier.com/locate/Phonetics

Research Article

Vowel perception by native Media Lengua, Quichua, and Spanish speakers

Jesse Stewart

Department of Linguistics, University of Saskatchewan, 916 Arts Building, Saskatoon, SK S7N 5A5, Canada

ARTICLE INFO

Article history:
Received 20 August 2016
Received in revised form 19 August 2018
Accepted 22 August 2018

Keywords:
Media Lengua
Quichua
Spanish
Mixed language
Vowel perception
Ecuador

ABSTRACT

This study explores mid and high vowel perception in and across Ecuadorian Spanish, Quichua, and Media Lengua (a mixed language containing Quichua systemic elements and Spanish lexicon). Quichua and Media Lengua were originally considered three vowel systems comprised of /i, u, a/. However, recent production results reveal that mid vowels /e, o/ may have entered these languages through Spanish lexical borrowings. The aim of the present study is to test listener perception with minimal pairs containing different mid and high vowels to determine how listeners identify them. A two-alternative forced choice (2AFC) identification task experiment with paired stimuli, gradually modified along 10-step continua, revealed that listeners of all three languages demonstrate a relatively high degree of consistent response patterns with the exception of older Quichua listeners. The results of this study coupled with the 'intermixed' acoustic spaces in which the vowels are produced also call into question the predictions that might be made in theoretical models of L2/non-native speech perception.

1. Introduction

In the Ecuadorian highlands, a well-defined language contact continuum exists between Standard Spanish and Unified Quichua. In the middle of this continuum, an intertwined 'mixed language' known as Media Lengua, or Chaupi-shimi (literally translated as "half-language"), was formed through various processes of lexification[1] and code-switching (see Gómez-Rendón, 2005; Muysken, 1980, 1981, 1997; Shappeck, 2011; Stewart, 2011). Impressionistic observations and traditional phonological analyses of Media Lengua and Quichua's vowel systems have produced varied claims regarding the degree of integration of historically Spanish origin vowels through lexical borrowings (see Cole, 1982; Gómez-Rendón, 2005, 2008; Muysken, 1997; van Gijn, 2009). However, a recent phonetic analysis of Media Lengua and Quichua vowel production revealed that historically Spanish origin vowels have been integrated into Media Lengua and Quichua, yet they exist in highly overlapping acoustic spaces with native Quichua vowels (Stewart, 2014). The first section of this paper introduces Quichua (1.1), Media Lengua, and its position as a mixed language (1.2), the vowel inventories of Media Lengua's source languages (1.3), previous analyses of Quichua and Media Lengua vowels (1.4), and background information on speech perception as it pertains to Quichua and Media Lengua (1.5).

1.1. Quichua

Imbabura Quichua (ISO 639-3: qvi), a highly agglutinating language with SOV word order, is a member of the Quechuan language family, which extends from southern Colombia to northern Argentina, primarily along the Andean mountain range. Quechua was the language of the former Incan empire, yet it was only introduced into what is now modern-day Ecuador just a couple of generations before the Spanish invasion in 1532 CE. Therefore, Quechua in this region has been in constant contact with Spanish for nearly 500 years (Adelaar & Muysken, 2004, p. 167). It is documented that nearly every semantic field, "from kinship and household to religion, education and administration" is influenced by Spanish lexical borrowings (Gómez-Rendón, 2008, p. 517). Contact with Spanish has also had a large influence at the syntactic level of Quichua.[2]

E-mail address: stewart.jesse@usask.ca

[1] The principle processes involved in creating the lexicon of Media Lengua include: relexification (the transfer of just the phonological shell of the lexifier language on to the semantic representation of the systemic language), translexification (the transfer of two or more elements from the lexifier language into the substrate language, e.g., the phonological shell and syntactic features) (Muysken, 1981), and adlexification (the lexical item from both the lexifier and substrate language co-exist) (Shappeck, 2011).

[2] In Ecuador, Quechua is known as Quichua (or Kichwa).

https://doi.org/10.1016/j.wocn.2018.08.005
0095-4470/

1.2. Media Lengua

Media Lengua (ISO 639-3: mue) is an endangered mixed language spoken by approximately 2000 people in the Ecuadorian province of Imbabura. Media Lengua was originally documented in the 1970's by Pieter Muysken (see 1981; 1997) in the province of Cotopaxi, but recent surveys by Shappeck (2011) and Stewart (2011) suggest that it has since been replaced by Spanish. In Imbabura, Media Lengua is spoken in the community of Pijal, where this study was conducted, and in four interlinked communities near the town of San Pablo (Angla, Casco Valenzuela, Ugsha, and El Topo). Based on surveys and speaker testimonies, Media Lengua spread to the San Pablo communities from Pijal through inter-community marriages (Stewart, 2011) and merchant contact (Gómez-Rendón, 2007; Jarrín Paredes, 2014) in the 1960's and as recently as 2000 in Ugsha (Müller, 2011). In Pijal, Media Lengua is moribund and in the San Pablo communities it is threatened as it slowly loses ground to Spanish. While researchers have yet to uncover any written documentation about the development of Media Lengua in the region, statements from elder speakers regarding the language of their parents and grandparents suggest that Media Lengua developed or was introduced to Pijal by the beginning of the 20th century and was used as an L1 during the 1910s (Stewart, 2011).

Mixed languages, as described in Meakins (2013) and Meakins and Stewart (2019), as distinct from other forms of language contact such as jargons, pidgins, creoles, or lingua francas. They often appear, not out of communicative necessity, but rather as a way of marking a new ethnic identity. Therefore, the originators of mixed languages are often competent bilinguals of the source languages. This aspect is often reflected in a mixed language's structure in that there is no reduced vocabulary nor simplification in the morphosyntactic structure. Moreover, mixed languages are often only used internally among members of a speech community (Bakker, 1997) while the source languages are used externally. Media Lengua is the result of mixed linguistic elements from two typologically unrelated languages, Spanish and Quichua. Nearly its entire lexical base (~90%) is of Spanish origin while its morphosyntax is essentially Quichua (see Deibel, 2017 accepted, for experimental evidence of Media Lengua's lexical-functional split). Example 1 illustrates a sentence in Media Lengua, with the Spanish origin elements displayed in italics.

(1) *Yo*taka *kiri*nimi *nieve*kunaka *cayi*chun.
jo-ta-ka *kiri*-ni-mi *nieβe*-kuna-ka *kaji*-ʧun.
1-ACC-TOP like-1S-VAL snow-PL-TOP fall-DS.SUBJ
Me gusta que caiga la nieve. (Spanish)
Ñukaka munanimi rasukunaka urmachun. (Quichua)
'I like it when the snow falls.'

1.3. Vowel inventories of Media Lengua's source languages

This section introduces the vowel inventories of Media Lengua's source languages. Section 1.3.1 discusses the Quichua vowel system and Section 1.3.2 discusses the Spanish system. Section 1.3.3 compares the two source languages.

1.3.1. Quichua

All Quechuan languages make productive use of three vowels consisting of /i/, /u/, and /a/, which are preserved from Proto-Quechua (Adelaar & Muysken, 2004, p. 196). Additionally, southern varieties have an allophonic rule that lowers the high vowels to [e] and [o] when preceded by a uvular consonant (/q/) (e.g., Cuzco [kuz**qo**]). However, in the northern Ecuadorian varieties, /q/ merged with /k/ nullifying the allophonic lowering rule (Adelaar & Muysken, 2004, p. 196). Because of this merger, Imbabura Quichua has no evidence of allophonic mid vowels.

1.3.2. Spanish

Like nearly every dialect of Spanish, Ecuadorian Spanish makes productive use of five vowels consisting of /i/, /u/, /e/, /o/, and /a/. However, Spanish spoken by Quichua-dominant bilinguals show mid vowel raising (Guion, 2003) and the use of a broad acoustic space with no bimodal clustering (i.e., categorical separation) for mid and high vowels (Lipski, 2015).

1.3.3. Comparison of Quichua and Spanish

On the surface, both source languages appear to have relatively small vowel inventories. Spanish with its five-vowel inventory contains two additional mid vowels not found in Quichua's three corner vowel system. From a phonological standpoint, Media Lengua speakers should have either adopted the Spanish mid vowels into the Quichua system or simply maintained Quichua's three vowel system by assimilating Spanish origin mid vowels. However, claims presented in Section 1.4 suggest that these scenarios may not be adequate to describe what is happening at the phonetic level.

1.4. Analyses of Media Lengua and Quichua vowels

This section introduces claims put forth by numerous researchers who have described Media Lengua and Quichua's vowel systems. Section 1.4.1 details impressionistic claims and 1.4.2 details acoustic studies of Media Lengua and Quichua vowel inventories.

1.4.1. Impressionistic claims

Two impressionistic observations of Media Lengua show mixed results in terms of the realisation of historically Spanish mid vowels. In the seminal paper on Media Lengua by Pieter Muysken (1997), he stated that, "...mid vowels /e/ and /o/ are collapsed with the high vowels /i/ and /u/, respectively..." (p. 383). This description suggests that Media Lengua vowels are essentially Quichua-like. However, Gómez-Rendón (2005) stated that, "...vowels [e] and [o], appear in Media Lengua almost exclusively in relexified roots and interjections...[yet there are] cases [where] one vowel is raised (*vendi-*, *ofreci-*)" (p. 48).[3]

For Quichua, impressionistic observations are more consistent. Cole (1982) stated that, "...the mid vowels /e/ and /o/ are borrowed from Spanish. [However], the mid vowels are found only in unassimilated Spanish words. Monolingual speakers generally pronounce borrowed mid vowels as high vowels"

[3] Translated by the author.

(p. 203). Similarly, Gómez-Rendón (2008) stated that, "Spanish medial vowels are raised (/e/>/i/, /o/>/u/) or otherwise pronounced as close as possible to their Quichua equivalents" (p. 106). However, he also provides detailed explanations for irregularities including, partial assimilation for words with several mid vowels, the environment of the mid vowel, word frequency, and level of bilingualism (p. 106–7).

1.4.2. Acoustic analyses

Three studies have examined the Quichua vowel system using acoustic methods for analysis. Guion (2003) and Lipski (2015) provided insights into the vowel systems of Quichua-Spanish bilinguals and Stewart (2014) examined the vowel inventories of both Media Lengua and Quichua speakers. In Stewart's (2014) analysis of source vowels in Media Lengua and Quichua speech, he revealed that historically Spanish mid vowels exist in both languages. Yet, these mid vowels occupy a region of acoustic space that highly overlaps with Quichua origin high vowels (see Fig. 1). However, the differences in F1 frequencies between Media Lengua mid and high vowels were significant with a mean average distance of 41 Hz (0.36 Bark)—a value just above the threshold of 0.3 Bark for formant discrimination suggested by Kewley-Port (2001) for values between 200 and 3000 Hz.[4] Like the Media Lengua results, the differences between Quichua mid and high vowels were also significant but at negligible distances, with a mean difference of just 26 Hz (0.23 Bark)—approximately half that of Media Lengua and just below the threshold of 0.3 Bark suggested by Kewley-Port. These findings point to substantial raising in the mid vowel series without complete merger.

For bilingual speech, Lipski (2015) revealed that L2 speakers of Spanish (Quichua-dominant) produce front and back vowels within a wide range of F1 and F2 frequencies, with no clear indication of systematic separation between mid and high vowels. Lipski's study reinforces Guion's (2003) earlier findings that Spanish mid vowels merge with Spanish high vowels in L2 speech of late-Quichua-dominant bilinguals. These studies also reflect Quichua and Media Lengua speech (Stewart, 2014) as the overlapping vowel spaces essentially look like a large single gradient category with no clear systematic division between mid and high vowels (see Fig. 1). Therefore, it appears Quichua-dominant/Media Lengua-dominant bilingual speakers are operating the same vowel categories when speaking both L1 Media Lengua/L1 Quichua and L2 Spanish.

With mid and high vowels overlapping in acoustic space and no bimodal clustering in bilingual speech, identifying mid and high vowels based solely on impressionistic observations might be exceedingly difficult by non-native Quichua or Media Lengua researchers. This premise returns us to the original research question, are native Quichua and/or Media Lengua listeners able to identify differences between such mid and high vowels? Or do mid and high vowels in Quichua and/or Media Lengua simply exist as near-mergers (i.e., as overlapping categories whose differences are not distinguished auditively and only be teased out though acoustic analysis)?

[4] Bark values calculated using Traunmüller (1990). A special thanks to an anonymous reviewer and Rob Hagiwara for pointing this out.

1.5. Speech perception

According to Labov (1994), near-mergers are a common occurrence in language contact scenarios. If Media Lengua and Quichua listeners do not identify consistent differences between mid and high vowels, this might suggest that the overlapping vowel configurations fit well within the standard description of near-merger (see Hickey, 2004; Labov, Yaeger, & Steiner, 1972; Labov, Karen, & Miller, 1991) or incomplete neutralisation (see Mitleb, 1981; Port & Crawford, 1989; Port & O'Dell, 1985; Roettger, Winter, Grawunder, Kirby, & Grice, 2014; Winter & Roettger, 2011). However, if Media Lengua and/or Quichua listeners are able to consistently identify differences between mid and high vowels, a question arises as to why the overlapping categories have not dispersed in the directions predicted by models of adaptive dispersion (see Johnson, 2000; Liljencrants & Lindblom, 1972; Lindblom, 1986, 1990; Livijn, 2000). Models of L2/non-native speech perception, such as the Perceptual Assimilation Model (Best, 1993, 1995; Best, Hallé, Bohn, & Faber, 2003) and the Speech Learning Model (Flege, 1995; Flege, Schirru, & MacKay, 2003) would predict that in Media Lengua and/or Quichua, Spanish origin mid vowels should either be collapsed into a single category (merger) with high vowels if listeners perceive them as the same or the mid vowel should be added to the Quichua vowel inventory as separate categories if listeners perceive them as different from high vowels. If Media Lengua and Quichua listeners are indeed able to identify differences between overlapping mid and high vowels, how do models of L2/non-native speech perception and adaptive dispersion account for this?

Age of acquisition of an L2 affects which acoustic cues are available to a learner. For early L2 acquisition, studies involving Spanish-Catalan bilinguals reveal that while early exposure to Catalan (typically before the age of 3) is the norm for Spanish dominant bilinguals, listeners struggle to discriminate Catalan low-mid vowels /ɛ/ and /ɔ/ from their mid vowel counterparts, /e/ and /o/ (see Navarra, Sebastian Galles, & Soto-Faraco, 2005; Pallier, Bosch, & Sebastian Galles, 1997; Pallier, Colome, & Sebastian Galles, 2001; inter alia). These results even appear when speakers maintain robust contrasts in their production (Amengual, 2014).

Such findings are also reflected in Guion's (2003) results, which showed that the acoustic spaces of mid and high vowels for simultaneous, early, and late Quichua-Spanish bilinguals are all formed differently due to how vowel input was processed auditively. Simultaneous bilinguals made use of two separate vowel systems, early bilinguals merged Spanish high vowels with Quichua high vowels, and late bilinguals merged Spanish mid vowels into approximately the same space as Quichua high vowels. Her findings revealed that distinct vowel arrangements are linked to the developmental differences related to a speaker's age of L2 acquisition. More specifically, the earlier a person is exposed to their L2, the greater the likelihood that they will acquire the necessary perceptual information required to produce native-like vowels.

The primary goal of this paper is to determine whether Media Lengua and/or Quichua listeners can identify differences between mid and high vowels within the overlapping acoustic spaces described in Stewart (2014). It is hypothesised that

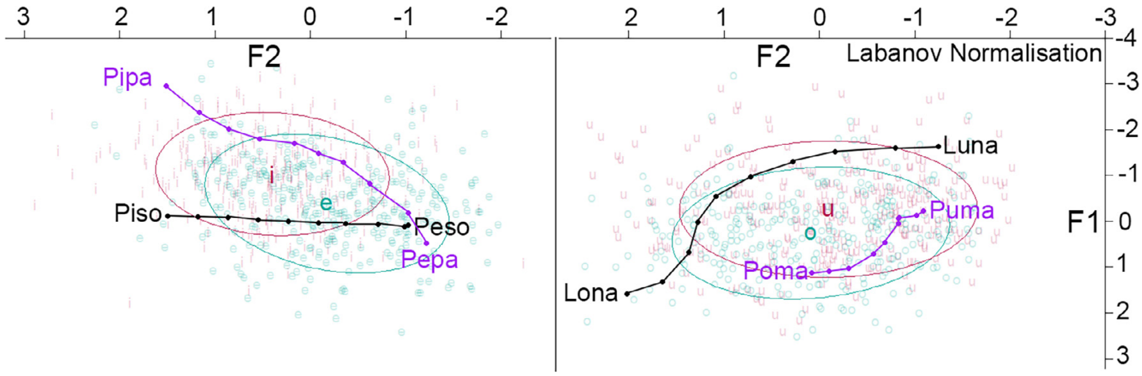

Fig. 1. Both plots represent the normalised F1 and F2 frequencies of the Media Lengua front and back vowel clusters detailed in Stewart (2014). Standard deviations of each Spanish origin vowel are shown with ellipses. The 10-step continua of each minimal pair used in this study are superimposed and their paths through the clusters are highlighted by the connected points. Each vowel pair has one continuum which begins and ends outside one standard deviation while the other pair has a shorter trajectory with smaller distances between the steps.

Media Lengua listeners will be able to identify such differences for two reasons (a) the physical differences in F1 frequencies between Media Lengua mid and high vowels are beyond that of 0.3 Bark for formant discrimination suggested by Kewley-Port (2001) and (b) cognitively, the presence of mid and high vowel contrasts in the phonology could be beneficial for managing the Spanish origin vocabulary. On the other hand, it is hypothesised that Quichua listeners will produce less consistent results as (a) differences in their mid and high vowel spaces are below 0.3 Bark and (b) native Quichua phonology does not make use of mid and high vowels for contrastive purposes.

An additional inquiry considers how the aforementioned theoretical models of L2/non-native speech perception might fall short at predicting how listeners perceive mid and high vowels in Quichua and/or Media Lengua. If mid and high vowels in Quichua and/or Media Lengua are indeed identified as different, it is hypothesised that such theoretical models will most often fail to predict the distinction due to the mid and high vowel's articulatory similarities. However, if Quichua and/or Media Lengua listeners do not identify differences between mid and high vowels, it is hypothesised that such theoretical models will most often accurately predict the merger due to the mid and high vowel's articulatory similarities.

2. Method

The experiment conducted in the present study examined mid and high vowel perception of Imbabura Quichua native listeners from Chirihuasi and Imbabura Media Lengua native listeners from Pijal. Results from Urban Andean Spanish native listeners from Quito were used as a control group as Spanish mid and high vowels are known to be clearly dispersed from each other (see e.g., Chládková, Escudero, & Boersma, 2011).

2.1. Stimuli

2.1.1. Minimal pairs

To collect vowel perception data, a two-alternative forced choice (2AFC) identification task experiment with modified minimal pair words differing by word-internal mid and high vowels (e.g., *p<u>e</u>so* 'weight' and *p<u>i</u>so* 'floor') was conducted. Each of the words was of Spanish origin and nativised in Media Lengua and co-exist alongside native Quichua words in Quichua. Table 1 includes the breakdown of the data used in this experiment.

Table 1
List of minimal pairs used during the identification task experiments.

[i]-[e]	[u]-[o]
piso [ˈpi.so] 'floor'	*lona* [ˈlo.na] 'canvas/tarp'
peso [ˈpe.so] 'weight'	*luna* [ˈlu.na] 'moon'
pipa [ˈpi.pa] 'pipe'	*poma* [ˈpo.ma] 'jug'
pepa [ˈpe.pa] 'seed'	*puma* [ˈpu.ma] 'puma'

While listeners tend to have a perceptual bias toward words with higher frequencies of usage (see Broadbent, 1967; Goldiamond & Hawkins, 1958), word frequency was not considered due to the lack of corpora data for the languages under analysis (Quichua, Media Lengua, and Ecuadorian Spanish). However, results from an international Spanish corpus (Corpus de Referencia del Español Actual (CREA), 2018), suggest that *pipa* 'pipe' and *pepa* 'seed' are both low frequency in addition to *poma* 'jug' and *puma* 'puma'. For *piso* 'floor' and *peso* 'weight', the latter is slightly higher in frequency, but this difference may be caused by references to the currency by the same name, which is not used in Ecuador (Ecuador uses the US Dollar and formally the Sucre). These factors may reduce the frequency of *peso* in the Ecuadorian context as this word usually just refers to 'weight'. The largest discrepancy in word frequency was found in the *lona* 'tarp' and *luna* 'moon' pair, in which the latter fell in the mid frequency range and the former in the lower frequency range. However, in the rural Ecuadorian context, *lona* 'tarp' is a common material with multiple uses; a context that is likely to increase its frequency.

Additional analyses, based on response frequency to the stimuli, suggest that Quichua and Media Lengua listeners had marginal differences in word choice between the minimal pairs containing /o/ and /u/ (the latter being slightly preferred). However, Spanish listeners preferred *poma* 'jug' over *puma* 'puma'. Similar trends were found between the front vowel series—marginal differences between minimal pairs containing /e/ and /i/ (the latter being preferred for Quichua and Media Lengua), while Spanish listeners had a slight preference toward /e/. It is speculated that the preference for the high vowel series in the Quichua group is more likely due to the familiarity of /i/ and /u/ over their mid vowel counterparts rather than word frequency. Contrarily, for Spanish listeners, /e/ is more frequent than /i/ in Spanish, which may explain the preference toward /e/.

A female native Spanish speaker from Quito was recorded reading each word listed in Table 1. A TASCAM DR-1 portable digital recorder with a NEXXTECH unidirectional dynamic microphone (50–13000 Hz response) was used for the recording. The words were recorded in 16-bit Waveform Audio File Format (WAV) with a sample rate of 44.1 kHz. As all words were of high quality and no issues were encountered during the manipulation phase, a single recording was sufficient.

2.1.2. Continuum construction

Rather than using purely synthetic audio samples, semi-synthetic stimuli based on natural speech were created. This step was taken to minimise issues of segmental quality often attributed to synthetic speech (see Vainio, Järvikivi, Werner, Volk, & Välikangas, 2002). To create the stimuli, the formants of the original sound tokens' vowels were modified and then combined with the following syllable of the original mid vowel token to create a more naturalistic sound sample. The mid vowel token add-ons were used because the perception of mid vowels is the focus of this study and reducing the number of high vowel correlates in the mid vowel-like stimuli was thought to reduce the chances of possible misidentification.

A 10-step vowel continuum was chosen to transition from one phoneme group to the next to cover a relatively large range of samples. The F1 and F2 formant frequencies of the vowels under analysis transitioned through or within one standard-deviation of the normalised formant values (Labanov) based on the Media Lengua vowel data used in Stewart (2014) (Fig. 1).

As values drift towards the centre of the continua, it was hypothesised that responses would be more random due to the greater amount of mid and high vowel mixing near the centre (as observed in the individual vowels in Fig. 1). For each stimulus used in the continua, the vowel's F1, F2, and F3 formant points were evenly spaced (using a linear scale in Hertz) based on Media Lengua formant frequencies from Stewart's 2014 data (see Table 1). However, before editing the vowel formants, the tokens were resampled at 11000 Hz and the fundamental frequency (f0) across the entire vowel was flattened to reduce prosodic cues that might affect language mode during the experiment. Both the f0 and vowel duration were then adjusted along the continua at equal steps. The final product resulted in a single continuum for each minimal pair, which contained modifications to all three acoustic cues (pitch, formant frequencies, and duration) of the vowels in question. Therefore, if a listener contrasts mid and high vowels perceptually, then the words at both poles (steps #1 and #10) should have more consistent responses as there is less vowel mixing in the peripheries. If a listener does not contrast mid and high vowels, it is predicted that only one pole (step #10) should have more consistent responses.

After formant and pitch data were modified, they were passed through an inverse filter, producing a WAV file with the modified values. It should be noted that this step slightly alters the pitch and formant adjustments. This results in continua which are not perfectly spaced at each interval. For example, Fig. 1 illustrates this effect in *peso*-*piso* where step 1 and 2 are nearly identical while the rest of the steps are more evenly spaced. While this phenomenon did not affect a listener's ability to identify the overall mid and high vowel contrasts, it may slightly alter where a possible categorical boundary between the two vowels appears. A breakdown of the formant values (centre point of the vowel only), along with pitch (f0) and duration (ms) from each minimal pair along each continuum is given in Table 2.[5]

Once the pitch, vowel duration, and vowel formants were modified and filtered, the unaltered portion of each token was removed and replaced with the same section from the original mid vowel phoneme resampled to 11000 Hz.[6] This sample provided more naturalistic stimuli compared to synthesised tokens. Fig. 2 shows a side-by-side comparison of two spectrograms of the first and last steps along the *peso* and *piso* continuum where the contrasts in F2 height and duration in the first vowel stand out.

2.2. Data collection procedure

This section describes the user interface used to present the experiment to the participants. The 10 audio stimuli created between each minimal pair, described in Section 2.1.2 were integrated into a PowerPoint presentation for the experimental task. Images corresponding to each minimal pair were loaded into the presentation (Fig. 3).

To achieve more precise results, the presentation was designed to contain more repeats of stimuli before the sixth step to aide in identifying the approximate location of the categorical boundary. It was expected that the high vowel-like stimuli would provide more consistent responses as it is a known phoneme in all three languages. Therefore, less stimuli repetitions were provided after step 5 and more were added from steps 1–5 to help decrease the chances that responses to more mid vowel-like stimuli were incidental. In total, the participants listened to 21 tokens within each continuum, providing a total of 84 data points (when considering all four minimal pair series). Table 3 provides an example of the repeated stimuli along the continuum.

The PowerPoint presentation was configured to play the audio sample 50 ms after each slide appeared on the screen, and participants had the option to repeat the stimulus by clicking on the speaker icon at the bottom of the screen.[7] The presentation was configured to use 'Kiosk' mode, which only allowed the participants to proceed to the following slide after

[5] An anonymous reviewer kindly pointed out that amplitude should have also been modified as per Lehiste and Peterson (1959). Unfortunately, this was not considered during the stimuli building phase. However, after passing the modified data through the filter, amplitude for the most part increases or decreases in a similar fashion compared to the original tokens, albeit to a lesser extent. For example, the difference between the original token intensity average between *pepa* and *pipa* was 7 dB (56 vs. 49 dB). However, the average difference between the *pepa-pipa* stimuli in the continuum was 2 dB (80 vs. 78 dB). The increase in the intensity is a result of the filtering process as well; however, volume was controlled during the experiment.

[6] It should be noted that a slight bias may exist in the stimuli containing the mid-like vowels since the natural speech add-on is from the word containing the mid vowel (e.g., the [na] from *lona* is used for all the stimuli in the *lona-luna* continuum).

[7] The number of repetitions was not recorded in the experiment. Informally, however, we did not observe anyone clicking more than three times on the speaker icon.

Table 2
Vowel formants, f0, and duration of each minimal pair analysed in this study.

Mid		Continuum values for [e]-[i]										High
		1	2	3	4	5	6	7	8	9	10	
peso 'weight'	F1	542	520	507	497	489	484	489	486	470	456	*piso* 'floor'
	F2	2530	2572	2623	2674	2720	2751	2769	2795	2861	2900	
	F3	3215	3254	3306	3382	3422	3454	3462	3490	3531	3588	
	f0	230	231	232	233	234	235	236	237	238	239	
	Ms	102	99	96	93	90	88	85	82	79	76	
pepa 'seed'	F1	509	477	446	424	414	404	399	389	371	343	*pipa* 'pipe'
	F2	2229	2263	2335	2384	2431	2475	2540	2598	2652	2714	
	F3	2953	2982	3017	3072	3106	3127	3150	3187	3216	3241	
	f0	207	206	206	205	204	203	203	203	202	201	
	Ms	94	92	90	89	89	88	86	83	81	81	
Mid		Continuum values for [o]-[u]										High
poma 'jug'	F1	590	588	584	563	547	520	513	511	508	502	*puma* 'puma'
	F2	1034	1006	972	930	911	889	888	887	858	847	
	F3	2393	2393	2492	2525	2547	2483	2533	2529	2457	2526	
	f0	197	202	208	213	219	224	230	235	241	246	
	Ms	150	143	137	130	123	117	110	103	97	90	
lona 'tarp'	F1	620	603	561	518	481	453	432	418	413	410	*luna* 'moon'
	F2	1345	1286	1241	1226	1196	1138	1067	995	894	821	
	F3	3230	3125	3040	2918	2731	2736	2667	2600	2570	2463	
	f0	214	213	213	212	212	211	211	210	210	209	
	Ms	164	159	154	149	144	138	133	128	123	118	

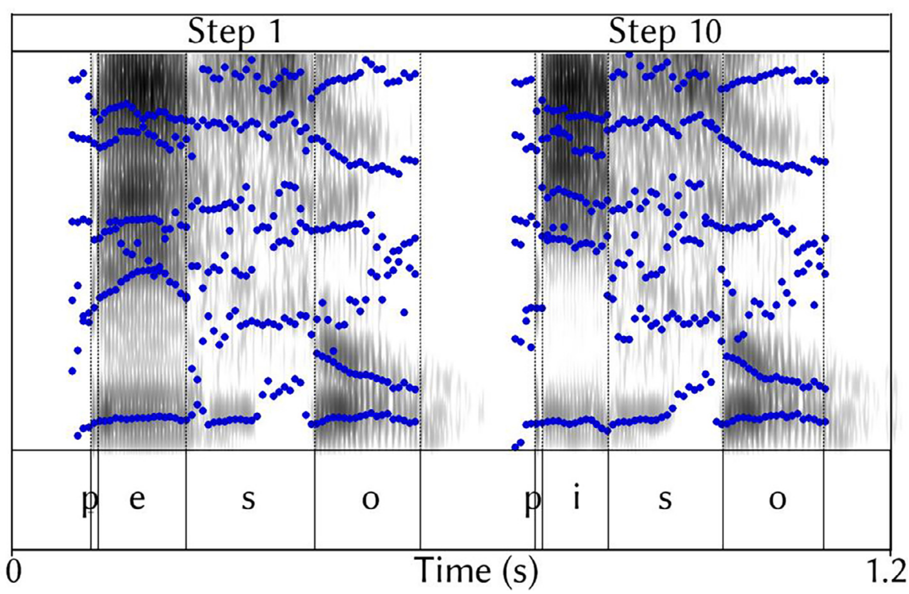

Fig. 2. Spectrograms and formants of step 1 (*peso* 'weight') and step 10 (*piso* 'floor').

selecting either of the two images. Each image was programmed using the Visual Basic for Applications (VBA) add-on in PowerPoint to record participant responses. The slides were randomised using a macro (Reilly, 2011) and further adjusted to ensure no two trials contained the same images consecutively. After randomization, one slide containing an audio sample from step 10 was inserted at the beginning of the presentation to provide the participants with stimuli furthest away from the intermixing of mid and high vowels in the centre of the vowel space (see Fig. 1) to orient themselves before being presented with other forms at random. Eight additional minimal pair continua with transitions between word-initial stops were added to the experiment as distractors to help reduce repetition.

2.3. Participants

Eleven Media Lengua trilingual listeners (L1 Quichua/Media Lengua, and L2 Spanish) participated in the experiment. This group consisted of seven women and four men, and all were from the community of Pijal. All listeners acquired Quichua and Media Lengua simultaneously from birth, though two younger listeners (#60 and #61) did not actively speak the languages. These younger speakers understood Quichua and Media Lengua but preferred to answer back in Spanish when spoken to by their parents or other speakers. When asked to speak Media Lengua or Quichua, they would state they could not or would reluctantly do so briefly before switching back to

Fig. 3. The slides containing pictures of allowed responses in the word identification task experiment. The top slide contains pictures of *lona* 'canvas/tarp' and *luna* 'moon'; the second contains pictures of *poma* 'jug' and *puma* 'puma'; the third contains pictures *pepa* 'seed' and *pipa* 'pipe'; and the fourth contains *peso* 'weight' and *piso* 'floor'.

Table 3
Example of the repeated steps along an /e/ to /i/ or an /o/ to /u/ continuum.

1	2	3	4	5	6	7	8	9	10	Steps
2	3	4	4	3	1	1	1	1	1	# of stimuli repeats

Spanish, suggesting that their language performance was more akin to that of a passive bilingual. These passive bilingual participants acquired Spanish as their L1 (see Appendix B, Table B1). Younger active speakers (typically in their 30's and early 40's) began learning Spanish upon entering primary school, typically at the age of six to seven. Most of the older generation of Media Lengua listeners (typically in their 50's and 60's) who took part in the present study did not acquire Spanish until they entered the workforce, as such acquisition was typically 'unguided'.

Ten Quichua listeners, six women and four men, also participated in the present study. Participants were all bilingual (L1 Quichua and L2 Spanish). Five older participants (43+) were deemed to have a lesser command of Spanish (mid-level) compared to four others deemed to have a high-level while one of the younger participants was a simultaneous bilingual. The younger participants acquired Spanish upon entering primary school, typically at the age of six to seven. Similar to the Media Lengua group, older participants often did not acquire Spanish until entering the workforce and under 'unguided' conditions. Participants were born, raised, and lived in the community of Chirihuasi at the time of this study.

Since providing a standardised Spanish assessment test or written questionnaire was not practical as several consultants from both groups struggled to read and the topic of language proficiency is delicate due to colonial resentment, Spanish proficiency for the Quichua group was judged using several factors: (1) feedback from our Quichua-speaking assistant, who was familiar with the participants; (2) our own interactions in Spanish with the participants after the experiment; (3) informal oral self-assessment loosely based on a number of questions from the Bilingual Language Profile (BLP)[8] (Appendix A) (Birdsong, Gertken, & Amengual, 2012); and (4) our own

[8] It should be noted that BLP does not necessarily assess language proficiency but rather language dominance. It is also worth noting that 'language attitude' was not formally assessed; however, it is clear from discussions with the Media Lengua speakers that Quichua and Spanish are considered more prestigious. Discussions with the Quichua participants also revealed that Spanish is considered more prestigious. In addition, there is a clear shift in language usage in both communities with children only acquiring Spanish.

familiarity with the participants. To provide a general metric on a scale from *low* to *native* each of the listed items above were weighted. As a leader in the Quichua community, our Quichua speaking assistant was already familiar with the participants' level of Spanish. His assessment of *native, high, mid,* and *low* was our first indicator. In every instance, his observations reflected our interactions with the participants. In our conversations with the participants in Spanish, we assessed their grammar, fluency, and vocabulary usage. We did not classify anyone as *low,* which we deemed as a speaker who struggled to hold a conversation in Spanish, lacked knowledge of verbal conjugations, and/or constantly asked our assistant for clarification or translation. Those classified as *mid* could hold a conversation, yet it was apparent that they struggled with specific grammatical structures and fluent speech was interrupted due to L1 interference. *Native* and *high* were sometimes difficult to differentiate; therefore, we would ask the participants questions about their language use (e.g., "When and where did you learn Spanish?"; "What languages did your parents speak to you growing up?") (see Appendix A for an exhaustive list of questions). This self-assessment led to the classification of *native* (acquired from birth or before the age of 2) and *high* (learned in school and grammar, fluidity, and vocabulary usage showed little to no signs of L1 interference). A nearly identical strategy for gauging language proficiency was used for the Media Lengua group. Our Media Lengua assistant was well-known to all the participants and provided accurate feedback regarding language proficiency that also matched our observations.

The Spanish variety used in the present study was from the capital city of Quito and was used to create a baseline for Media Lengua and Quichua vowel judgements. From the Spanish group, 18 participants (11 women and 7 men) participated in the experiment. This group consisted of Spanish monolinguals with little or no knowledge of Quichua. All participants were primary school teachers, apart from one law student and one orthodontist, and were born, raised, and lived in Quito, with the exception of one participant who was born in Tulcán, Carchi. Appendix B contains information pertaining to the participants who took part in the 2AFC identification task. These data include participant code, age at the time of the experiment, gender, and demographic descriptors including education, fluency level of Spanish, Quichua, and Media Lengua, and the usage of Quichua or Media Lengua,[9] and place of residence. It should be noted that due to fieldwork limitations (the participant recruitment method by which data were gathered for this study), the data might not have been controlled as desired (e.g., variability in passive/active bilingualism and level of education). Therefore, readers should bear in mind the limitations of the study in interpreting the results.

2.4. Experimental procedures

Prior to the experiment, the participants met together to allow them to speak amongst themselves in their L1 while the author and assistants set up the task. After set up, each participant was individually informed that they would hear a variety of words, known to them in their L1 but borrowed from Spanish, and that their task was to click on the picture that corresponded to the word they heard. The experiment was an individual activity conducted away from the other participants. Instructions were given in the native language of the participant by the author (an L2 speaker of all three languages) or by assistants who were L1 speakers of each language. Participants were given the option to repeat the stimuli if they chose to do so by clicking on a speaker icon at the bottom of the screen. However, participants were asked to react with their first instinct. The participants were also informed that the words would be repeated many times and that, despite the fact that some of the words might be harder to perceive than others, they should try their best to identify the word. Before beginning the task, the possible answers were reviewed from a print out with the participants to avoid confusion during the task. The experimenter modelled the words contrasting the mid and high vowels. Before giving written consent, the participants were informed that the entire task would last approximately 15 minutes and that there were no right or wrong answers. Participants were monetarily compensated for their time.

The participants were provided with a PC laptop and noise cancelling headphones to prevent distractions by random ambient noise. All participants reported normal hearing, and the audio stimuli were presented at a comfortable and consistent volume level of three-fifths of the maximum volume (for all participants). They were asked to inform the experimenters if the volume was too high or too low, but none did. Participants who did not feel comfortable using the laptop mouse to click on the pictures were asked to point at the picture they heard in the audio sample, and the experimenter would click on the image for them. If the participant was interested in seeing the results after the experiment, the output response data were loaded into Excel and automatically graphed to display the results. Participants were informed whether they were able to identify differences in the minimal pairs based on their performance in the task, and, if so, at what point they identified one over the other.

3. Results

3.1. Data analysis

Figs. 4 and 5 display line plots illustrating the mean trajectories of the responses from each language along the continua, specifically Spanish (dotted, blue), Quichua (solid, green), and Media Lengua (dashed, red). Error bars consisting of the 95% confident intervals were also used to estimate where the mean averages at each step would land 95% of the time were the experiment to be run repeatedly with different random sample groups. The percentage of responses containing the mid vowel is presented on the y-axes in Figs. 4 and 5. Specifically, the closer a response is to the top of the y axis, the higher the percentage of mid vowel responses, while the closer a response is to the bottom of the y axis, the lower the percentage the mid vowel was chosen (i.e., the higher the percentage a high vowel token was selected). The x-axes of each graph

[9] 'Daily' = used at home or with a close family member. 'Intermittently' = may be used during brief exchanges at home or with specific people outside the home who speak it. 'Infrequently' = Spanish is typically the language spoken at home and with friends, but s/he is not opposed to using Media Lengua with others. 'Rarely' = only used when told to do so, which happens from time to time between young adults (18–25) and their grandparents.

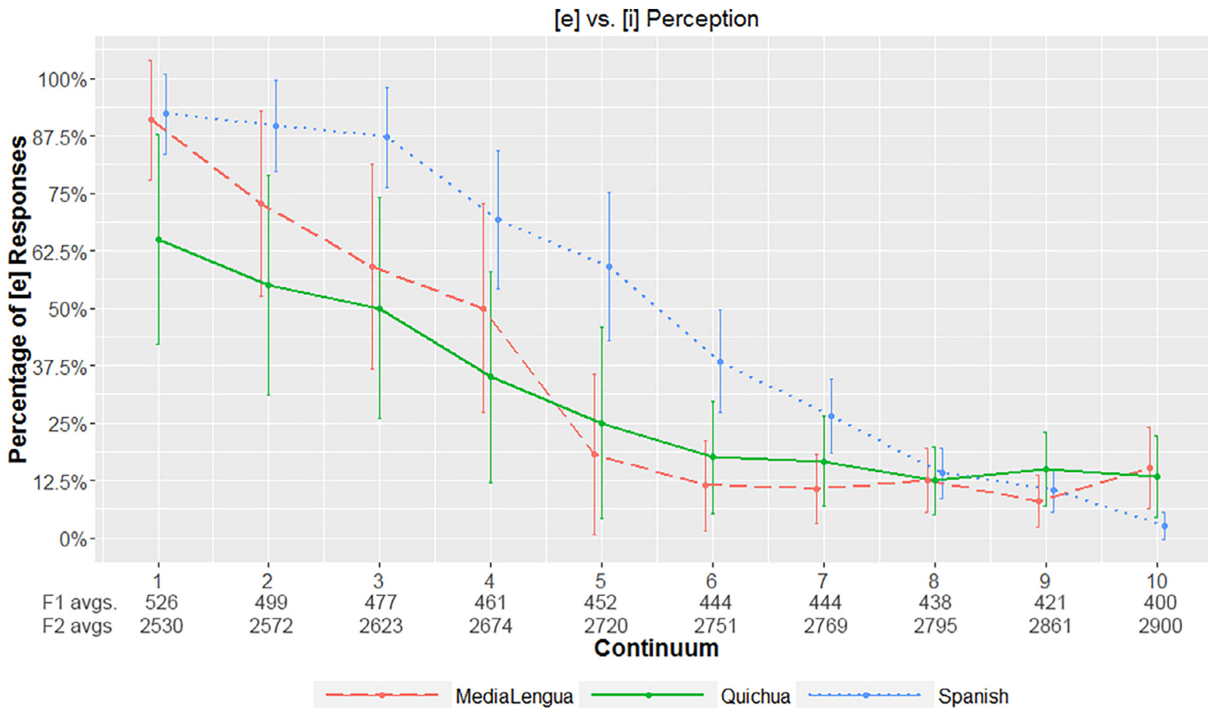

Fig. 4. Front vowel perception from Spanish (dotted), Media Lengua (dashed), and Quichua (solid).

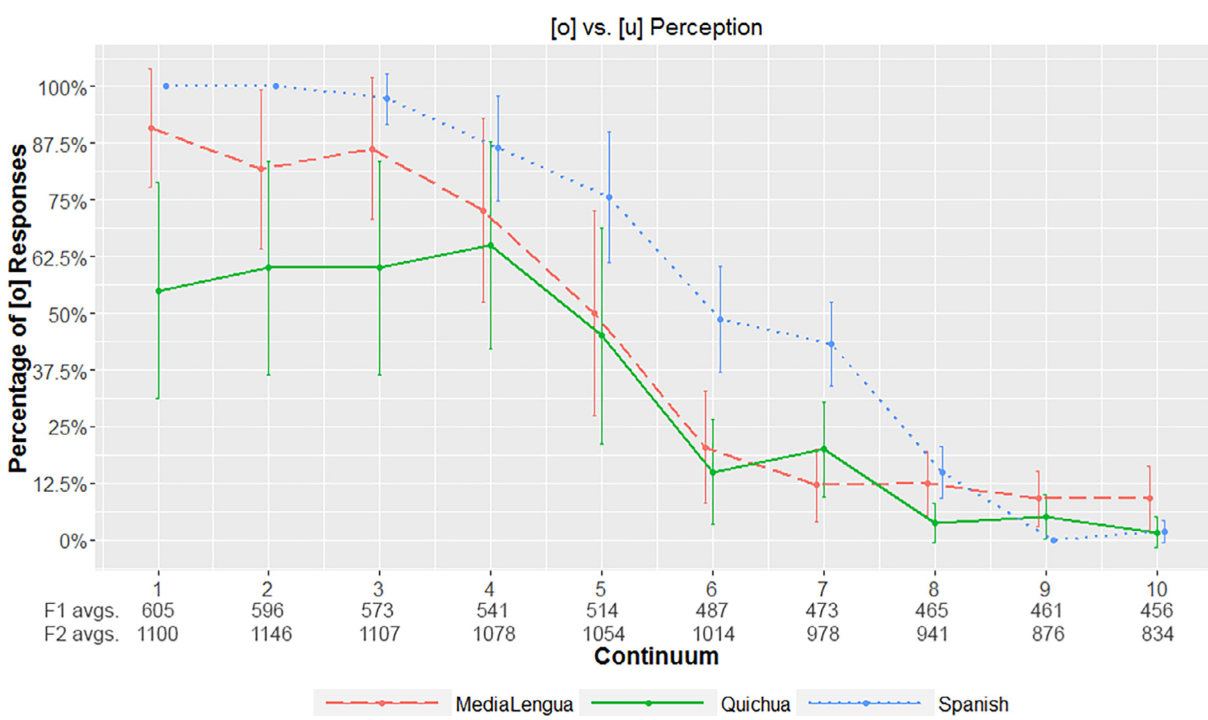

Fig. 5. Back vowel perception from Spanish (dotted), Media Lengua (dashed), and Quichua (solid).

indicate the steps along the continuum (1–10) and the average F1 and F2 frequencies from the centre point for each vowel as detailed in Table 2. The line plots displayed in Figs. 4 and 5 for the Spanish and Quichua participants are slightly staggered from the centre of each continuum step for visualization purposes only.

The raw data in Figs. 4 and 5 can be interpreted with two metrics, (1) where the categorical boundary falls for each language group (i.e., the continuum step on the x-axis where the trendline crosses the 50% line on the y-axis) and (2) the percentage of mid vowel responses for the first mid vowel stimulus (i.e., the percentage point at the first step of the continua).

In Fig. 4, the categorical boundary for Media Lengua listeners appears directly at step 4, and the percentage of mid-vowel responses at step 1 reaches 90%. For the Quichua listener results, the categorical boundary appears directly at step 3 with the percentage of mid-vowel responses at step 1 reaching just 65%. For the Spanish listeners, the categorical boundary appears between steps 5 and 6 with their percentage of mid-vowel responses at step 1 reaching 92%.

In Fig. 5, the categorical boundary for Media Lengua listeners appears directly at step 5, and their percentage of mid-vowel responses at step 1 is again 90%. For the Quichua listener group, the categorical boundary also appears between steps 4 and 5 but with the percentage of mid-vowel responses at step 1 reaching just 55%. For the Spanish listeners, the categorical boundary appears between steps 5 and 6 with their percentage of mid-vowel responses reaching 100%.

The trends in Figs. 4 and 5 reveal that the location of the categorical boundary is different among the three language groups for both the front and back vowels. More precisely, the Quichua boundary was found furthest leftward and the Spanish boundary most rightward. This means Quichua listeners identify more stimuli as high vowels than Spanish listeners, who identify more mid vowels—Media Lengua listeners fall in between. Regarding the differences in response percentage to mid vowel stimuli at step 1, Spanish listeners clearly identified both [e] and [o] as such with little variation (92% & 100%, respectively). Media Lengua speakers also showed a high degree of consistent responses to the mid vowel stimuli at step 1 (90% for each) while Quichua speakers had more variable responses to the mid vowel stimuli at step 1 (65% for [e] and 55% for [o]).

The trends in Figs. 4 and 5 also suggest that for all three language groups, the boundaries for the front vowel continua were further to the left than for the back vowel continuum. This means that when comparing responses between the front and back continua, participants identify more stimuli as the mid vowel than the high vowel in the latter, compared to the front series where participants identify more stimuli as the high vowel than the mid vowel. Otherwise stated, comparing identification in the front vs. back series, there is a preference for /i/ in the former, while in the latter that preference for the high (/u/) vowel is not as strong. One reason for this discrepancy might be attributed to the higher frequency of /i/ in Spanish (and thus Media Lengua due to the relexified Spanish lexicon) compared to /u/ (Guirao & García Jurado, 1990).

After analysing individual responses, it became apparent that the decrease in /e/ and /o/ responses on the left side of the continua in the Quichua group (the green, solid lines illustrated in Figs. 4 and 5), came from older listeners who overwhelmingly preferred either the high vowel series or who had variable responses during the first 4 steps along the continua. The details of each individual listener in the Quichua group are presented in Figs. 6 and 7. Individual Media Lengua responses are also presented in these figures as a basis for comparison with the Quichua responses.

Results by Quichua participant in Fig. 6 suggest that the mid high vowel contrast was only robustly identified by younger participants (#82, #79, #78, and #84). The age range of the participants who consistently contrasted the front vowels was between 26 and 34 while three of the four were in the higher education bracket, and all were deemed to have a high level of Spanish. For the older group of listeners, with an age range of 43 to 68, four of the five (#63, #81, #83, and #85) overwhelmingly chose /i/ over /e/, and listeners #86 and #80 displayed varied responses with a higher response rate to /e/, that did not surpass the 50% mark toward /i/.

Results from the Media Lengua participants revealed greater response consistency compared to Quichua listeners. The vowels /e/ and /i/ were identified as different to varying degrees, for participants #56, #41, #55, #43, #58, #60, and

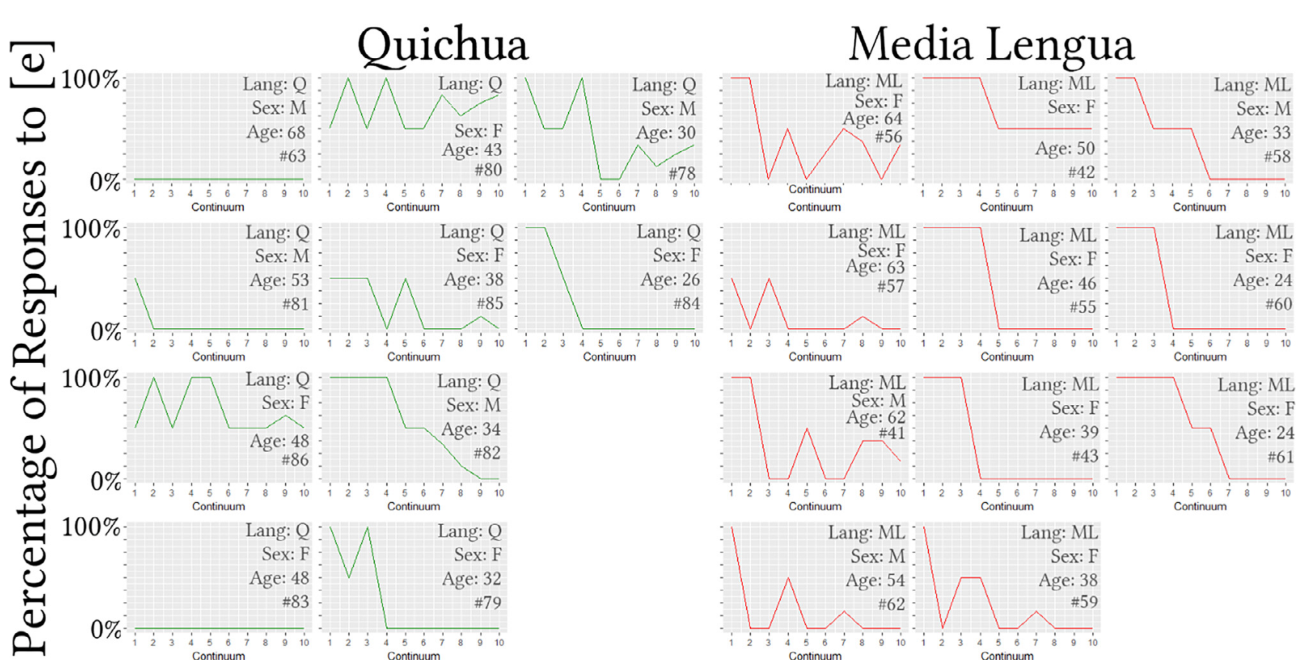

Fig. 6. Quichua (left) and Media Lengua (right) front vowel perception by participant. Each individual chart contains the participant's language, sex, age, and consultant number. Graphs are ordered by age with the oldest in the top left corner and youngest in the bottom right corner.

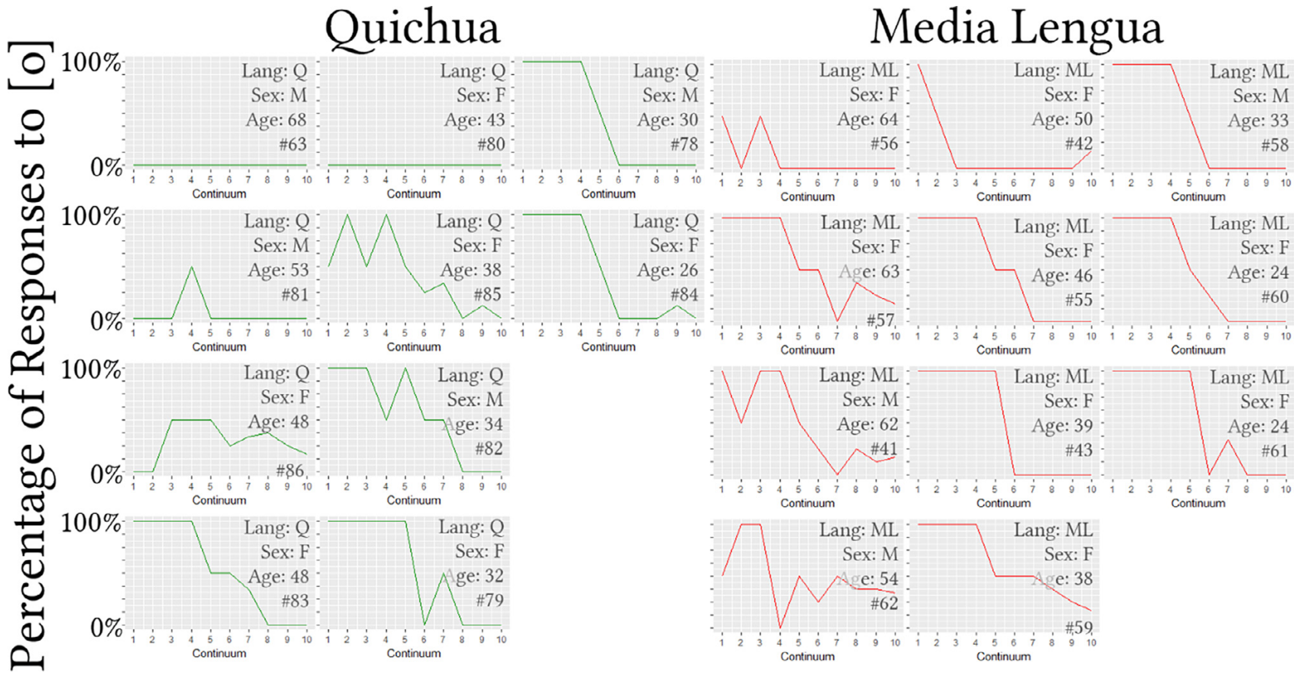

Fig. 7. Quichua (left) and Media Lengua (right) back vowel perception by participant. Each individual chart contains the participant's language, sex, age, and consultant number. Graphs are ordered by age with the oldest in the top left corner and youngest in the bottom right corner.

#61; participants #62 and #59 only demonstrating a high degree of consistent responses to the /e/ token at the first step; #57 showed a preference toward /i/, while #42 demonstrated varied responses on the right of the continuum.

Results by Quichua participants in Fig. 7 suggest that, like front vowel perception, only young participants (#85, #82, #79, #78, and #84) identified differences between back vowels /o/ and /u/, apart from participant #83 (48 years of age), who had mid-level fluency in Spanish. Results by Media Lengua participants suggest that /o/ was also identified as different from /u/ by all participants apart from participant #56. Participant #42 demonstrated only a high degree of consistent responses for /o/ when presented with the stimuli at the first step. Similar to the front vowel series, the individual results from the Quichua group to the back vowel stimuli also suggest that the level of Spanish and/or age were responsible for the greater number of responses to /u/ in the older Quichua cohort.

To delve further into the experimental results, a generalised mixed effects model was created in R x64 3.2.1 with the *glmer* function of the *lme4* package (Bates, 2012). This model considered all listener responses from all four continua. Ninety-five percent confidence intervals (CI_{95}) were computed using the *confint* function from the *lmerTest* package (Kuznetsova, Brockhoff, & Bojesen, 2014). The model included *listener* as a random effect and *response* (levels: *mid vowel*/ *high vowel*) as the dependent variable. The following predictors (fixed effects) were considered for each model: *continuum* (levels: steps 1–10); *language group* (levels: *Spanish*, *Quichua*, and *Media Lengua*); *age* (normalised ordinal variable[10] with 31 levels); *education* (levels: *low*, which included none and primary; *high*, which included secondary and university); *level of Spanish* (Levels: *mid, high,* and *native*); and *place of articulation* (levels: *front* and *back*). If a single factor in a categorical variable (with 3 + levels) was shown to be non-significant during the model building phase, the individual factors were converted to binary variables (levels: *true* and *false*) to amalgamate the non-significant factor with the intercept. This was the case for the *language* predictor as Quichua was shown to be non-significant in the categorical *language* variable. In this case the new binary variables, *Spanish* and *Quichua* were created and tested (*Spanish*, with *true* and *false* factors; *Quichua* with *true* and *false* factors). During the model building phase, interactions between predictors were also tested and non-significant predictors were removed from the model one-by-one based on the closest z-value to zero, until only significant predictors remained.

The model reported contains no non-significant predictors. For the significant results, the coefficient estimate (β), a conservative estimate of the average difference in log-odds response (a measure of probability) between the predictors in question, is of particular interest. For example, a negative log-odd result for *continuum* (levels: steps 1–10) means the likelihood of a participant choosing mid vowel stimuli decreases *x* amount per step while a positive log-result for a given *language* (levels: *Spanish, Quichua* and *Media Lengua*) simply means a mid vowel was identified significantly more often by one group than another at step 1. Because more mid vowel responses are predicted on the left of the continua than on the right, the *continuum* result should be negative. To avoid overfitting the model due to possible correlated predictors (e.g., *age*, *education*, and *level of Spanish*) two diagnostics were considered (1.) models were compared using the Bayesian Information Criteria (BIC) score (which also correlated with the Akaike Information Criteria (AIC) score) and the model with the lowest BIC where all the predictors were significant was chosen and (2.) the *testUniformity* function from

[10] Age was normalised to avoid unequal gaps in the age range with the following equation: $z_i = \frac{x_i - \min(x)}{\max(x) - \min(x)}$

Table 4
Generalised linear mixed effects model results.

	Coef. estimate β	Std. error	2.5%	97.5%	z-value	Pr(>|z|)	
(Intercept)	2.15	0.53	1.12	3.19	4.076	4.58E−05	***
Continuum	−0.56	0.04	−0.64	−0.48	−14.276	<2e−16	***
Spanish: True	2.84	0.53	1.8	3.87	5.392	6.97E−08	***
Quichua: True *ns*	0.075	0.88	−1.65	1.8	0.085	0.9325	
Age (31 factors)	−0.014	0.016	−0.045	0.017	−0.886	3.76E−01	
Continuum * Spanish: True	−0.41	0.064	−0.53	−0.28	−6.422	1.35E−10	***
Continuum * Quichua: True	0.14	0.056	0.032	0.25	2.526	0.012	*
Quichua: True * Age (31 Factors)	−0.07	0.036	−0.14	−0.001	−1.99	0.0461	*

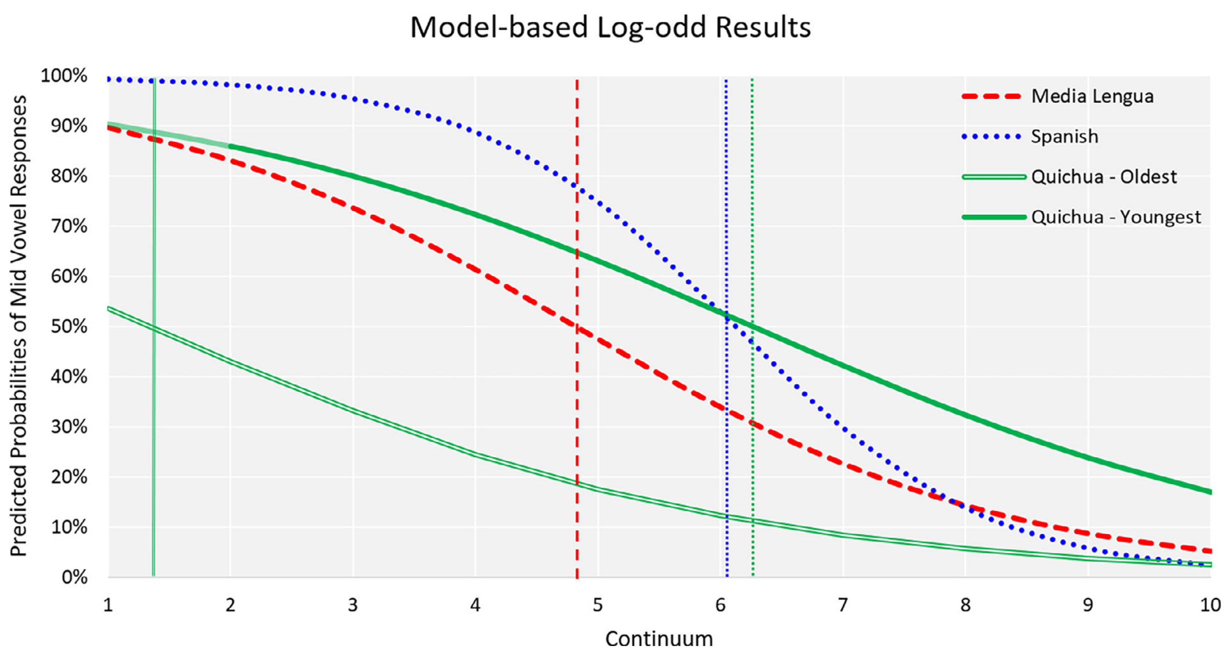

Fig. 8. Log-odd results of the predicted probabilities of a listener choosing a mid vowel response for each language along the 10-step continuum. Perceptual boundaries are indicated with the vertical lines.

the DHARMa R-package (Hartig, 2017), which uses a “simulation-based approach to create readily interpretable scaled (quantile) residuals for fitted generalized linear mixed models”, was used to assure model ‘goodness’ (p. 1). The model results are presented in Table 4.

The intercept, with a ‘base’ value of 2.15 log-odds, suggests that Media Lengua participants selected the image containing the mid vowel over the high vowel 90%[11] of the time at step 1. The probability of selecting a mid vowel (/o, e/) decreased, on average, by 0.56 log-odds per additional step along the continuum. For both the Spanish and Quichua groups, there were significant interactions with *continuum*. This result suggests that Spanish participants had a steeper slope (0.97 log-odd decrease per step (*continuum* + *continuum & Spanish*)) and the Quichua participants had a shallower slope (0.42 log-odd decrease per step (*continuum* + *continuum & Quichua*)) compared to Media Lengua participants (0.56 log-odd decrease (*continuum*)). The significant *Spanish* result correlates to a greater number of mid vowel responses at step 1 for the Spanish group compared to the Media Lengua group (*Spanish intercept* = 4.99 (99%) (*intercept* + *Spanish*)). For the Quichua group, there was a significant interaction with *age*, resulting in a gradient intercept with the oldest Quichua participant identifying mid vowel tokens just 54% of the time at step 1 (*intercept for the oldest Quichua participant* = 0.14 (*intercept* + *Quichua* + *age* + (Quichua & age *31))) and with the youngest participant identifying mid vowels 90% of the time at step 1 (*Quichua intercept* = 2.2 (*intercept* + *Quichua*)). This age-based variation in the Quichua group is not directly available in the raw data plotted in Figs. 4 and 5, though it appears to be responsible for the greater random responses and larger error bar for this group at step 1 compared to the Spanish and Media Lengua groups.

Fig. 8 details the model results for each language along the continua; Media Lengua (red, dashed), Quichua (green, solid (youngest) /double solid (oldest)), and Spanish (blue, dotted). The predicted perceptual boundary (were the curve crosses the 50% mark) for each language is marked with a vertical line.

According to the statistical model presented in Table 4 and illustrated in Fig. 8, the Spanish and Media Lengua groups appear to have more stable response patterns compared to the Quichua group, which varied based on the age of the participant. The categorical boundaries in Fig. 8 reveal that the older a Quichua listener is, the more vowels they identified

[11] Percentages based on log-odds are converted using the following equation: $p = \frac{\exp(LO)}{1+\exp(LO)}$

as high while, overall, younger Quichua and Spanish listeners identified a roughly similar number of stimuli as high and mid vowels with a preference toward the latter. Media Lengua participants show a more equal split between mid and high vowel identification. Based on the statistical model, the Media Lengua perceptual boundary (50% crossing point) appears before the youngest Quichua and Spanish boundaries. While mid and high vowels were identified as different in all three languages (save older Quichua participants), both Media Lengua and the younger Quichua listeners demonstrated more varied responses to the stimuli at step 1 compared to Spanish listeners.

4. Discussion

The first research question asked whether Media Lengua and/ or Quichua listeners can consistently identify differences between mid and high vowels within the formant ranges for Media Lengua (and Quichua by proxy) outlined in Stewart (2014). In Section 4.1, it is argued that evidence from Section 3 supports the presence of mid and high vowels in all three languages with varying degrees of consistency. Section 4.2 returns to the second research goal and provides some possible theoretical implications for models of L2/non-native speech perception, and in Section 4.3 remaining questions are considered including limitations of this study.

4.1. Mid and high vowel perception

This study sought to determine whether differences between mid and high vowels can be identified in minimal pair words by listeners of Imbabura Media Lengua and Imbabura Quichua. In order to investigate the phonological claims presented in Section 1.4, a 2AFC identification task experiment was used. It was hypothesised that Media Lengua listeners would be able to identify such differences due to the physical distances in F1 frequencies between Media Lengua mid and high vowels and because the vowel contrast could be beneficial for managing the Spanish origin vocabulary. For Quichua, it was hypothesised that listeners would produce less consistent results due to the greater overlap between mid and high vowels in acoustic spaces compared to in Media Lengua and because native Quichua phonology does not make use of mid and high vowels for contrastive purposes. As predicted, the Media Lengua and Spanish results revealed that most listeners from both language backgrounds were able to identify consistent differences in the stimuli. However, for Quichua, younger listeners were able to identify consistent differences as well. This was not the case for the older Quichua listeners, who overwhelmingly identified the stimuli as high vowels or showed random response patterns.

Tentatively, these results, in conjunction with those from Stewart (2014), suggest that Media Lengua participants contrast mid and high vowels both in production and perception. However, the results for Quichua are not as clear due to the discrepancies between older and younger listeners. In general, older Quichua and older Media Lengua speakers form part of the last generation to have a relatively late age of Spanish acquisition and to have lesser proficiency in the language compared to those of the younger generations. Because of this, the older Quichua participants provided an ideal point of comparison with the older Media Lengua participants regarding mid and high vowel perception. The fact that older Media Lengua listeners showed similar response patterns to those of the younger listeners suggests that the mid and high contrast could be an integral part of Media Lengua phonology. Contrarily, the older Quichua group, by in large, identified the majority of the stimuli as high vowels or showed random responses to the stimuli.

One interpretation of this finding could be that Media Lengua speakers benefit by maintaining the historically Spanish mid and high vowel contrast for reasons yet to be investigated (e.g., functional load, phonological optimisation, etc.). On the other hand, the production contrast between mid and high vowels in Quichua, revealed by Stewart (2014), appears to act more like a near-merger since there is little evidence that older Quichua listeners identify consistent differences between mid and high vowels in this study. This might suggest that the mid and high vowel contrast is not as important to Quichua phonology since the number of Spanish borrowings is substantially smaller than in Media Lengua (∼45% vs. ∼90% as described in Stewart, 2011). Additionally, many Spanish borrowings in Quichua took place long before Media Lengua developed, which may have resulted in the mid and high vowel contrasts disappearing overtime. Results from the Spanish listeners revealed clear differences between mid and high vowels even within the shorter range between prototypical Media Lengua mid and high vowel formants. This is most likely the result of the stimuli at steps 1 and 10 being situated within normal Spanish mid and high vowel categories.

4.2. Possible theoretical implications

Media Lengua and Quichua vowel spaces provide an interesting testing ground for the concept of vowel category. Findings from Guion (2003) revealed that Quichua dominant bilinguals nearly merge mid and high vowels when speaking Spanish. This was also the case in Stewart's (2014) analysis of Spanish borrowings and native Quichua words in Quichua and Media Lengua speech. Additionally, Lipski (2015) revealed that Quichua dominant bilinguals also make use of a broad acoustic space with no systematic separation between mid and high vowels. Therefore, if Media Lengua and younger Quichua listeners contrast mid and high vowels productively and perceptually, why have the categories not dispersed as predicted by models of adaptive dispersion? Instead of well-defined categories, Media Lengua and Quichua appear to have a large single gradient category with high concentrations of both mid and high vowels near the centre and more consistent high and mid vowels in the peripheries (as illustrated in Fig. 1)—the opposite from what would be expected in a system with two separate categories for mid and high vowels.

The lack of dispersion may be due to the fact that Media Lengua was created by late bilinguals who assimilated borrowed words based on Quichua phonology. In such broad *intermixed*[12] vowel categories, listeners may typically ignore the mid and high vowel distinctions unless a vowel is produced

[12] I use this term to refer to two overlapping vowel categories arranged in such a way that their average F1/F2 production values are nearly-identical in the centre while vowels on the extreme ends of the categories are less mixed (see Fig. 1).

atypically high or low (i.e., outside a specific standard deviation or concentration). This atypicality could cue the listener that a difference is being expressed for contrastive purposes. Speakers may also benefit from an intermixed category as it might allow for greater articulatory flexibility during production (minimising articulatory effort in the dispersion theory literature e.g., see Flemming, 1995), though with the trade-off of reduced articulatory precision—a reason as for why the categories do not disperse as would be expected in models of adaptive dispersion. Evidence for this flexibility vs. precision trade off can be observed in the more random response patterns to the mid vowel stimuli at step 1 by Media Lengua and Quichua listeners compared to Spanish listeners. Additionally, in natural speech, differentiating between mid and high vowels in Media Lengua simply may not be a high priority as pragmatic context and familiarity with individual interlocutors' speech patterns may provide sufficient cues to accurately process the majority of ambiguous input.

The second research goal of this study was to consider how theoretical models of L2/non-native speech perception (e.g., Best's Perceptual Assimilation Model and Flege's Speech Learning Model) might fall short at predicting listener perception of mid and high vowels in Quichua and/or Media Lengua. It was hypothesised that if listeners could indeed identify consistent differences in the stimuli, such theoretical models would most often fail to predict separate categories due to mid and high vowel articulatory similarities. Instead predictions from such models would suggest listeners would merge mid and high vowels into a single category since the highest concentration of both vowels is in the centre of what appears to be a single intermixed vowel category. However, such an arrangement does not mean listeners are unable to consistently identify differences since periphery vowels have a higher probability of being identified as different (as revealed in this study). Therefore, such models might falsely predict a merger of mid and high vowel categories in an overlapping acoustic space when in fact speakers are able to produce and interpret vowels outside the centre of the category for contrastive purposes. The older generations of Media Lengua and Quichua speakers, with late acquisition of Spanish, are most likely to have these types of intermixed gradient vowel categories, with no systematic division. However, the older Media Lengua cohort were able to consistently identify differences between mid and high vowels while the older Quichua cohort showed more random responses. This may be because the Media Lengua vowel categories are slightly less merged in the Media Lengua group (see Stewart, 2014)—just beyond the threshold of 0.3 Bark for formant discrimination suggested by Kewley-Port (2001). For the younger generation of Quichua listeners, the predominant use of Spanish and younger age of acquisition may have allowed them to establish more Spanish-like mid and high vowel categories allowing for more consistent identification of the differences in the stimuli.

4.3. Limitations

There were two primary limitations to this study. Firstly, it should be noted that due to Media Lengua's position as a moribund language in the community of Pijal, the participant recruitment method, through which data were gathered, might not have been controlled as desired (e.g., variability in bilingualism and level of education). However, the passive bilingual results (participants #60 and #61 in Figs. 6 & 7) show the same overall pattern as other younger native Media Lengua listeners (#43, #59, #58, & #55). Future investigators may want to run this experiment in the San Pablo communities where a more homogenous group of listeners might be recruited.

Finally, sizable Media Lengua and Quichua corpora need to be created to identify any influences from word frequency that might have affected word choice. Such corpora are also needed to analyse functional load as a possible driving force for adopting Spanish origin phonemic contrasts through lexical borrowings.

5. Conclusion

Overall, Media Lengua, young Quichua, and Spanish listeners are able to identify differences in mid and high vowel stimuli presented at random in a 2AFC identification task experiment. Contrarily, older Quichua listeners overwhelmingly identified the stimuli as high vowels. The large single overlapping configuration of mid and high vowels in acoustic space in Quichua and Media Lengua (Stewart, 2014) and in L2 Spanish (Lipski, 2015) provides a new test case for models of L2/non-native speech perception. In more typical vowel categories, the centre is the most dense and homogenous portion of the distribution and vowels produced in this region allow for optimal contrastability with centre vowels produced in other categories. However, in Quichua and Media Lengua, the dense centre of these vowel spaces are the most heterogeneous, thus affecting potential contrastability. Conversely, vowels along the periphery are less 'mixed', allowing for greater contrastability if the system allows for it, despite these vowels being less archetypical. These atypical overlapping/intermixed configurations call into question how models of L2/non-native speech perception might predict the formation of a category in these languages.

Acknowledgements

This research was funded by the University of Manitoba Graduate Fellowship (UMGF). I would like to thank Isabel Deibel, Christiani Thompson, Nicole Rosen, Kevin Russell, and the anonymous reviewers for their insightful comments and suggestions on the drafts of this paper. Finally, I would like to thank Gabriela Prado, Cecilia Ayala, Antonio Maldonado, Lucía Gonza, Elvis Tuquerres, and the participants for taking part in this investigation. Any remaining errors are my responsibility.

Appendix A:. Self-assessment questions

English translations

1) How old are you?
2) Where did you grow up?
3) Have you lived outside the community? If so, for how long?
4) When and where did you learn Spanish? Quichua? Media Lengua?
5) What languages did your parents speak to you growing up?
6) Which language do you feel most comfortable using?
7) Do you understand Quichua/Media Lengua but can't speak it?
8) How many years of school did you finish?

a. *Note: Spanish was the only language of instruction when the consultants were going to school.*
9) What do you do for a living? And how long have you had this profession?
b. *Virtually any job outside of the community uses Spanish as the language of communication.*
10) Which language do you use when you gather with your friends? family? co-workers?
11) When you think to yourself, what language do you think in?
12) Do people compliment you on your Spanish? Quichua?

Spanish

1) *¹Cuántos años tiene?*
2) *¹Dónde creció?*
3) *¹Ha vivido fuera de la comunidad? Si lo ha hecho, ¹por cuánto tiempo?*
4) *¹Dónde y cuándo aprendió español? ¹kichwa? ¹media lengua?*
5) *¹Qué idiomas le hablaban sus padres cuando crecía?*
6) *¹Qué idioma se siente más confortable al usarlo?*
7) *¹Usted entiende kichwa/media lengua, pero no puede hablarlo?*
8) *¹Cuántos años de escuela terminó?*
9) *¹Cuál es su ocupación y cuánto tiempo viene trabajando en esto?*
10) *¹Qué idioma habla cuando se reúne con sus amigos, familia y compañeros de trabajo?*
11) *¹Cuándo estás pensando, en qué idioma lo haces?*
12) *¹La gente lo felicita cuando habla español? kichwa?*

Appendix B

Tables B1–B3.

Table B1
This table provides information on the Media Lengua group including, age at the time of the experiment, gender, level of formal education, level of Spanish, Media Lengua, and Quichua, frequency of Media Lengua usage, and place of residency.

Media Lengua listener Code	Age	Gender	Formal education	Spanish level	Media Lengua/Quichua level	Media Lengua usage	Place of residence
041	62	M	Primary	High	Native	Intermittently	Pijal Bajo
043	42	F	Secondary	High	Native	Daily	Pijal Bajo
055	46	F	Primary	Mid	Native	Intermittently	Pijal Bajo
056	64	F	None	Mid	Native	Intermittently	Pijal Bajo
057	63	F	None	Mid	Native	Intermittently	Pijal Bajo
058	33	M	University	High	Native	Infrequently	Pijal Bajo
059	38	M	University	High	Native	Infrequently	Pijal Bajo
060	24	F	Secondary	Native	Passive	Rarely	Pijal Bajo
061	24	F	University	Native	Passive	Rarely	Pijal Bajo
062	54	M	Primary	High	Native	Intermittently	Pijal Bajo
063	50	F	None	Mid	Native	Daily	Pijal Bajo
Average:	45						

Table B2
This table provides information on the Quichua and L2 Spanish speaking groups including: age at the time of recording, gender, level of formal education, level of Spanish and Quichua, frequency of Quichua usage, and place of residency.

Quichua listener Code	Age	Gender	Formal education	Spanish level	Quichua Lengua level	Quichua usage	Place of residence
063	68	M	Primary	Mid	Native	Daily	Chirihuasi
078	30	M	University	High	Native	Daily	Chirihuasi
079	32	F	Primary	High	Native	Daily	Chirihuasi
080	43	F	NA	Mid	Native	Daily	Chirihuasi
081	53	M	Primary	Mid	Native	Daily	Chirihuasi
082	34	M	Secondary	High	Native	Daily	Chirihuasi
083	48	F	NA	Mid	Native	Daily	Chirihuasi
084	26	F	University	Native	Native	Daily	Chirihuasi
085	38	F	Primary	High	Native	Daily	Chirihuasi
086	48	F	NA	Mid	Native	Daily	Chirihuasi
Average	42						

Table B3
This table provides demographic information on the L1 Spanish speaking group from Quito including: age at the time of recording, gender, level of formal education, profession, and place of birth/residency.

Spanish listener Code	Age	Gender	Formal education	Profession	Place of birth/residence
97	22	F	University	Student	Quito
112	51	F	Teaching cert.	Teacher	Quito
114	55	M	Medical degree	Orthodontist	Tulcán/Quito
115	58	M	University	Teacher	Quito
116	41	F	University	Teacher	Quito

(*continued on next page*)

Table B3 (*continued*)

Spanish listener Code	Age	Gender	Formal education	Profession	Place of birth/residence
117	47	M	Teaching cert.	Teacher	Quito
118	40	M	University	Teacher	Quito
119	28	F	University	Teacher	Quito
120	58	M	University	Teacher	Quito
121	57	F	Teaching cert.	Teacher	Quito
122	41	F	Teaching cert.	Teacher	Quito
123	45	F	University	Teacher	Quito
124	31	F	University	Teacher	Quito
125	40	F	Teaching cert.	Teacher	Quito
126	65	M	Teaching cert.	Teacher	Quito
127	31	F	University	Teacher	Quito
129	38	F	University	Teacher	Quito
132	37	M	Teaching cert.	Teacher	Quito
Average	44				

Appendix C. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.wocn.2018.08.005.

References

Adelaar, W., & Muysken, P. (2004). *The languages of the Andes*. Cambridge UK: Cambridge University Press.

Amengual, M. (2014). The perception and production of language-specific mid-vowel contrasts: Shifting the focus to the bilingual individual in early language input conditions. *International Journal of Bilingualism, 20*(2), 1–20. https://doi.org/10.1177/1367006914544988.

Bakker, P. (1997). *A Language of Our Own : The Genesis of Michif, the Mixed Cree-French Language of the Canadian Metis* (Vol. 10) New York: Oxford University Press.

Bates, M. (2012). lme4: Mixed-effects modeling with R. Retrieved from http://lme4.r-forge.r-project.org/book/.

Best, C. (1993). Emergence of language-specific constraints in perception of nonnative speech: A window on early phonological development. In B. de Boysson-Bardies, S. de Schonen, P. W. Jusczyk, P. McNeilage, & J. Morton (Eds.), *Developmental neurocognition: Speech and face processing in the first year of life* (pp. 289–304). Dordrecht: Kluwer Academic.

Best, C. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 171–204). Timonium, MD: York Press.

Best, C., Hallé, P., Bohn, O.-S., & Faber, A. (2003). Cross-language perception of non-native vowels: Phonological and phonetic effects of listeners' native languages. In M. Solé, D. Recasens, & J. Romero (Eds.), *Presented at the proceedings of the 15th international congress of phonetic sciences* (pp. 2889–2892). Causal Productions.

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

Broadbent, D. E. (1967). Word-frequency effect and response bias. *Psychological Review, 74*(1), 1–15. https://doi.org/10.1037/h0024206.

Chládková, K., Escudero, P., & Boersma, P. (2011). Context-specific acoustic differences between Peruvian and Iberian Spanish vowels. *The Journal of the Acoustical Society of America, 130*(1), 416–428. https://doi.org/10.1121/1.3592242.

Cole, P. (1982). *Imbabura Quichua*. Amsterdam: North-Holland Publishing Company.

CREA, R. A. E. B. (2018). Corpus de referencia del español actual. Retrieved from: http://www.rae.es.

Deibel, I. (accepted - 2017). Adpositions in Media Lengua: Quichua or Spanish? – Evidence of a lexical-functional split. Journal of Language Contact.

Flege, J. (1995). Second-language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues* (pp. 233–272). Timonium, MD: York Press.

Flege, J., Schirru, C. B., & MacKay, I. R. A. (2003). Interaction between the native and second language phonetic subsystems. *Speech Communication, 40*, 467–491.

Flemming, E. (1995). *Auditory Representations in Phonology* (PhD Dissertation). Los Angeles: UCLA.

Goldiamond, I., & Hawkins, W. F. (1958). Vexierversuch: The log relationship between word-frequency and recognition obtained in the absence of stimulus words. *Journal of Experimental Psychology, 56*, 457–463.

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

Gómez-Rendón, J. (2007). Grammatical borrowings in Imbabura Quichua. In Y. Matras (Ed.). *Grammatical borrowing in cross-linguistic perspective* (Vol. 1, pp. 481–521). New York: Walter de Gruyter.

Gómez-Rendón, J. (2008). Spanish lexical borrowing in Imbabura Quichua: In search of constraints on language contact. In T. Stolz, D. Bakker, & R. Salas Palomo (Eds.), *Hispanisation: The impact of Spanish on the lexicon and grammar of the indigenous languages of Austronesia and the Americas* (pp. 95–120). Berlin: Mouton de Gruyter.

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.

Guirao, M., & García Jurado, M. A. (1990). Frequency of occurence of phonemes in American Spanish. *Revue Québécoise de Linguistique, 19*(2), 135–149.

Hartig, F. (2017). DHARMa: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models (Version R Package Version 0.1.5). Retrieved from http://florianhartig.github.io/DHARMa/.

Jarrín Paredes, G. (2014). *Estereotipos Lingüísticos en Relación al Kichwa y a la Media Lengua en las Comunidades de Angla, Casco Valenzuela, El Topo y Ucsha de la Parroquia San Pablo del Lago*. Quito: Pontificia Universidad Católica Del Ecuador.

Johnson, K. (2000). Adaptive disperion in vowel perception. *Phonetica, 57*, 181–188.

Kewley-Port, D. (2001). Vowel formant discriminaton II: Effects of stimulus uncertainty, consonantal context, and training. *Journal of the Acoustical Society of America, 85*, 1726–1740.

Kuznetsova, A., Brockhoff, P. B., & Bojesen, R. (2014). lmerTest (Version 2.0-11) [R library]: Cran.r Project. Retrieved from http://cran.r-project.org/web/packages/lmerTest/index.html.

Labov, William (1994). *Principles of linguistic change* (Vol. 1). MA: Blackwell: Oxford & Cambridge.

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

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

Lehiste, I., & Peterson, G. (1959). Vowel amplitude and phonemic stress in American English. *Journal of Phonetics, 31*(4), 428–435.

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 J. Ohala & J. Jaeger (Eds.), *Experimental phonology* (pp. 13–44). Orlando, Florida: 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). Dordrecht: Springer Science & Business Media.

Lipski, J. (2015). Colliding vowel systems in Andean Spanish. *Linguistic approaches to bilingualism, 5*(1), 91–121. https://doi.org/10.105/lab.5.1.04lip.

Livijn, P. (2000). Acoustic distribution of vowels in differently sized inventories – hot spots or adaptive dispersion? *PERILUS, 23*.

Meakins, F. (2013). Mixed languages. In Y. Matras & P. Bakker (Eds.), *Contact languages: A comprehensive guide* (pp. 159–228). Berlin: Mouton de Gruyter.

Meakins, F., & Stewart, J. (2019). Mixed Languages. In S. Mufwene & A. M. Escobar (Eds.), *Cambridge Handbook of language contact*. Cambridge: Cambridge University Press. to appear.

Mitleb, F. M. (1981). *Segmental and non-segmental structure in phonetics: Evidence from foreign accent* (Ph.D. Dissertation). Bloomington: Indiana University.

Müller, A. (2011). *La media lengua en comunidades semi-rurales del Ecuador: Uso y significado social de una lengua mixta bilingüe* (Ph.D. Dissertation). Zürich: Universität Zürich.

Muysken, P. (1980). Sources for the study of Amerindian contact vernaculars in Ecuador. *Amsterdam Creole Studies, 3*, 66–82.

Muysken, P. (1981). Halfway between Quechua and Spanish: The case for relexification. In A. R. Highfield (Ed.). *Historicity and variation in Creole studies* (57–78). Ann Arbor: Karoma Publishers.

Muysken, P. (1997). Media Lengua. In S. G. Thomason (Ed.), *Contact languages: A wider perspective* (pp. 365–426). Amsterdam: J. Benjamins Pub. Co.

Navarra, J., Sebastian Galles, N., & Soto-Faraco, S. (2005). The perception of second language sounds in early bilinguals: New evidence from an implicit measure. *Journal of Experimental Psychology: Human Perception and Performance, 31*, 912–918.

Pallier, C., Bosch, L., & Sebastian Galles, N. (1997). A limit on behavioral plasticity in speech perception. *Cognition, 64*, B9–B17. https://doi.org/10.1016/S0010-0277(97)00030-9.

Pallier, C., Colome, A., & Sebastian Galles, N. (2001). The influence of native-language phonology on lexical access: Exemplar-based versus abstract lexical entries. *Psychological Science, 12*, 445–449.

Port, R., & Crawford, P. (1989). Pragmatic effects on neutralization rules. *Journal of Phonetics, 16*(4), 257–282.

Port, R., & O'Dell, M. (1985). Neutralization of syllable-final voicing in German. *Journal of Phonetics, 13*(4), 455–471.

Hickey, R. (2004). Mergers, near-mergers and phonological interpretation. In *New perspectives on english historical linguistics* (pp. 125–137). Amsterdam: John Benjamins Co.

Sort_rand. Retrieved from http://www.pptfaq.com/FAQ00429_Randomize_the_order_of_a_PowerPoint_presentation.htm.

Roettger, T. B., Winter, B., Grawunder, S., Kirby, J., & Grice, M. (2014). Assessing incomplete neutralization of final devoicing in German. *Journal of Phonetics, 43*, 11–25. https://doi.org/10.1016/j.wocn.2014.01.002.

Shappeck, M. (2011). *Quichua-Spanish language contact in Salcedo, Ecuador: Revisiting Media Lengua syncretic language practices* (Ph.D. dissertation). 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). MSpace: University of Manitoba.

Stewart, J. (2014). A comparative analysis of Media Lengua and Quichua vowel production. *Phonetica, 71*(3), 159–182. https://doi.org/10.1159/000369629.

Traunmüller, H. (1990). Analytical expressions for the tonotopic sensory scale. *Journal of the Acoustical Society of America, 88*(1), 97–100.

Vainio, M., Järvikivi, J., Werner, S., Volk, N., & Välikangas, J. (2002). Effect of prosodic naturalness on segmental acceptability in synthetic speech. In Paper presented at the Proceedings of 2002 IEEE workshop on speech synthesis, Santa Monica, CA.

van Gijn, R. (2009). The phonology of mixed languages. *Journal of Pidgin and Creole Languages, 24*, 91–117. https://doi.org/10.1075/jpcl.24.1.04gij.

Winter, B., & Roettger, T. B. (2011). The nature of incomplete neutralization in German: Implications for laboratory phonology. *Grazer Linguistische Studien, 76*, 55–74.