

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

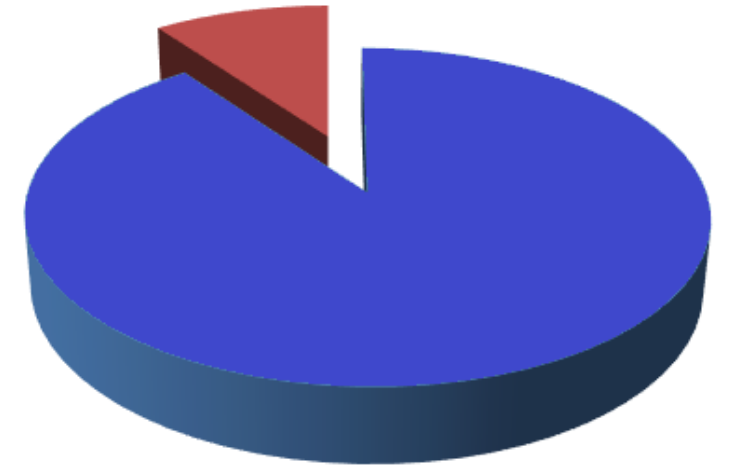
January 3rd, 2020

Media Lengua

Mixed Language (Quichua-Spanish)

Quichua – phonological, morphological, and syntactic system

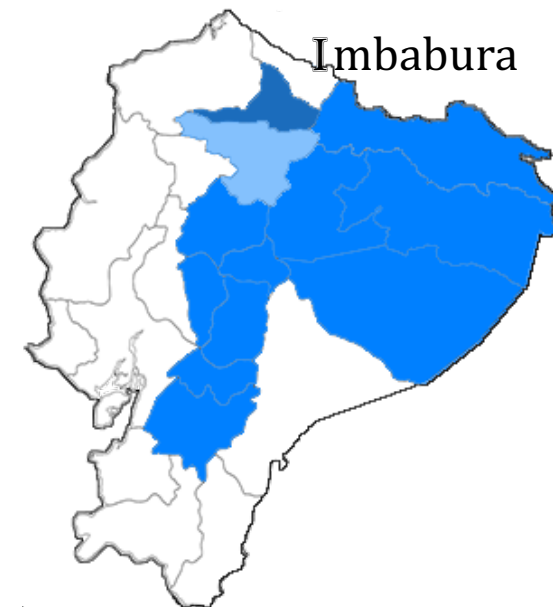
Spanish – Lexicon



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

(Jones, Meakins, and Buchan
2011)

/æ/ & /e/

/u:/ &

/o:/

Show more overlap in Gurindji Kriol than in
Kriol

Michif

(Rosen, Stewart, & Sammons 2020)

/ɛ/ & /ɔ/

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

Media Lengua

(Stewart
2014)

/e/ & /i/

/o/ &

/i/, /u/ &

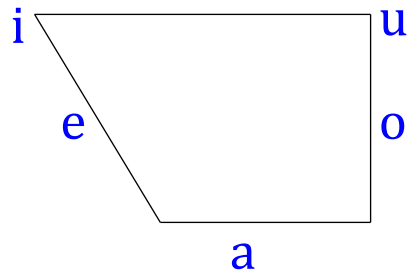
/ɔ/

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

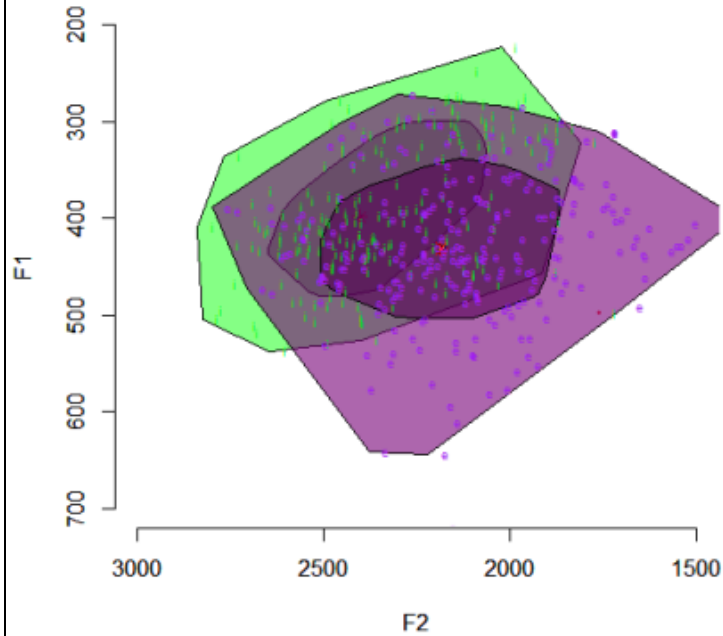
Media Lengua's Vowel System

Production

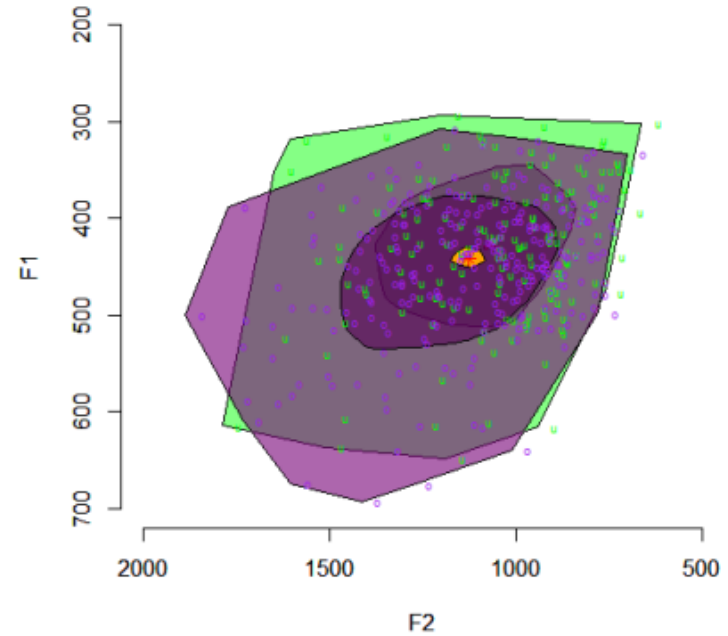
Spanish



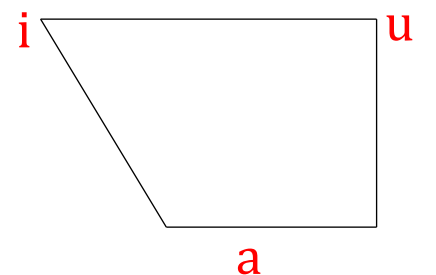
IQ: Spanish-derived /i/ Vs. Spanish-derived /e/



IQ: Spanish-derived /u/ Vs. Spanish-derived /o/



Quichua



sarten-pimi

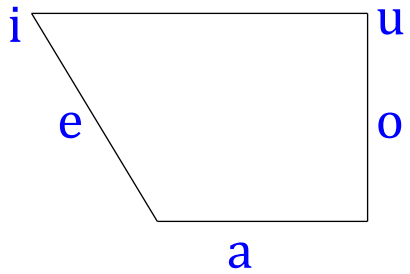
pan-LOC-VAL

'in the pan'

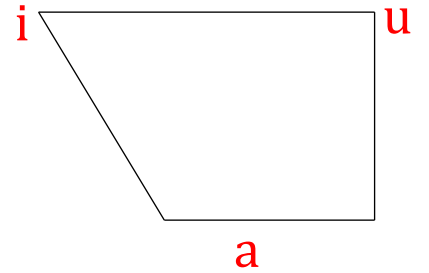
Media Lengua's Vowel System

Production

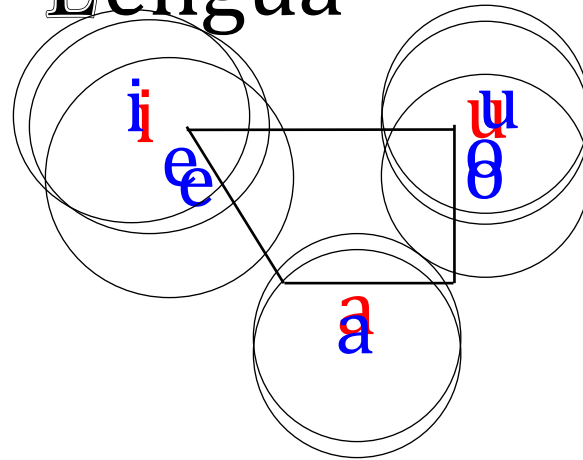
Spanish



Quichua



Media
Lengua



Media Lengua's Vowel System

2AFC

Identification

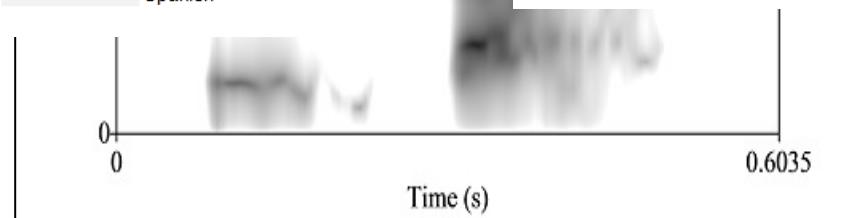
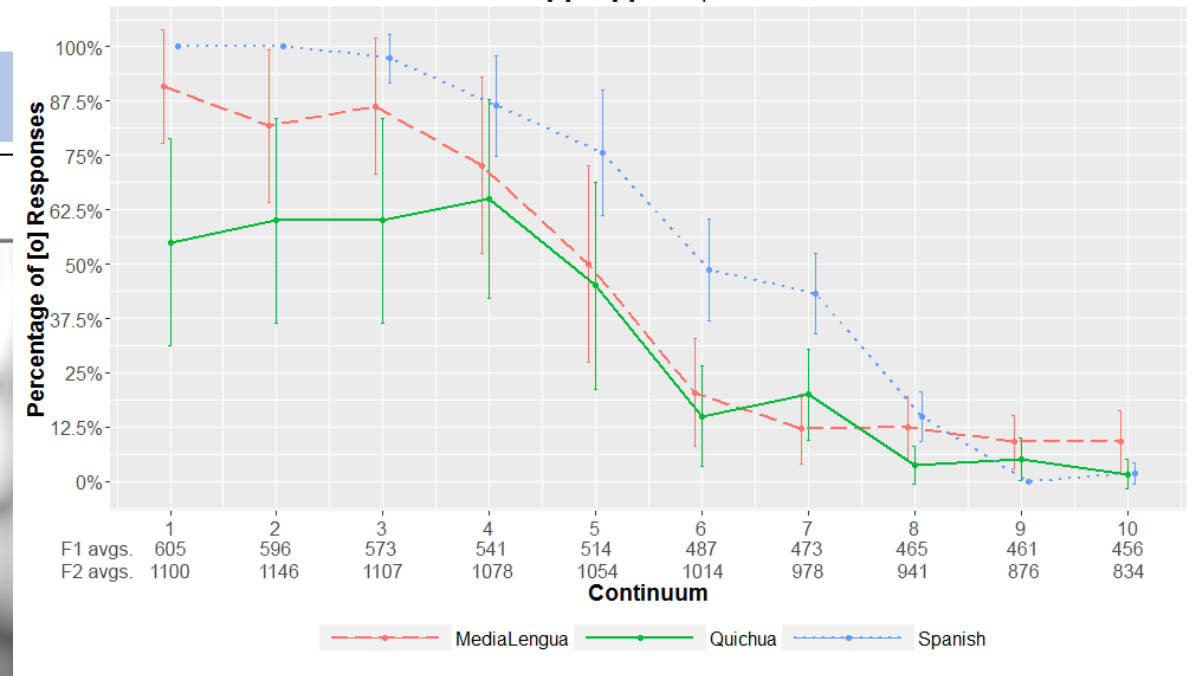
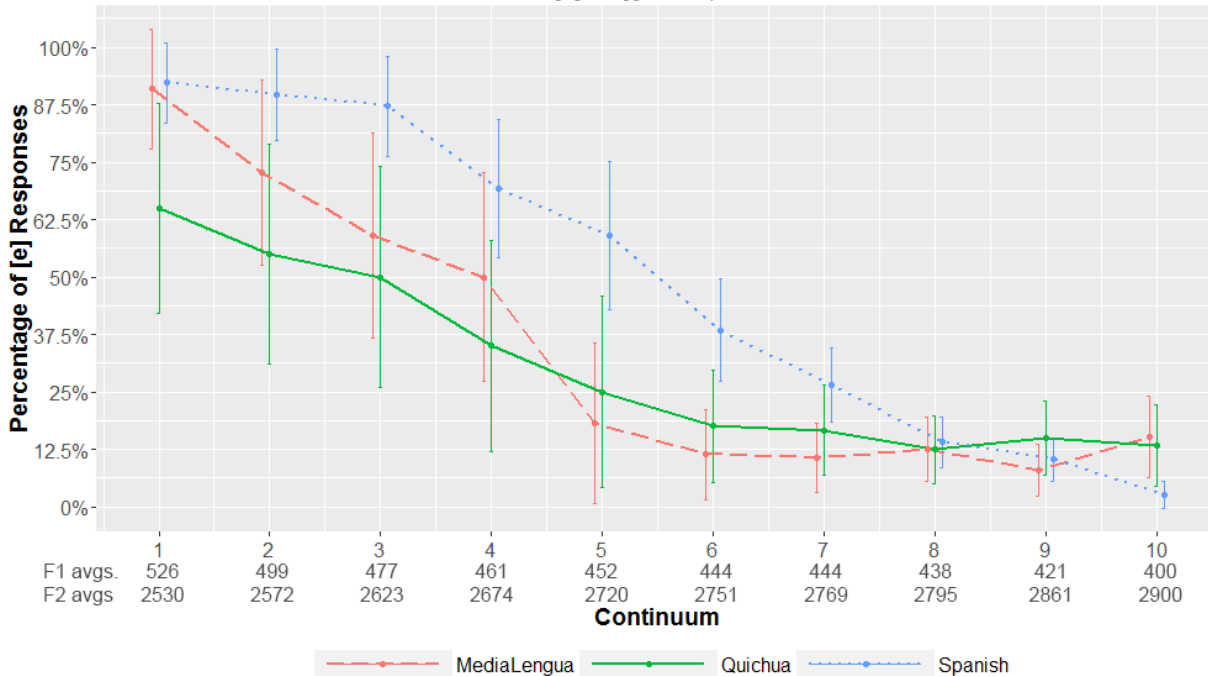
Task Experiment

Perception

Spanish

542	520	486	497	489	484	486	507	470	456	F1 – mid point
2530	2572	2795	2674	2720	2751	2795	2623	2861	2900	F2 – mid point
3215	3254	3490	3382	3422	3454	3490	3306	3531	3588	F3 – mid point
[e] vs. [i] Perception										

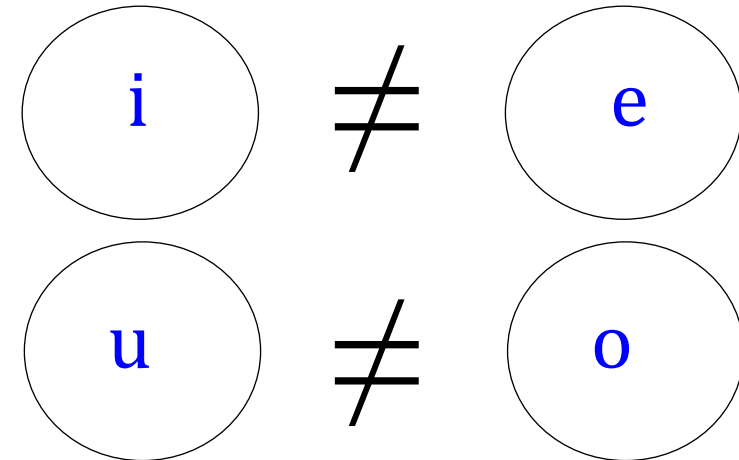
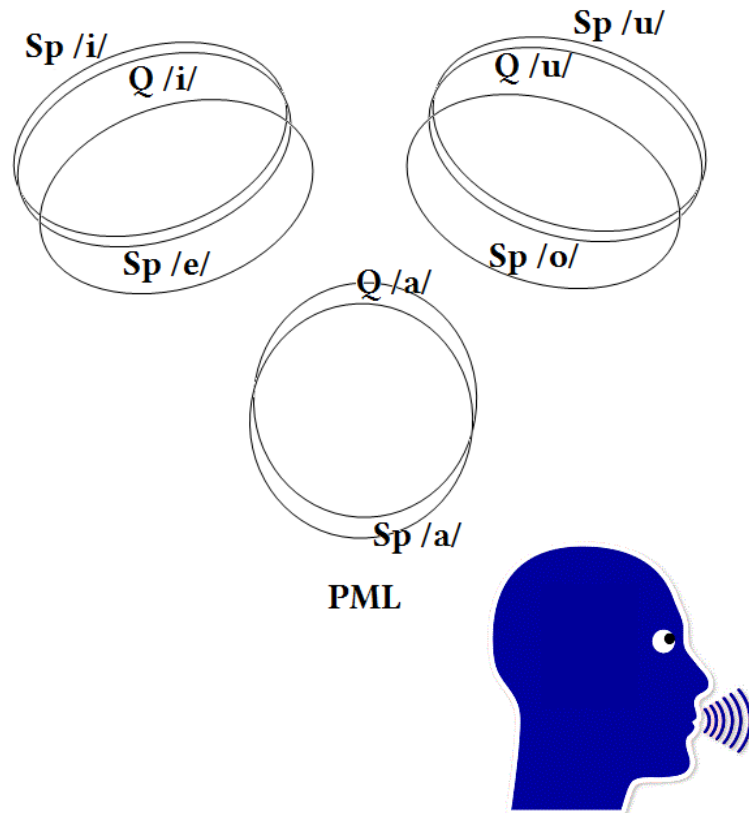
Quichua



Media Lengua's Vowel System

Conclusions

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



Media Lengua's Vowel System

Research question

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

Media Lengua's Vowel System *Data*

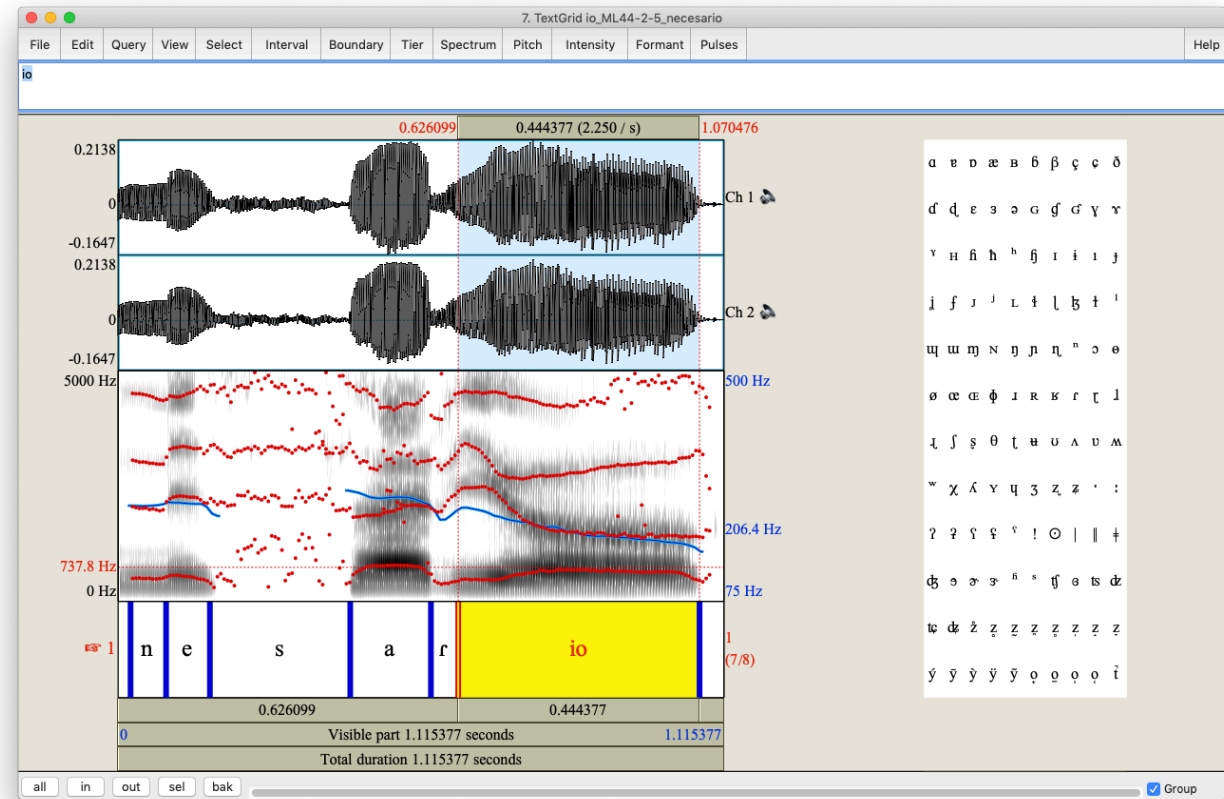
Media Lengua: 23 (14F/ 9M)
 Quichua: 10 (6F/ 4M)
 Spanish: 14 (8F/ 6M)
 Total: 47 (28F/ 19M)

2010-2019
 Wordlists
 Elicitations
 Spontaneous speech

Diphthong	Media Lengua	Quichua	Spanish
ae	25	--	26
ai	27	68	30
ao	9	0	17
au	24	17	31
ea	20	--	49
ei	15	--	23
eo	11	--	26
eu	13	--	25
ia	32	20	44
ie	85	21	73
io	34	--	55
oa	5	--	14
ua	29	11	23
ue	78	12	58
ui	13	17	20
uo	6	--	11
Total: 1117	426	116	525

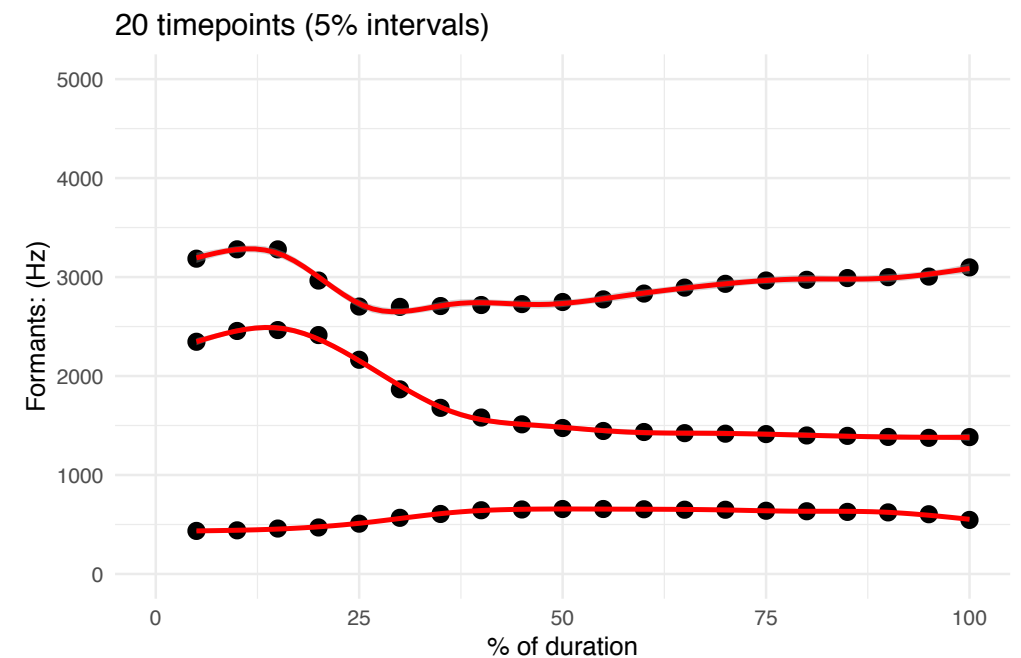
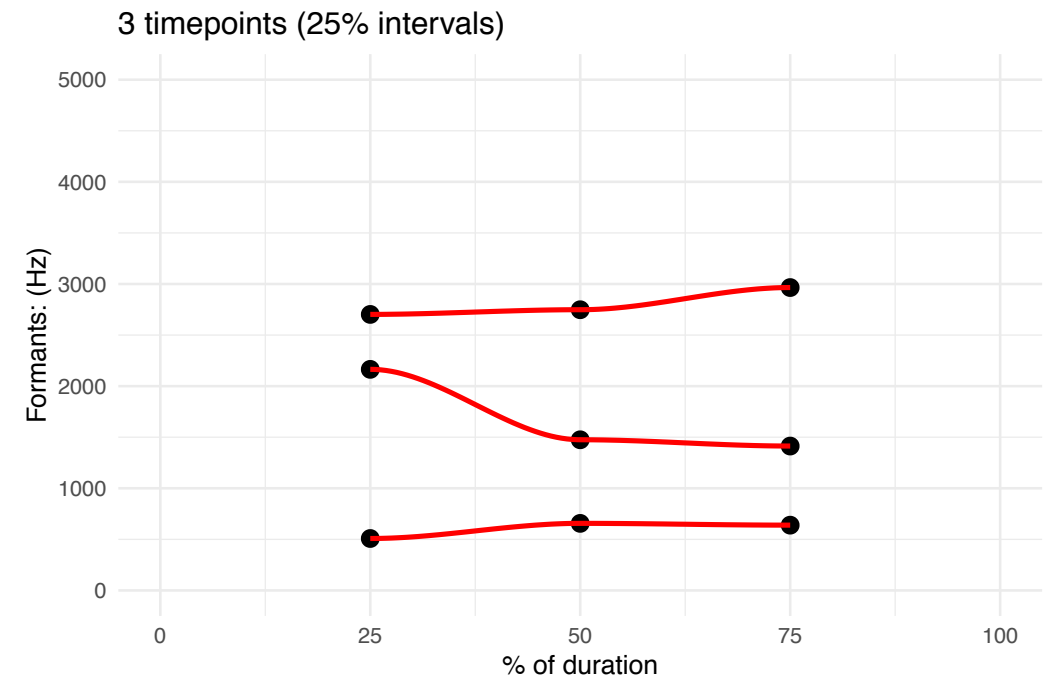
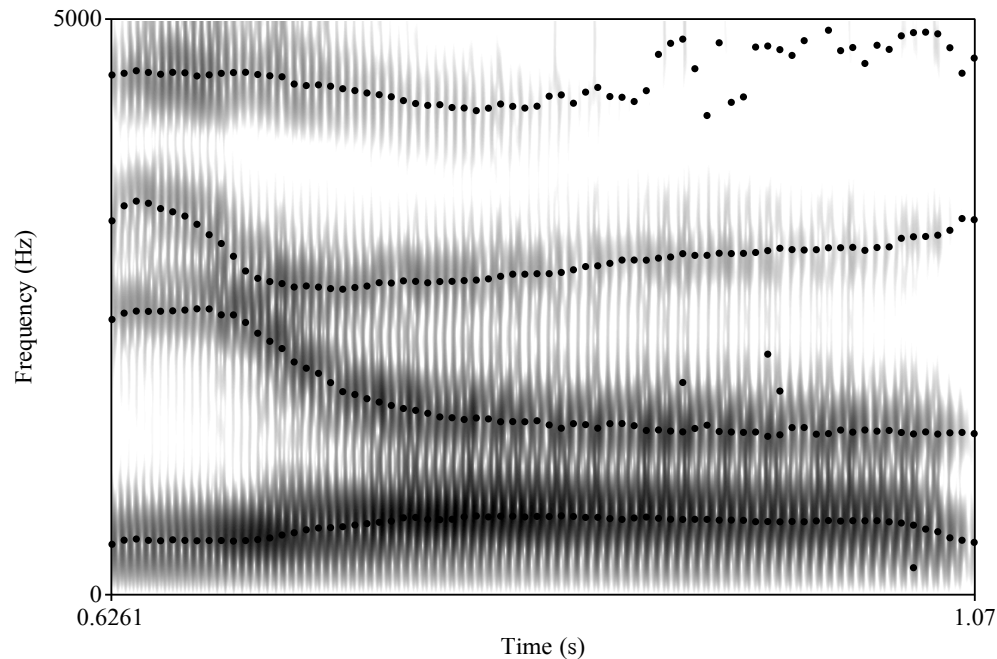
Acoustic Measurement

- 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



Acoustic Analysis

- 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)

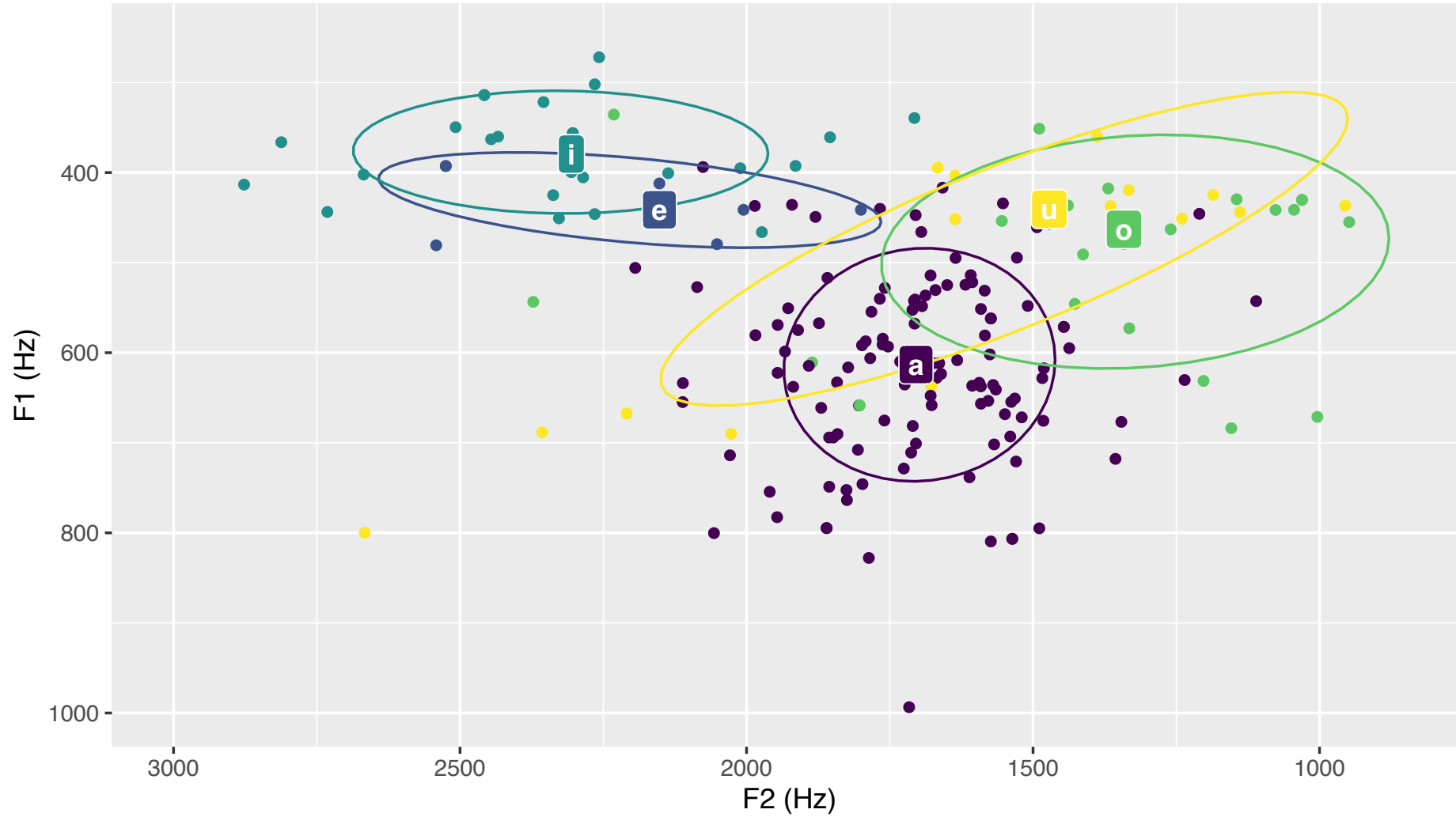


GAMMs comparisons

- **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

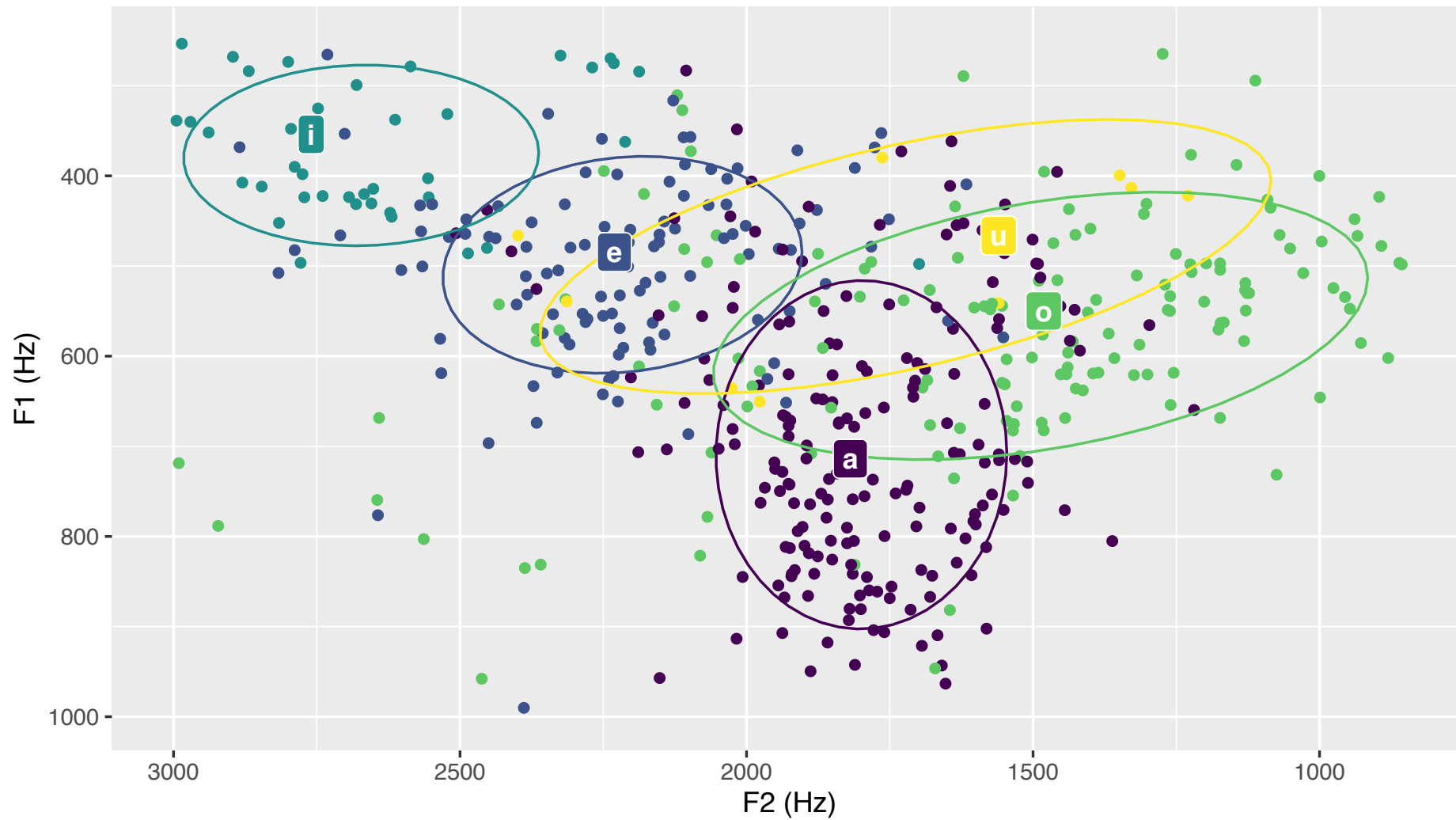
Quichua vowel space

Ellipses indicate 2 standard deviations; monophthong n = 634



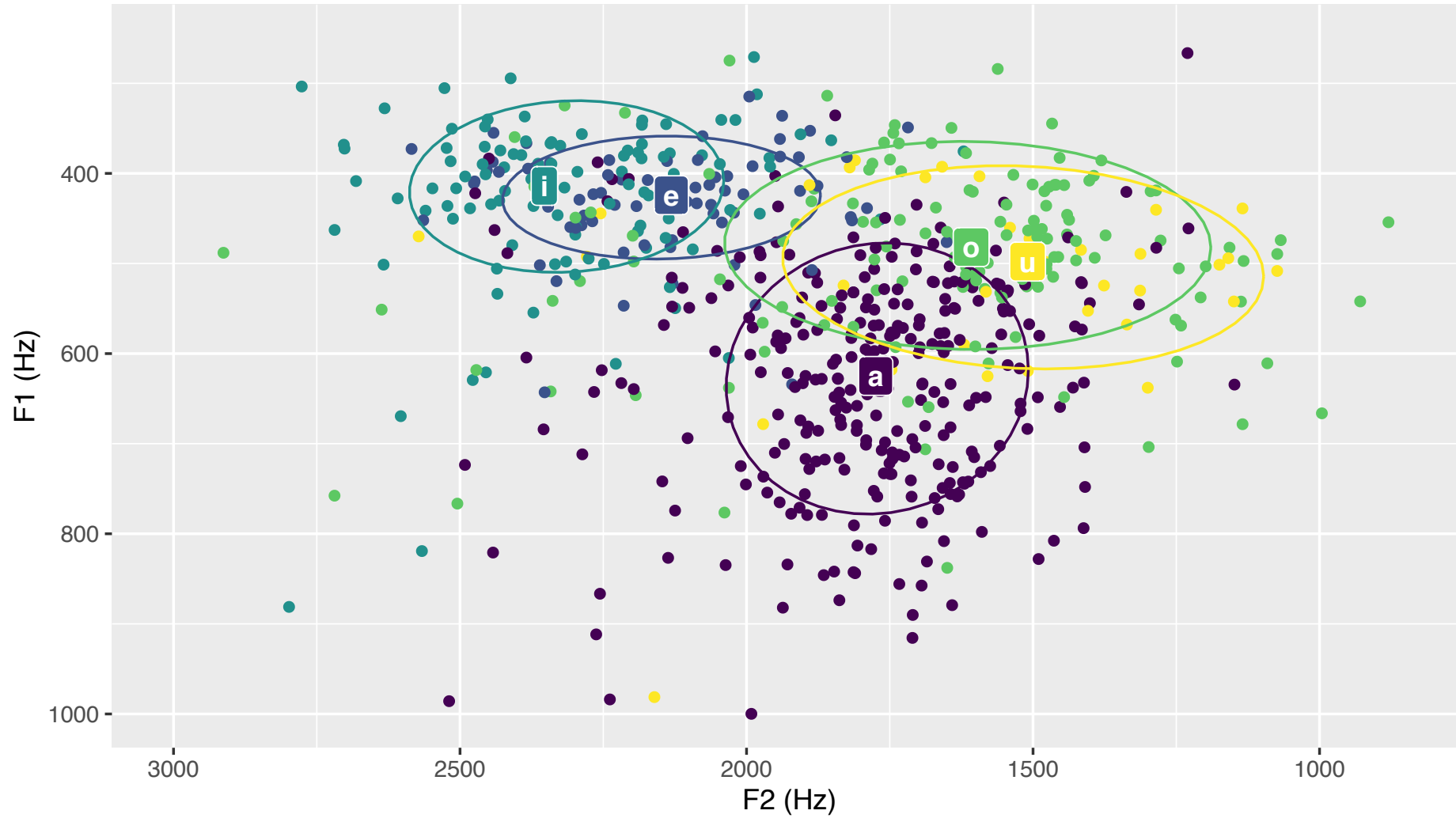
Spanish vowel space

Ellipses indicate 2 standard deviations; monophthong n = 1929



Media Lengua vowel space

Ellipses indicate 2 standard deviations; monophthong n = 2040



Cross-Language Diphthong Comparisons

- 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

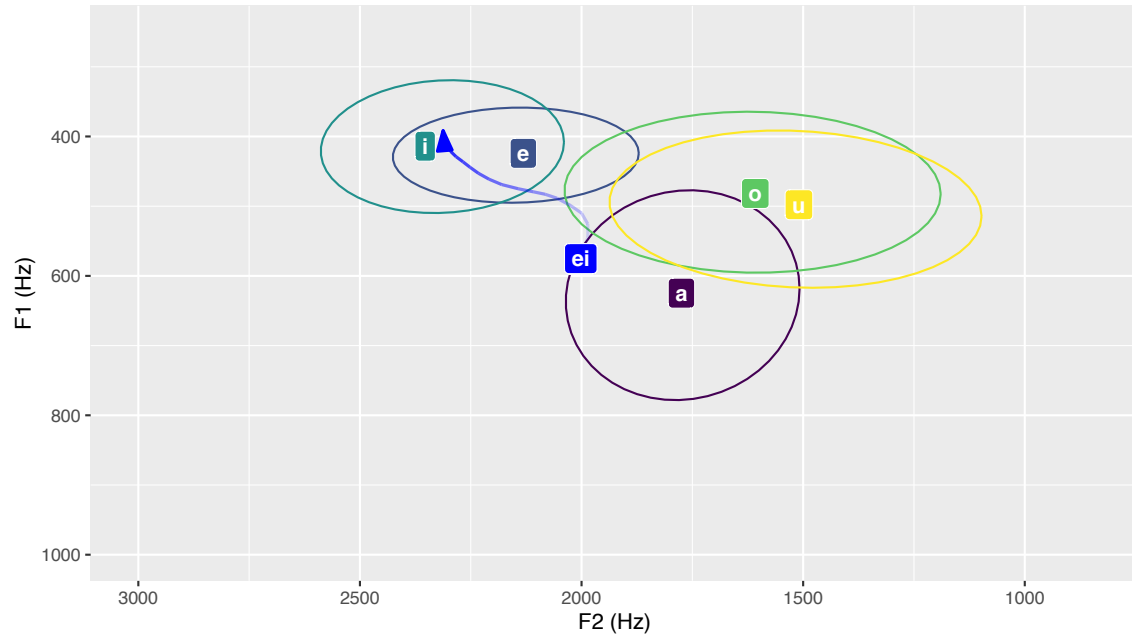
	ae	ai	ao	au	ea	ei	eo	eu
F1	n.s.	n.s.	n.s.	<i>p=0.02</i>	n.s.	<i>p=0.011</i>	n.s.	<i>p=0.005</i>
F2	n.s.	n.s.	n.s.	n.s.	n.s.	<i>p≈0</i>	n.s.	<i>p=0.006</i>
Langs	ML, S	ML, Q, S	ML, S	ML, Q, S	ML, S	ML, S	ML, Q, S	ML, S
	ia	ie	io	oa	ua	ue	ui	uo
F1	<i>p≈0</i>	n.s.	<i>p≈0</i>	n.s.	n.s.	n.s.	n.s.	<i>p=0.002</i>
F2	<i>p≈0</i>	<i>p≈0</i>	<i>p=0.023</i>	n.s.	n.s.	<i>p≈0</i>	n.s.	n.s.
Langs	ML, Q, S	ML, Q, S	ML, S	ML, S	ML, Q, S	ML, Q, S	ML, Q, S	ML, S

Cross-Language Differences in both F1 and F2

- Media Lengua vs. Spanish: /ei, eu, io/
- Media Lengua vs. Quichua vs. Spanish: /ia/

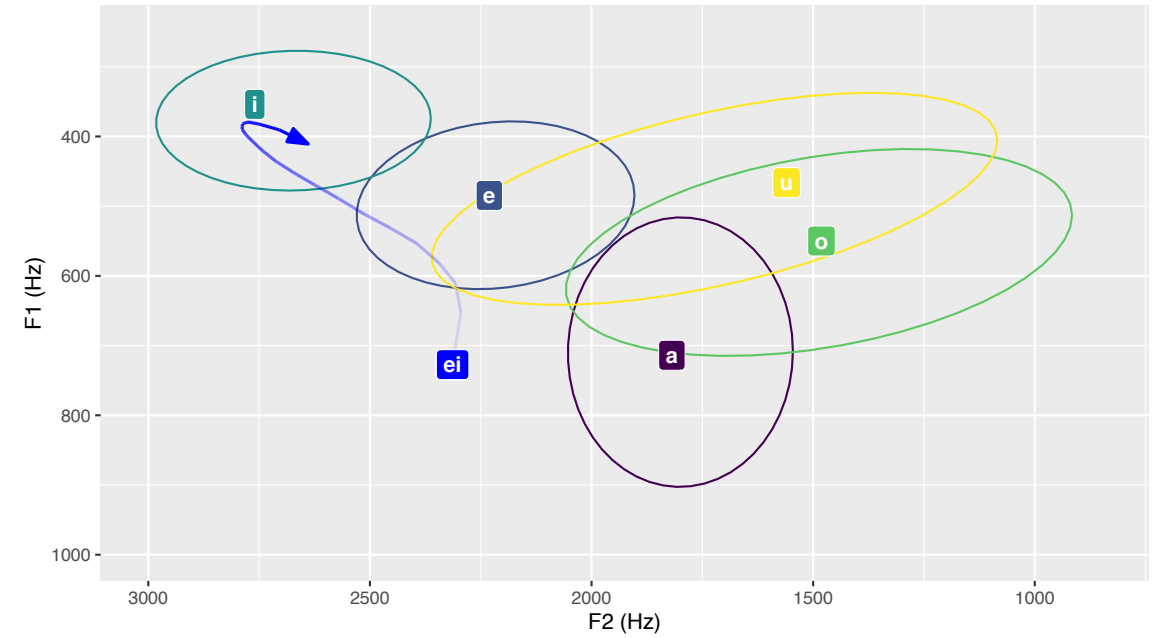
Media Lengua vowel space with /ei/ trajectory

Ellipses indicate 2 standard deviations; monophthong n = 2040

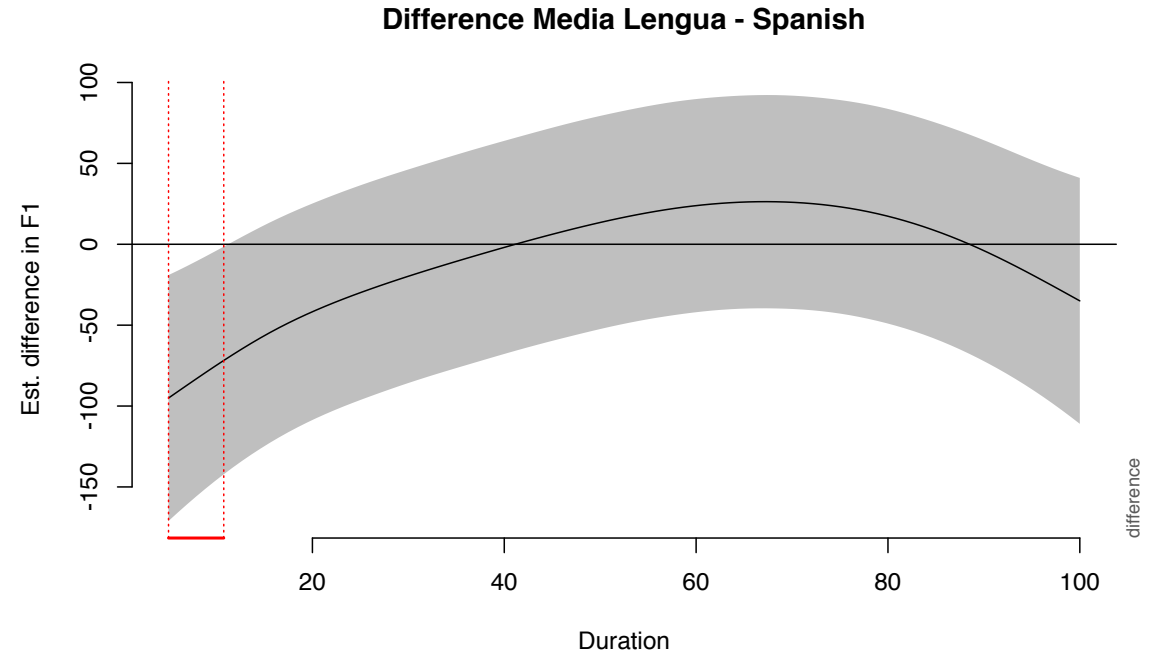
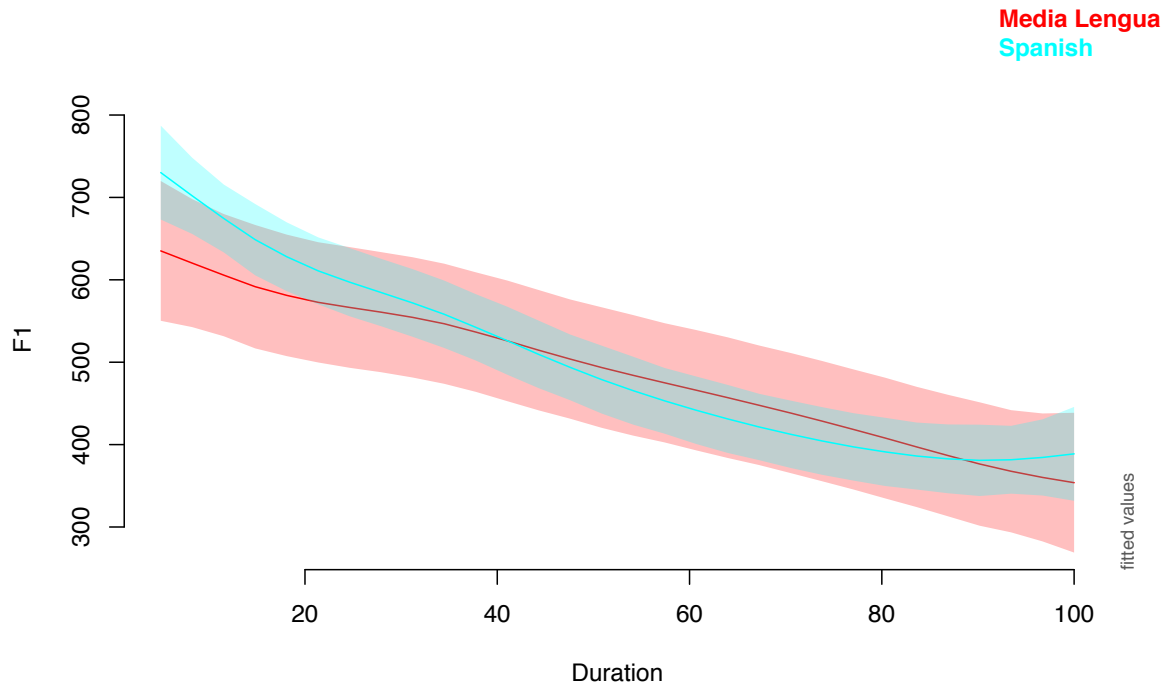


Spanish vowel space with /ei/ trajectory

Ellipses indicate 2 standard deviations; monophthong n = 1929

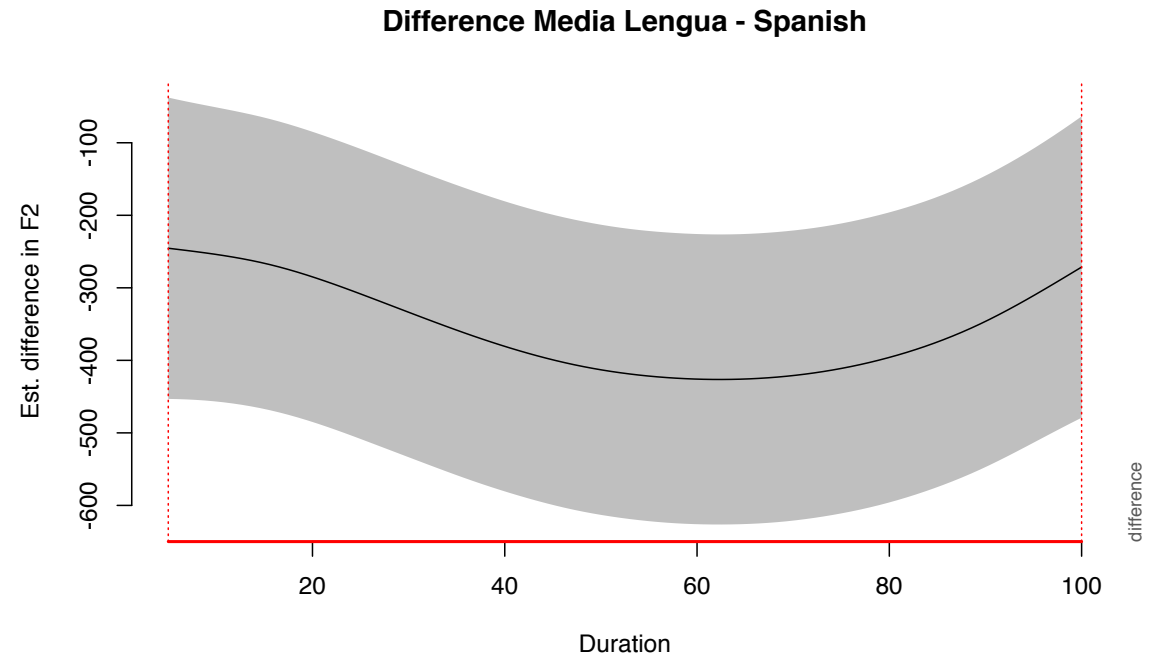
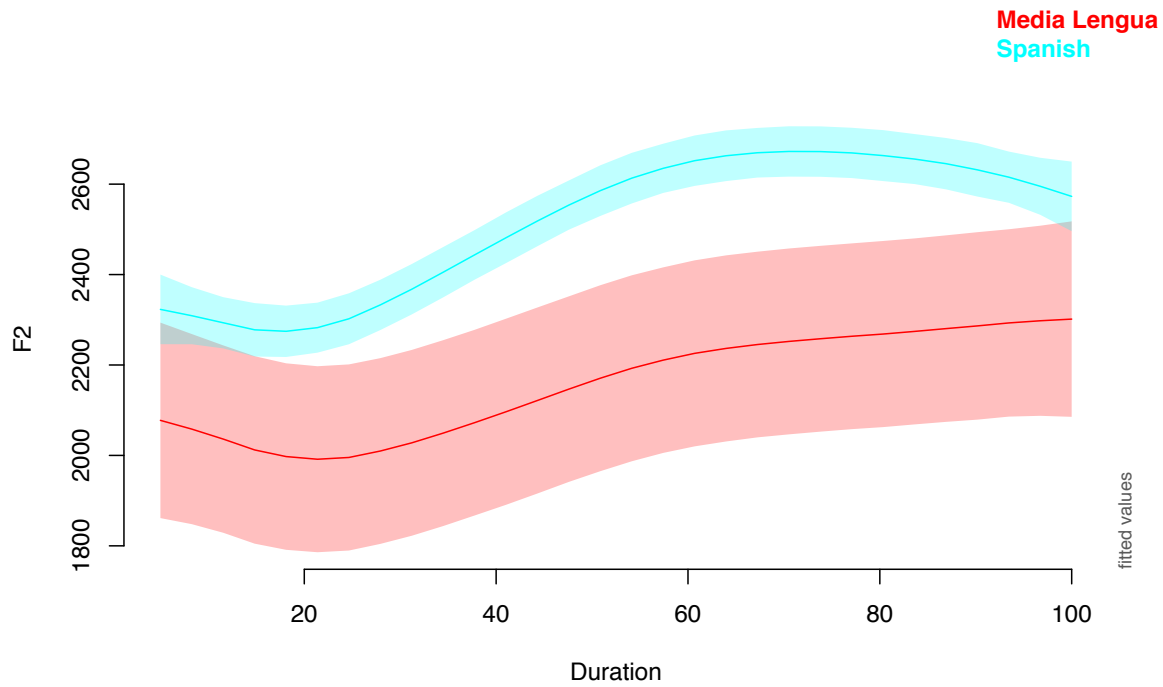


/ei/: F1 & F2



/ei/ F1

- Very small differences, localized to onset
- Media Lengua: lower F1 = higher articulatory position

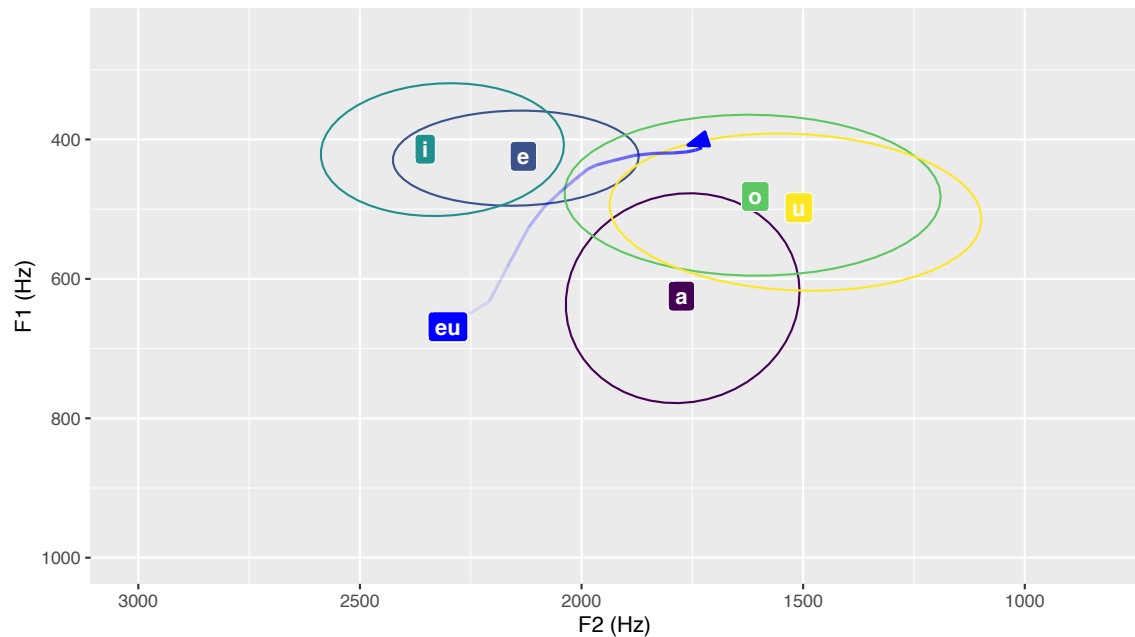


/ei/ F2

- Significant differences across full trajectory
- Media Lengua: F2 much lower = retracted position

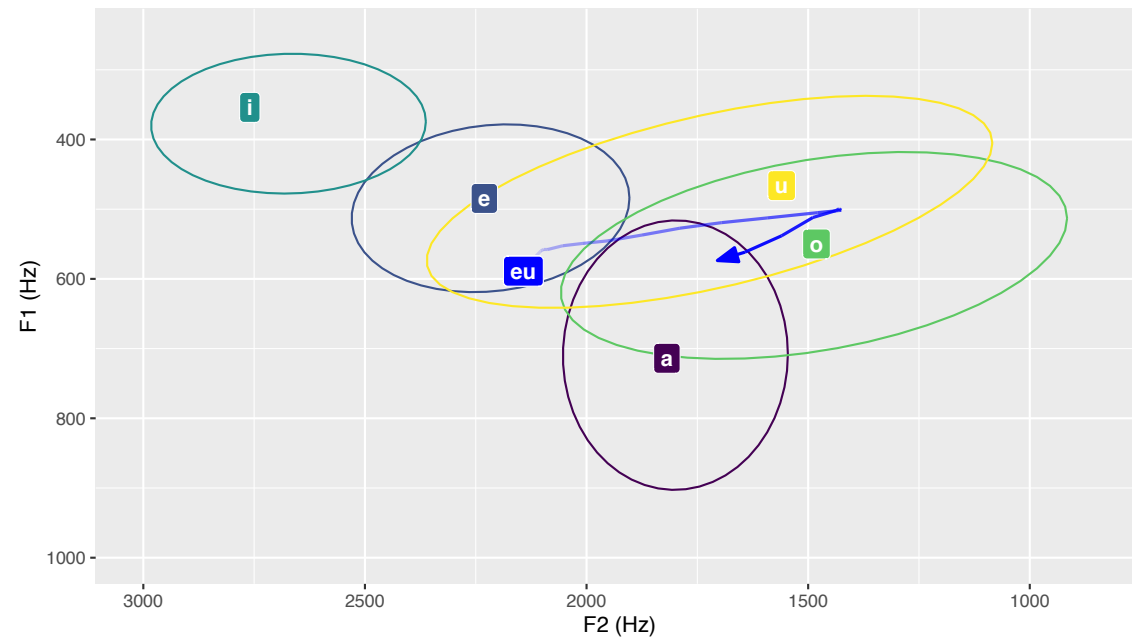
Media Lengua vowel space with /eu/ trajectory

Ellipses indicate 2 standard deviations; monophthong n = 2040

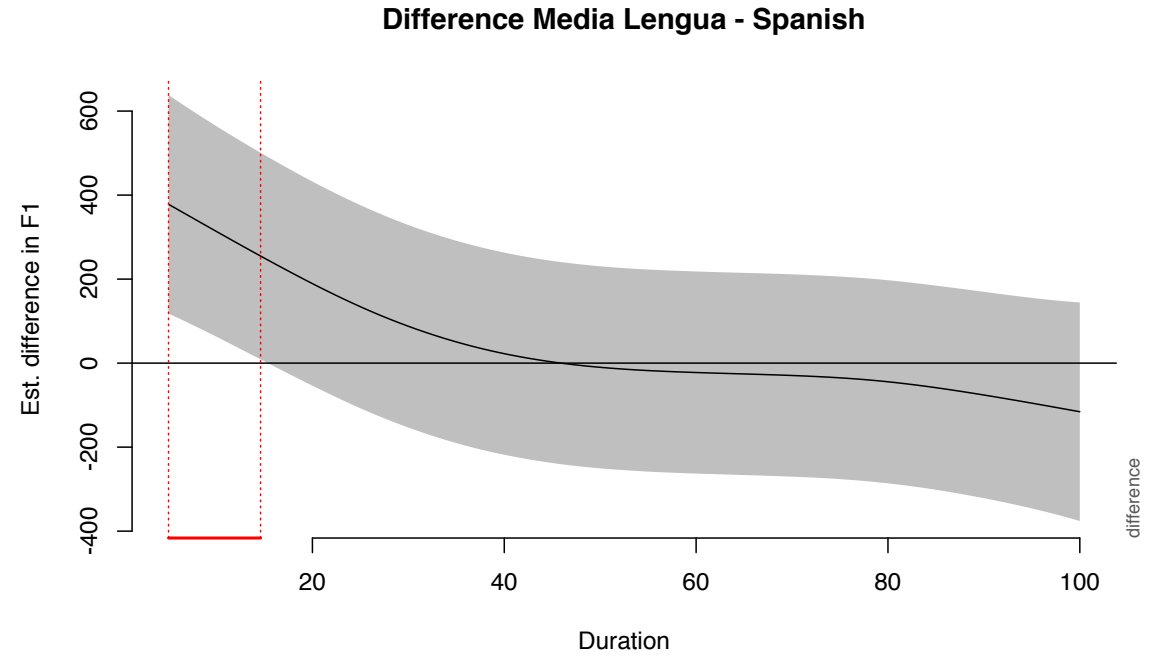
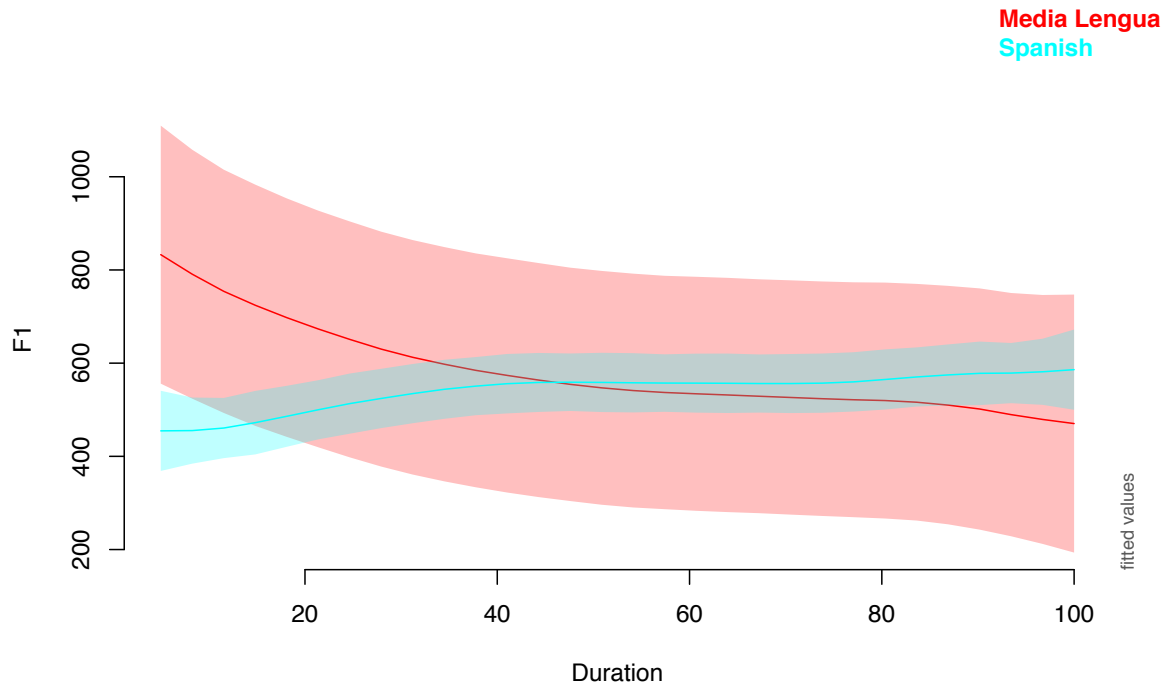


Spanish vowel space with /eu/ trajectory

Ellipses indicate 2 standard deviations; monophthong n = 1929

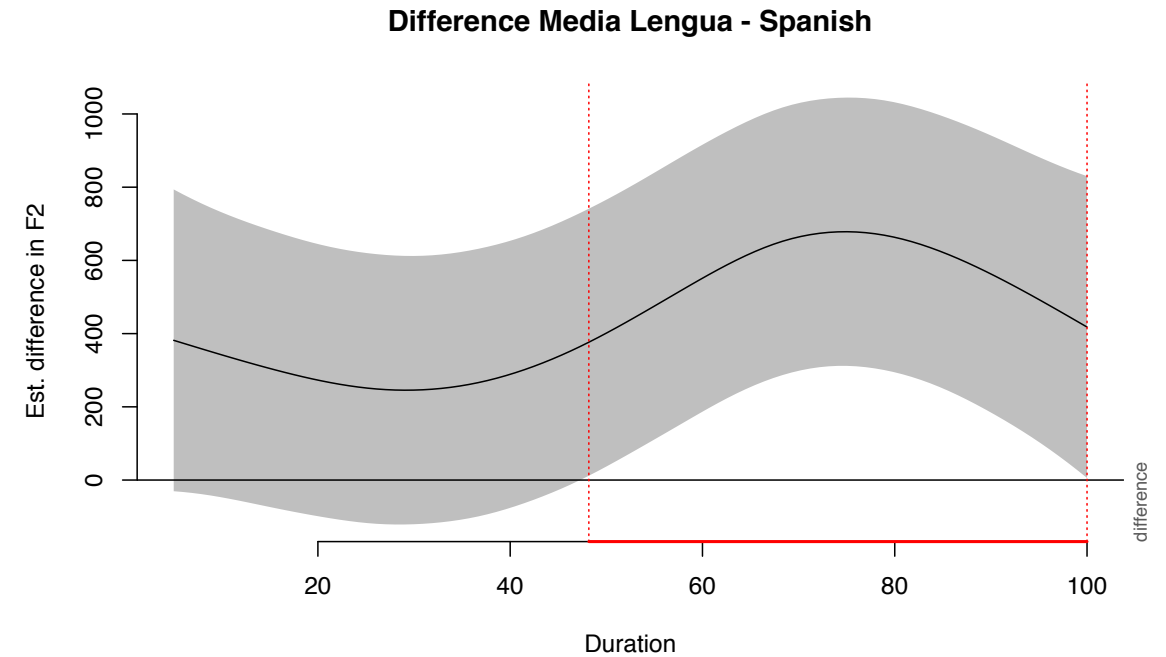
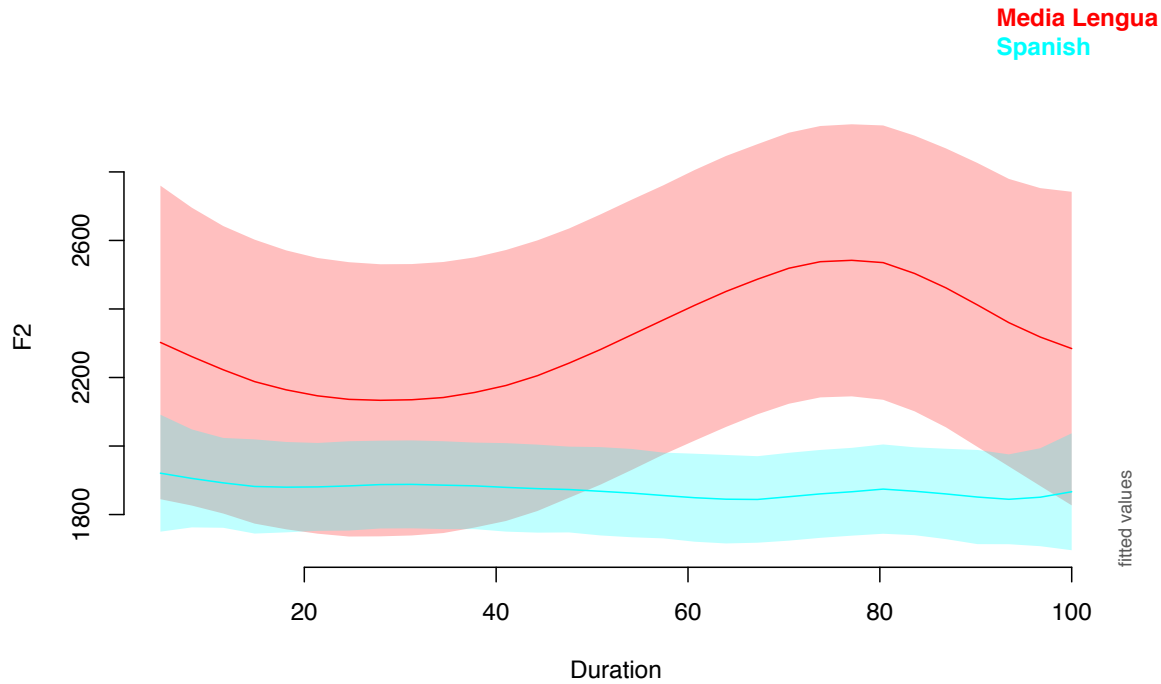


/eu/: F1 & F2



/eu/ F1

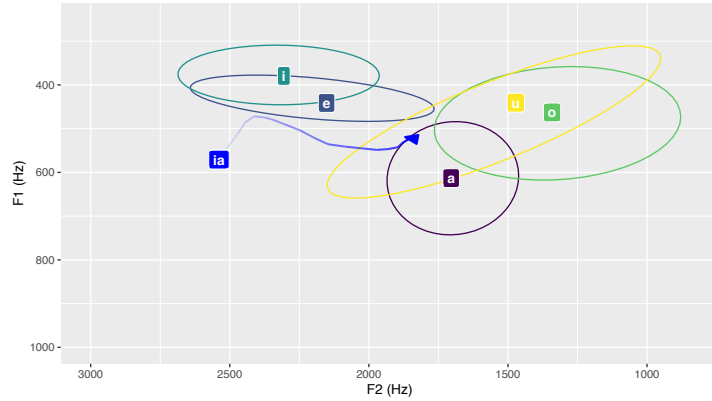
- Small, localized difference in onset
- Media Lengua: higher F1 = lower articulation



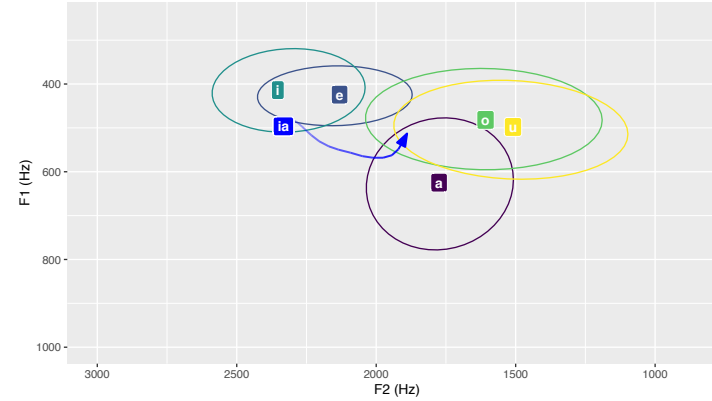
/eu/ F2

- 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

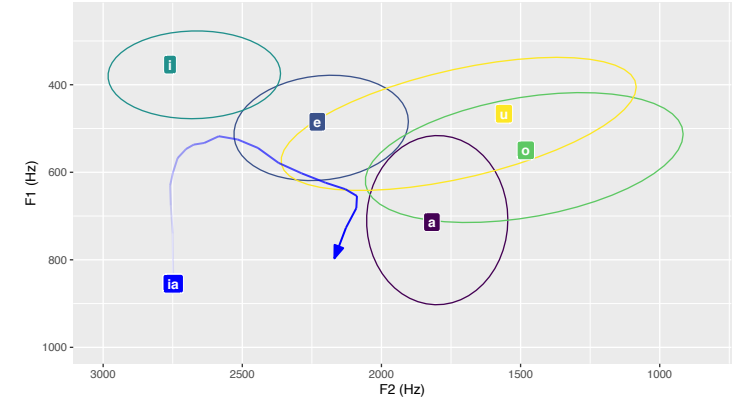
Quichua vowel space with /ia/ trajectory
Ellipses indicate 2 standard deviations; monophthong n = 634



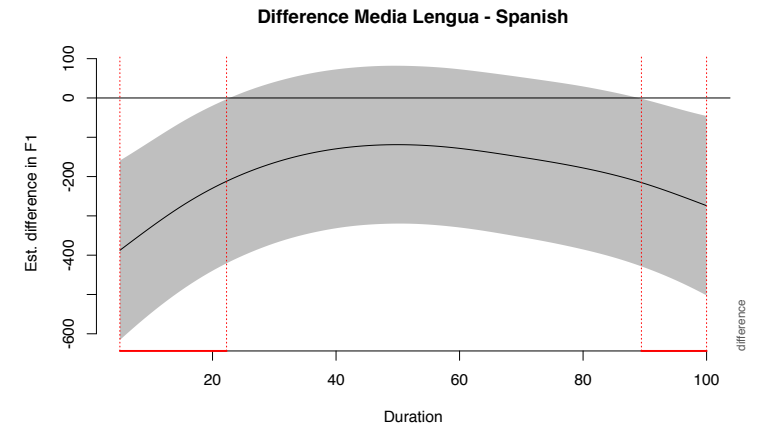
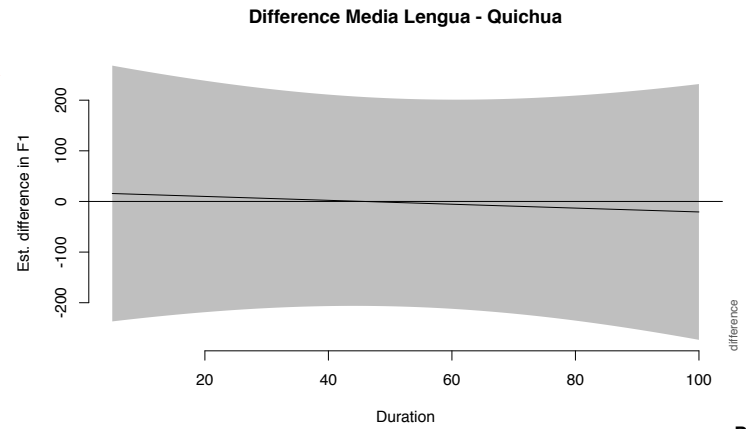
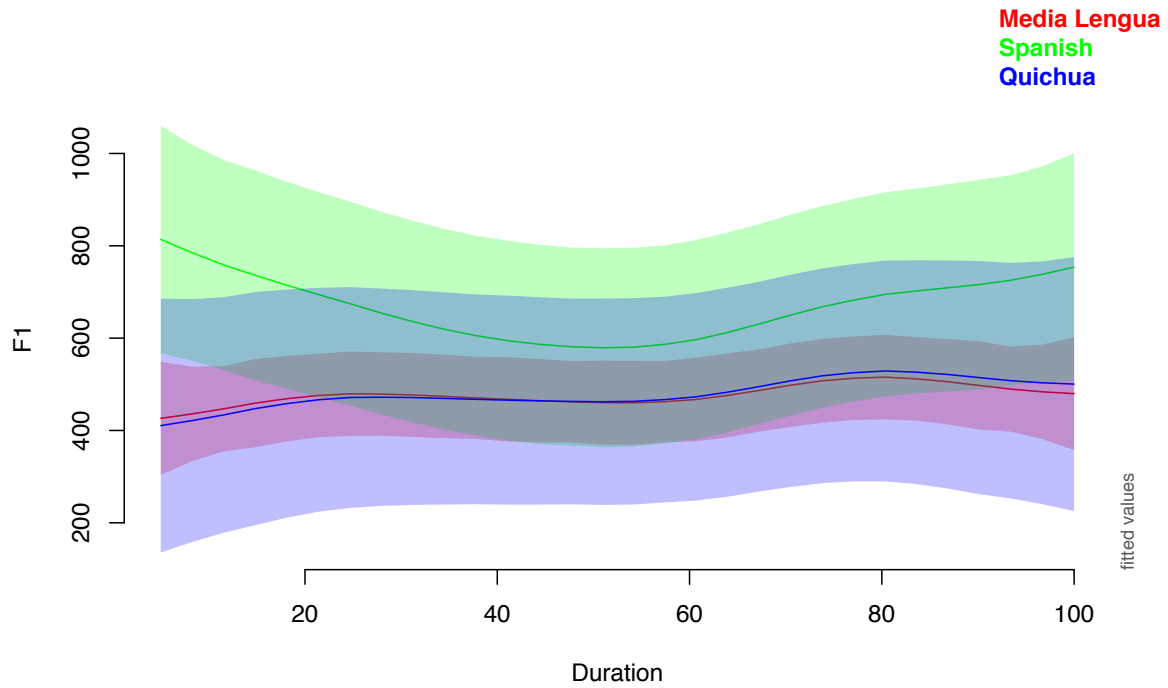
Media Lengua vowel space with /ia/ trajectory
Ellipses indicate 2 standard deviations; monophthong n = 2040



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

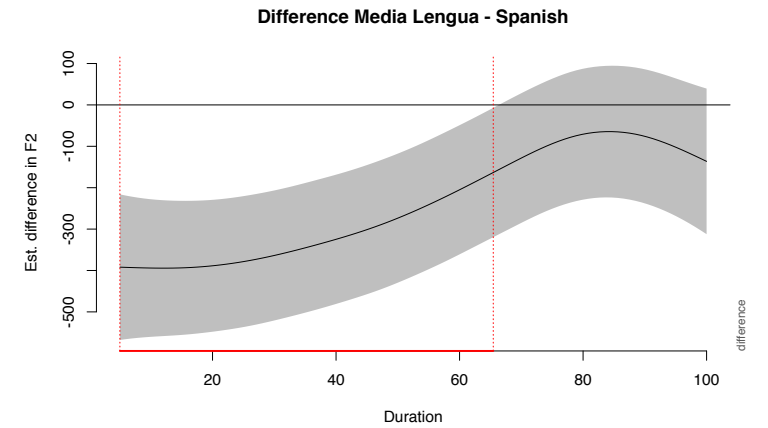
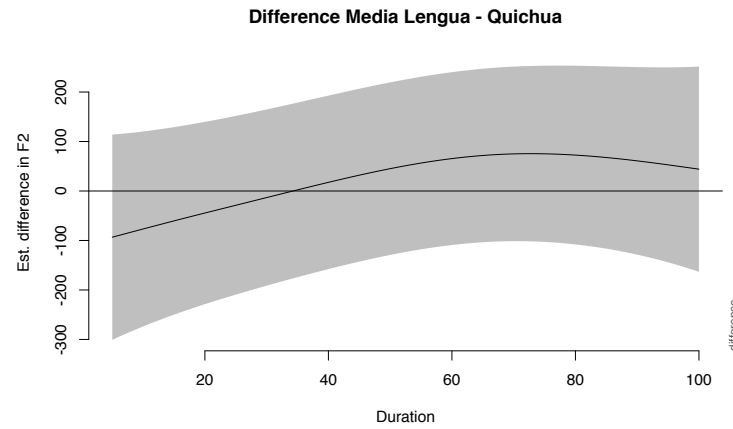
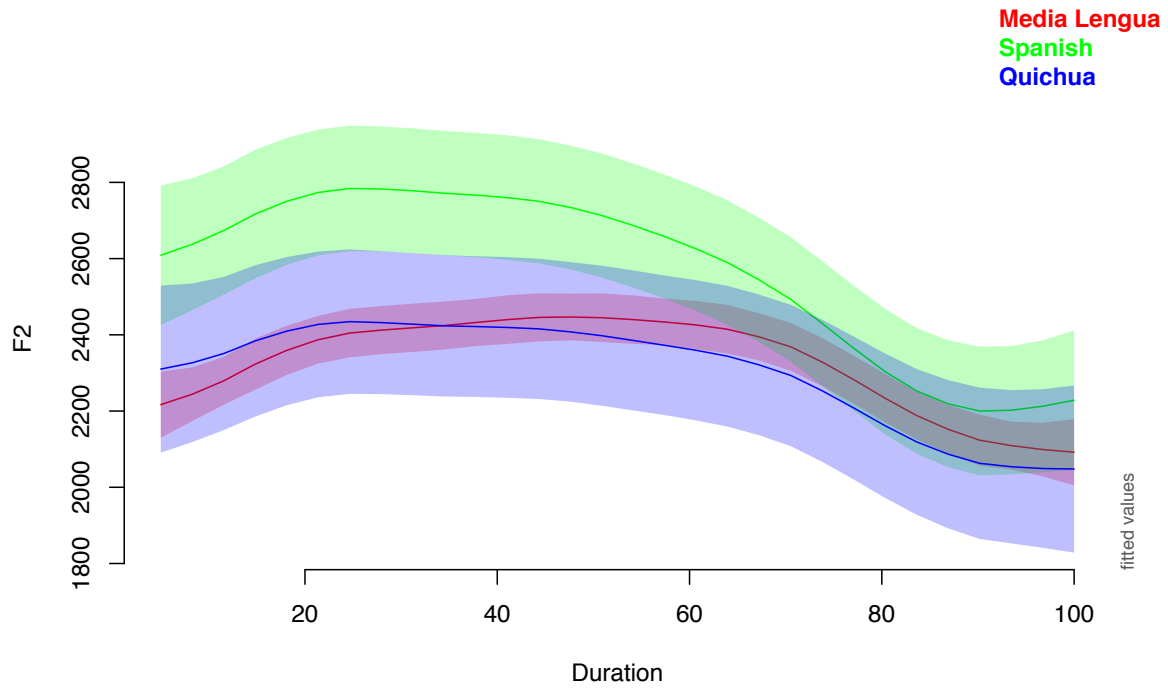


/ia/: F1 & F2



/ia/ F1

- Media Lengua vs. Quichua: no sig. difference
- Spanish: somewhat higher F1 = lower articulation, sig. difference at onset and offset

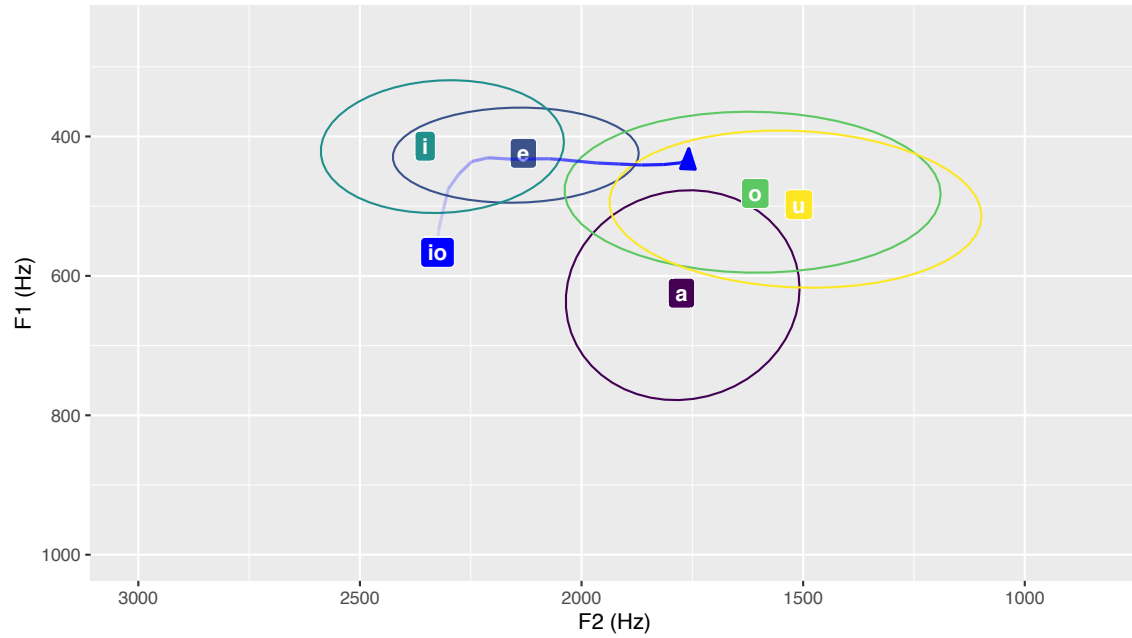


/ia/ F2

- Media Lengua vs. Quichua: no sig. difference
- Spanish: higher F2 = advanced articulation, over initial 2/3 of duration

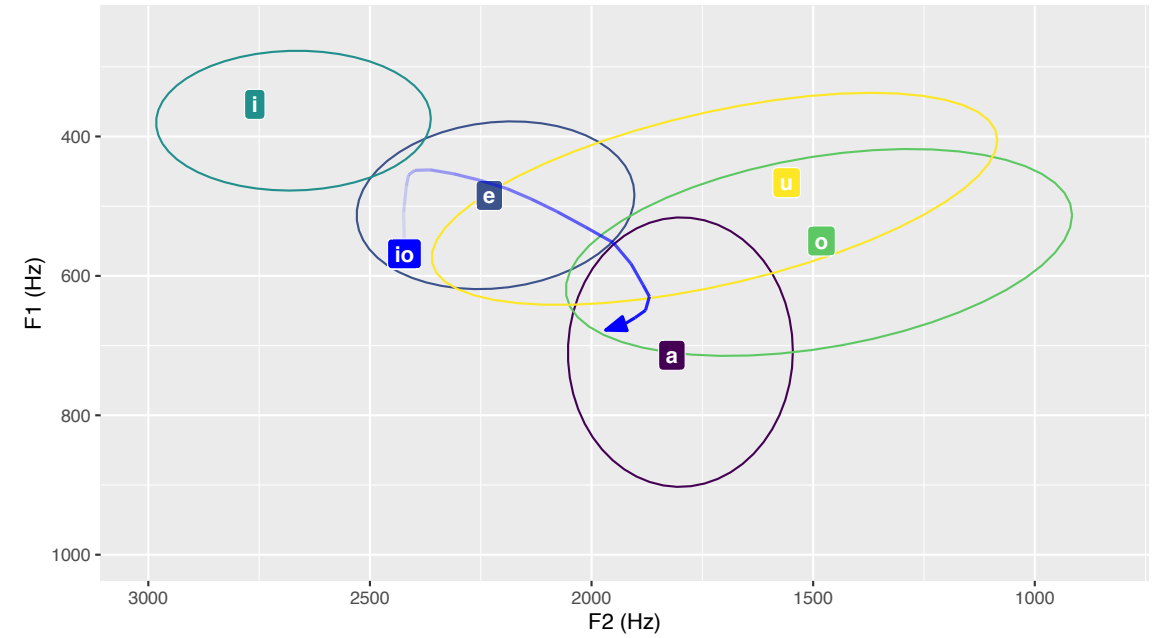
Media Lengua vowel space with /io/ trajectory

Ellipses indicate 2 standard deviations; monophthong n = 2040

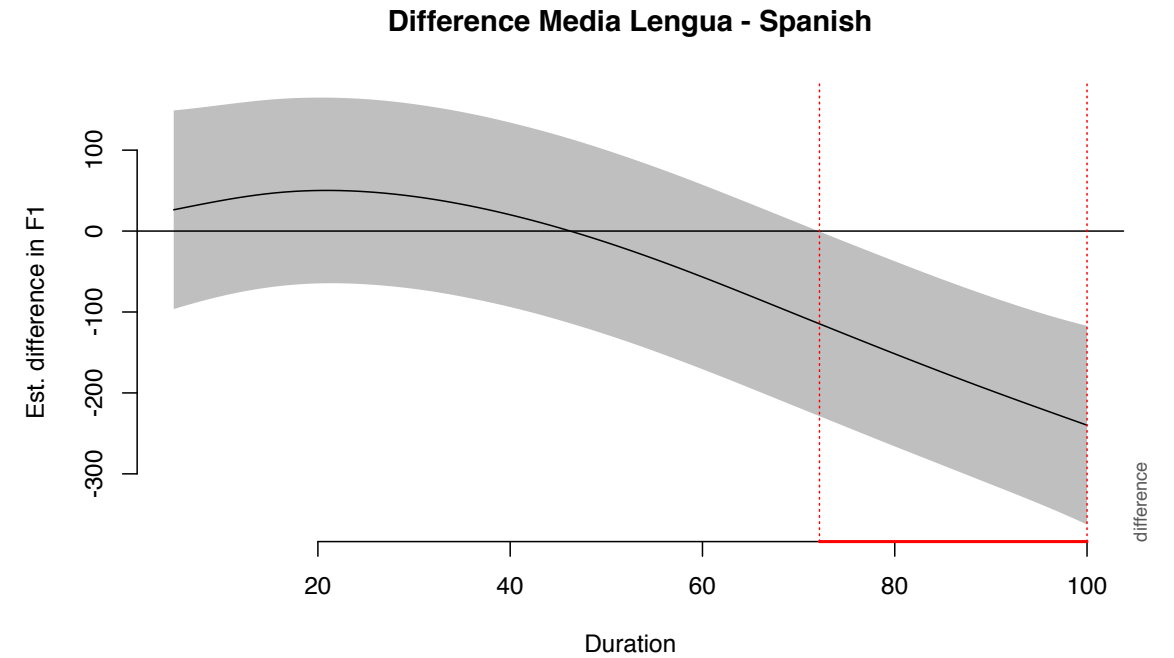
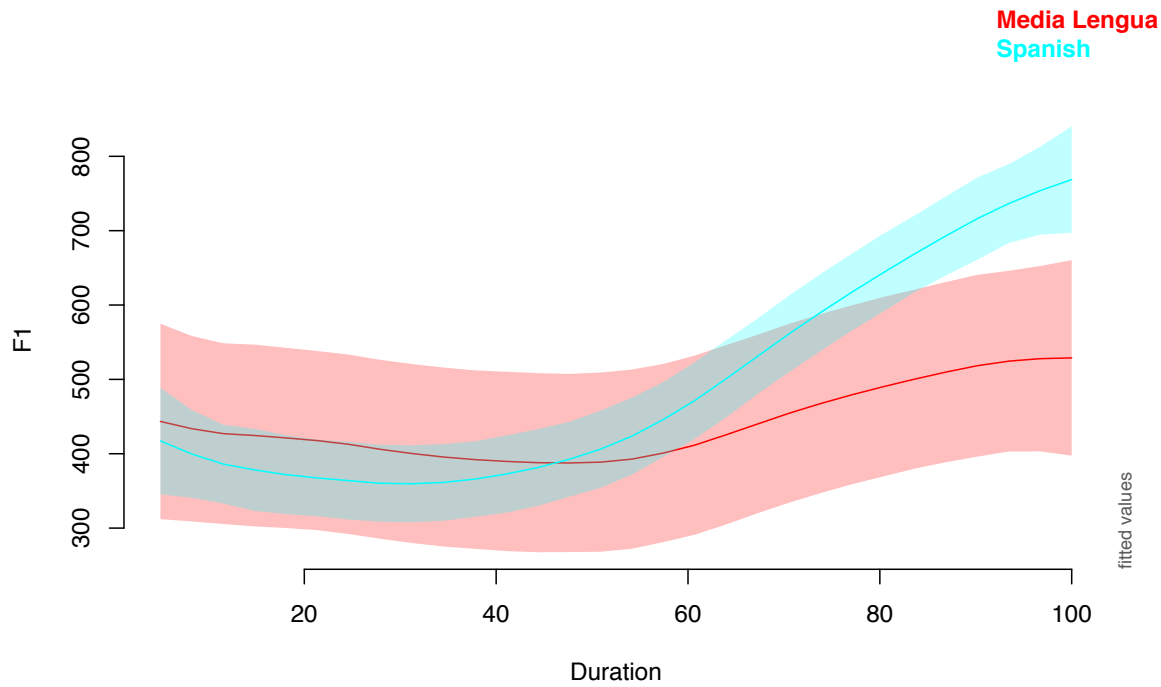


Spanish vowel space with /io/ trajectory

Ellipses indicate 2 standard deviations; monophthong n = 1929

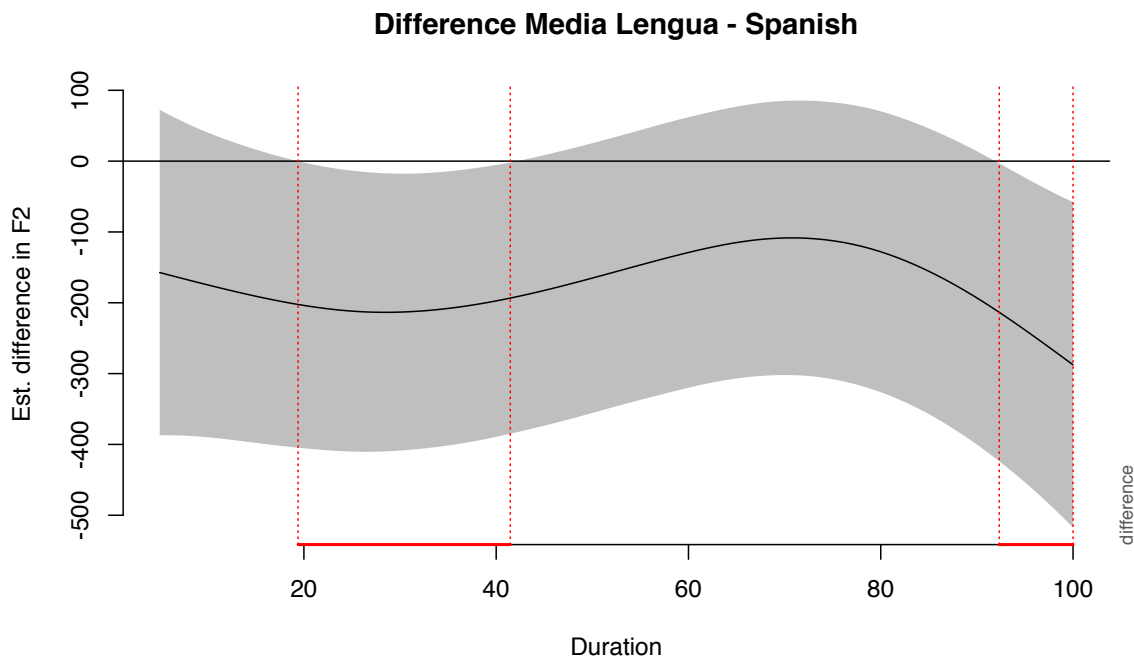
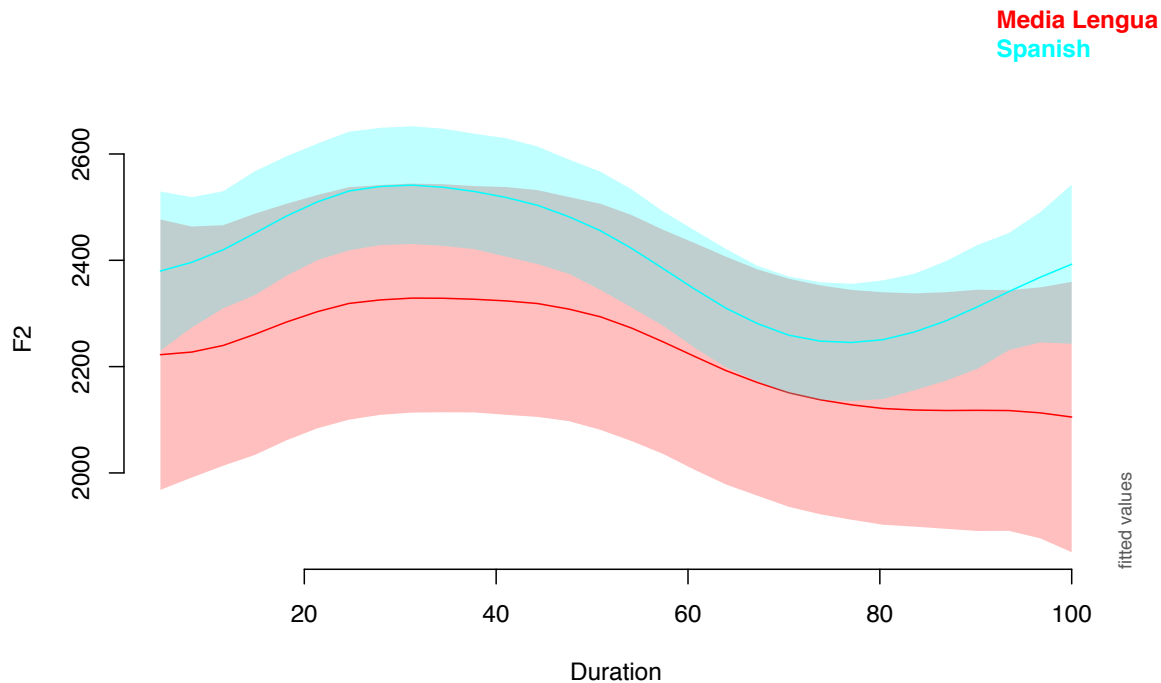


/io/: F1 & F2



/iɔ/ F1

- Media Lengua: F1 lower (higher articulation) during final 25% of duration



/iɔ/ F2

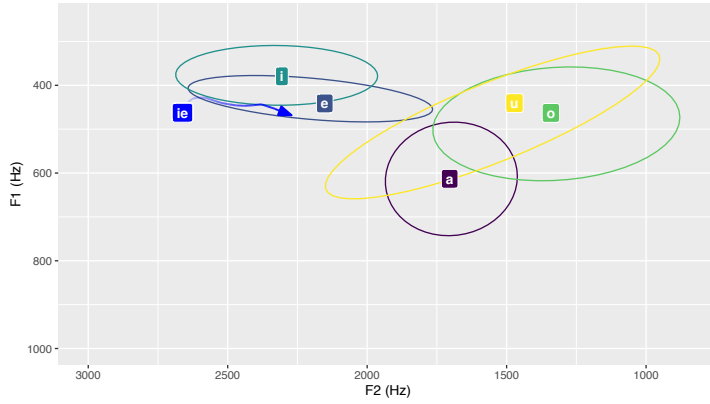
- Spanish: slightly higher F2 = advanced articulation, from 20–40% duration and at offset

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/

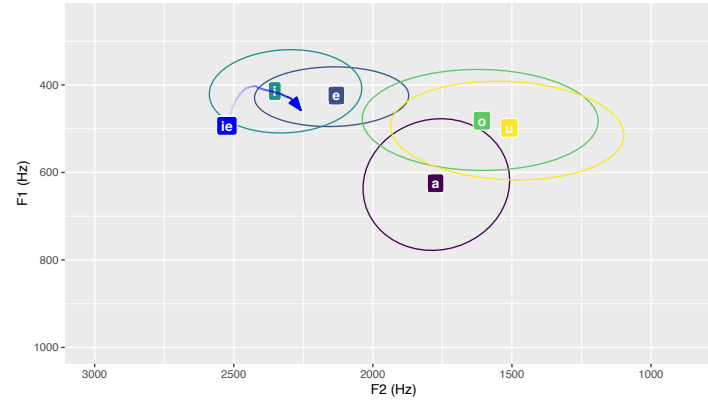
Quichua vowel space with /ie/ trajectory

Ellipses indicate 2 standard deviations; monophthong n = 634



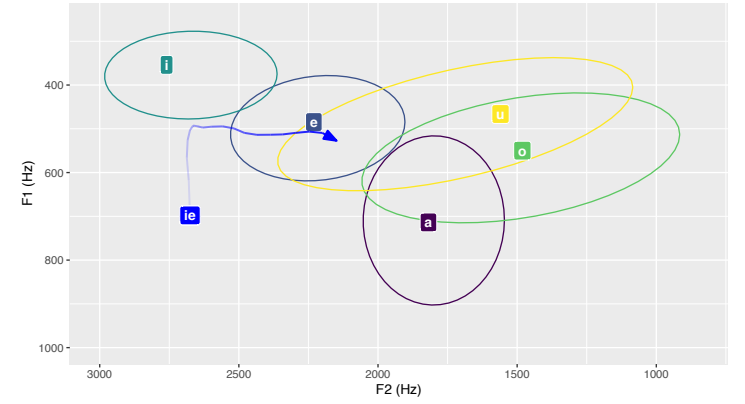
Media Lengua vowel space with /ie/ trajectory

Ellipses indicate 2 standard deviations; monophthong n = 2040

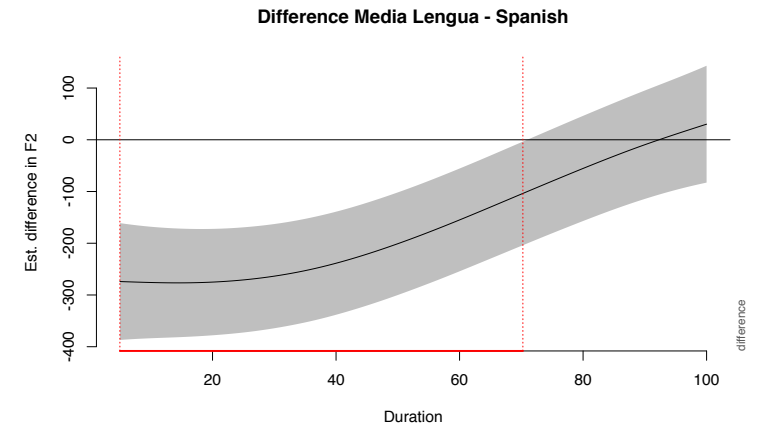
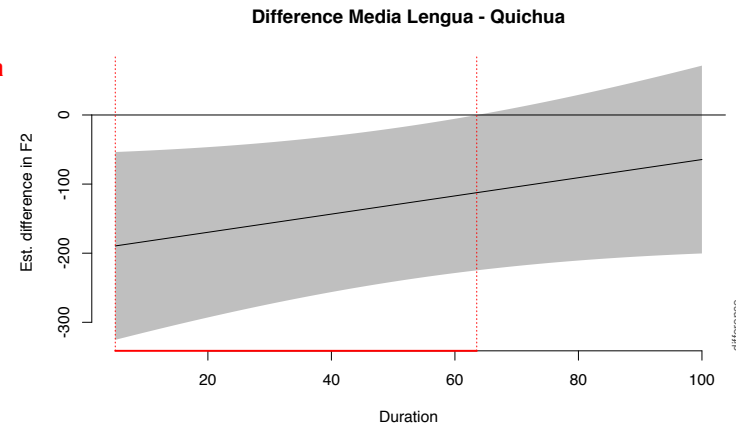
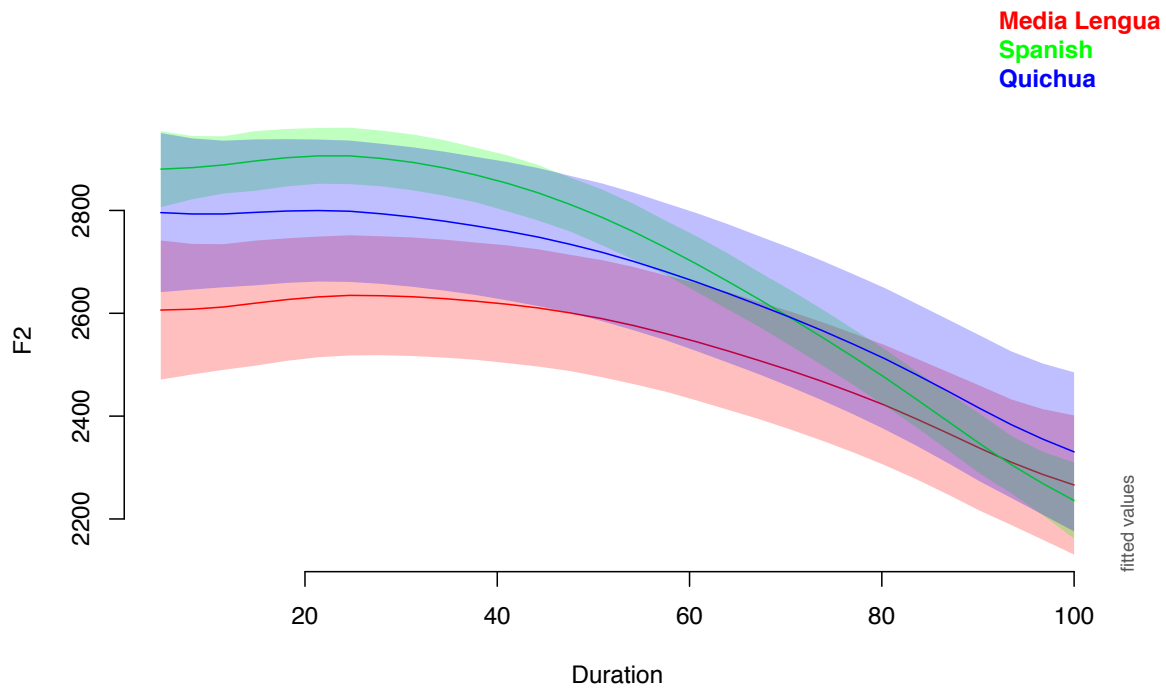


Spanish vowel space with /ie/ trajectory

Ellipses indicate 2 standard deviations; monophthong n = 1929



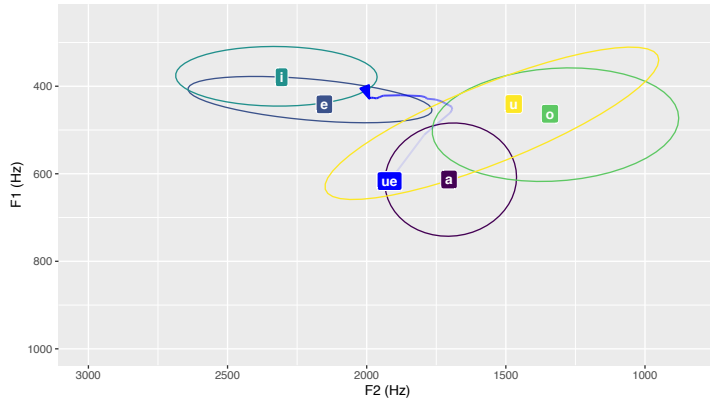
/ie/: F2



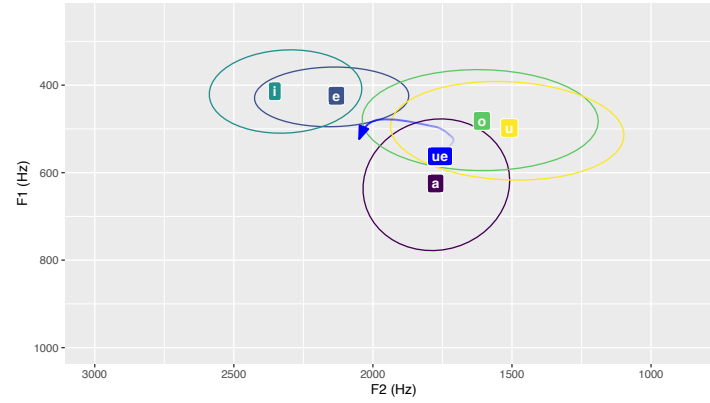
/ie/ F2

- 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

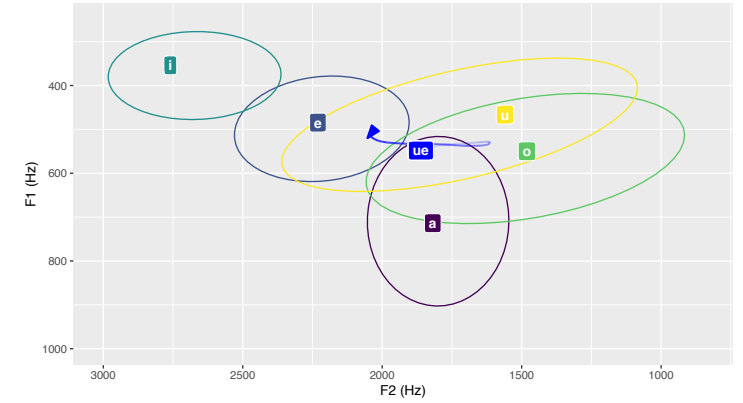
Quichua vowel space with /ue/ trajectory
Ellipses indicate 2 standard deviations; monophthong n = 634



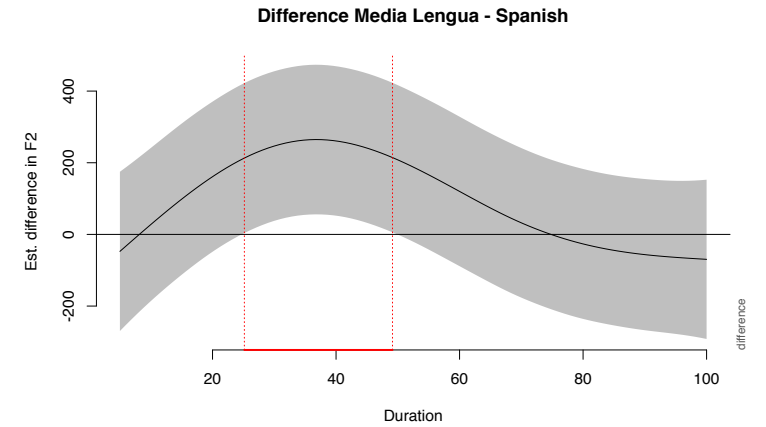
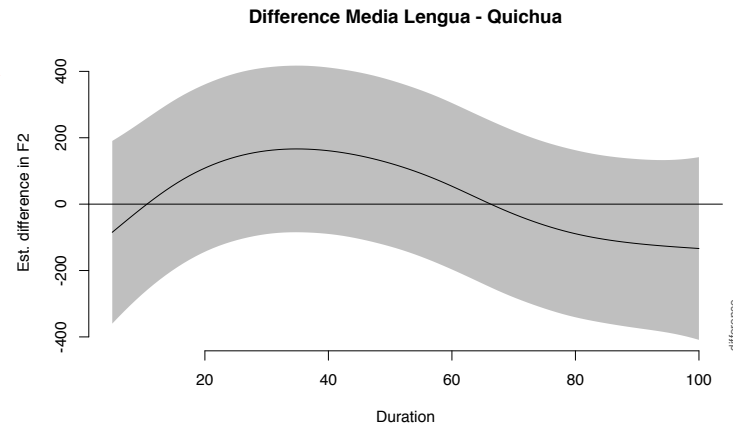
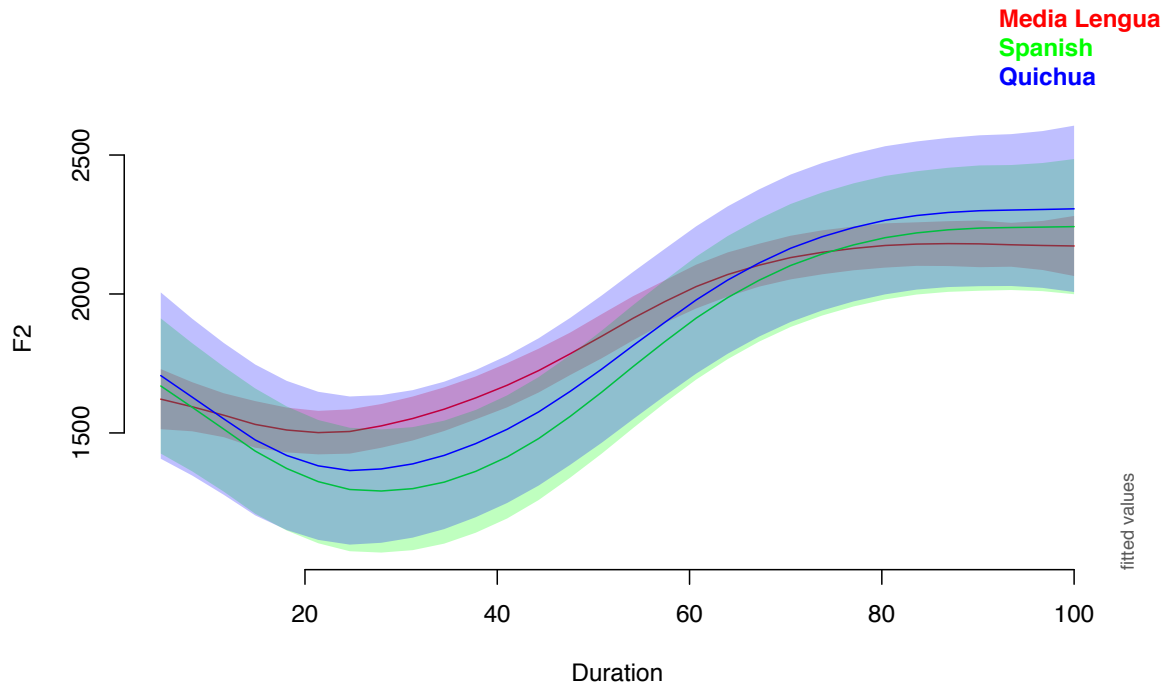
Media Lengua vowel space with /ue/ trajectory
Ellipses indicate 2 standard deviations; monophthong n = 2040



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



/ue/: F2



/ue/ F2

- 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-contrasting
 - 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*

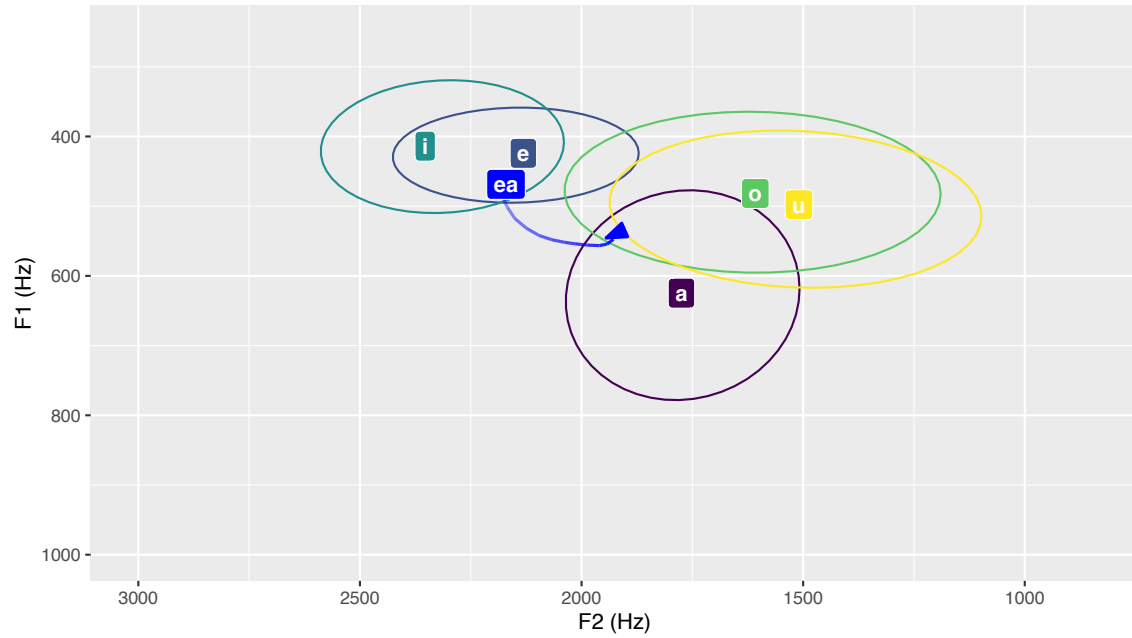
Media Lengua-Internal Diphthong Comparisons

- 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

	ai vs. ei	ea vs. ia	eo vs. eu	ue vs. ui
F1	$p \approx 0$	n.s.	$p \approx 0$	$p = 0.008$
F2	$p \approx 0$	$p \approx 0$	$p \approx 0$	$p = 0.002$

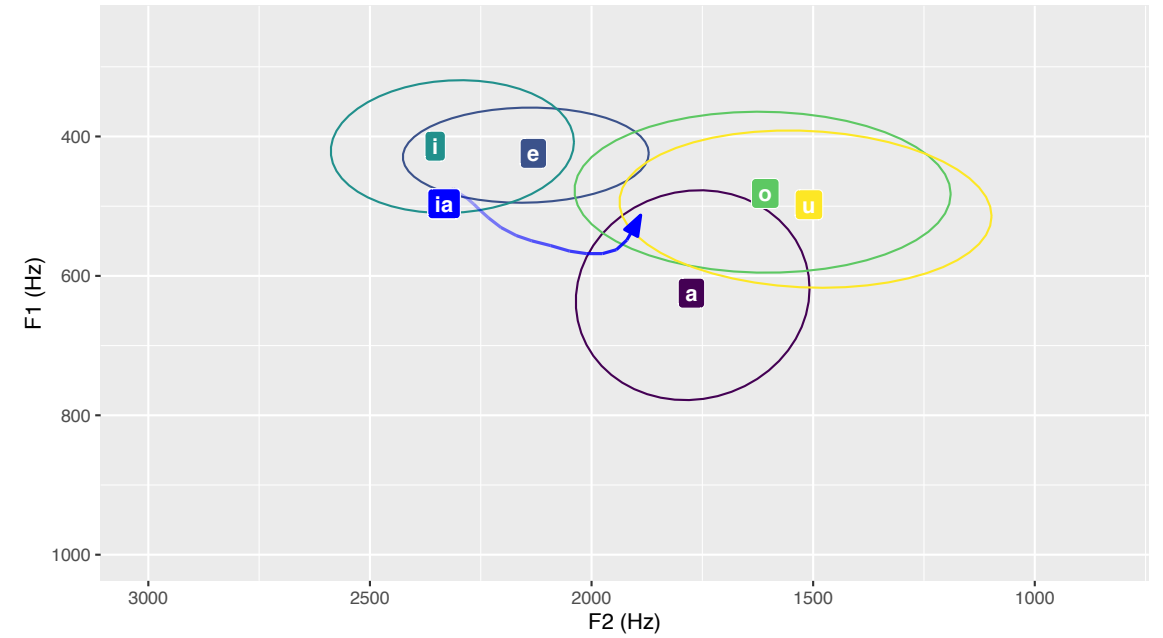
Media Lengua vowel space with /ea/ trajectory

Ellipses indicate 2 standard deviations; monophthong n = 2040

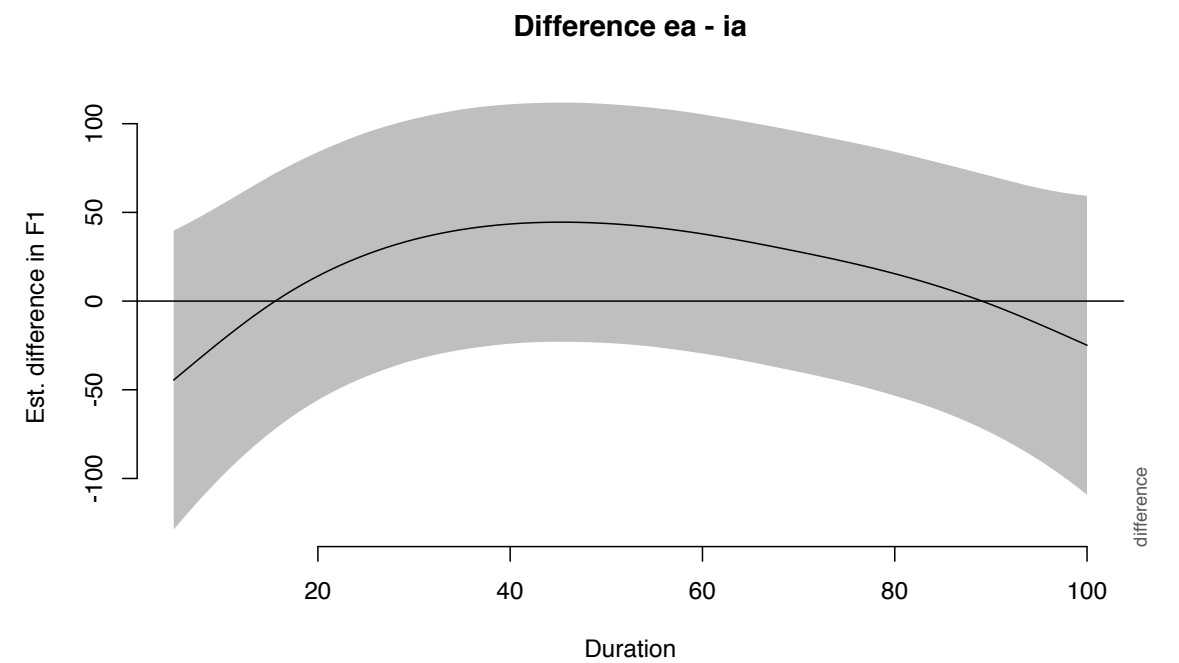
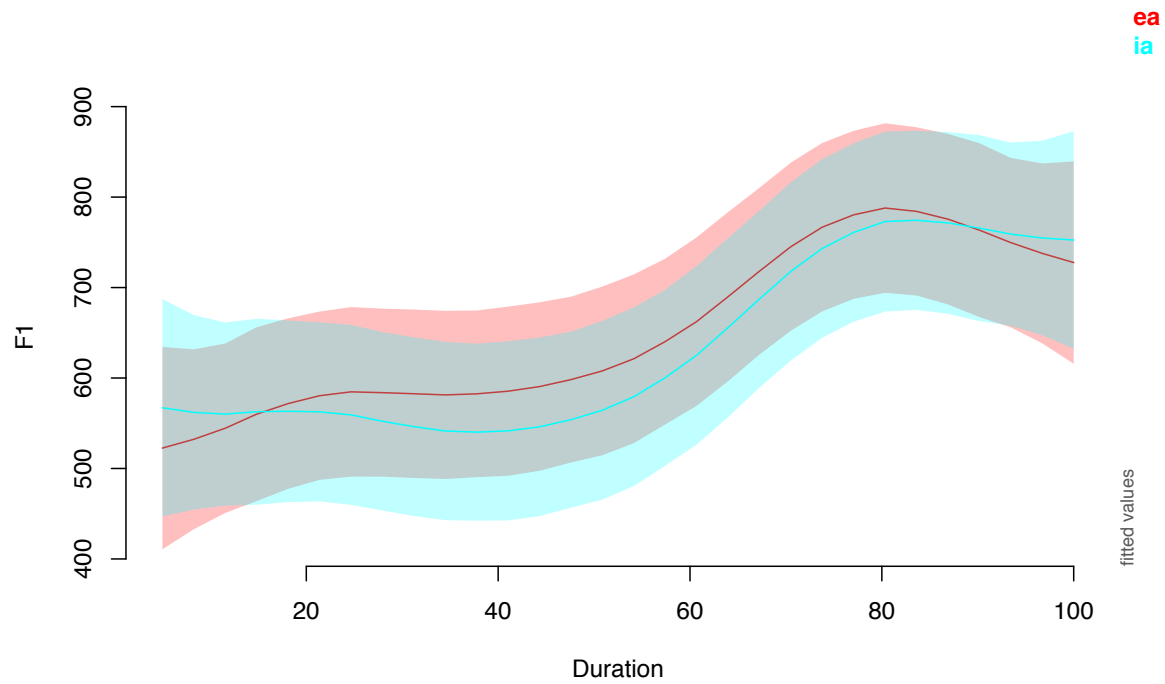


Media Lengua vowel space with /ia/ trajectory

Ellipses indicate 2 standard deviations; monophthong n = 2040



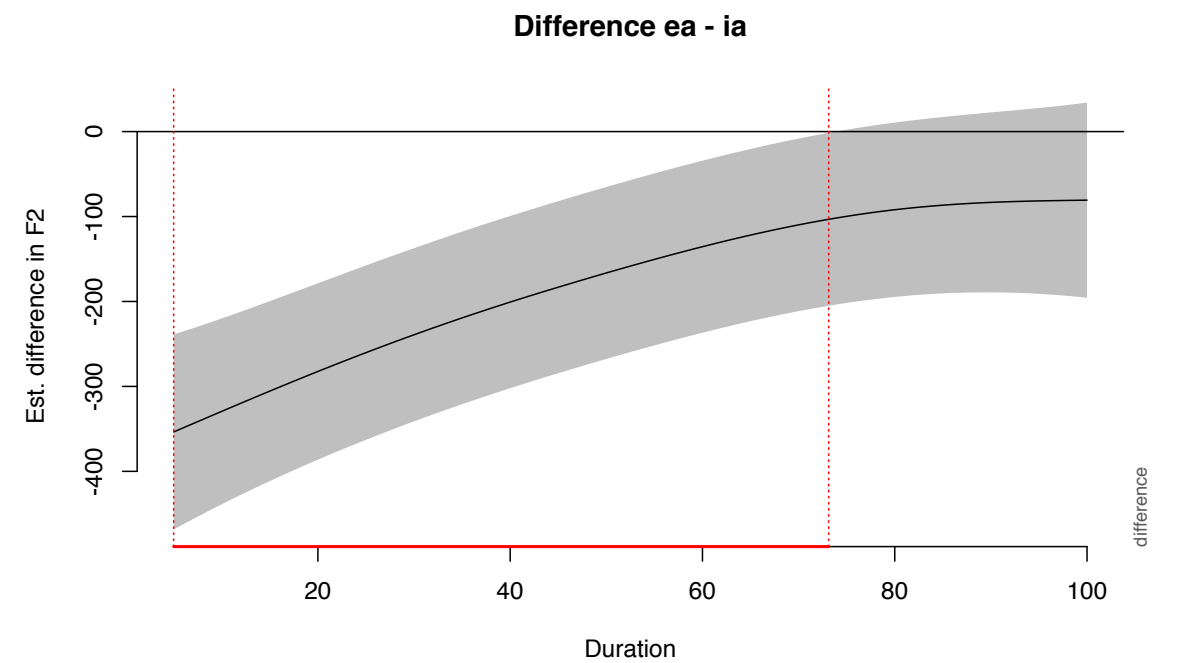
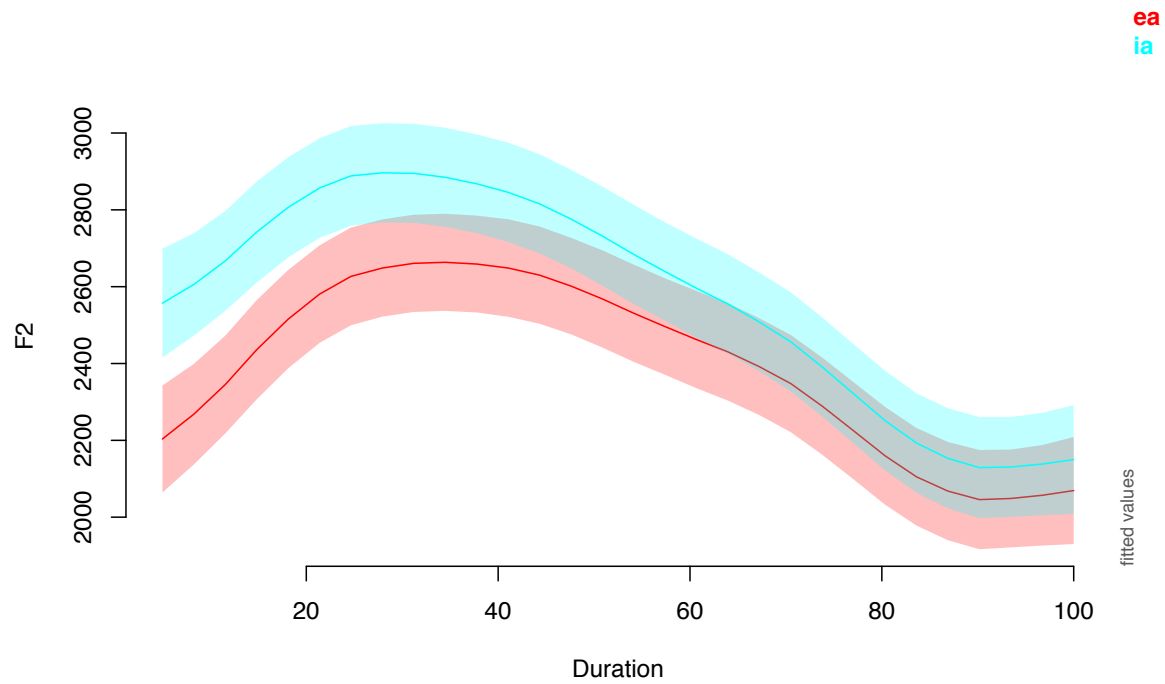
Media Lengua /ea/ vs. /ia/



Media Lengua

/ea, ia/ F1

- 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

Media Lengua-Internal Diphthong Comparisons

- 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

Media Lengua Comparisons by Source Language

- 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

Discussion: Why?

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).

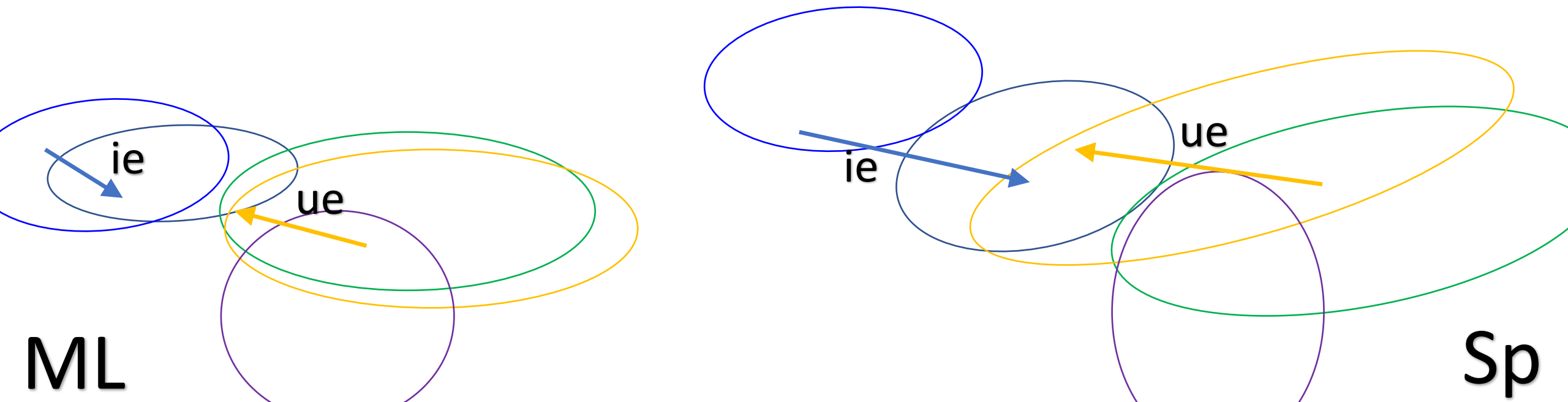
- Then the vowel system could have been subsequently nativized with the overlaps ‘frozen’ in place.
- Phonological ‘stresses’ from relexification (e.g., high functional loads of contrastive non-native phonemes) from the mid-vowels may have been driving forces for maintaining/creating contrasts with the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

1,415 ML Word Sample

Vowel	e	o	ea	ae	eo	eu	ue	oi	io	ei	ie	ao	oa	uo
Count	461	524	9	0	5	4	33	1	55	6	37	3	1	2
%	33%	37%	0.6%	0%	0.4%	0.3%	2.3%	0.07%	4%	0.4%	2.6%	0.2%	0.1%	0.1%
Count	985		147											
%	70%		10%											

Discussion: Diphthongs

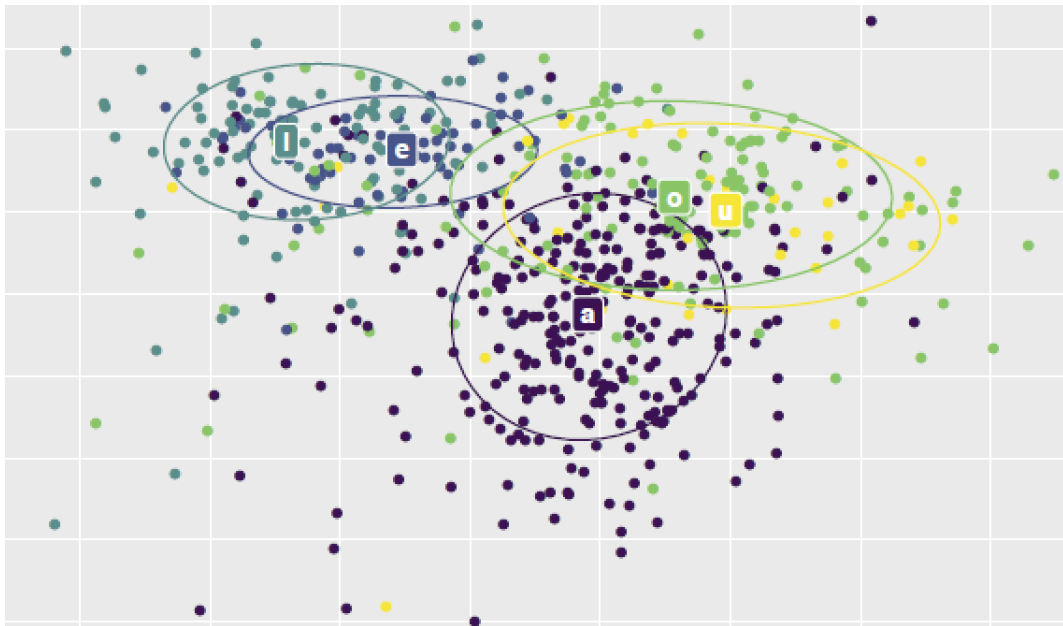
Diphthongs also appear to follow this tendency given that the formant trajectories in Media Lengua look strikingly similar to Quichua diphthongs (e.g., size, shape, variability). However, given that the mid- and high-vowels have ‘solidified’ into highly overlapping acoustic spaces, more reminiscent of 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

Vowel	e	o	i	u
Count	461	524	588	286
%	33%	37%	42%	20%
Count	985		788	
%	70%		62%	



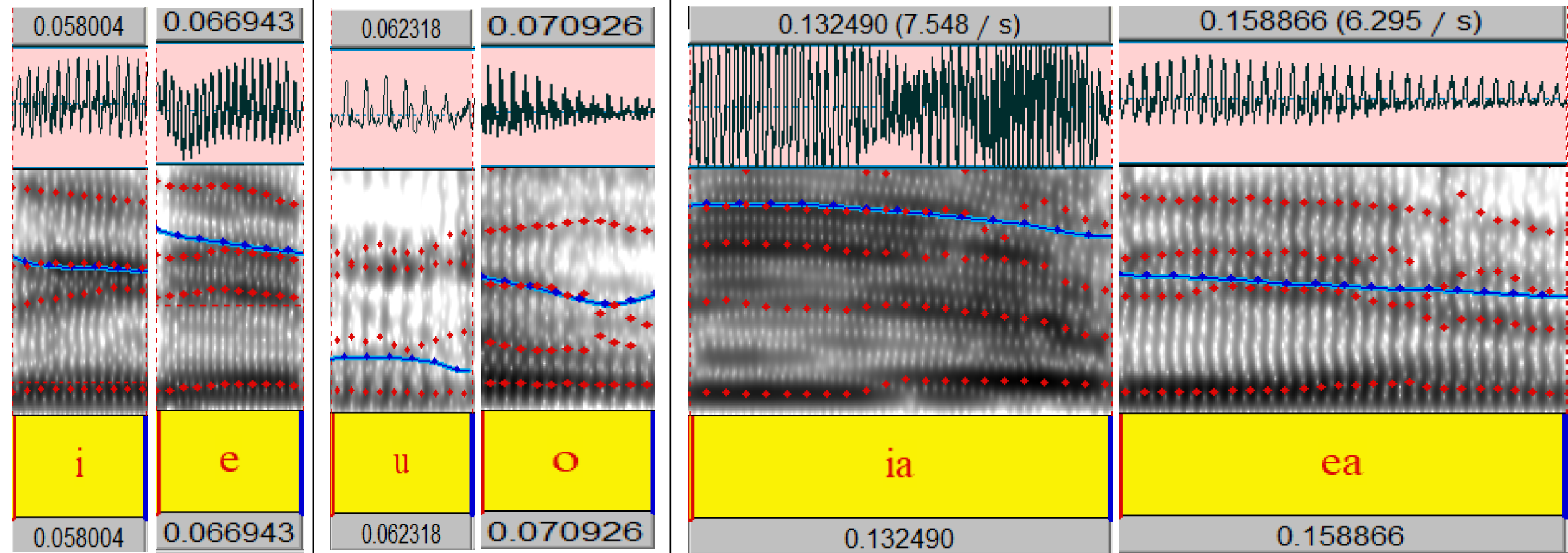
Cognitive load reaction time/ eye-tracking experiment



piscadu  pescado
piscado  piscadu

Future research

Vowel length



References

- Boersma, P., & Weenink, D. (2019). *Praat: doing phonetics by computer*. <http://www.fon.hum.uva.nl/praat/>
- Flege J (2007): Language contact in bilingualism: phonetic system interactions; in Cole J, Hualde J (eds): *Laborator Phonology 9*. Berlin, Mouton de Gruyter, pp 353–380.
- Hastie, T. J., & Tibshirani, R. J. (1990). *Generalized Additive Models*. New York: Chapman and Hall.
- Jones, C., Meakins, F., & Buchan, H. (2011). Comparing vowels in Gurindji Kriol and Katherine English: Citation speech data. *Australian Journal of Linguistics*, 31, 305–326.
- Livijn P (2000): Acoustic distribution of vowels in differently sized inventories – hot spots or adaptive dispersion? PERILUS 23
- Onosson, S., & Bird, S. (2019). Differences in Vowel-Glide Production Between L1 and L2 Speakers of Hul'q'umi'num'. *ICPhS19*, pp. 1–5. https://assta.org/proceedings/ICPhS2019/papers/ICPhS_1033.pdf
- Rosen, N., Stewart, J., & Sammons, O. (accepted). A comparative analysis of Michif, Métis French, and Cree vowel spaces. *Journal of the Acoustical Society of America*.
- R Core Team. (2019). *R: A language and environment for statistical computing*. <https://www.r-project.org/>
- RStudio Team. (2019). *RStudio: Integrated Development for R*. <https://www.rstudio.com/>
- Sóskuthy, M. (2017). *Generalised Additive Mixed Models For Dynamic Analysis In Linguistics: A Practical Introduction*. https://github.com/soskuthy/gamm_intro
- Stewart, J. (2014). A comparative analysis of Media Lengua and Quichua vowel production. *Phonetica*, 7, 159–182. <https://doi.org/10.1159/000369629>
- Stewart, J. (2018). Vowel perception by native Media Lengua, Quichua, and Spanish speakers. *Journal of Phonetics*, 71, 177–193. <https://doi.org/10.1016/j.wocn.2018.08.005>
- Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.
- Wood, S. N. (2017) *Generalized Additive Models: An Introduction with R* (2nd edition). Chapman and Hall/CRC.
- Xu, Y., & Gao, H. (2018). FormantPro as a Tool for Speech Analysis and Segmentation. *Revista De Estudos Da Linguagem*, 26(4), 1435-1454. <https://doi.org/10.17851/2237-2083.26.4.1435-1454>