VOWEL RAISING IN COTOPAXI QUICHUAN LANGUAGES

Sky Onosson[1] and Jesse Stewart[2]

[1]University of Victoria, [2]University of Saskatchewan
Sky@Onosson.com , Stewart.Jesse@usask.ca

ABSTRACT

This study investigates a phenomenon of morpheme-specific vowel raising that occurs in some Ecuadorian dialects of Quichua as well as the associated mixed language Media Lengua, where certain forms of /a/ are pronounced as [u]. Our data reveal that this process is region-specific, occurring in the Quichua dialect spoken in Cotopaxi province, but not in the relatively more well-known dialect of Imbabura province. Moreover, the dialects of Media Lengua in these respective regions also exhibit different realizations of the forms involved in this raising process which have both similarities to and differences from the patterns seen in their local Quichua dialects. These results may provide some insight into the relationships between these and other cross-regional dialects of Quichua, and how these may relate to similar dialects of Media Lengua, both contemporaneously and historically.

Keywords: Quichua, Media Lengua, vowels, acoustic phonetics, dialectal variation

1. INTRODUCTION

This paper offers a comparative analysis of vowels found in a set of morphemes in two dialects of Ecuadorian Quichua and two dialects of Media Lengua that are known to vary impressionistically based on the region where they are spoken. This analysis primarily focuses Cotopaxi Quichua (Glottocode: COTO1253) and Cotopaxi Media Lengua (Glottocode: MEDI1245). The former is a Quechuan language that forms part of the highland Quichua dialect continuum that extends throughout the Andean region of Ecuador. It is considered one of the central varieties and is spoken in Cotopaxi province by an estimated 50k speakers [1]. Like all Quechuan languages, Cotopaxi Quichua is agglutinating with SOV word order and contains three phonemic vowels (/i, u, a/). Cotopaxi Media Lengua is an endangered mixed language [2] which is spoken by an estimated 1703 people [3]. Media Lengua is essentially Quichua grammar, semantics, and phonology overlaid with Spanish vocabulary through the process of relexification [4].

To date, there are no phonetic studies of Cotopaxi Media Lengua and the only study that analyses phonetic data in Cotopaxi Quichua is a 2010 phonological description by Kohlberger ([5]). Included in his study is a brief overview of the vowel system which includes an F1xF2 plot with an estimated 50 vowels from six speakers (3 males & 3 females). This plot shows clear separation among /i/, /u/, and /a/ [5, p. 44]. Kohlberger's description also mentions that /a/ behaves differently in several morphemes when compared with other varieties in Ecuadorian Quichua. Specifically, the first /a/ in the suffixes, *-man* /man/ 'directional', *-manta* /manta/ 'ablative', and *-pac* /pak/ 'genitive' can be raised to [u] in spontaneous speech resulting in [muŋ], [munda], and [pux ~ puk], respectively while remaining as [a] in careful speech [5, p. 43]. Henceforth we refer to /a/ in these morphemes as /A/ to distinguish it from canonical /a/. /A/-raising appears to be a characteristic unique to central Quichua varieties (Chimborazo, Tungurahua, & Cotopaxi) as /A/ in Imbabura Quichua (north) and Loja Quichua (south) is documented as remaining canonically low ([a]) ([6, p. 18]). Additionally, multiple phonetic studies of Imbabura Quichua make no mention of /A/-raising in the previously mentioned suffixes ([7]–[10]).

The primary goal of this study is to revisit Kohlberger's analysis of the Cotopaxi vowel system and specifically focus on the morphemes involved in /A/-raising. In addition to *-man, -manta*, and *-pac*, we also investigate the verbal morpheme *-ngapac* [ngabu̯] 'same subject purpose marker' and the evidential *-mari* [mu̯ri] 'affirmative', both of which also give the impression of /A/-raising in our data.

The use of data from both Cotopaxi Quichua and Cotopaxi Media Lengua will further reveal whether the mixed language's vowel system patterns with those of Cotopaxi Quichua, makes use of its own innovations, or is more similar to other dialects of Media Lengua. To further the aims of this secondary goal and more fully understand the movement of /A/, this study compares Cotopaxi vowel data with previously analysed vowel data from both Imbabura Quichua and Imbabura Media Lengua (see [8], [9], [11]).

Coincidentally, both Cotopaxi and Imbabura provinces are home not only to different dialects of Quichua but also to different dialects of the mixed language, making them an ideal set of regions for comparison of the vowel raising process, among other

linguistic features. Example (1) provides a phrase from Cotopaxi Quichua (CQ) with parallel translations in Cotopaxi Media Lengua (CML), Imbabura Quichua (IQ), and Imbabura Media Lengua (IML) along with their respective glosses and Spanish and English translations. Bolded elements are of Spanish origin.

(1)

CQ	***Plaza**mun apangaboga na tarpanchic.*
Gloss	plaza-mu̯n apa-ngabu̯-ga na tarpa-ntʃix
CML	***Plaza**mun **llev**angaboga **no sembra**nchi.*
Gloss	plaza-mu̯n ʒeβa-ngabu̯-ga no sembra-ntʃi
IQ	***Plaza**man apangapaga na tarpanchic.*
Gloss	plaza-man apa-ngapa-ga na tarpa-ntʃik
IML	***Platsa**man **llev**angapaca **no sembra**nchi.*
Gloss	plaza-man ʒeβa-ngapa-ka no sembra-ntʃi
Sp:	Plaza-DIR llevar-PURP-TOP no sembrar-1P
En:	market-DIR bring-PURP-TOP no plant-1P
Trans	*No sembramos* [*comida*] *para llevar a la plaza.*
Trans	We don't plant [food] to bring to the market.

2. METHOD

2.1. Participants

Given the recent and ongoing documentation of Cotopaxi Media Lengua, for the first time in approximately 40 years [3], conversational data for this study was limited to two female speakers. Therefore, data from all languages and dialects in this study were also restricted to female speakers to avoid known differences in vowel production based on anatomical differences between males and females.

Cotopaxi Quichua data were gathered from three speakers in 2009 in the community of Quilotoa. A Spanish phrase list was used for elicitation. At the time of recording, the participants were 19, 28, and 38 years of age. For Cotopaxi Media Lengua conversational data comes from two speakers, aged 29 and 37 at the time of recording in 2022. Both were born and raised in the community of Yacubamba.

For comparative purposes, previously analysed data from Imbabura Quichua and Imbabura Media Lengua were also used [8], [11]. For Imbabura Quichua, data come from six female speakers from the community of Chirihuasi between the ages of 21 and 55 at the time of recording. For Imbabura Media Lengua, data come from five female speakers between the ages of 39 and 50 at the time of recording. Data from both Imbabura languages were recorded in 2011.

In terms of speech style or register, the Quichua data are primarily from elicitations and the Media Lengua data are primarily from conversations.

2.2. Data extraction and analysis

F1 and F2 frequencies were extracted from the centre points of vowels that were marked on a point tier in Praat [12] at their most steady state. A Praat script was used to automatically extract the data points. Table 1 provides a breakdown of the token counts by language and vowel. /A/ represents the vowel tokens extracted from the morphemes under analysis i.e., those potentially susceptible to the /a/-raising process, whereas the /a/, /o/, and /u/ columns reference vowels taken from full words and non-target morphemes. The libraries *ggplot2* [13] and *LmerTest* [14] were used to analyse the data in R version 4.2.2 [15].

	/A/	/a/	/o/	/u/	**Totals**
Cotopaxi Q	27	38		53	**118**
Cotopaxi ML	36	69	135	61	**301**
Imbabura Q	42	193		198	**433**
Imbabura ML	43	351	186	338	**918**
Totals	**148**	**651**	**321**	**650**	**1,770**

Table 1: Vowel token counts under analysis by language and region.

2.3. Procedures

Linear mixed-effects regression (LMER) models were run on a per-formant basis for F1 and F2, with *Vowel*, *Region* (Cotopaxi or Imbabura) and *Language* (Media Lengua or Quechua) as fixed effects including interactions between the fixed effects, and random intercepts for both *Speaker* and *Word*.

3. RESULTS

The LMER results for F1 and F2 are shown from the perspective of Cotopaxi Quichua /A/, which is set as the intercept in Table 2, and then from the perspective of its Imbabura Quichua counterpart which is set as the intercept in Table 3. To conserve space, Table 3 is simplified by removing rows involving *Region* as a single or interaction effect, as these are identical to the values in Table 2 except for the direction (sign) of the effect.

For Cotopaxi Quichua (Table 2), both formants of /A/ differ significantly from /a/ by a substantial amount, with a small (46 Hz) but significant difference from /u/ in terms of F1 only (there is also a non-significant difference from /o/, but this is irrelevant as the /o/ tokens are exclusively from Media Lengua). For Imbabura Quichua (Table 3), nearly the inverse is true; /A/ differs from /u/ in both formants, and from /a/ by a relatively large (308 Hz) difference in F2.

The effect of *Region* shows a significant difference for both formants. On a per-vowel basis, this shows up as a significant regional F1 difference

in /a/ (relative to /A/), and a regional difference in both formants of /o/ and /u/. *Language* turns out to have a relatively small role to play in the model, showing up in a two-way interaction with *Vowel* as a significant F1 difference between /A/ and /a/ in Cotopaxi (Table 2), and a corresponding F2 difference in Imbabura (Table 3). There is also a three-way interaction with *Region*, *Language* and *Vowel* involving F1.

	F1		F2	
Predictors	*Est.*	*p*	*Est.*	*p*
(Intercept)	507.78	**<0.001**	1306.06	**<0.001**
Vowel [a]	128.92	**<0.001**	439.33	**<0.001**
Vowel [o]	2.90	0.866	76.24	0.075
Vowel [u]	-45.73	0.038	-106.63	0.050
Region [Imbabura]	183.19	**<0.001**	209.04	**0.010**
Lang. [Media Lengua]	77.43	0.180	98.77	0.288
Vowel [a] * Region [Imbabura]	-120.85	**<0.001**	-131.76	0.061
Vowel [o] * Region [Imbabura]	-227.17	**<0.001**	-384.73	**<0.001**
Vowel [u] * Region [Imbabura]	-171.28	**<0.001**	-303.09	**<0.001**
Vowel [a] * Lang. [Media Lengua]	90.49	**0.003**	-67.92	0.360
Vowel [u] * Lang. [Media Lengua]	38.53	0.187	97.00	0.179
Region [Imbabura] * Lang. [Media Lengua]	-51.33	0.451	-22.05	0.846
(Vowel [a] * Region [Imbabura]) * Lang. [Media Lengua]	-112.76	**0.002**	-73.58	0.423
(Vowel [u] * Region [Imbabura]) * Lang. [Media Lengua]	-73.76	**0.043**	-44.18	0.625

Table 2: Linear mixed-effects regression results, intercept = Cotopaxi Quichua /A/.

	F1		F2	
Predictors	*Est.*	*p*	*Est.*	*p*
(Intercept)	690.96	**<0.001**	1515.10	**<0.001**
Vowel [a]	8.07	0.625	307.56	**<0.001**
Vowel [o]	-224.27	**<0.001**	-308.49	**<0.001**
Vowel [u]	-217.01	**<0.001**	-409.71	**<0.001**
Lang. [Media Lengua]	26.10	0.474	76.72	0.260
Vowel [a] * Lang. [Media Lengua]	-22.27	0.320	-141.51	**0.014**
Vowel [u] * Lang. [Media Lengua]	-35.22	0.117	52.82	0.358

Table 3: Linear mixed-effects regression results, intercept = Imbabura Quichua /A/; rows involving the effect of *Region* are excluded.

The regional differences in Quichua vowel production are visualized in Figure 1. In Imbabura (lower plot) while the F2 difference between /a/ and /A/ appears in the form of a slightly retracted distribution, the central regions of their respective distributions are nearly contiguous. In Cotopaxi (upper plot), contrastingly, there is virtually no overlap between /a/ and /A/. This is not the only difference, however, as /A/ also displays a clearly bimodal distribution, with one more retracted set of tokens which fully overlap the distribution of /u/, and another set which is more centralized, occurring in an intermediary position between /a/ and /u/. All /A/ realizations are substantially higher than /a/.

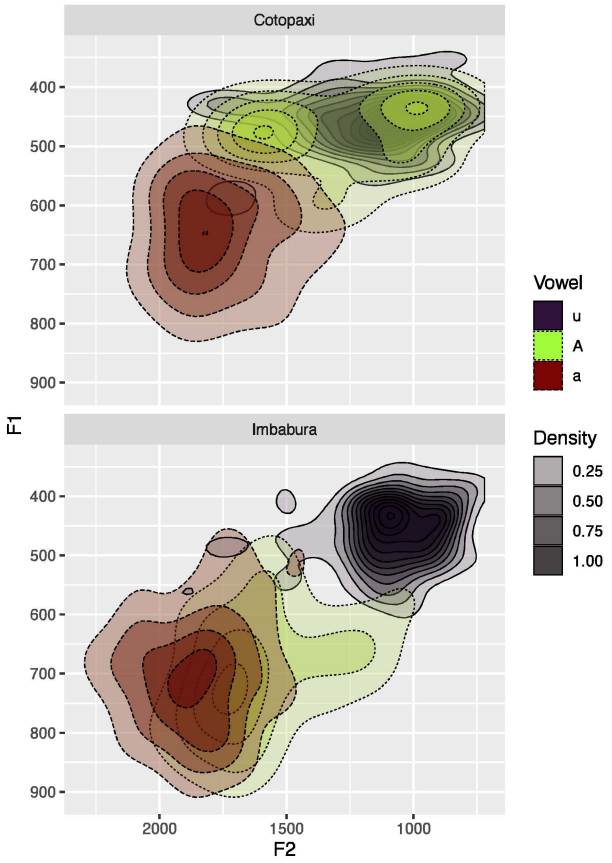

Figure 1: Quichua vowels by region.

For Media Lengua (Figure 2), the distributions of /A/ are similarly distinguished on a regional basis; however, the region-specific Media Lengua patterns do not precisely reflect those found for Quichua, and are moreover complicated by the presence of the additional back vowel /o/ (which itself overlaps substantially with /u/, a topic not focused on in this paper). Imbabura Media Lengua (lower plot) shows the most similarity with its regional Quichua variety: the distributions of /a/ and /A/ are largely overlapping, but the latter is slightly retracted. In Cotopaxi (upper plot), Media Lengua /A/, as in Quichua, appears to have a bimodal distribution, with one set of tokens largely overlapping /u/ (and, incidentally, /o/). The second set of /A/ tokens however, unlike Cotopaxi Quichua, are realized in a nearly identical manner as Imbabura's Media Lengua /A/, appearing as a slightly retracted /a/.

Note that the internal makeup of the distribution of Cotopaxi Media Lengua /A/ is not as readily discernible in Figure 2 compared with some of the

other plots. This is due to the combination of its bimodal pattern along with its relatively low token quantity, the fewest of any vowel token set (Table 1).

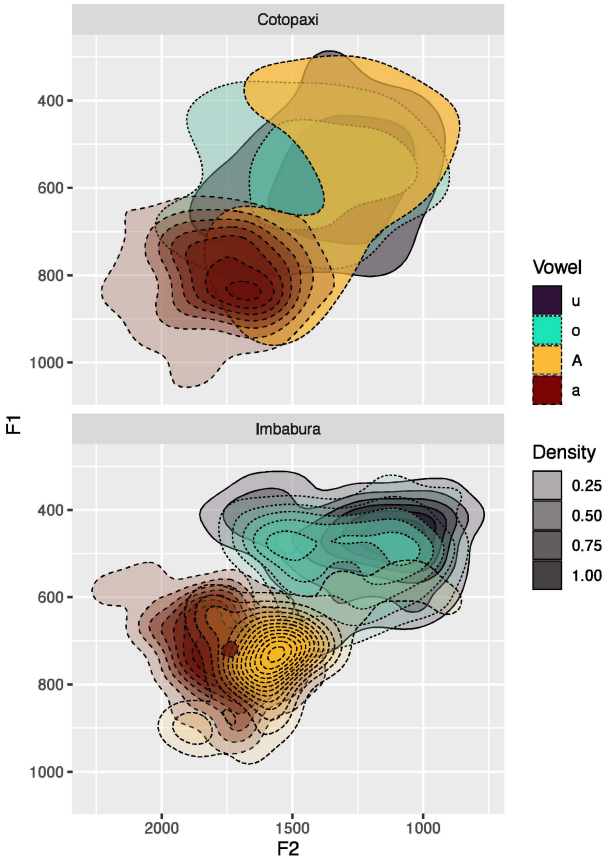

Figure 2: Media Lengua vowels by region.

4. DISCUSSION

This study supports Kohlberger's [5] observations of /A/-raising in Cotopaxi Quichua morphemes *-man* /man/ 'directional', *-manta* /manda/ 'ablative', and *-pac* /pak/ 'genitive', and further confirms that *-ngapac* 'same subject purpose marker' and the evidential *-mari* 'affirmative' also fit this paradigm. Kohlberger also observed some variation where /A/ remained [a] in careful speech but was raised to [u] in spontaneous speech. Although our Cotopaxi Quichua data were not set up to test differences in register, we found some variation in the production of /A/. While many of the /A/ tokens analyzed were both raised and retracted to the position of [u], a subset of /A/ tokens were only raised. These non-retracted vowels may give the perception of [a] as they are located directly above, and overlap in the front-back dimension with, the non-raising /a/ cluster, while remaining structurally distinct in that the average F1 frequency of /A/ remains far below that of /a/ (i.e., they remain raised). Based on this observation, an additional clustering analysis of each /A/-raising morpheme was conducted. While there was some variability between morphemes, the results were too inconsistent to suggest that any specific morpheme was responsible for the front-back variation. It may also be the case that the subset of non-retracted /A/s may be simply weakening to [ə] as this cluster is quite centralised in acoustic space. However, as our data are not currently set up to analyse other correlates of weakening (intensity, length, etc.), we must leave this question to future research.

Cotopaxi Media Lengua offers a somewhat different pattern. Like Cotopaxi Quichua, the majority of /A/s are raised. However, these raised realizations are generally not quite as retracted as their Quichua counterparts. A smaller cluster of /A/ tokens also occur in a position with both F1 and F2 averages close to that of canonical /a/. Once again, additional clustering analysis did not reveal that any specific morpheme was responsible for /A/ remaining unraised ([a]).

The analyses of both Cotopaxi languages, along with previous impressionistic observations ([6]), suggest that /A/-raising is a regional feature of central Quichuan. Viewing /A/-raising as an areal feature, it is plausible that register (careful vs. spontaneous) may not in fact be a robust predictor of /A/ production (low vs. high) as Kohlberger suggests, given that Cotopaxi Quichua data were elicited while Media Lengua data were conversational. Further investigation is warranted to validate this claim.

Comparing the Cotopaxi to Imbabura languages, it becomes clear that there is regional variation between the north (Imbabura) and central (Cotopaxi) varieties in terms of /A/ production. In both Imbabura Quichua and Imbabura Media Lengua, not a single token of /A/ is raised, and the production differences between /A/ and canonical /a/ are marginal. It is, however, noteworthy that both Imbabura Quichua and Imbabura Media Lengua /A/, are slightly retracted (but not raised) compared to canonical /a/ (307 Hz for Quichua & 166 Hz for Media Lengua). We do not expect that this entails an overall change in vowel quality for several reasons (1) /a/ and /A/ are still highly overlapping in both languages, (2) there is no other phonemic vowel occupying that region of acoustic space, and (3) both languages are known to occupy relatively large acoustic spaces making the precise articulation of low vowels inconsequential.

A final point of consideration is the similarity between the small subset of retracted-but-unraised Cotopaxi Media Lengua /A/ tokens, which positionally align with the distribution of Imbabura Media Lengua /A/. On this point, the cross-regional varieties of Media Lengua are in accordance, indicating that the areal differences are not wholesale, and that there is still some cross-regional influence between the Media Lengua dialects themselves.

5. REFERENCES

[1] Chisaguano, S. 2006. *La Población Indígena del Ecuador: Análisis de Estadísticas socio-demográficas*. Quito: Instituto Nacional de Estadísticas y Censos.

[2] Muysken, P. 1997. Media Lengua. In: Thomason, S. (ed) *Contact languages: A Wider Perspective*. J. Benjamins Pub. Co.

[3] Stewart, J., Gonza Inlago, L., Prado Ayala, G. Accepted. Cotopaxi Media Lengua is still very much alive. In: Language Documentation and Conservation.

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

[5] Kohlberger, M. 2010. *Cotopaxi Quichua: A phonological description and an analysis of stops and affricates in Central Highland Ecuadorian Quichua*. University of Edinburgh.

[6] Orr, C. 1978. *Dialectos Quichuas del Ecuador*. Instituto Lingüístico de Verano.

[7] 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.

[8] Stewart, J. 2014. A comparative analysis of Media Lengua and Quichua vowel production. *Phonetica* 7, 159–182.

[9] Onosson, S., Stewart, J. 2021. The Effects of Language Contact on Non-Native Vowel Sequences in Lexical Borrowings: The Case of Media Lengua. *Language and Speech PaPE Special Issue*, 1–30.

[10] Lipski, J. 2015. Colliding vowel systems in Andean Spanish. *Linguistic approaches to bilingualism*, vol. 5, 91–121.

[11] Onosson, S., Stewart, J. 2021. A multi-method approach to correlate identification in acoustic data: The case of Media Lengua. *Laboratory Phonology* 12(1), 1–30.

[12] Boersma, P., Weenink, D. 2016. *Praat*. Institute of Phonetic Sciences, University of Amsterdam.

[13] Wickham, H. 2016. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag.

[14] Kuznetsova, A., Brockhoff, P. B., Bojesen, R. lmerTest. 2016. CRAN.

[15] R Development Core Team. 2017. *R*. R Foundation for Statistical Computing.