The effects of language contact on non-native diphthongs in lexical borrowings: The case of Media Lengua

Sky Onosson¹ & Jesse Stewart²
sky@onosson.com stewart.jesse@usask.ca

University of Manitoba¹ University of Saskatchewan²

Society for Pidgin and Creole Languages
New Orleans
January 3rd, 2020
Mixed Language (Quichua-Spanish)

Quichua – phonological, morphological, and syntactic systems
Spanish – Lexicon

Si no aceiti-ta ocupa-kpi-ka huevo-ka sarten-pi-mi pega-shpa queda-n.
If not oil-ACC use-DS.COND-TOP egg-TOP pan-LOC-VAL stick-SS remain-3.PRES

‘If you don’t use oil, the eggs will stick to the pan.’
Mixed language Vowel Systems

Gurindji Kriol
(/æ/ & /e/)
(/u:/ & /o:/)

Show more overlap in Gurindji Kriol than in Kriol

Michif
(/ɛ/ & /ɔ/)

Merged French and Cree vowels into a Cree-like arrangement with the exception of French origin /ɛ, ɔ/

Media Lenga
(/e/ & /i/)
(/o/ & /u/)

Considerable overlap in Quichua/ Spanish high vowels with Spanish mid vowels. Substantial overlap in Spanish and Quichua high and low vowels

(Rosen, Stewart, & Sammons 2020)

(Jones, Meakins, and Buchan 2011)
**Media Lengua’s Vowel System**

**Production**

**Spanish**

- i
- u
- o
- a

**Quichua**

- i
- u
- a

---

**sarten-pimi**

**pan-LOC-VAL**

‘in the pan’

(Stewart, 2014)
Media Lengua’s Vowel System

Production

Spanish

i
e
a
u
o

Quichua

i
a
u

Media Lengua

i

e

u

θ

a

(Stewart, 2014)
Media Lengua’s Vowel System

Perception

2AFC Identification Task Experiment

Spanish

Quichua

F1 – mid point
F2 – mid point
F3 – mid point

(e) vs. (i) Perception

([o] vs. [u] Perception

Stewart, 2018
Conclusions

Media Lengua’s vowel system is highly overlapping yet vowels are perceptually distinct.

(Sewart 2014, 2018)
Research question

How do diphthongs in Spanish origin words adapt to Media Lengua’s overlapping vowel space?
Media Lengua’s Vowel System

<table>
<thead>
<tr>
<th>Diphthong</th>
<th>Media Lengua</th>
<th>Quichua</th>
<th>Spanish</th>
</tr>
</thead>
<tbody>
<tr>
<td>ae</td>
<td>25</td>
<td>--</td>
<td>26</td>
</tr>
<tr>
<td>ai</td>
<td>27</td>
<td>68</td>
<td>30</td>
</tr>
<tr>
<td>ao</td>
<td>9</td>
<td>0</td>
<td>17</td>
</tr>
<tr>
<td>au</td>
<td>24</td>
<td>17</td>
<td>31</td>
</tr>
<tr>
<td>ea</td>
<td>20</td>
<td>--</td>
<td>49</td>
</tr>
<tr>
<td>ei</td>
<td>15</td>
<td>--</td>
<td>23</td>
</tr>
<tr>
<td>eo</td>
<td>11</td>
<td>--</td>
<td>26</td>
</tr>
<tr>
<td>eu</td>
<td>13</td>
<td>--</td>
<td>25</td>
</tr>
<tr>
<td>ia</td>
<td>32</td>
<td>20</td>
<td>44</td>
</tr>
<tr>
<td>ie</td>
<td>85</td>
<td>21</td>
<td>73</td>
</tr>
<tr>
<td>io</td>
<td>34</td>
<td>--</td>
<td>55</td>
</tr>
<tr>
<td>oa</td>
<td>5</td>
<td>--</td>
<td>14</td>
</tr>
<tr>
<td>ua</td>
<td>29</td>
<td>11</td>
<td>23</td>
</tr>
<tr>
<td>ue</td>
<td>78</td>
<td>12</td>
<td>58</td>
</tr>
<tr>
<td>ui</td>
<td>13</td>
<td>17</td>
<td>20</td>
</tr>
<tr>
<td>uo</td>
<td>6</td>
<td>--</td>
<td>11</td>
</tr>
</tbody>
</table>

Total: 1117

- Media Lengua: 23 (14F/9M)
- Quichua: 10 (6F/4M)
- Spanish: 14 (8F/6M)

Total: 47 (28F/19M)

2010-2019

Wordlists
Elicitations
Spontaneous speech

Data
• Tokens of diphthong-containing words isolated in Praat and manually segmented

• *FormantPro* script (Xu & Gao, 2018) used to extract formant measurements for F1, F2, F3 at 5% intervals across vowel duration
Formant measurements taken at 5% intervals (20 points) permit high-fidelity replication of original formant contours with minimal errors, versus more commonplace 25% intervals (3 points).
• Generalized Additive Mixed Models (GAMMs; Hastie & Tibshirani, 1990; Wood, 2017; Sóskuthy, 2017) for comparison of non-linear data:
  • E.g. comparisons of formant trajectories across different conditions:
    • Vowel A vs. vowel B
    • Vowel A before segment X and segment Y
    • Vowel A by L1 vs. L2 speakers (Onosson & Bird, 2019)

• Applied within our dataset to compare:
  • Production of the same diphthong across Media Lengua, Quichua, and Spanish
  • Production of Quichua- and Spanish-derived diphthongs within Media Lengua
Quichua vowel space
Ellipses indicate 2 standard deviations; monophthong n = 634
Spanish vowel space
Ellipses indicate 2 standard deviations; monophthong n = 1929
Media Lengua vowel space

Ellipses indicate 2 standard deviations; monophthong n = 2040
• Per-diphthong, per-formant GAMMs comparisons:

  • **Dependent variable:** – F1 or F2
  • **Main independent/fixed effect:** – Language (M.L. vs. Quichua vs. Spanish)
  • **2 random effects:**
    – Speaker (by Duration)
    – Word (by Duration)
## Cross-Language Diphthong Comparisons

<table>
<thead>
<tr>
<th></th>
<th>ae</th>
<th>ai</th>
<th>ao</th>
<th>au</th>
<th>ea</th>
<th>ei</th>
<th>eo</th>
<th>eu</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>F1</strong></td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
<td><strong>p=0.02</strong></td>
<td>n.s.</td>
<td><strong>p=0.011</strong></td>
<td>n.s.</td>
<td><strong>p=0.005</strong></td>
</tr>
<tr>
<td><strong>F2</strong></td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
<td><strong>p≈0</strong></td>
<td>n.s.</td>
<td><strong>p=0.006</strong></td>
</tr>
<tr>
<td><strong>Langs</strong></td>
<td>ML, S</td>
<td>ML, Q, S</td>
<td>ML, S</td>
<td>ML, Q, S</td>
<td>ML, S</td>
<td>ML, S</td>
<td>ML, Q, S</td>
<td>ML, S</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ia</th>
<th>ie</th>
<th>io</th>
<th>oa</th>
<th>ua</th>
<th>ue</th>
<th>ui</th>
<th>uo</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>F1</strong></td>
<td><strong>p≈0</strong></td>
<td>n.s.</td>
<td><strong>p≈0</strong></td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
</tr>
<tr>
<td><strong>F2</strong></td>
<td><strong>p≈0</strong></td>
<td><strong>p≈0</strong></td>
<td><strong>p=0.023</strong></td>
<td>n.s.</td>
<td>n.s.</td>
<td><strong>p≈0</strong></td>
<td>n.s.</td>
</tr>
<tr>
<td><strong>Langs</strong></td>
<td>ML, Q, S</td>
<td>ML, Q, S</td>
<td>ML, S</td>
<td>ML, S</td>
<td>ML, Q, S</td>
<td>ML, Q, S</td>
<td>ML, Q, S</td>
</tr>
</tbody>
</table>
• Media Lengua vs. Spanish: /ei, eu, io/
• Media Lengua vs. Quichua vs. Spanish: /ia/
/ei/: F1 & F2

Media Lengua vowel space with /ei/ trajectory
Ellipses indicate 2 standard deviations; monophthong n = 2040

Spanish vowel space with /ei/ trajectory
Ellipses indicate 2 standard deviations; monophthong n = 1929
• Very small differences, localized to onset

• Media Lengua: lower F1 = higher articulatory position
• Significant differences across full trajectory
• Media Lengua: F2 much lower = retracted position
/eu/: F1 & F2

Media Lengua vowel space with /eu/ trajectory
Ellipses indicate 2 standard deviations; monophthong n = 2040

Spanish vowel space with /eu/ trajectory
Ellipses indicate 2 standard deviations; monophthong n = 1929
/eu/ F1

- Small, localized difference in onset
- Media Lengua: higher F1 = lower articulation
/eu/ F2

- Sig. F2 differences across >50% of trajectory
- Media Lengua: F2 higher (advanced position) and generally more dynamic
- Spanish: F2 very flat, little front-to-back movement
/ia/: F1 & F2
• Media Lengua vs. Quichua: no sig. difference
• Spanish: somewhat higher F1 = lower articulation, sig. difference at onset and offset
• Media Lengua vs. Quichua: no sig. difference
• Spanish: higher F2 = advanced articulation, over initial 2/3 of duration
Media Lengua vowel space with /io/ trajectory
Ellipses indicate 2 standard deviations; monophthong n = 2040

Spanish vowel space with /io/ trajectory
Ellipses indicate 2 standard deviations; monophthong n = 1929

/io/: F1 & F2
Media Lengua: F1 lower (higher articulation) during final 25% of duration
• Spanish: slightly higher F2 = advanced articulation, from 20–40% duration and at offset
Cross-Language Differences in only F1 or F2

• F1:
  • Media Lengua vs. Spanish: /uo/
  • Media Lengua vs. Quichua vs. Spanish: /au/

• F2:
  • Media Lengua vs. Quichua vs. Spanish: /ie, ue/
/ie/: F2
• Media Lengua distinct from both Quichua & Spanish; lower F2 (retracted) over initial 2/3 duration
• Quichua: F2 slightly higher = advanced position
• Spanish: F2 substantially higher with steeper trajectory
/ue/: F2

Quichua vowel space with /ue/ trajectory
Ellipses indicate 2 standard deviations; monophthong n = 634

Media Lengua vowel space with /ue/ trajectory
Ellipses indicate 2 standard deviations; monophthong n = 2040

Spanish vowel space with /ue/ trajectory
Ellipses indicate 2 standard deviations; monophthong n = 1929
• Media Lengua vs. Quichua: no sig. diff.
• Spanish: lower F2 = retracted position, during approx. 30-50% of duration
Cross-Language Comparisons

- Spanish diphthongs generally show less internal variability:
  - Confidence intervals typically much narrower than other languages
  - Lower variation can’t be attributed to either speaker $n$, which is notably higher than Quichua, nor token $n$ which is the largest among all three languages

- Media Lengua vs. Quichua typically non-contrastning
  - Single identified contrast between ML~Q is F2 of /ie/;
    ML~Sp /ie/ F2 difference is even greater

- In general, Media Lengua production matches Quichua production
• Pairs of Media Lengua diphthongs selected for cross-comparison
  • /ai ~ ei/  
  • /ea ~ ia/  
  • /eo ~ eu/  
  • /ue ~ ui/

• GAMMs model uses main fixed effect of Vowel instead of Language

• ANOVA comparison of GAMM null vs. non-null models identifies where formant trajectories differ significantly between diphthongs

<table>
<thead>
<tr>
<th></th>
<th>ai vs. ei</th>
<th>ea vs. ia</th>
<th>eo vs. eu</th>
<th>ue vs. ui</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1</td>
<td>p≈0</td>
<td>n.s.</td>
<td>p≈0</td>
<td>p=0.008</td>
</tr>
<tr>
<td>F2</td>
<td>p≈0</td>
<td>p≈0</td>
<td>p≈0</td>
<td>p=0.002</td>
</tr>
</tbody>
</table>
Media Lengua /ea/ vs. /ia/

Ellipses indicate 2 standard deviations; monophthong n = 2040
• Substantial overlap in F1 confidence intervals = no significant difference in vowel height across entire trajectory
Media Lengua /ea, ia/ F2

• /ia/ F2 higher (advanced) until ~70% duration
Most diphthong pairs differ significantly across both F1 & F2
  - /ea, ia/ do not differ in F1, leaving F2 (i.e. front-back position) as the sole differentiating factor for this pair
  - Some pairs (e.g. /eo, eu/ and /ue, ui/) exhibit only subtle differences, and not always restricted to the expected portion of the trajectory
• Three diphthongs are represented by Media Lengua lexical items of both Quichua and Spanish origin:
  • /ai, au, ui/

• GAMMs comparison conducted for each diphthong using Media Lengua data with a fixed effect of Source Language

• *None* turned out to exhibit significant differences in formant trajectories
  • In other words, different-source-language lexical items are incorporated into a single phonological vowel system (with regards to these particular diphthongs)
  • This aligns with cross-linguistic comparison across diphthong inventory, where vast majority of ML~Q comparisons were non-distinctive
The Media Lengua vowel system reflects that of a late bilingual where interference from their L1 (Quichua) impedes native-like production in their L2 (Spanish).

• Then the vowel system could have been subsequently nativized with the overlaps ‘frozen’ in place.
• Phonological ‘stresses’ from relexification (e.g., high functional loads of contrastive non-native phonemes) from the mid-vowels may have been driving forces for maintaining/creating contrasts with the high-vowels in the predominately Quichua system.

1,415 ML Word Sample

<table>
<thead>
<tr>
<th>Vowel</th>
<th>e</th>
<th>o</th>
<th>ea</th>
<th>ae</th>
<th>eo</th>
<th>eu</th>
<th>ue</th>
<th>oi</th>
<th>io</th>
<th>ei</th>
<th>ie</th>
<th>ao</th>
<th>oa</th>
<th>uo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Count</td>
<td>461</td>
<td>524</td>
<td>9</td>
<td>0</td>
<td>5</td>
<td>4</td>
<td>33</td>
<td>1</td>
<td>55</td>
<td>6</td>
<td>37</td>
<td>3</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>%</td>
<td>33%</td>
<td>37%</td>
<td>0.6%</td>
<td>0%</td>
<td>0.4%</td>
<td>0.3%</td>
<td>2.3%</td>
<td>0.07%</td>
<td>4%</td>
<td>0.4%</td>
<td>2.6%</td>
<td>0.2%</td>
<td>0.1%</td>
<td>0.1%</td>
</tr>
<tr>
<td>Count</td>
<td>985</td>
<td>147</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>%</td>
<td>70%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Diphthongs also appear to follow this tendency given that the formant trajectories in Media Lengua look strikingly similar to Quichua diphthongs (e.g., size, shape, variability). However, given that the mid- and high-vowels have ‘solidified’ into highly overlapping acoustic spaces, more reminiscent of the 3-vowel Quichua system, Spanish origin diphthongs were forced to accommodate to these regions while maintaining contrastive. This might suggest why Media Lengua diphthongs may act more like what would be expected from Quichua, if such sounds existed, rather than Spanish.
Future research

Functional load

<table>
<thead>
<tr>
<th>Vowel</th>
<th>e</th>
<th>o</th>
<th>i</th>
<th>u</th>
</tr>
</thead>
<tbody>
<tr>
<td>Count</td>
<td>461</td>
<td>524</td>
<td>588</td>
<td>286</td>
</tr>
<tr>
<td>%</td>
<td>33%</td>
<td>37%</td>
<td>42%</td>
<td>20%</td>
</tr>
<tr>
<td>Count</td>
<td>985</td>
<td></td>
<td>788</td>
<td></td>
</tr>
<tr>
<td>%</td>
<td>70%</td>
<td></td>
<td>62%</td>
<td></td>
</tr>
</tbody>
</table>

Cognitive load reaction time/ eye-tracking experiment

Vowel: piscadu, piscado

Images:
- Fish
- Rabbit
- Dog
- Lizard
Future research

Vowel length

- i
- e
- u
- o

- ia
- ea

0.058004
0.066943
0.062318
0.070926

0.058004
0.066943
0.062318
0.070926

0.132490 (7.548 / s)

0.132490

0.158866 (6.295 / s)

0.158866
References


