

Documenting endangered varieties: Eastern Basque (Zuberoan & Mixean)

Ander Egurtzegi (CNRS-IKER)

Advanced course in Sound Change with a focus on Basque

University of Chicago, 2025/03/31

Phonological inventory

Standard Basque — Euskara Batua (Hualde 2003)

		labial		apical		laminal	predorsal		postdorsal
		bilabial	labio-dental	apico-dental	apico-alveolar		palato-alveolar	palatal	
stop	voiceless	p		t				c	k
	voiced	b		d				ɟ	g
	fricative		f		s	ʃ	ʃ		x
	affricate				ts		tʃ		
	nasal	m			n			ɲ	
	lateral				l			ʎ	
	tap				r				
	trill				r				
	front	central	back						
	high	i		u					
mid	e			o					
low		a							

Zuberoan Basque (Hualde 1993; Egurtzegi 2018)

		labial		apical		laminal	predorsal		postdorsal	laryngeal
		bilabial	labio-dental	apico-dental	apico-alveolar	alveolar	palato-alveolar	palatal		
stop	voiceless	p		t				c	k	
	aspirated	p ^h		t ^h				c ^h	k ^h	
	voiced	b		d				ɟ	g	
Fricative	voiceless		f		s	ʃ	ʃ			h/ħ
	voiced				z	ʒ	ʒ			
	voiceless				ts	tʃ	tʃ			
affricate	voiced				dʒ					
	nasal	m			n			ɲ		
	lateral				l			ɬ		
	tap				r					
	trill				r					
	glides							j	w	
high	front	i/ĩ	y/ỹ							
	mid	e								
	low		a/ã							

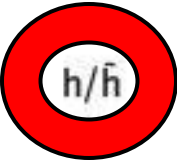
Zuberoan Basque — Zuberera

		labial		apical		laminal	predorsal		postdorsal	laryngeal
		bilabial	labio-dental	apico-dental	apico-alveolar	alveolar	palato-alveolar	palatal		
stop	voiceless aspirate	p ^h		t ^h				c ^h		k ^h
	voiced			d						
Fricative	voiceless		f		s	ʃ	ç			h/ħ
	voiced				z	ʒ	ʝ			
affricate	voiceless				tʃ	tʃ	tʃ			
	voiced				dʒ					
nasal	nasal	m			n			ɲ		
	lateral				l			ʎ		
	tap				r					
	trill				r					
	glides							j	w	
high	front	i/ĩ	y/ỹ							
	mid	e								
	low			a/ã		u/ũ				

Zuberoan Basque — Zuberera

		labial		apical		laminal	predorsal		postdorsal	laryngeal
		bilabial	labio-dental	apico-dental	apico-alveolar	alveolar	palato-alveolar	palatal		
stop	voiceless	p		t				c	k	
	aspirated	p ^h		t ^h				c ^h	k ^h	
	voiced	b		d				tʃ	g	
Fricative	voiceless		f		s		ʃ			h/ħ
	voiced				z	ʒ	ʒ			
affricate	voiceless									
	voiced				dz					
	nasal	m			n			ɲ		
	lateral				l			ʎ		
	tap				r					
	trill				r					
	glides							j	w	
high	front	i/ĩ	y/ỹ							
	mid	e								
	low		a/ã							

Zuberoan Basque — Zuberera

		labial		apical		laminal	predorsal		postdorsal	laryngeal
		bilabial	labio-dental	apico-dental	apico-alveolar	alveolar	palato-alveolar	palatal		
stop	voiceless	p		t				c	k	
	aspirated	p ^h		t ^h				c ^h	k ^h	
	voiced	b		d				ɟ	g	
Fricative	voiceless		f		s	ɬ	ʃ			
	voiced				z	ʒ	ʒ			
affricate	voiceless				ts	tɬ	tʃ			
	voiced				dʒ					
	nasal	m			n			ɲ		
	lateral				l			ɭ		
	tap				r					
	trill				r					
	glides							j	w	
high	front	i/ĩ	y/ỹ							
	mid	e								
	low		a/ã							

Zuberoan Basque — Zuberera

		labial		apical		laminal	predorsal		postdorsal	laryngeal
		bilabial	labio-dental	apico-dental	apico-alveolar	alveolar	palato-alveolar	palatal		
stop	voiceless aspirated	p ^h		t ^h				c ^h	k ^h	
	voiced	b		d				ɟ	g	
	Fricative	voiceless		f		s	ʃ			h/ħ
	voiced				z	ʒ				
affricate	voiceless				t͡s	t͡ʃ				
	voiced				d͡z	d͡ʒ				
	nasal	m			n			ɲ		
	lateral				l			ʎ		
	tap				r					
	trill				r					
	glides							j	w	
high mid low	front	i/ī								
		y								
	central									
					back					
					u/ũ					
					o					
			a/ã							

Zuberoan Basque — Zuberera

		labial		apical		laminal	predorsal		postdorsal	laryngeal
		bilabial	labio-dental	apico-dental	apico-alveolar	alveolar	palato-alveolar	palatal		
stop	voiceless	p		t				c	k	
	aspirated	p ^h		t ^h				c ^h	k ^h	
	voiced	b		d				tʃ	g	
Fricative	voiceless		f		s	ʃ	ʃ			h/ħ
	voiced				z	ʒ	ʒ			
affricate	voiceless				ts	tʃ	tʃ			
	voiced				dʒ					
nasal	nasal	m			n			ɲ		
	lateral				l			ʎ		
	tap				r					
	trill				r					
	glides							j	w	
high mid low		ɪ	ʏ			u				
		e				o				
			ã							

Bearnese Gascon (Mooney 2014)

	Bilabial	Labio-dental	Alveolar	Post-alveolar	Palatal	Velar	Glottal
Plosive	p b		t d			k g	
Affricate					$\widehat{t\zeta}$ $\widehat{d_j}$		
Nasal	m		n		ɲ	ŋ	
Trill			r				
Tap			ɾ				
Fricative		f v	s z	ʃ (ʒ)			h
Approximant					j	w	
Lateral			l		ʎ		

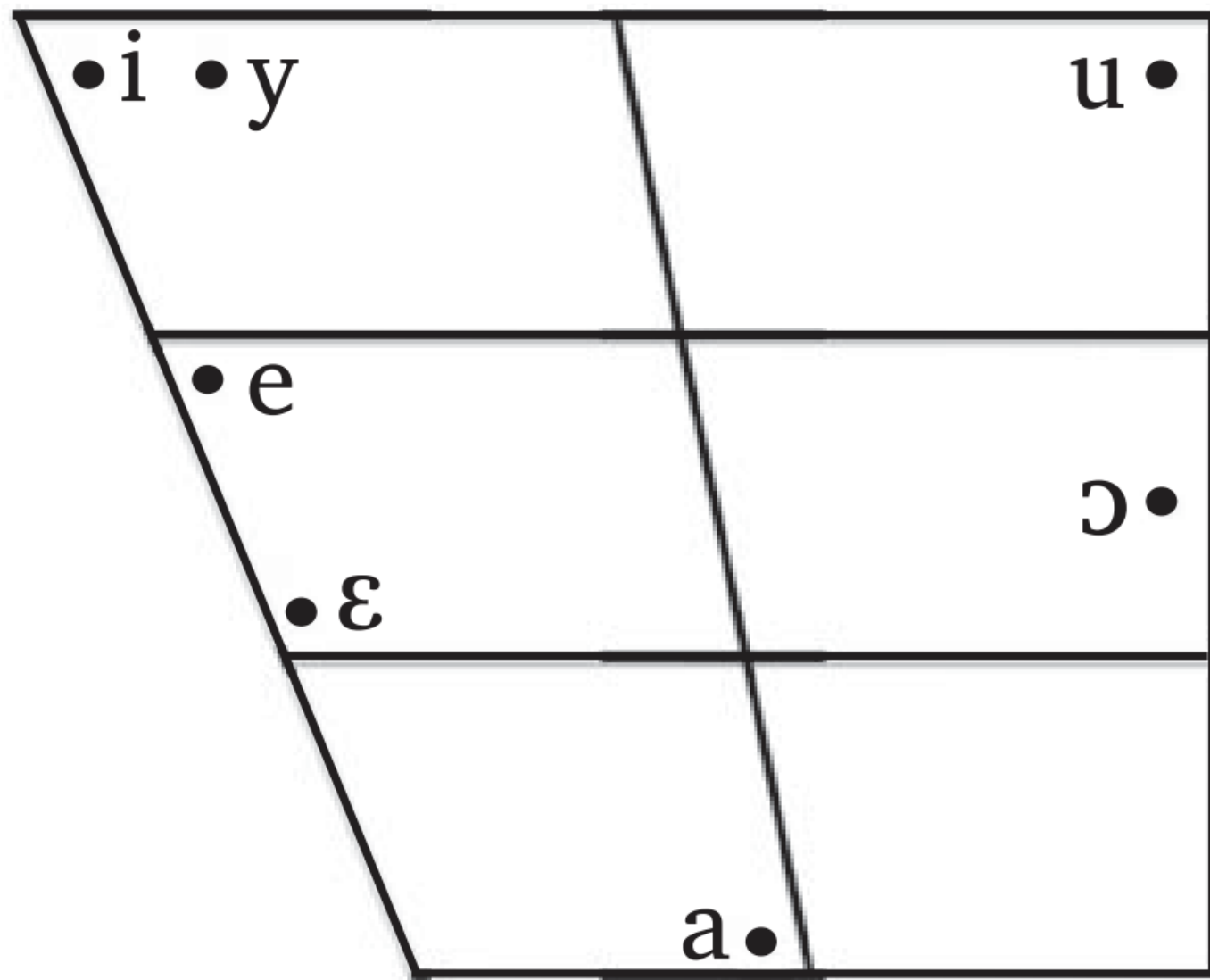
Bearnese Gascon (Mooney 2014)

	Bilabial	Labio-dental	Alveolar	Post-alveolar	Palatal	Velar	Glottal
Plosive	p b		t d			k g	
Affricate					t͡ʃ d͡ʒ		
Nasal	m		n		ɲ	ŋ	
Trill			r				
Tap			ɾ				
Fricative		f v	s z	ʃ ʒ			h
Approximant					j	w	
Lateral			l		ʎ		

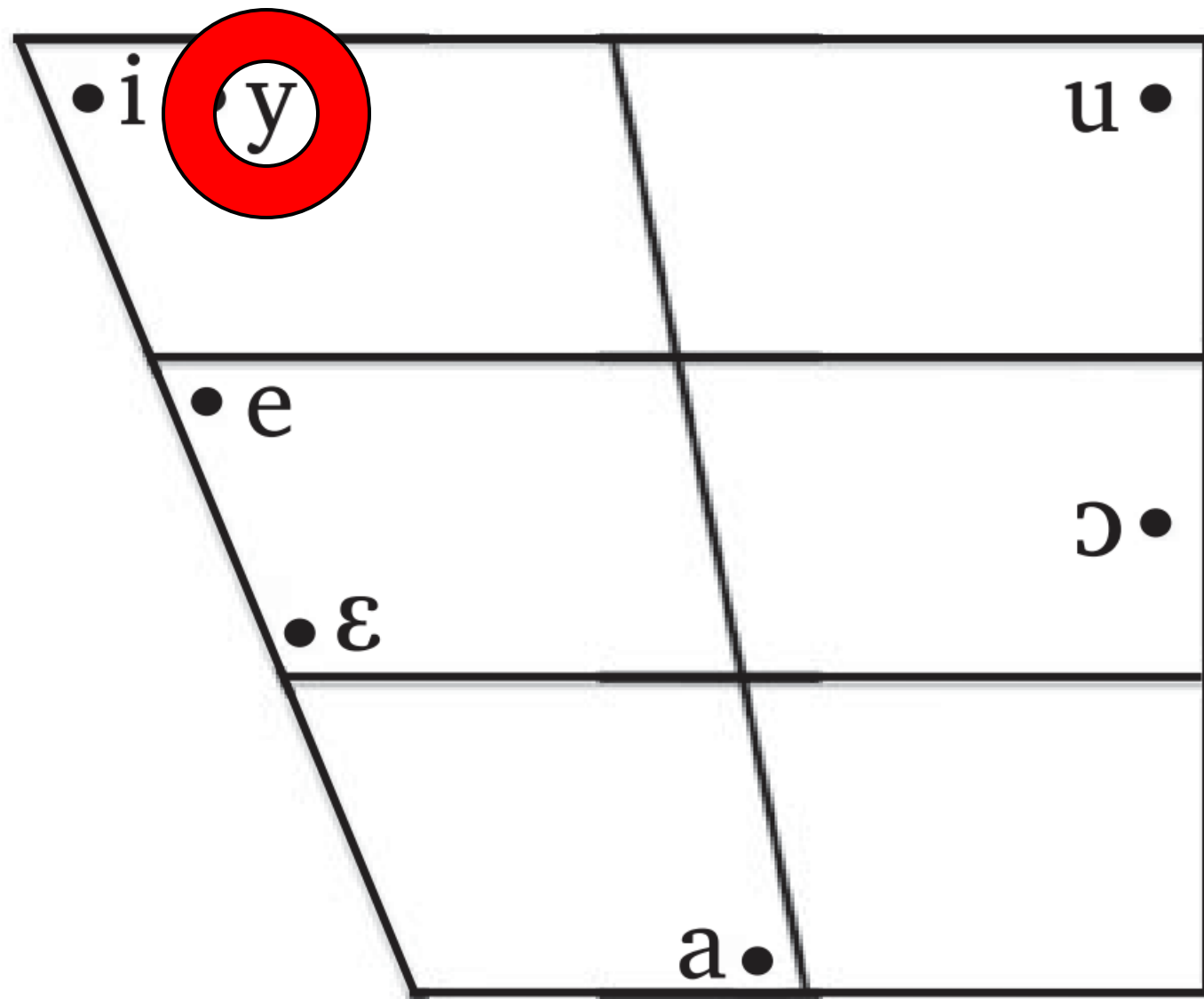
Bearnese Gascon (Mooney 2014)

	Bilabial	Labio-dental	Alveolar	Post-alveolar	Palatal	Velar	Glottal
Plosive	p b		t d			k g	
Affricate					t͡ʃ d͡ʒ		
Nasal	m		n		ɲ	ŋ	
Trill			r				
Tap			ɾ				
Fricative		f v	s z	ʃ (ʒ)			h
Approximant					j	w	
Lateral			l		ʎ		

Bearnese Gascon (Mooney 2014)



Bearnese Gascon (Mooney 2014)

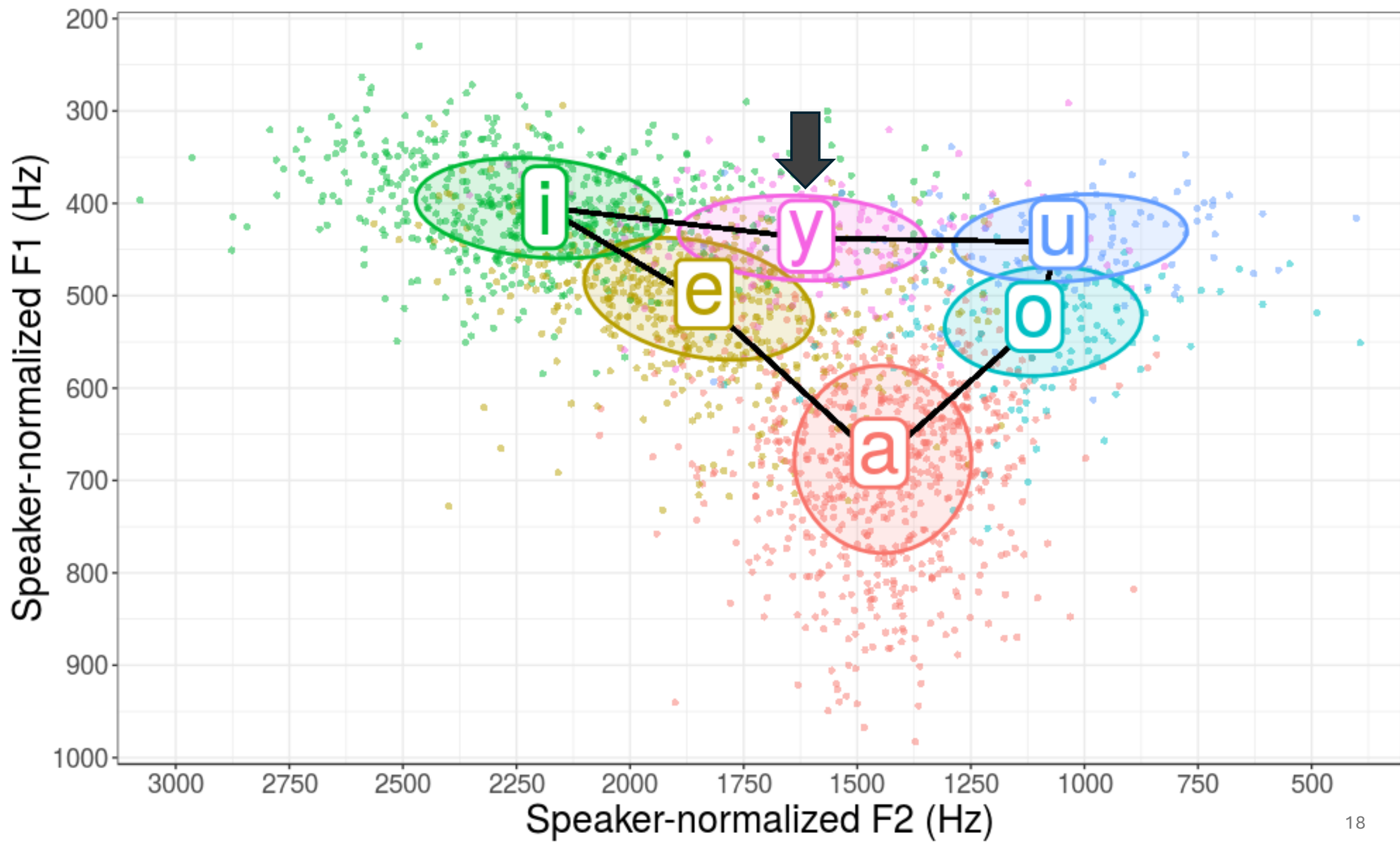


Phonetic analysis

Vowels

Oral vowels

- Six oral vowel qualities:
 - High: /i, y, u/
 - Mid: /e, o/
 - Low: /a/

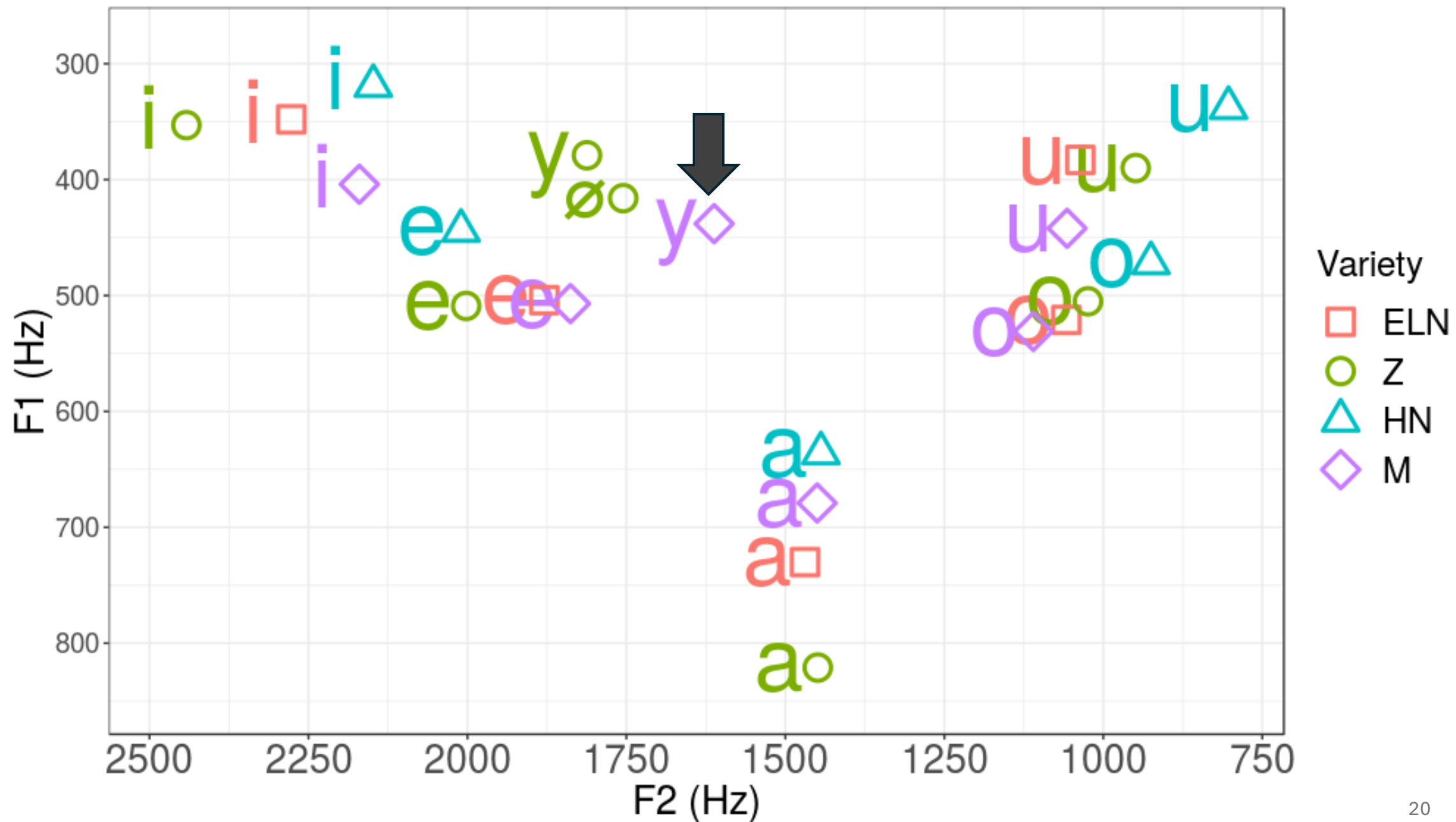


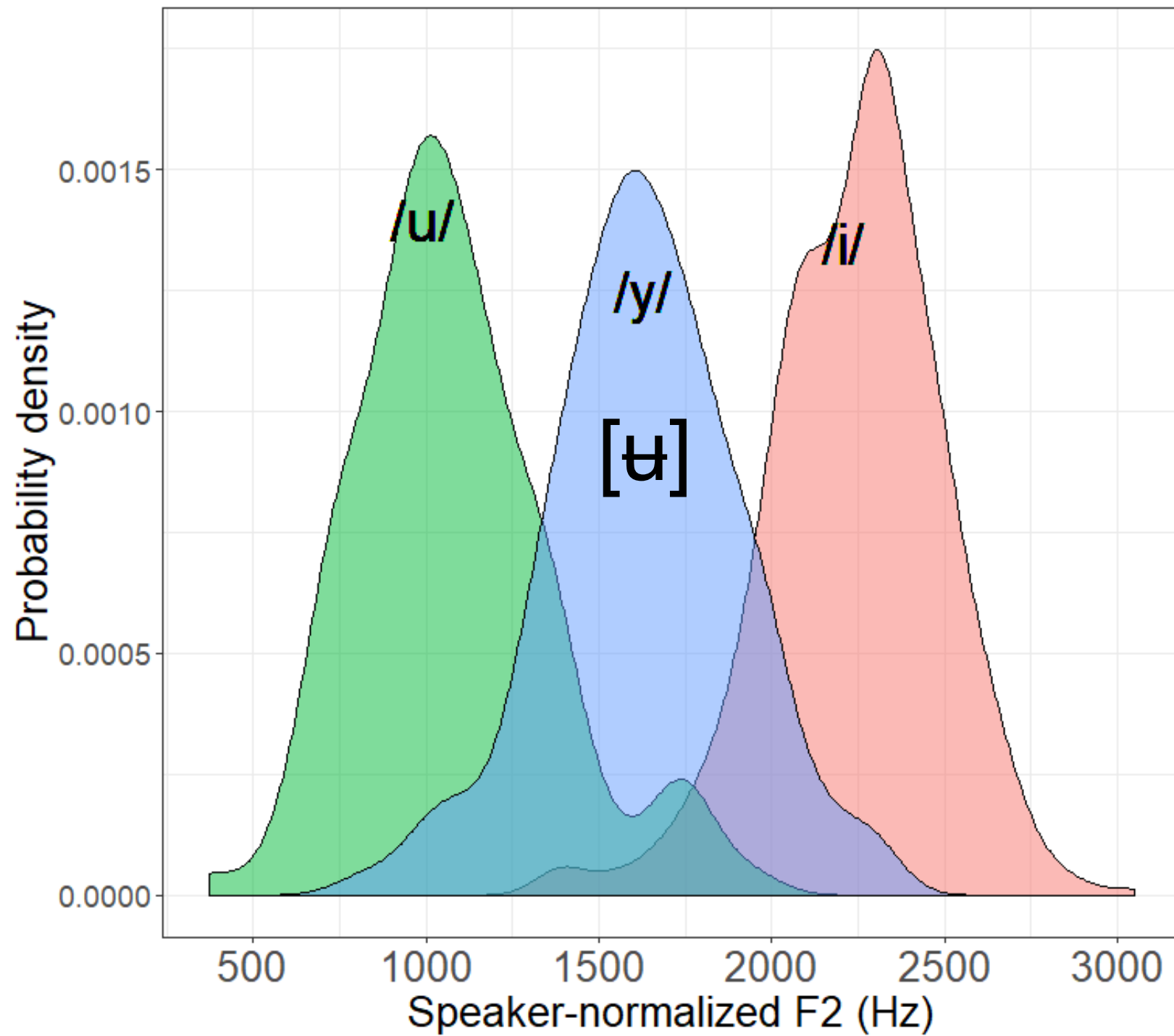
/y/ or /ɥ/ in Zuberoan and Mixean?

Table 1: Vowel formant summary from Urrutia et al. (1995), Pagola (1992), and the current study.

Variety	Vowel													
	/i/		/e/		/a/		/o/		/u/		/y/		/ø/	
	F1	F2	F1	F2	F1	F2	F1	F2	F1	F2	F1	F2	F1	F2
East. Low Navarrese (Urrutia et al., 1995)	348	2277	504	1879	730	1469	521	1058	383	1036				
Zuberoan (Urrutia et al., 1995)	353	2442	509	2002	821	1449	505	1024	390	949	379	1812	416	1755
High Navarrese (Pagola, 1992)	319	2148	444	2010	636	1444	472	925	338	803				
Mixean	404	2170	507	1838	679	1450	531	1110	442	1057	438	1612		

Acoustic vowel spaces in four varieties of Basque





Nasal vowels in Zuberoan (Egurtzegi 2015)

- 4 nasal vowel qualities:
 - High: /ĩ, ã, õ/
 - Low: /ã/
- No mid vowels:
 - /ẽ, õ/ raised to /ĩ, õ/

Table 5. Contrastively nasalized vowels in the Zuberoan inherited lexicon.

Zuberoan	Trans.	Std. Bsq.	Gloss
<i>ardũ</i>	/ar.ˈdũ/	<i>ardo</i>	‘wine’
<i>hügũ</i>	/hy.ˈgỹ/	<i>higuin</i>	‘repugnance’
<i>gorriũ</i>	/go.ˈrjũ/	—	‘reddish mushroom’
<i>hazkũ</i>	/haʂ.ˈkũ/	<i>azko(i)n</i>	‘badger’
<i>Xiberũ</i>	/ʃi.be.ˈrũ/	<i>Zuberoa</i>	‘(region)’

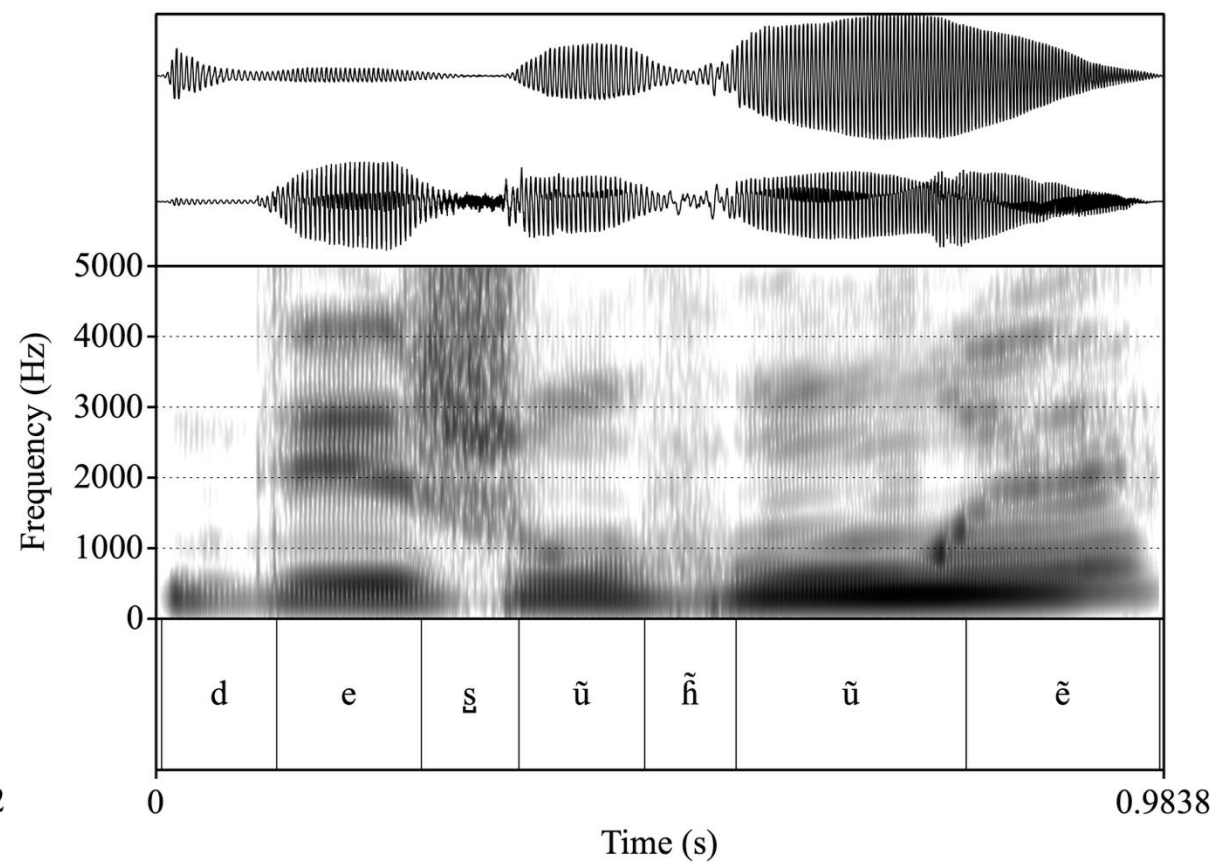
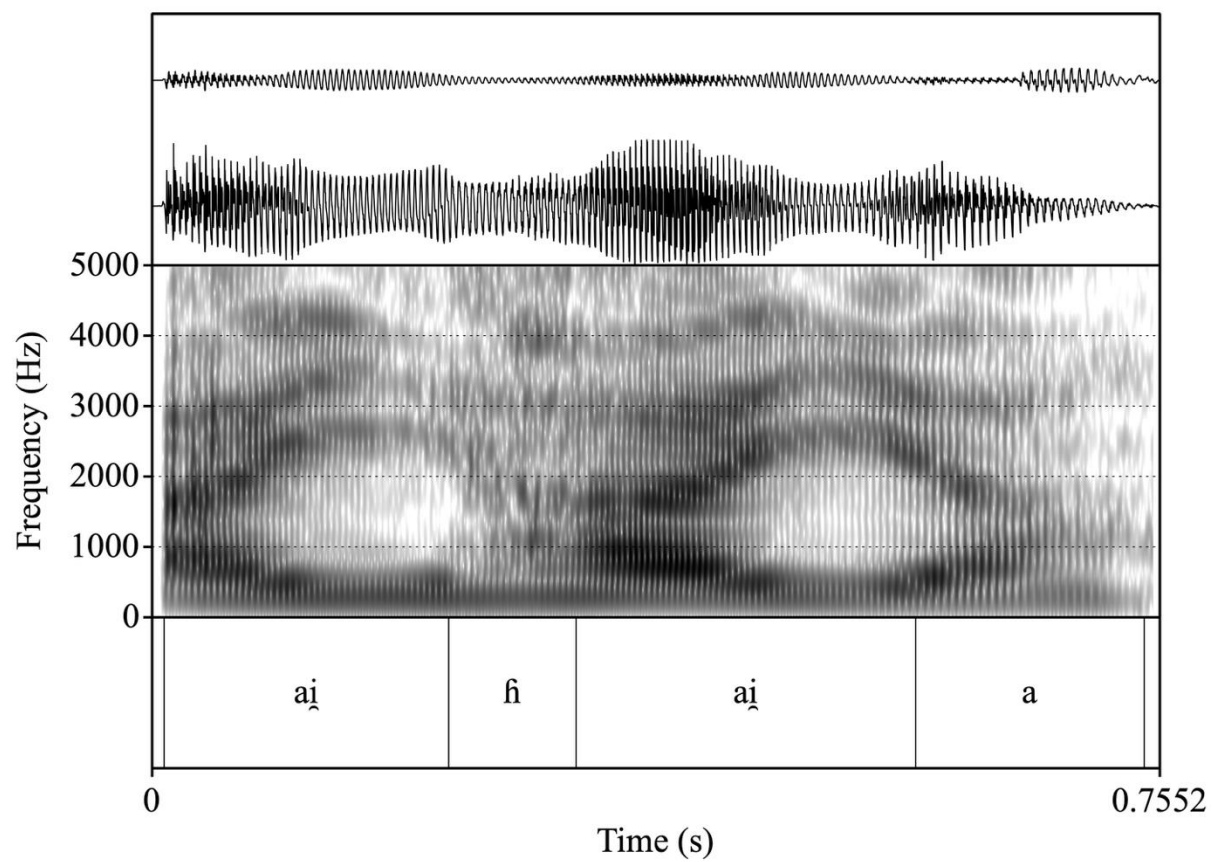
Nasal vowels in Bearnese Gascon

ĩ	<i>bî</i> [bĩ]	‘wine’	ã	<i>pâ</i> [pã]	‘bread’
ỹ	<i>û</i> [ỹ]	‘one’	ũ	<i>boû</i> [bũ]	‘good’
ẽ	<i>hé</i> [hẽ]	‘hay’			

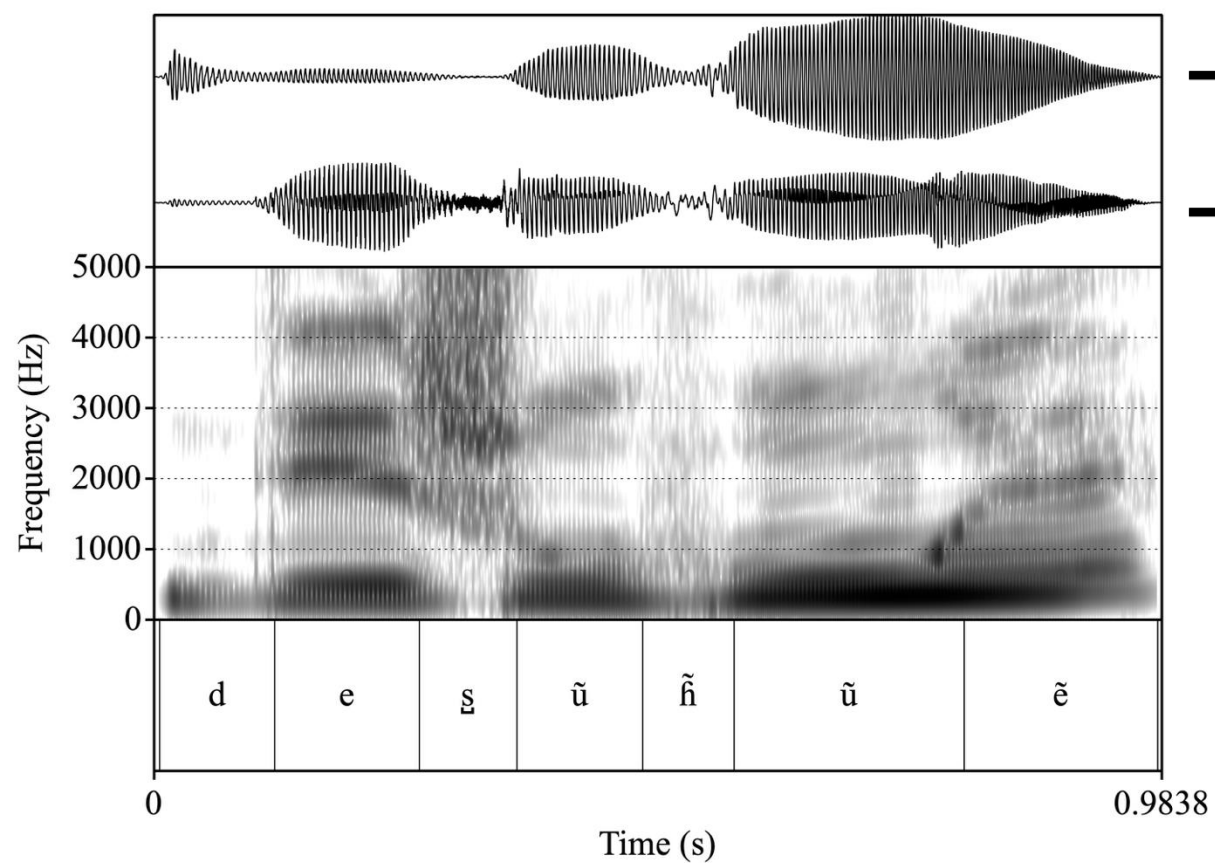
Vowel nasality is receding

- Mixean does not maintain it.
- Some Zuberoan varieties don't show it any longer.
- Do younger speakers produce contrastively nasalize vowels?
 - No studies.

Nasalance



Nasalance



Amplitude

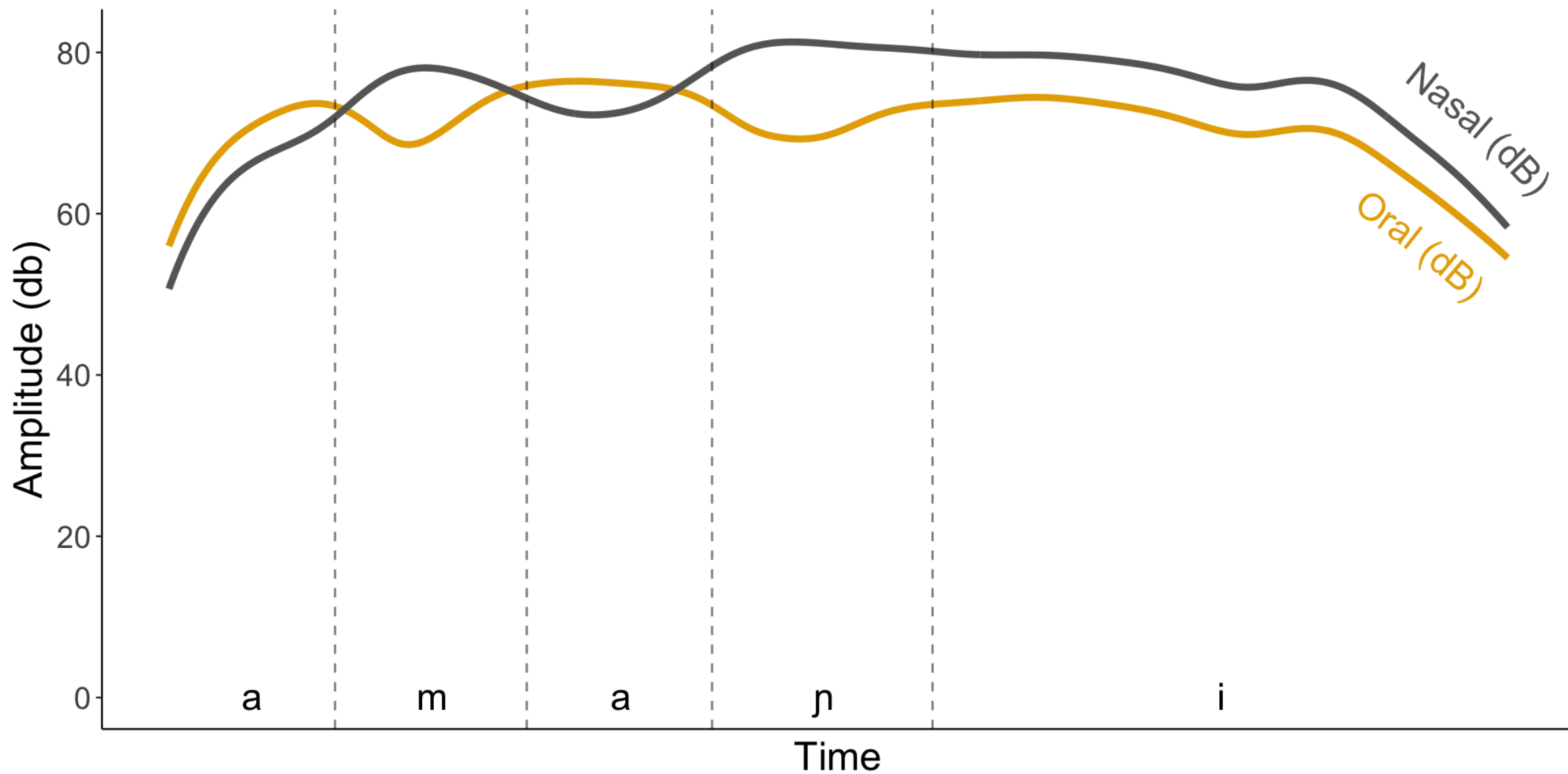
A_n

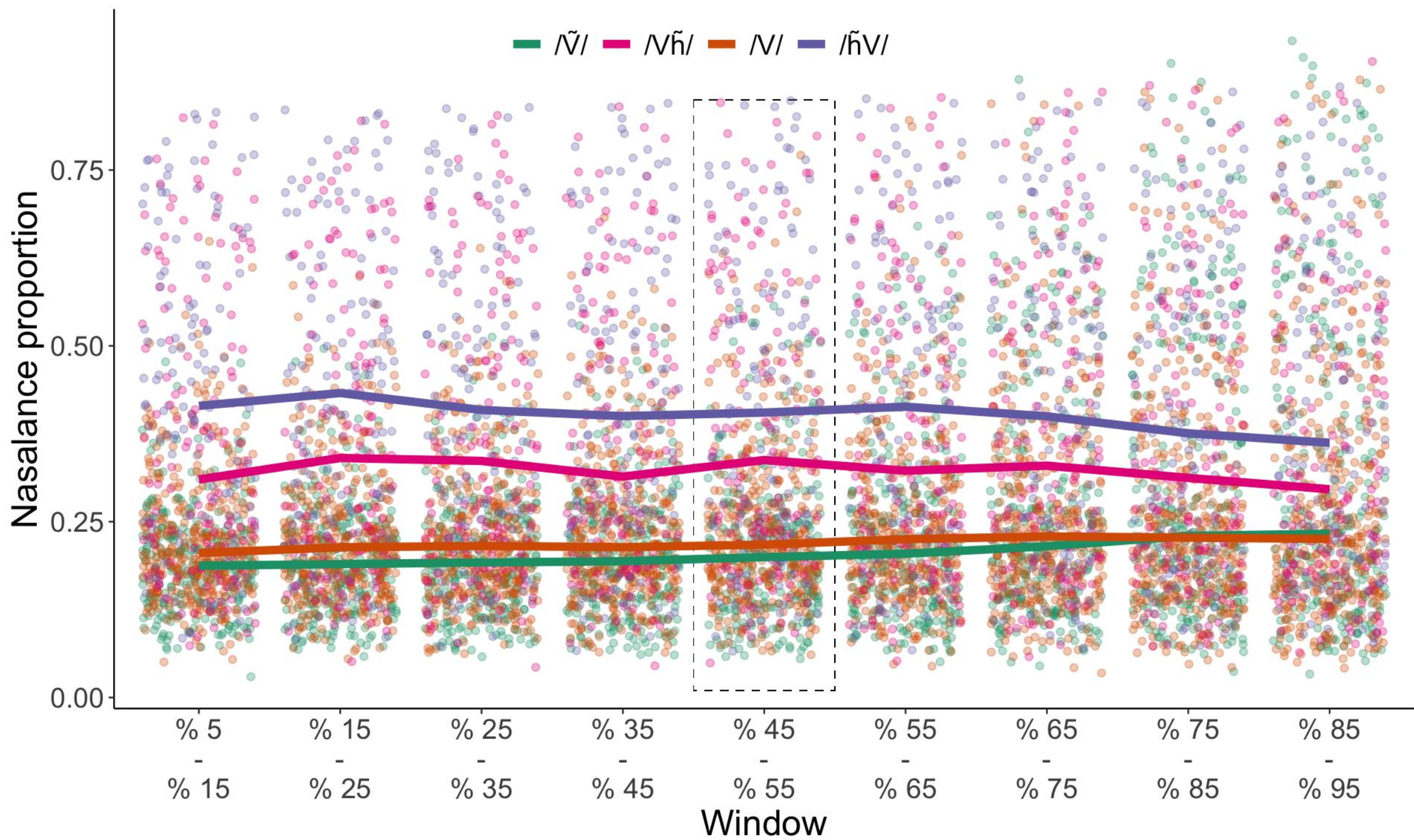
A_o

A_n

$A_n + A_o$

Nasalance proportion





Bayesian hierarchical beta regression model with brms

Nasalance $\sim$ 0 + group : place + (0 + group : place | speaker) + (1 | word)

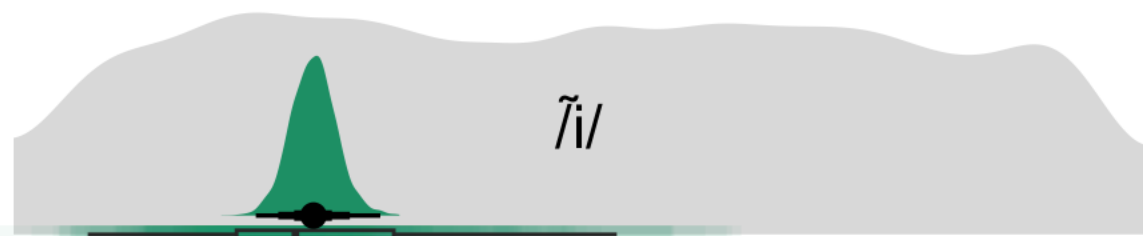
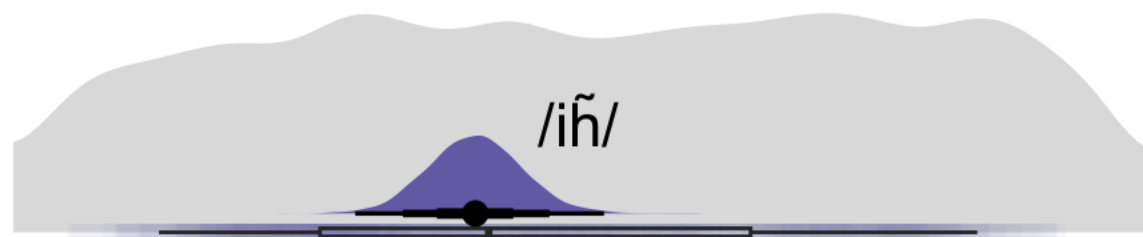
Priors:

$\beta \sim \text{Normal}(0, 1.5)$

$\varphi \sim \text{Normal}(0, 1.2)$

$\text{sd} \sim \text{Gamma}(0.01, 0.01)$

$\rho \sim \text{LKJ}(2)$

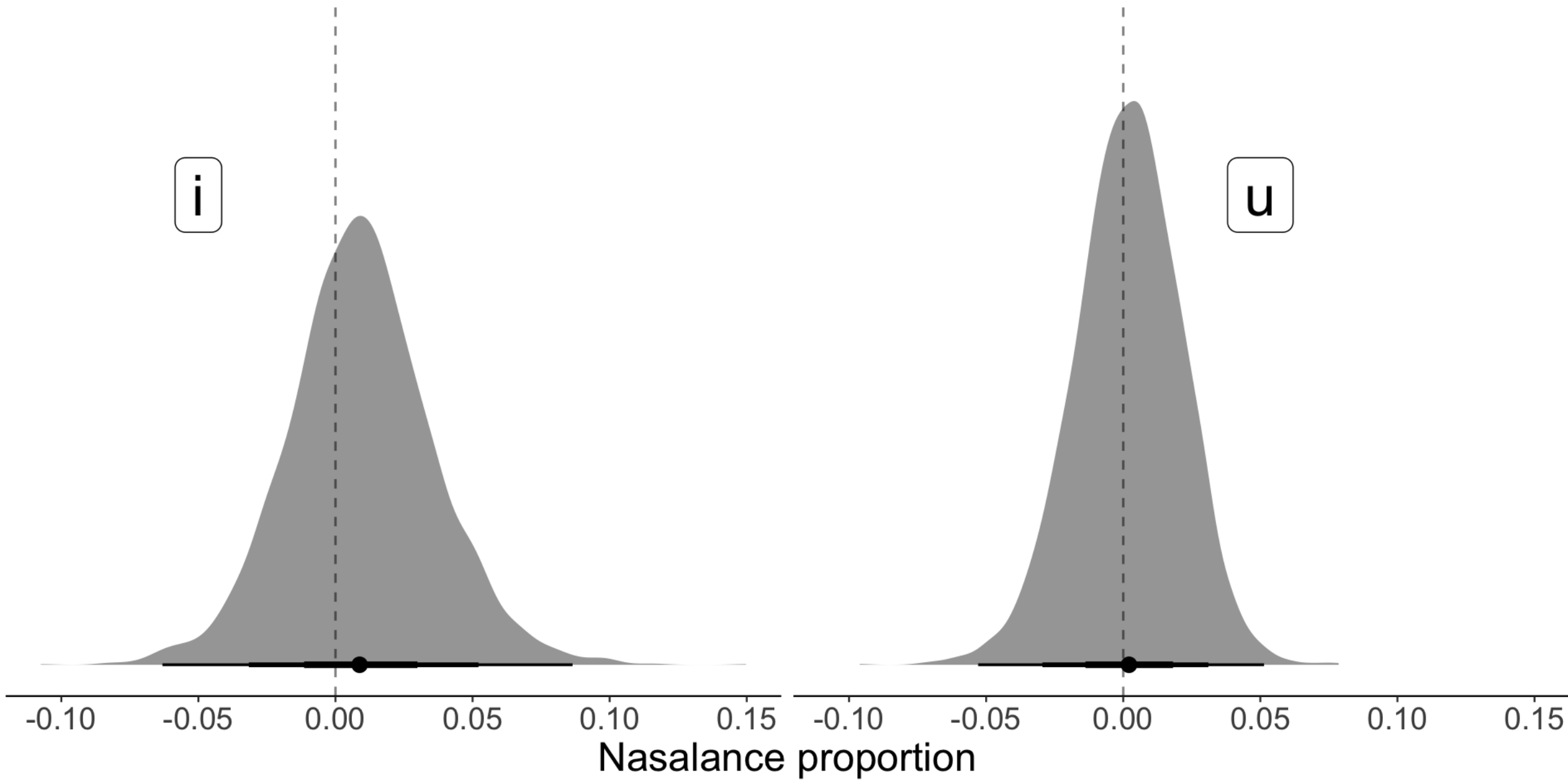


0.00 0.25 0.50 0.75 1.00 0.00 0.25 0.50 0.75 1.00

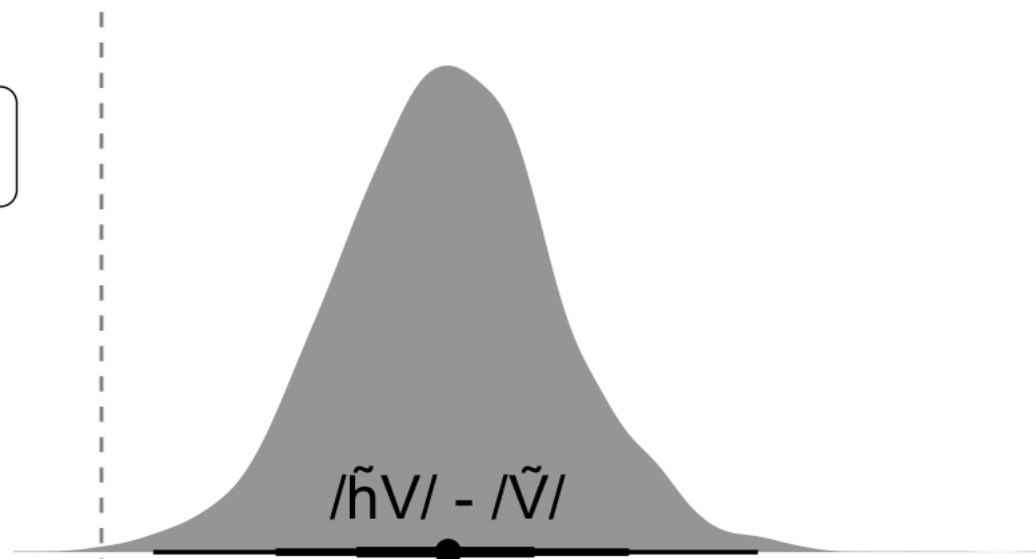
Nasalance proportion

i

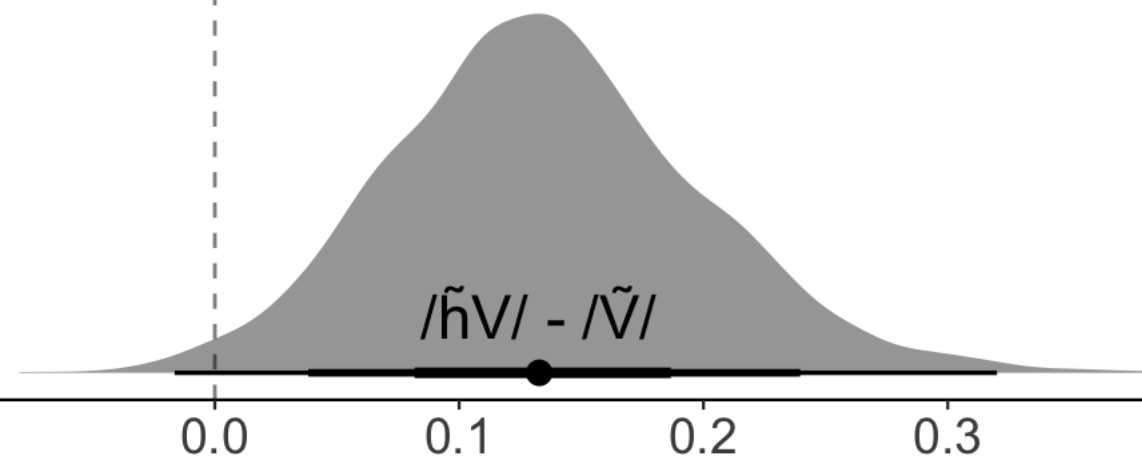
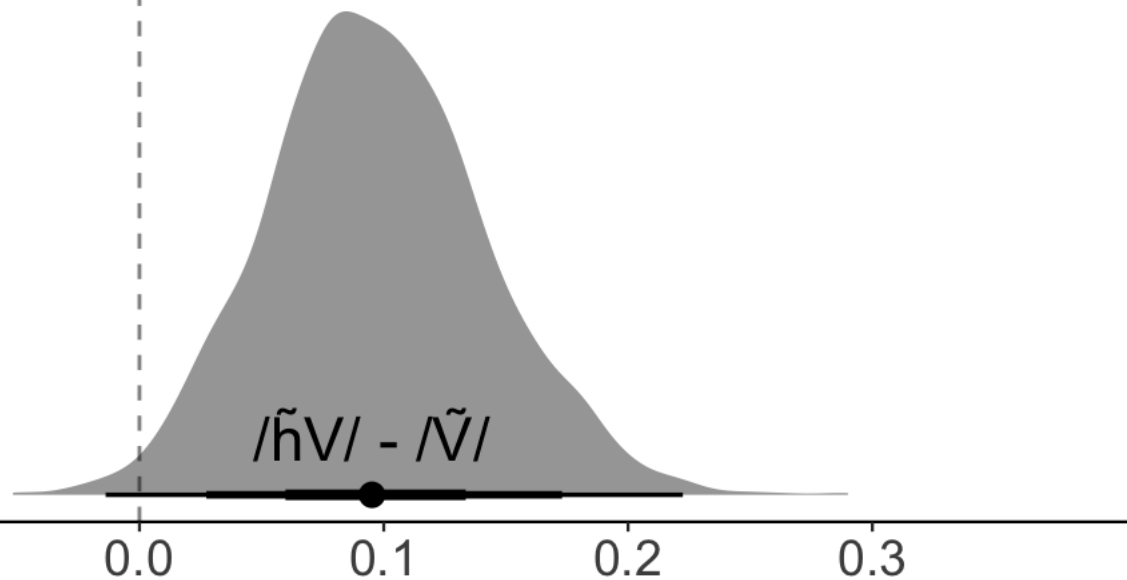
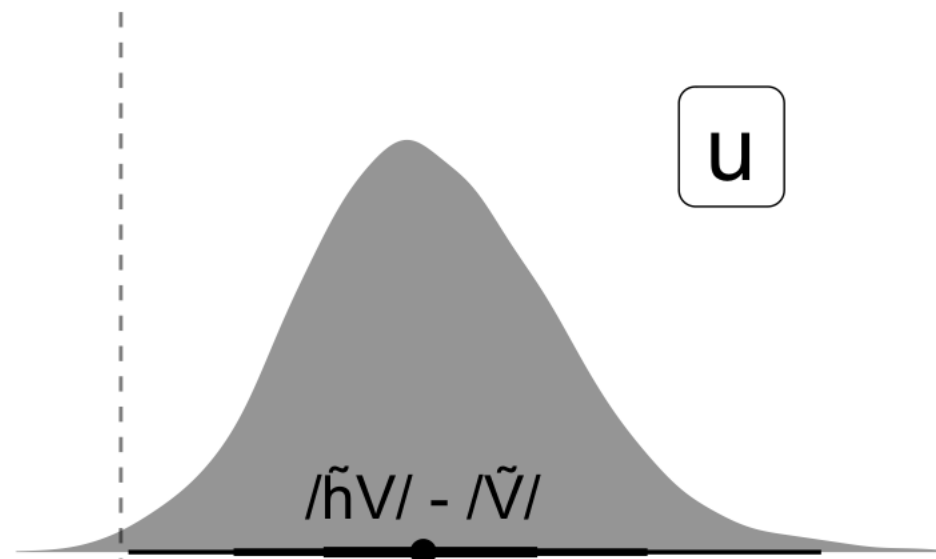
u



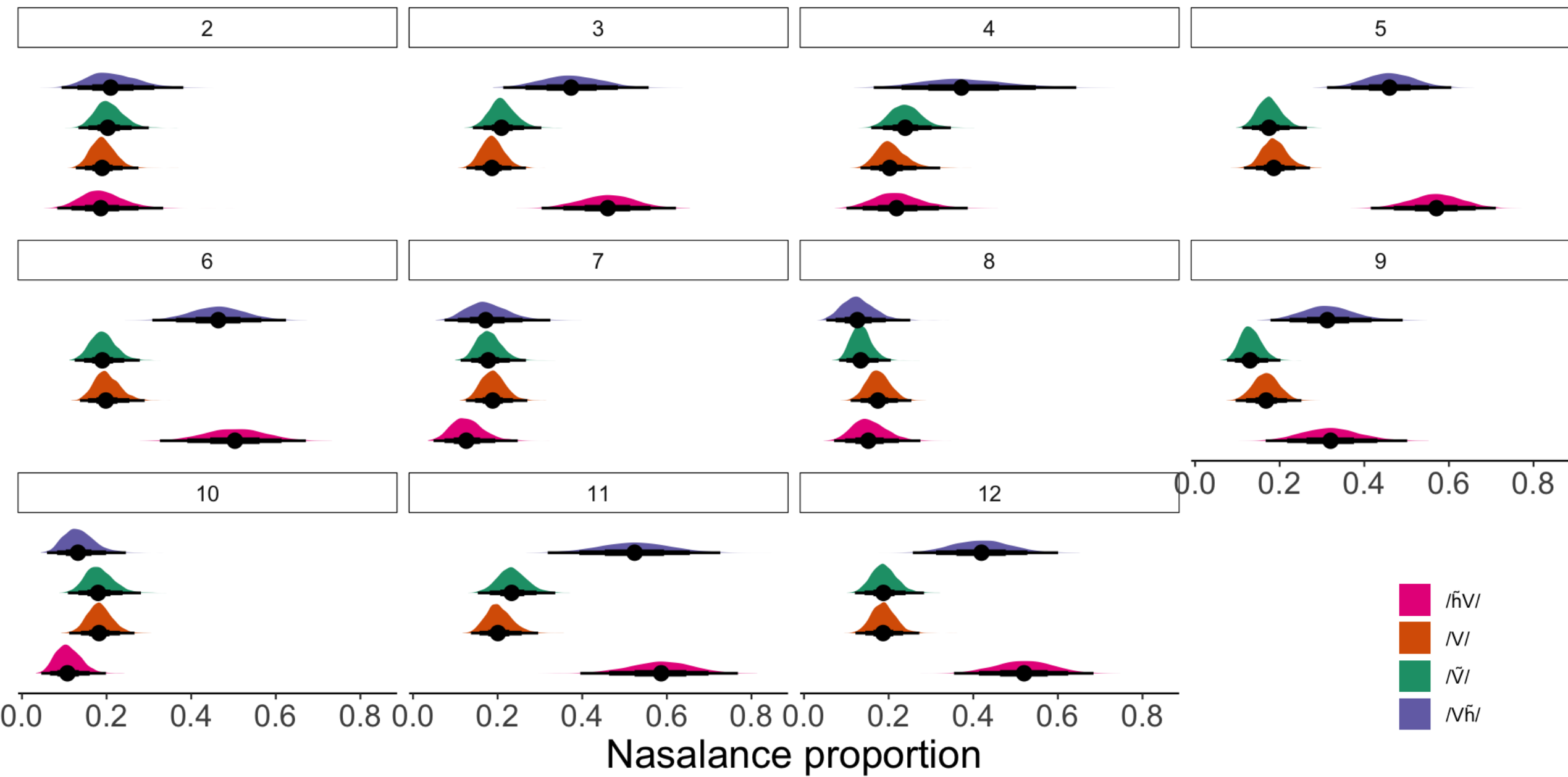
i



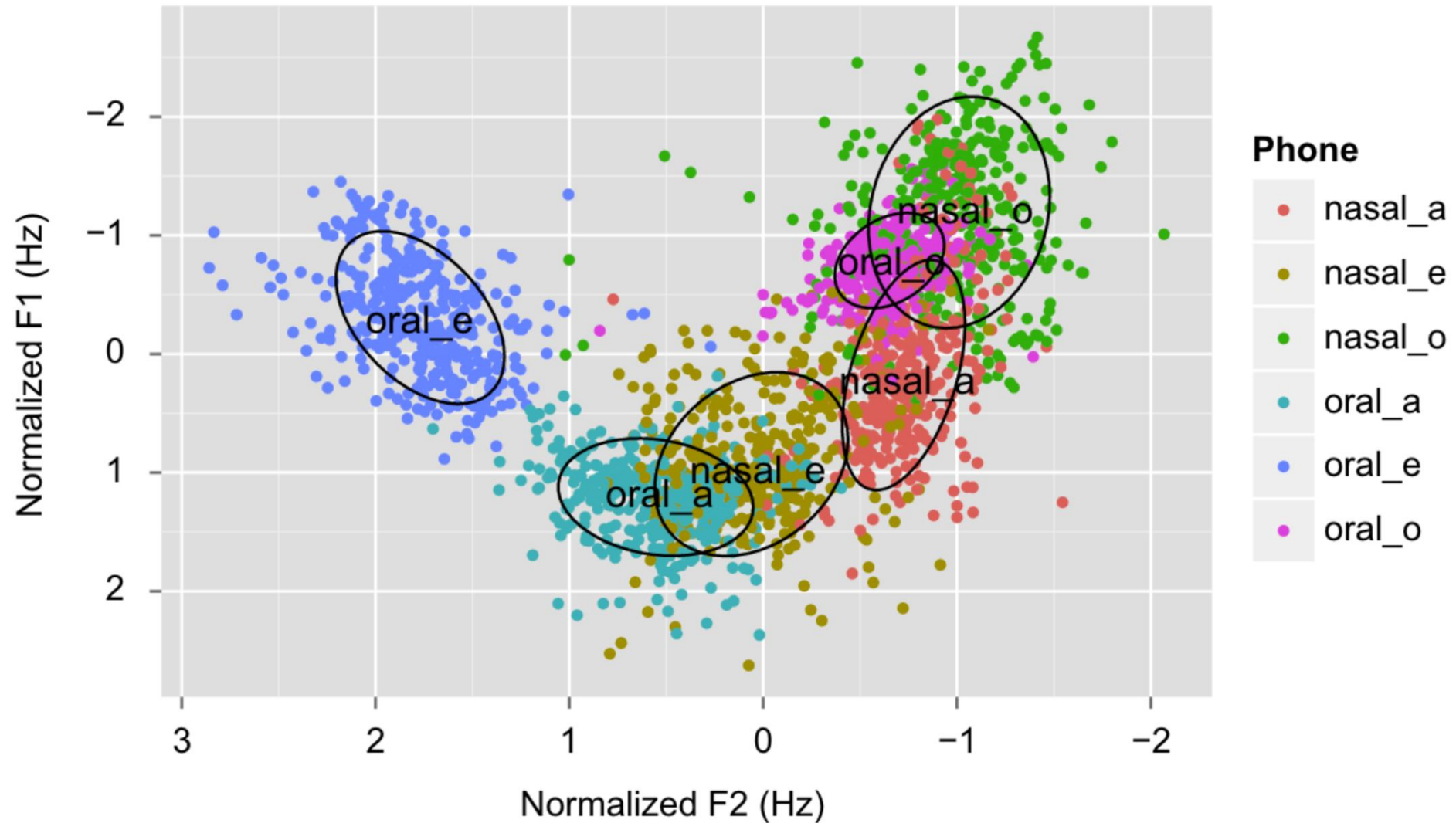
u



Nasalance proportion



/o/ ≠ /õ/, /e/ ≠ /ẽ/...



Potential transphonologization?

- Larraine Basque has lost the opposition fairly recently.
- No studies to confirm the merger between oral and (formerly) nasal vowels.

Bayesian hierarchical 4-variate Gaussian regression model with brms

$$\begin{pmatrix} \log(F1) \\ \log(F2) \\ \log(F3) \\ \log(\text{duration}) \end{pmatrix} \sim 0 + \text{group} : \text{place} + (0 + \text{group} : \text{place} \mid \text{speaker}, \text{word})$$

Bayesian hierarchical 4-variate Gaussian regression model with brms

Priors:

$$\beta_{F1} \sim \text{Normal}(5.5, 0.5)$$

$$\beta_{F2,i} \sim \text{Normal}(7.5, 0.5), \beta_{F2,u} \sim \text{Normal}(7, 0.5)$$

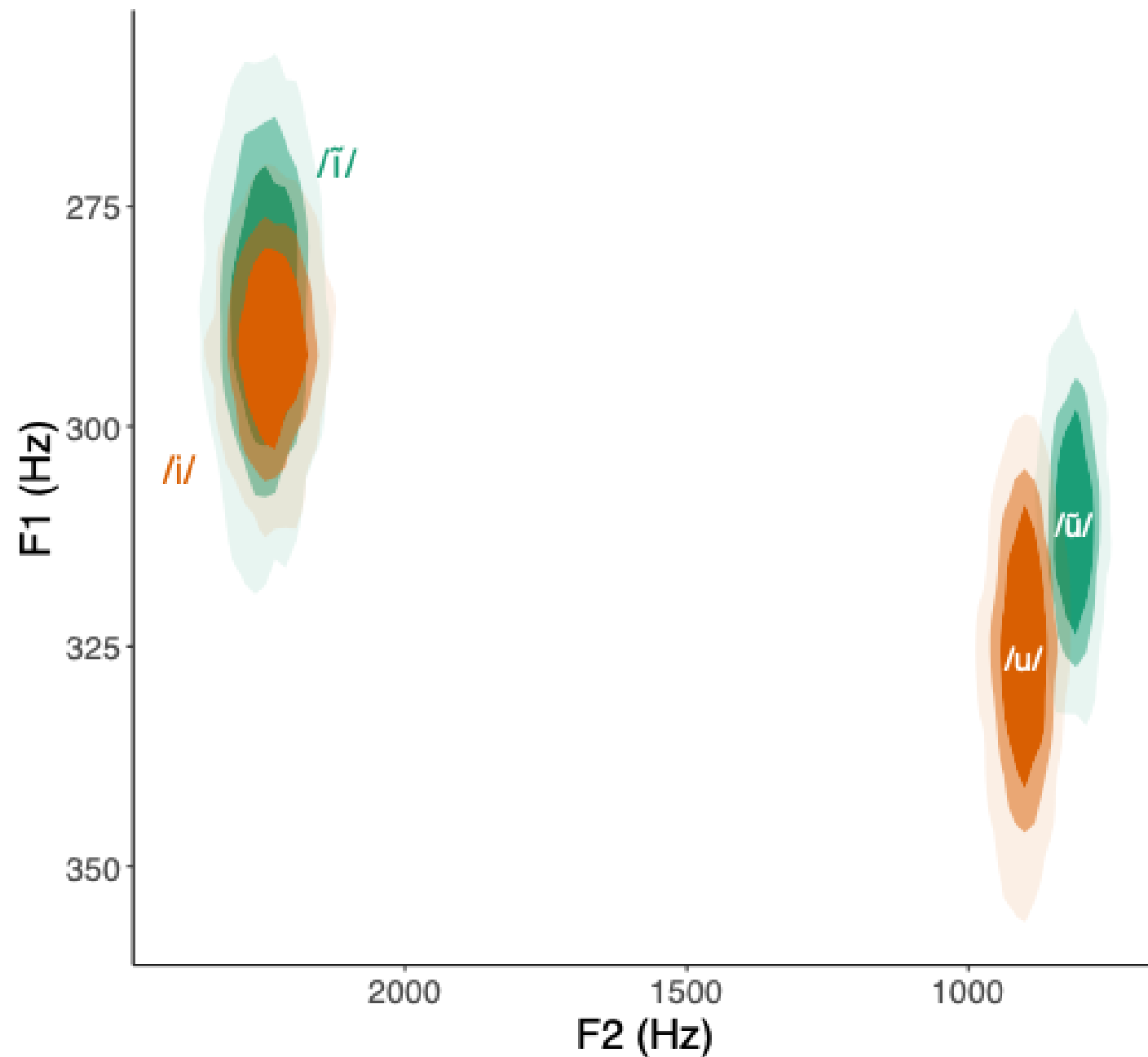
$$\beta_{F3,i} \sim \text{Normal}(8, 0.5), \beta_{F3,u} \sim \text{Normal}(7.9, 0.5)$$

$$\beta_{\text{duration}} \sim \text{Normal}(-2, 0.5)$$

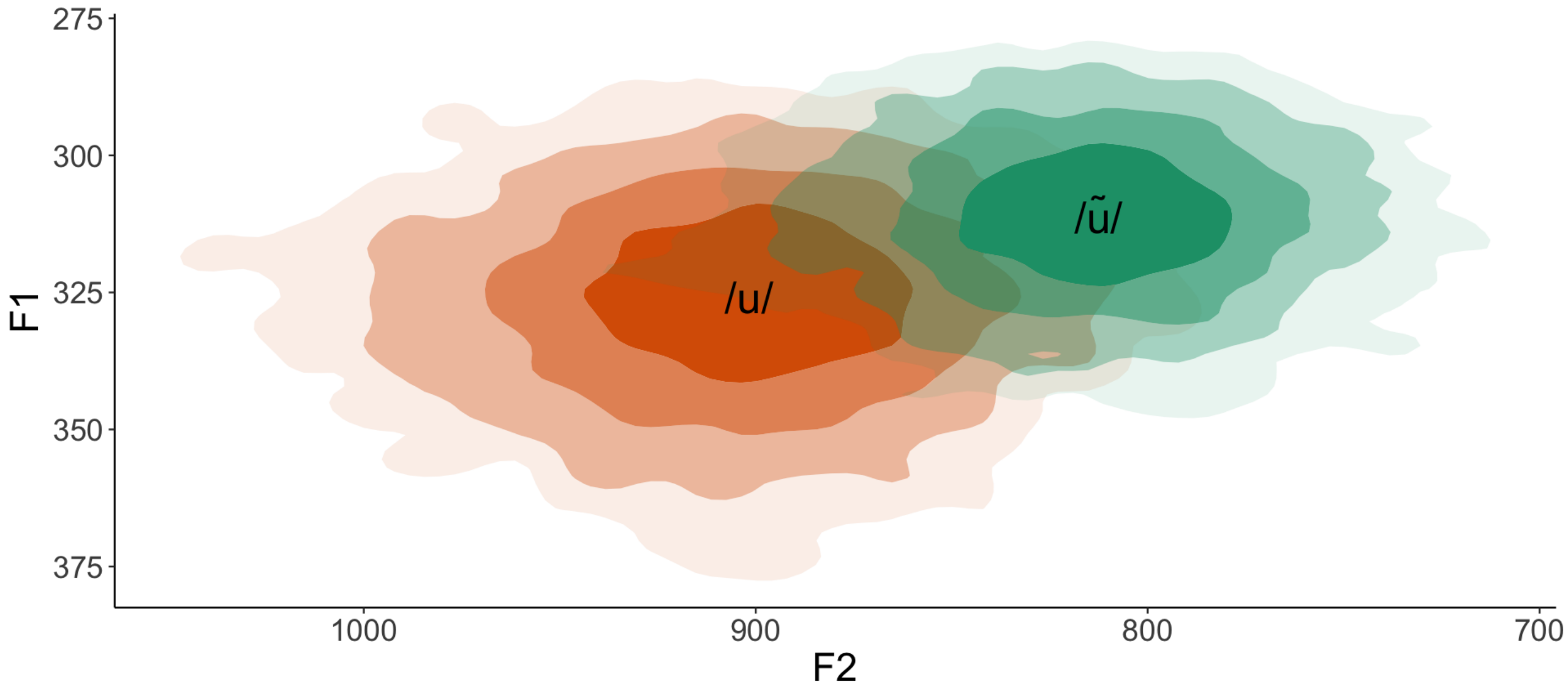
$$\sigma_{F1,F2,F3} \sim \text{Normal}(0, 0.5), \sigma_{\text{duration}} \sim \text{Normal}(0, 0.25)$$

$$\text{sd}_{F1,F2,F3} \sim \text{Normal}(0, 0.5), \text{sd}_{\text{duration}} \sim \text{Normal}(0, 0.25)$$

$$\rho \sim \text{LKJ}(2)$$



Posterior %CrI 88% 70% 50%



i

u

-50

0

50

Difference in F1 (Hz)

-50

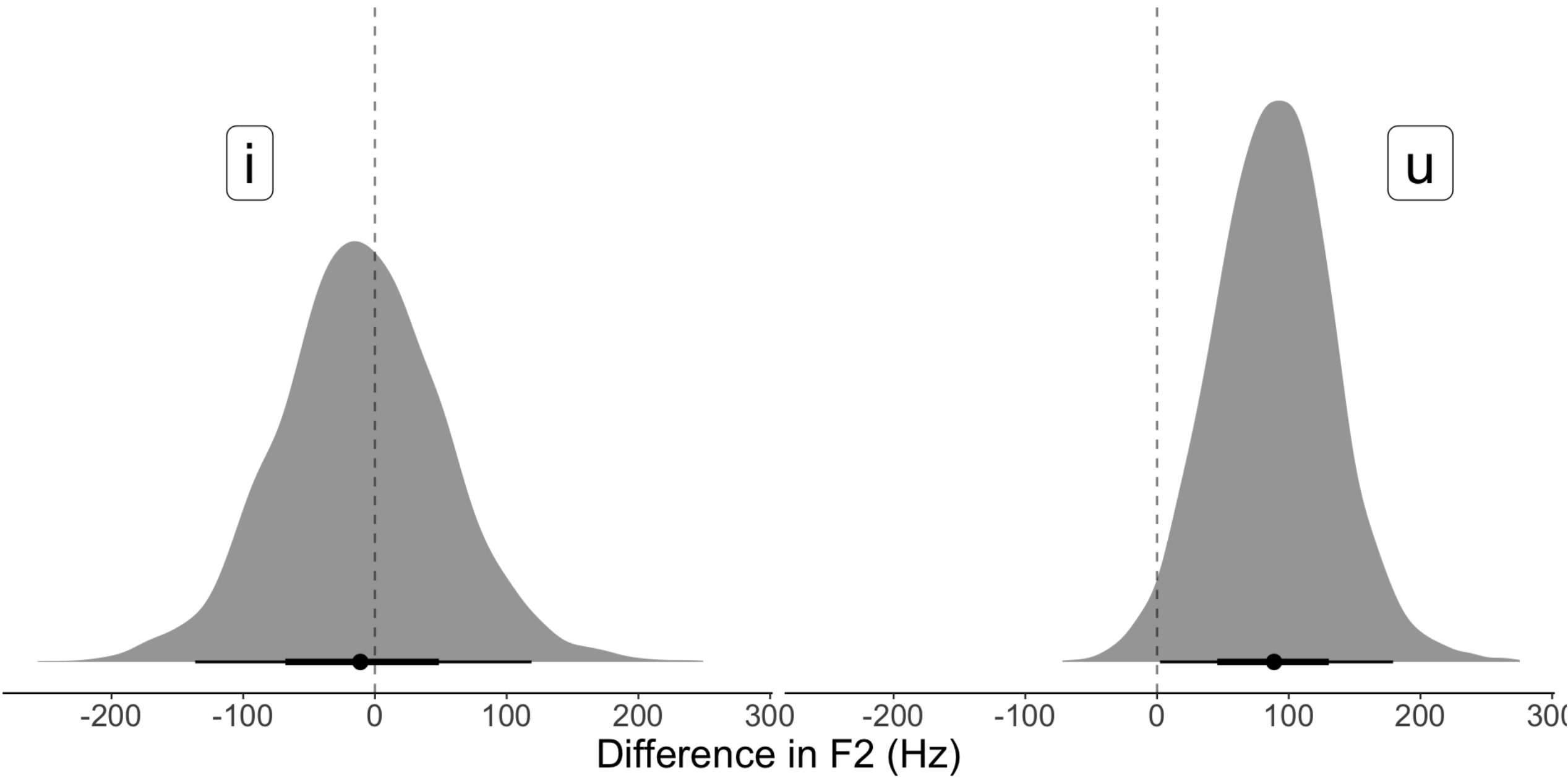
0

50

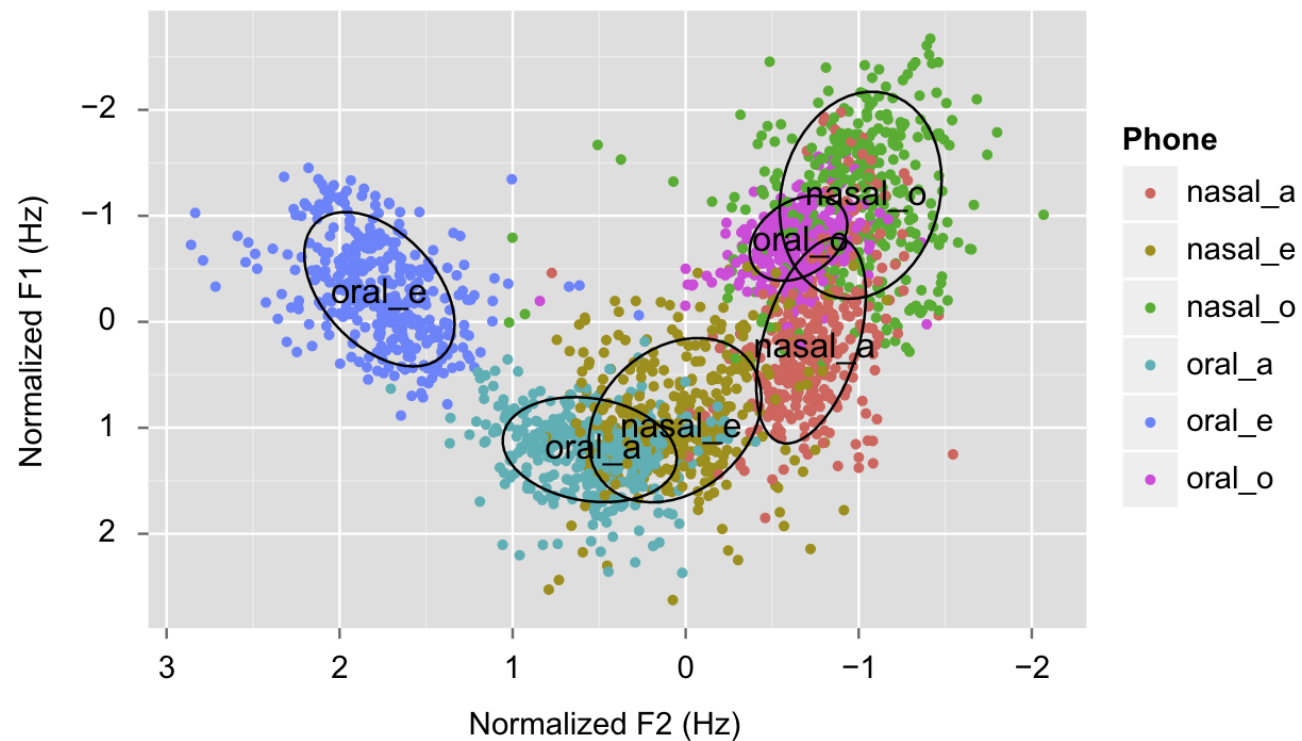
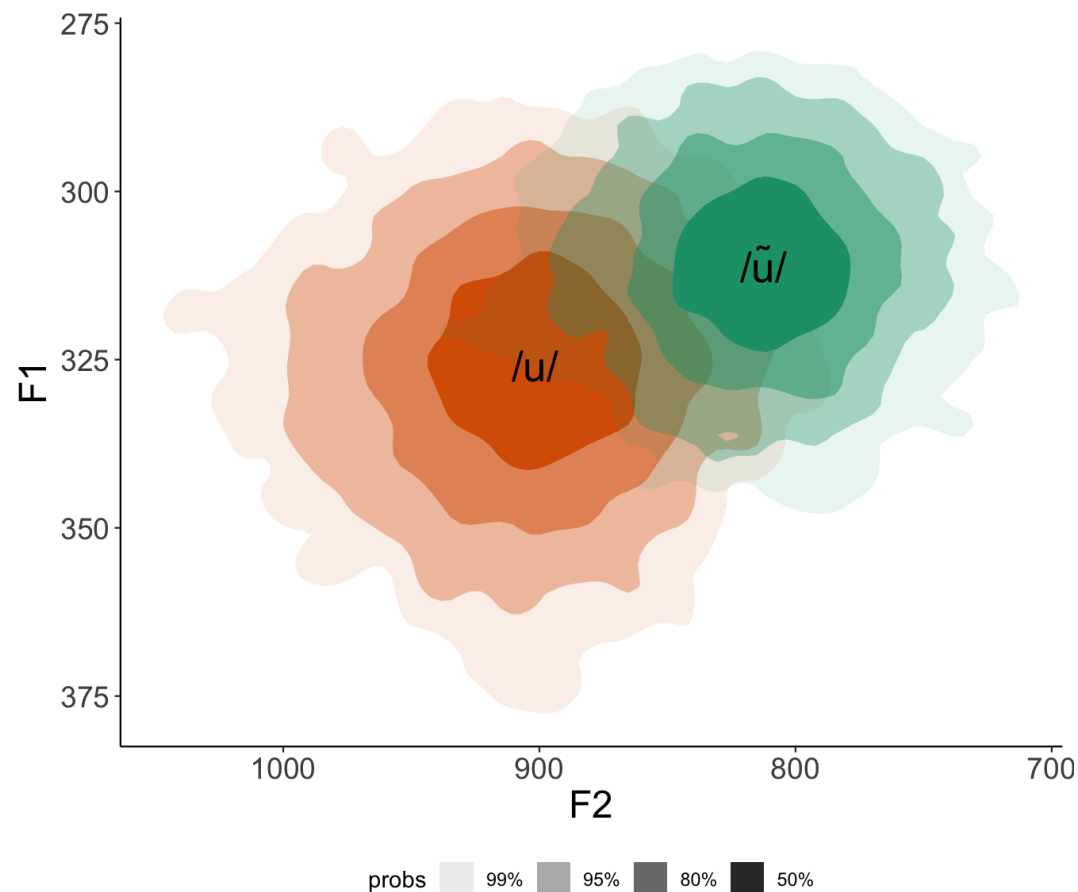
i

u

Difference in F2 (Hz)



Transphonologization of formerly nasalized vowels?



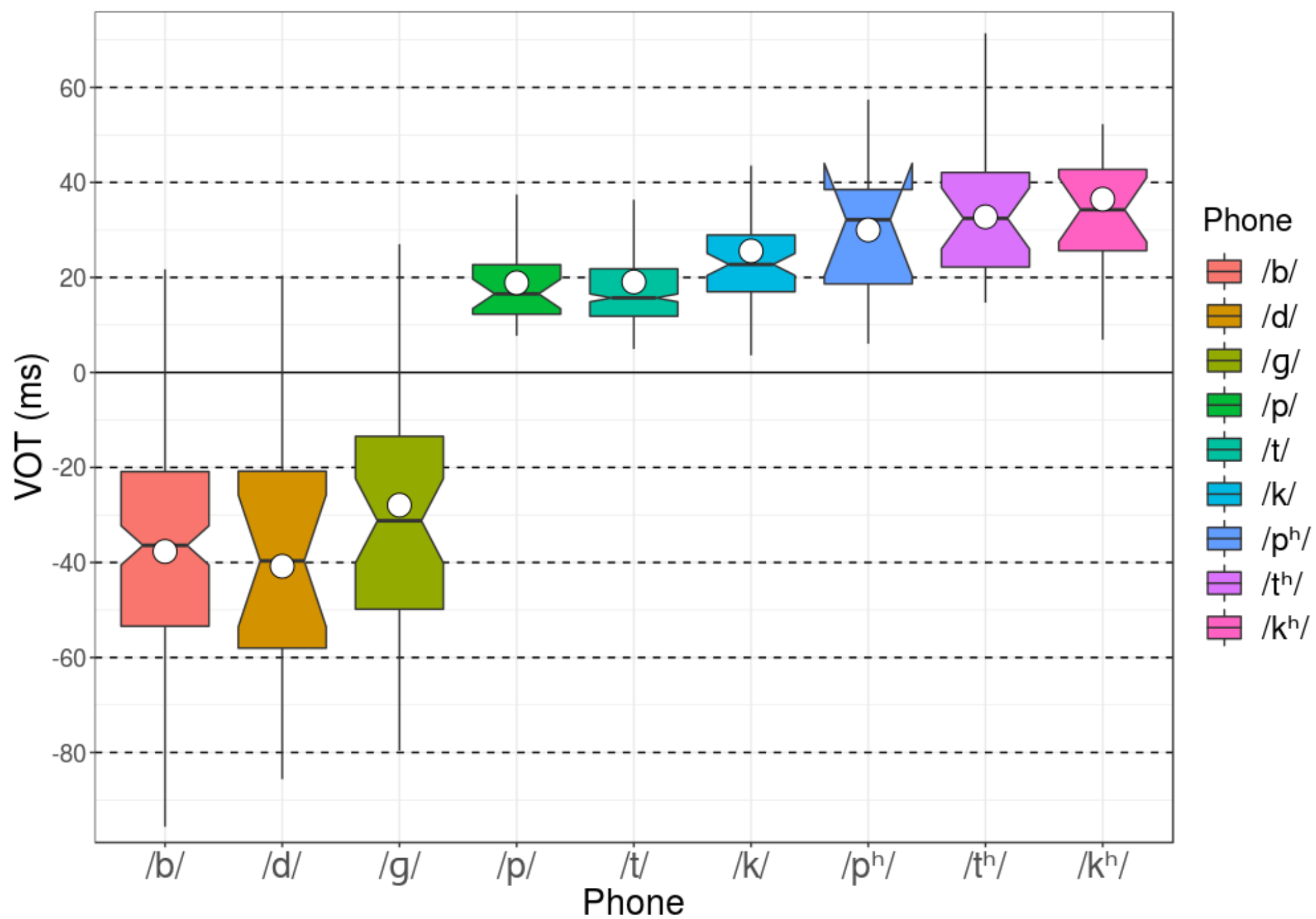
Stops

Stops

- 3 (+1) places of articulation:
 - Labial (/p, b/)
 - Alveolar (/t, d/)
 - Velar (/k, g/)
 - Palatal (/c/ *tt*, /ɟ/ *dd*)
- 3 series:
 - Voiced (common to all varieties)
 - Plain voiceless (common to all varieties)
 - Voiceless aspirated (limited to the North-eastern varieties).

Distribution of stops

- All stops are contrastive:
 - Intervocally
 - Word-initially
- Deaspiration:
 - No aspiration after the second syllable (with rare exceptions)
 - Only one aspirated stop in each word
- Neutralization to the plain voiceless:
 - Word-finally
 - After a sibilant

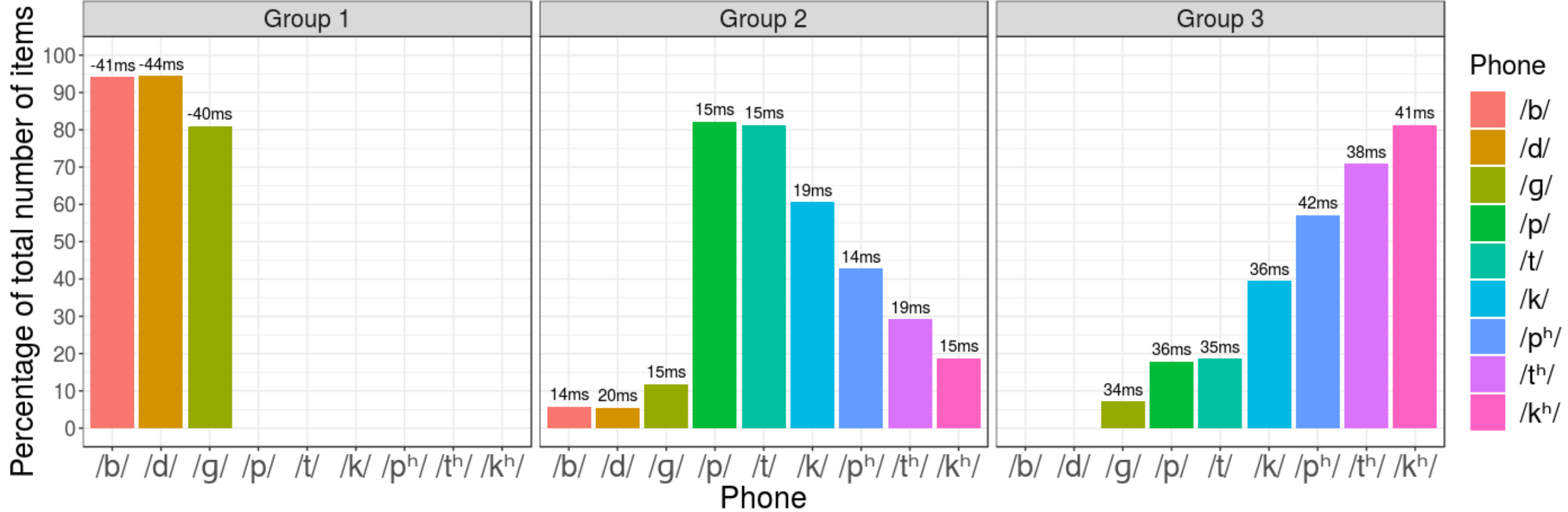


2 or 3 series of stops?

Table 3: Stop consonant voice onset time (VOT; ms) from Gaminde et al. (2002), Mounole (2004), and the current study.

Variety	Stop consonant								
	/b/	/p/	/p ^h /	/d/	/t/	/t ^h /	/g/	/k/	/k ^h /
Zuberoan (Gaminde et al., 2002)	-102	20	61	-105	24	67	-101	27	83
Zuberoan (Mounole, 2004)		14	47		20	52		28	67
Mixean	-38	19	30	-41	19	33	-28	26	36

Stop consonant grouped by Gaussian mixture clustering of VOT



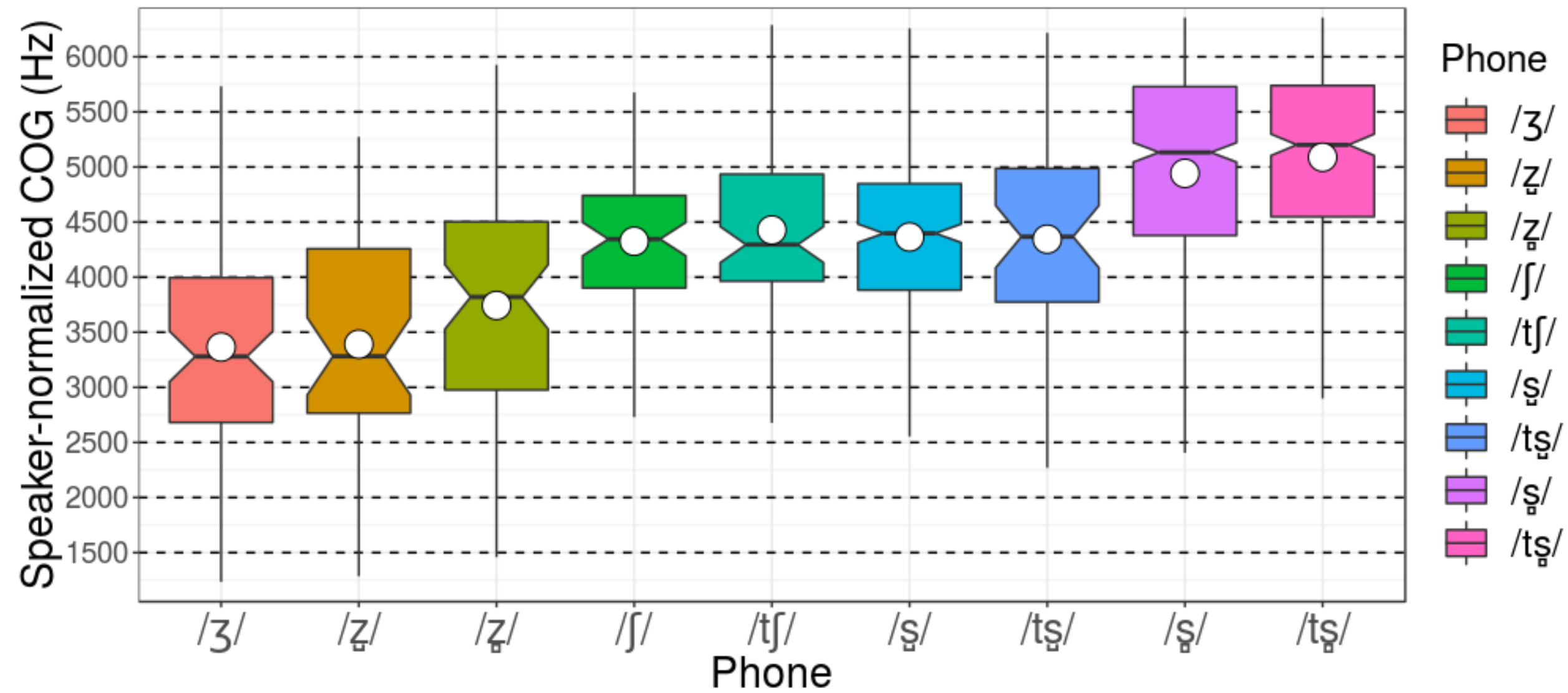
Sibilants

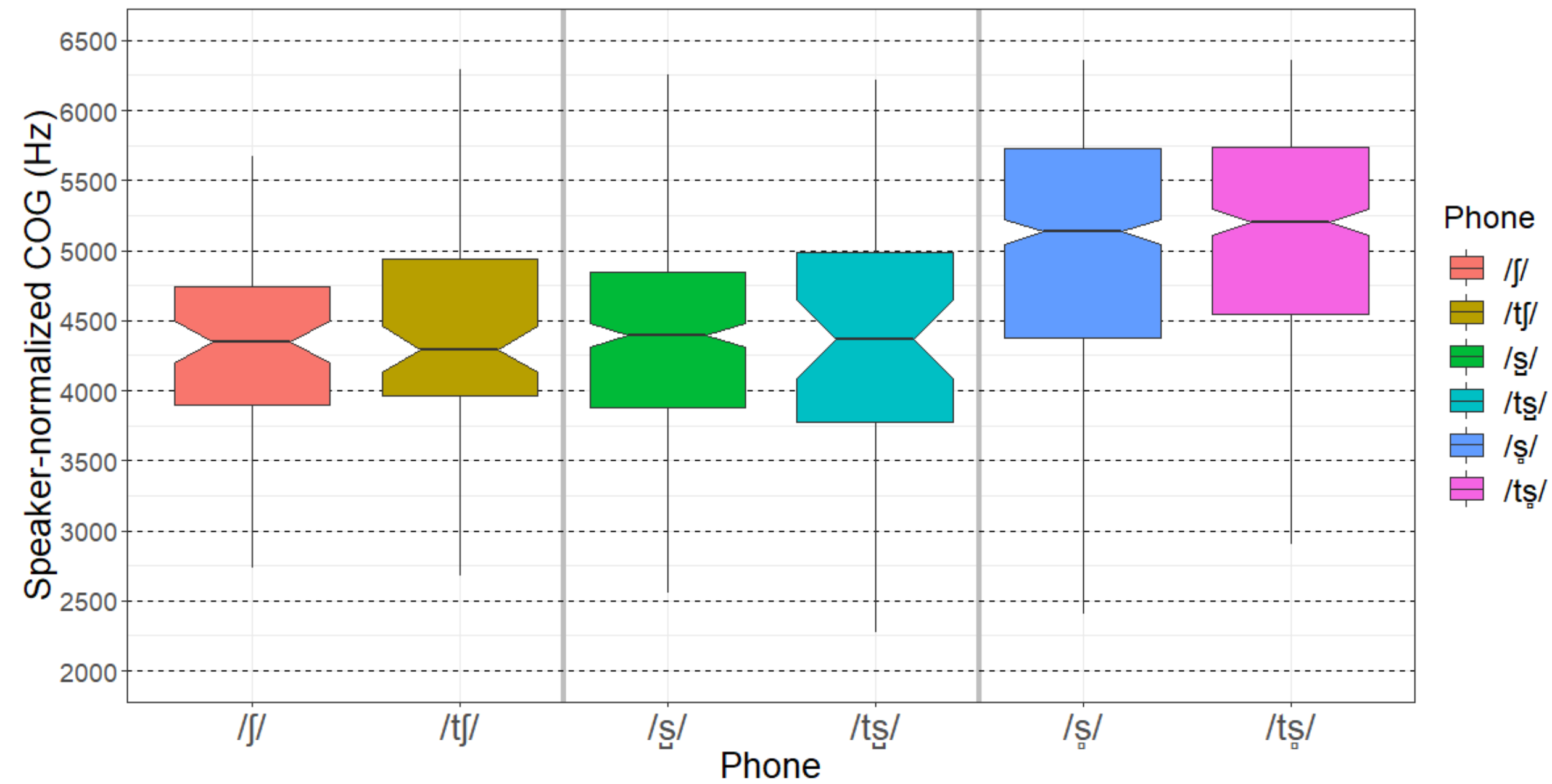
Sibilants

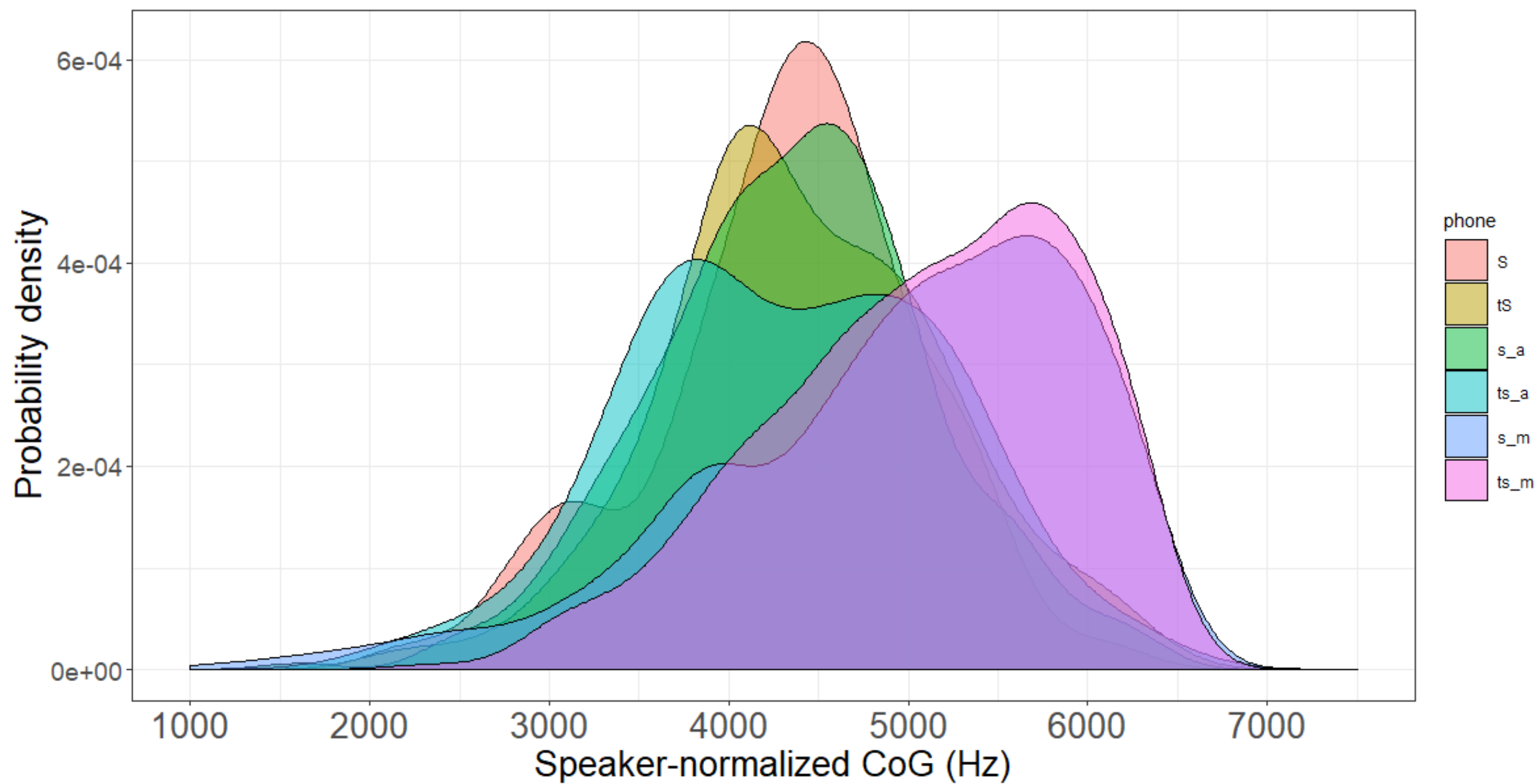
- As in other Basque varieties, 3 places of articulation:
 - Apico-alveolar: /s̺/, /t̺s̺/, /z̺/, /d̺z̺/
 - Lamino-alveolar: /s̺̪/, /t̺̪s̺̪/, /z̺̪/
 - Post-alveolar: /ʃ/, /tʃ/, /ʒ/
- 4 series:
 - Voiceless fricatives (common to all varieties)
 - Voiceless affricates (common to all varieties)
 - Voiced fricatives (only found in varieties historically in contact with Gascon)
 - At least one voiced affricate (only in varieties in contact with Gascon)

Voiced sibilants

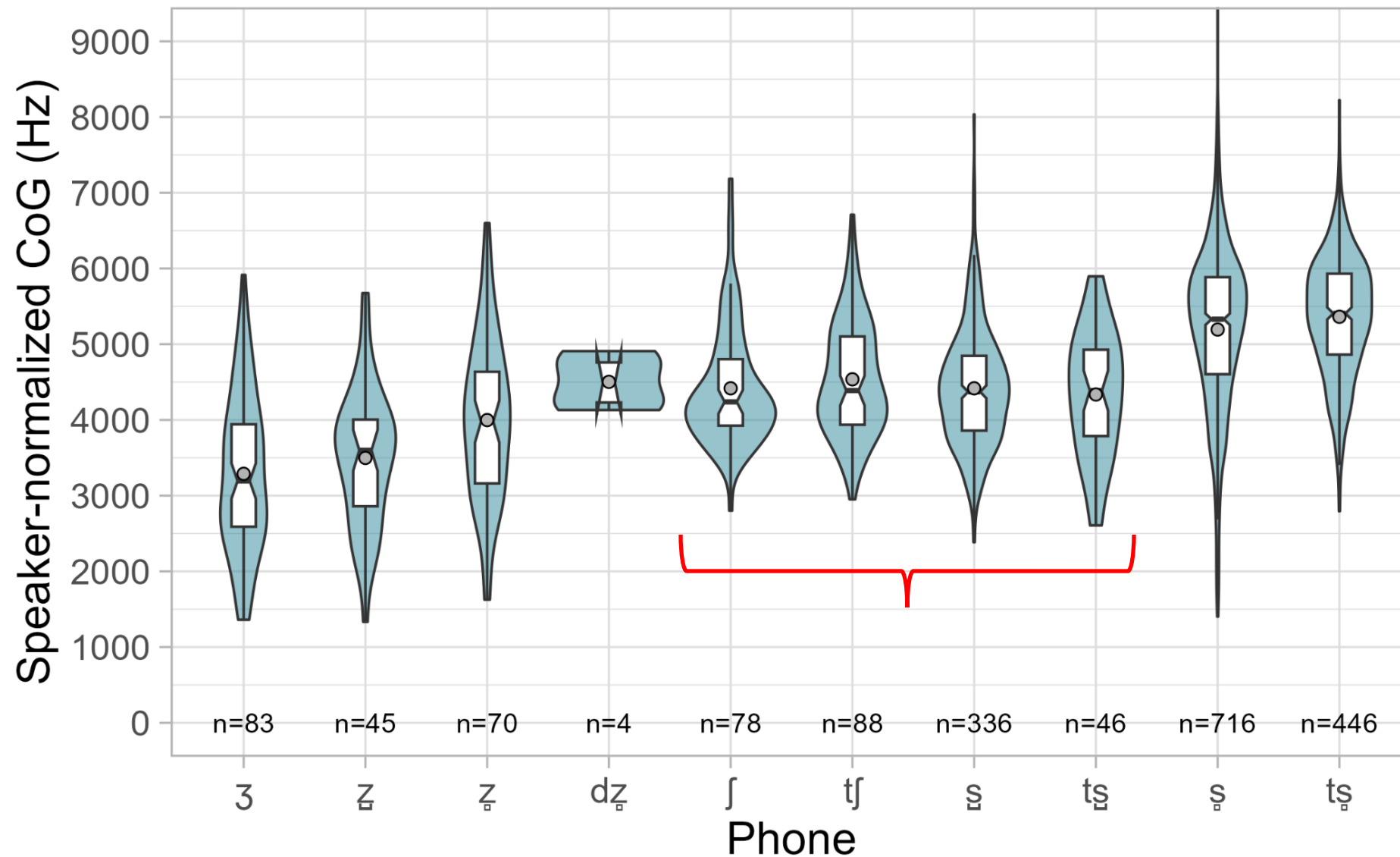
- Voiced sibilants are predictable allophones in most varieties of Basque:
 - Before a voiced consonant: *esne* [ez̞n̥ẽ], *desberdin* [dez̞βerð̞ĩn].
- Contrastive voiced sibilants are only found in the Eastern varieties.
 - Contact with Gascon / French.
 - Gascon has the contrast in alveolar & post-alveolar sibilants (Mooney 2014).
- Two main contexts of development:
 - In Bearnese loanwords: *bedezĩ* /bede'zĩ/, *aizina* /aiz̞ina/
 - Morpheme boundary: *deuse* /deuz̞e/, *ezaxola* /ez̞axola/







2 or 3 places of articulation in sibilants?



No CoG difference between alveolar sibilants

- $\text{CoG} \sim \text{phone} + (1 \mid \text{word}) + (\text{phone} \mid \text{speaker})$
- Weakly informative priors.
- Decision rule:
 $< 5\% \text{ HDI in ROPE}$

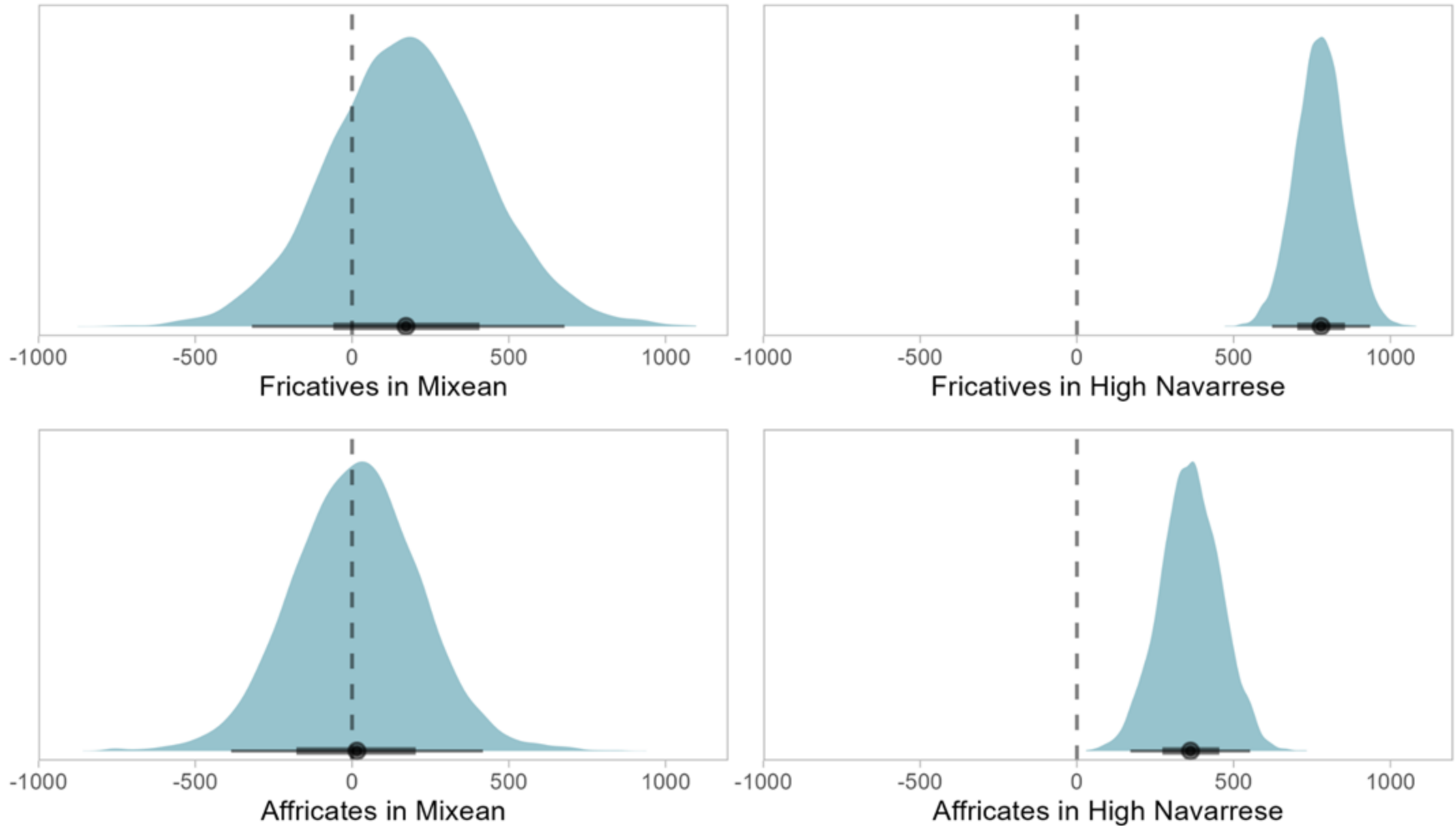


Strong evidence of a difference

Table 4. Contrasts between the values of PHONE within each manner category

	Estimate	Lower HDI	Upper HDI	ROPE
$/\underset{\square}{s}/ - /s/$	809.9	466.2	1156.4	0.00
$/\underset{\square}{s}/ - /s_{\square}/$	817.8	477.4	1145.3	0.00
$/s/ - /s_{\square}/$	7.0	-395.8	424.7	0.52
$/\widehat{ts}/ - /ts_{\square}/$	-864.5	-1322.0	-404.9	0.00
$/ts_{\square}/ - /ts_{\square}/$	-1029.1	-1466.2	-620.9	0.00
$/\widehat{ts}/ - /ts_{\square}/$	165.9	-316.5	700.0	0.35
$/\underset{\square}{z}/ - /z_{\square}/$	-618.4	-1227.4	-0.9	0.04
$/z/ - /z_{\square}/$	-34.6	-661.2	594.8	0.35
$/z/ - /z_{\square}/$	-653.8	-1198.0	-69.3	0.01

Contrast distributions of the CoG posteriors of apico-alveolar and postalveolar phones (Hz)



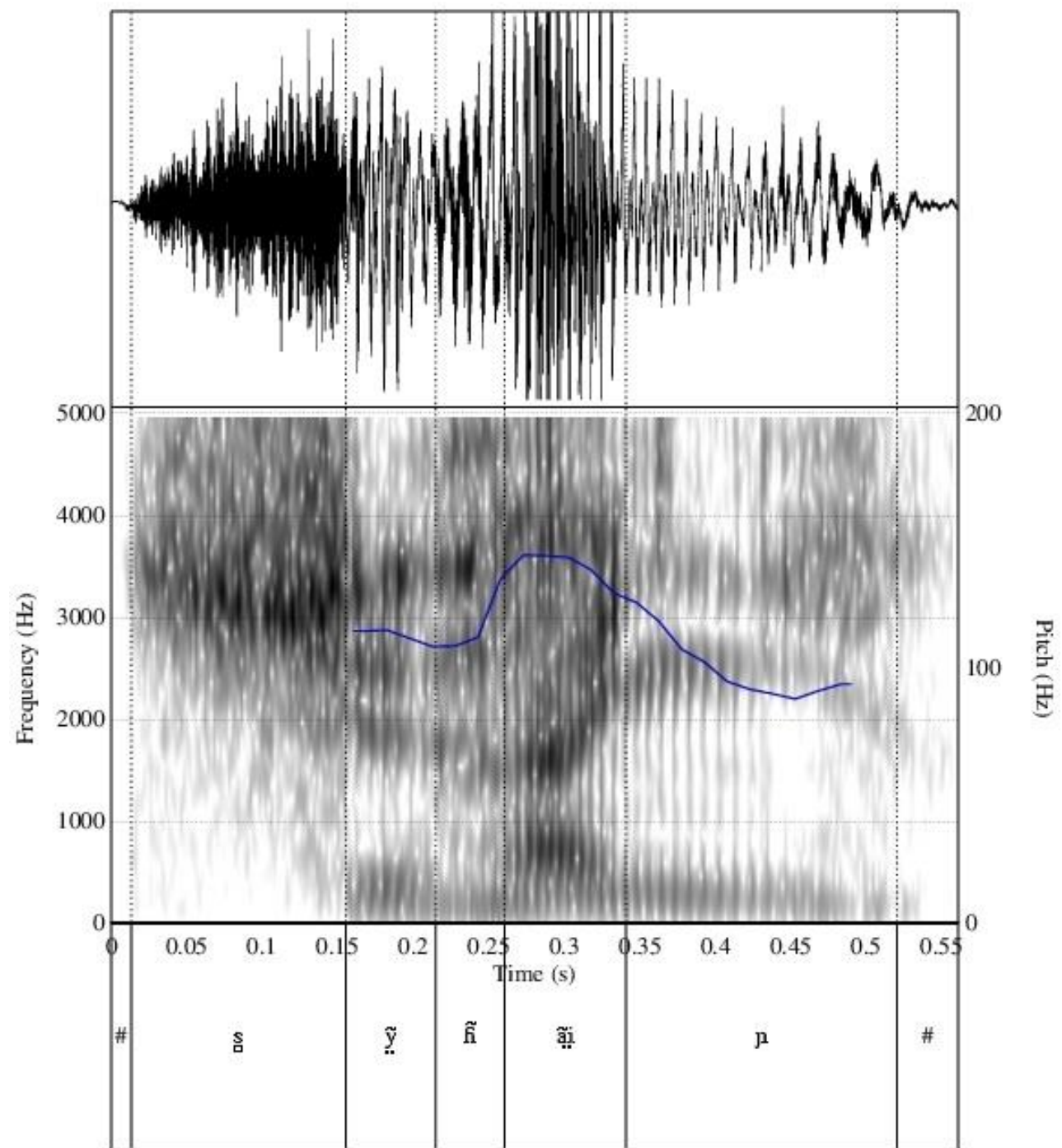
Laryngeals

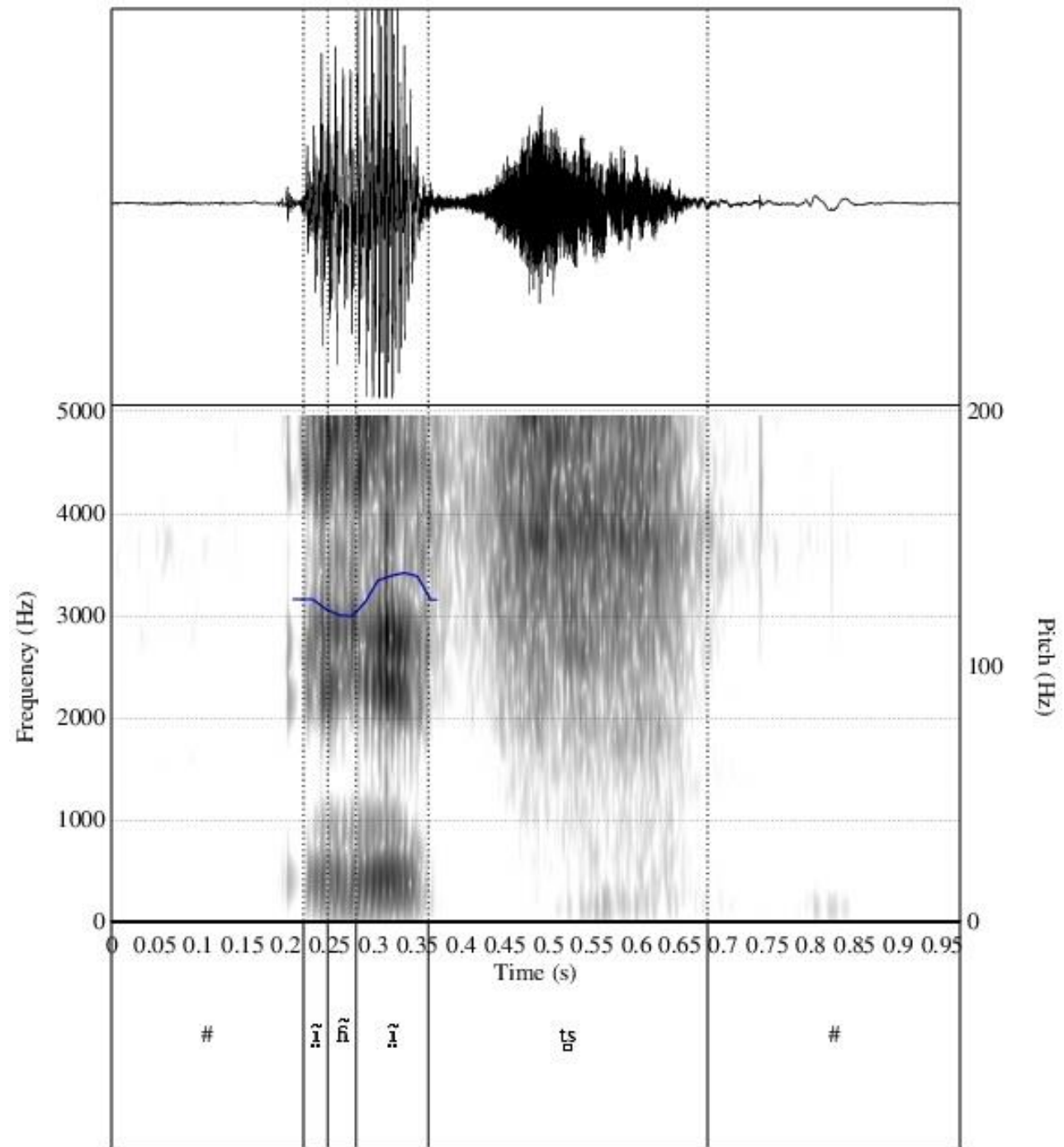
Laryngeal approximants

- Two aspirates / laryngeal approximants:
 - Oral /h/, potentially present in Proto-Basque
 - Nasalized /h̃/, from historical *n
- Geographical distribution:
 - /h/ -> Only in Eastern varieties
 - /h̃/ -> Only in Zuberoan & Mixean Basque
- Restrictions:
 - Onset only
 - First 2 syllables only
 - Only 1 aspirated segment in each word

Laryngeal approximants in Gascon

- One aspirate / laryngeal approximant:
 - Oral /h/
- It evolved from historical *f.
- No nasal aspirate, but same phonological context of change:
 - n > h/ø intervocalically (/V_V)





Nasal aspiration

Aquitanian *Se***n***i*- > Basque *se***h***i*

na*n***-i* > Basque *na***h***i*

Latin *a***n***āte(m)* > Basque *a***h***ate*

Zuberoan [sẽhĩ] ‘boy, servant’

Zuberoan [nãhĩ] ‘to want’

Zuberoan [ãhãte] ‘duck’

- /h/ vs. /h̃/ opposition
 - Phonetically unexpected.
 - Attested in just a handful of languages.
- What is the actual result of nasal aspiration?
- What is its status in the modern language?

Aspiration in Basque: Categories

Oral /h/

behi ‘cow’

bihotz ‘heart’

ehi ‘finger’

Nasal /h̃/

ahate ‘duck’

ihaute ‘carnival’

ehi ‘easy’

Assimilated /h/ [h̃]

uñhu ‘onion’

lehen ‘first, before’

nihaur ‘me, myself’

Posterior distribution

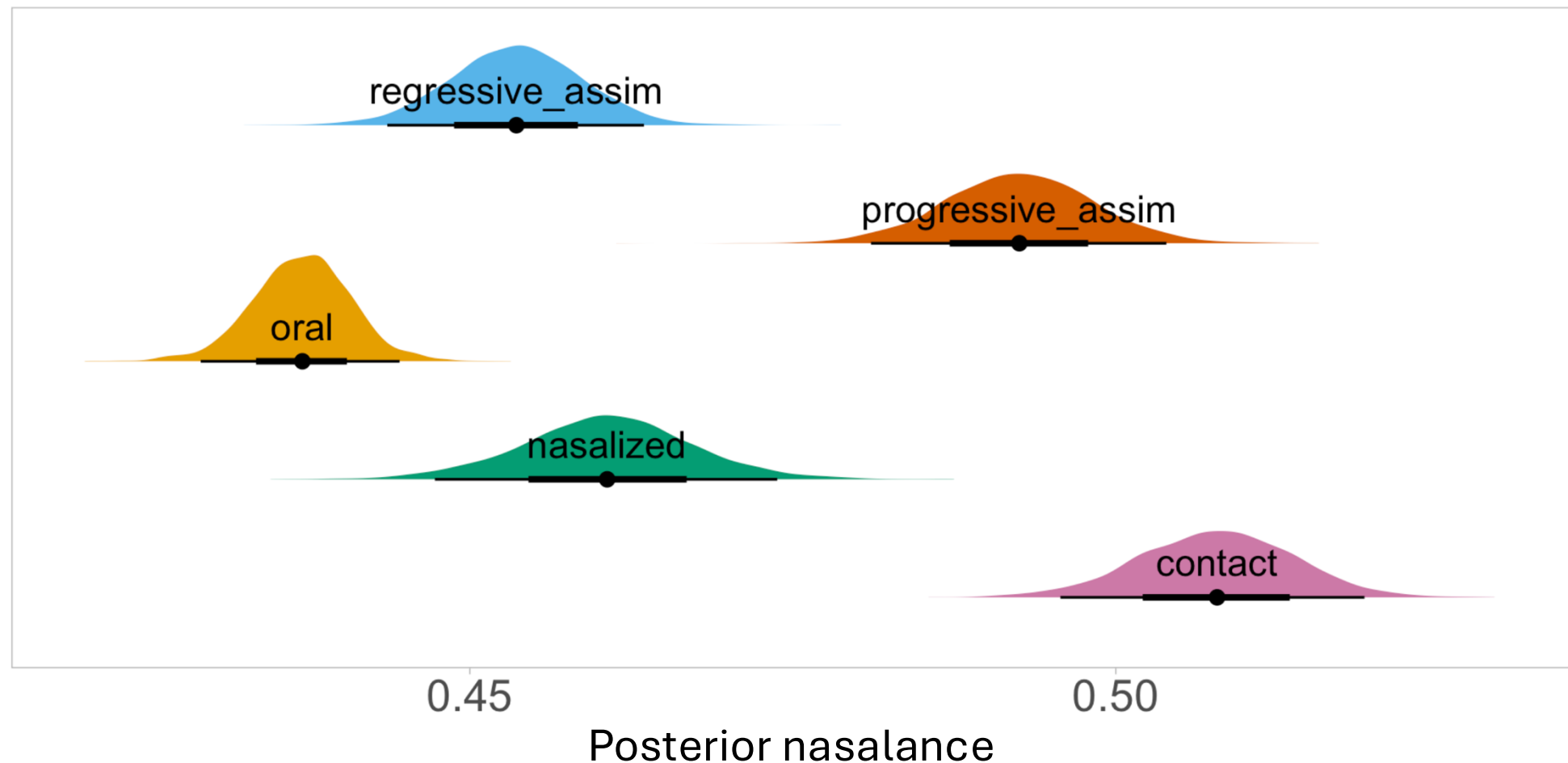
Dependent variable: **nasalance**

Independent variables: **aspirate category**

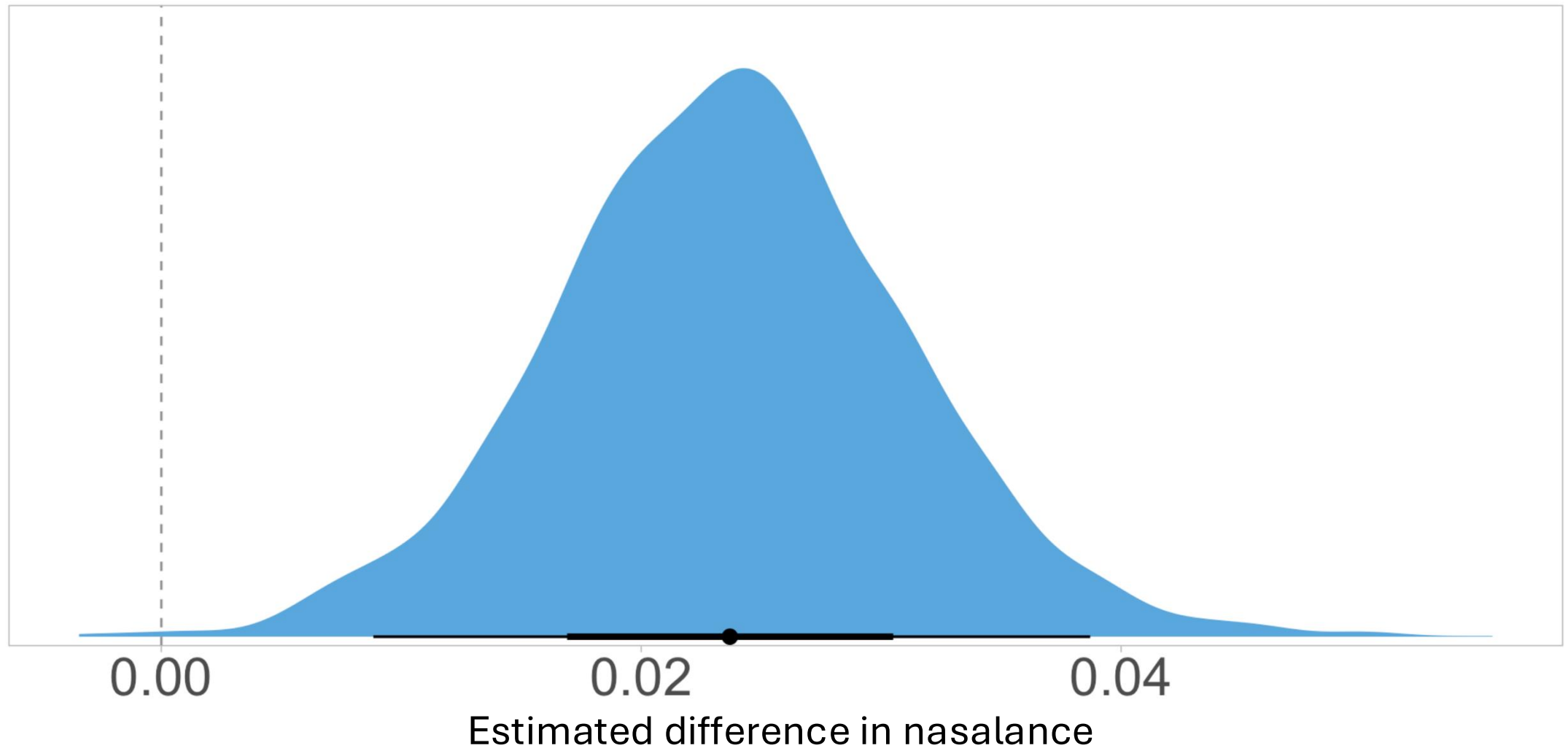
Family: **Beta**

Correlated varying intercept & slope: **speaker, word**

Weakly informative priors



Posterior contrast distribution: / $\tilde{h}$ / — /h/



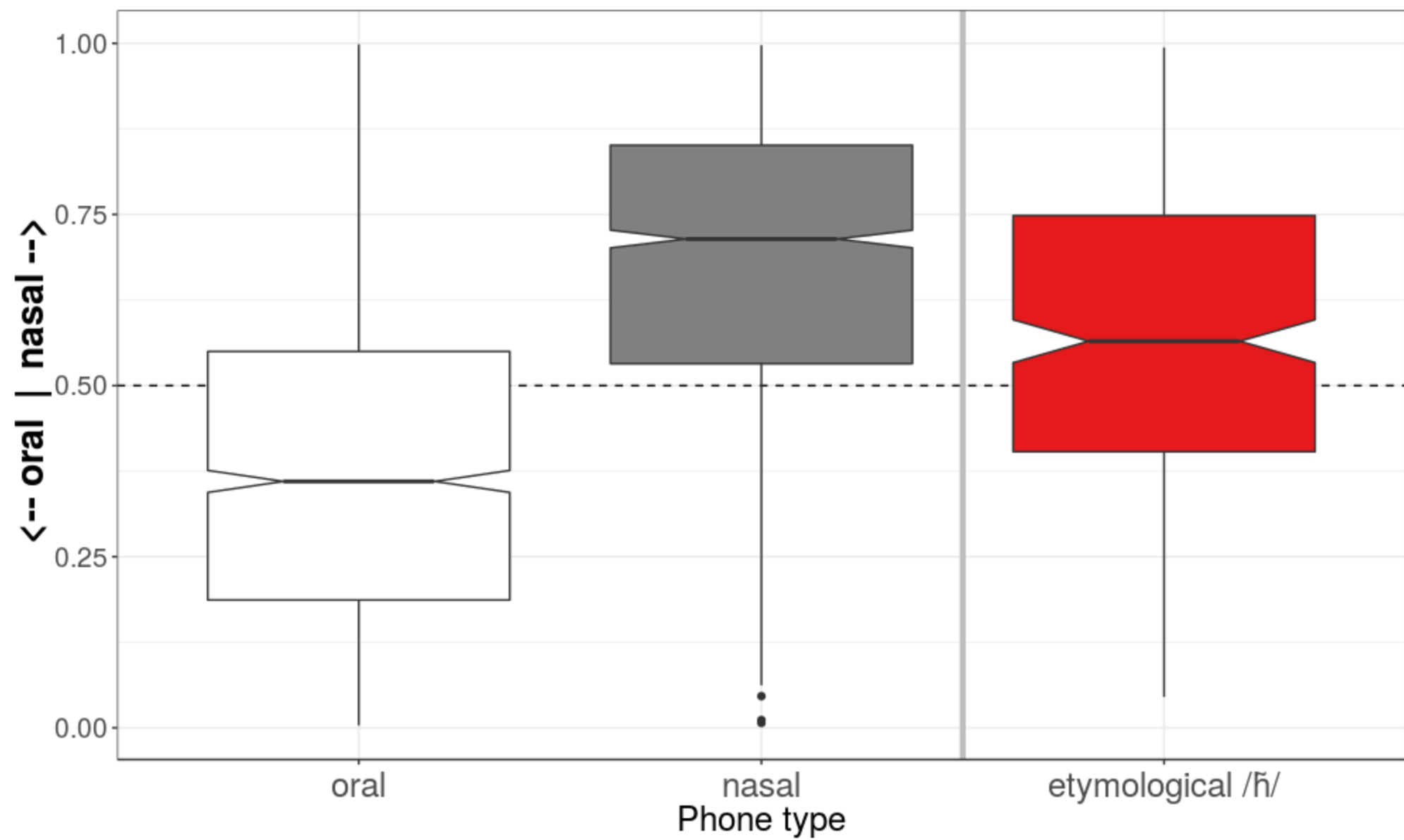
/h̃/ Phonetic experimentation

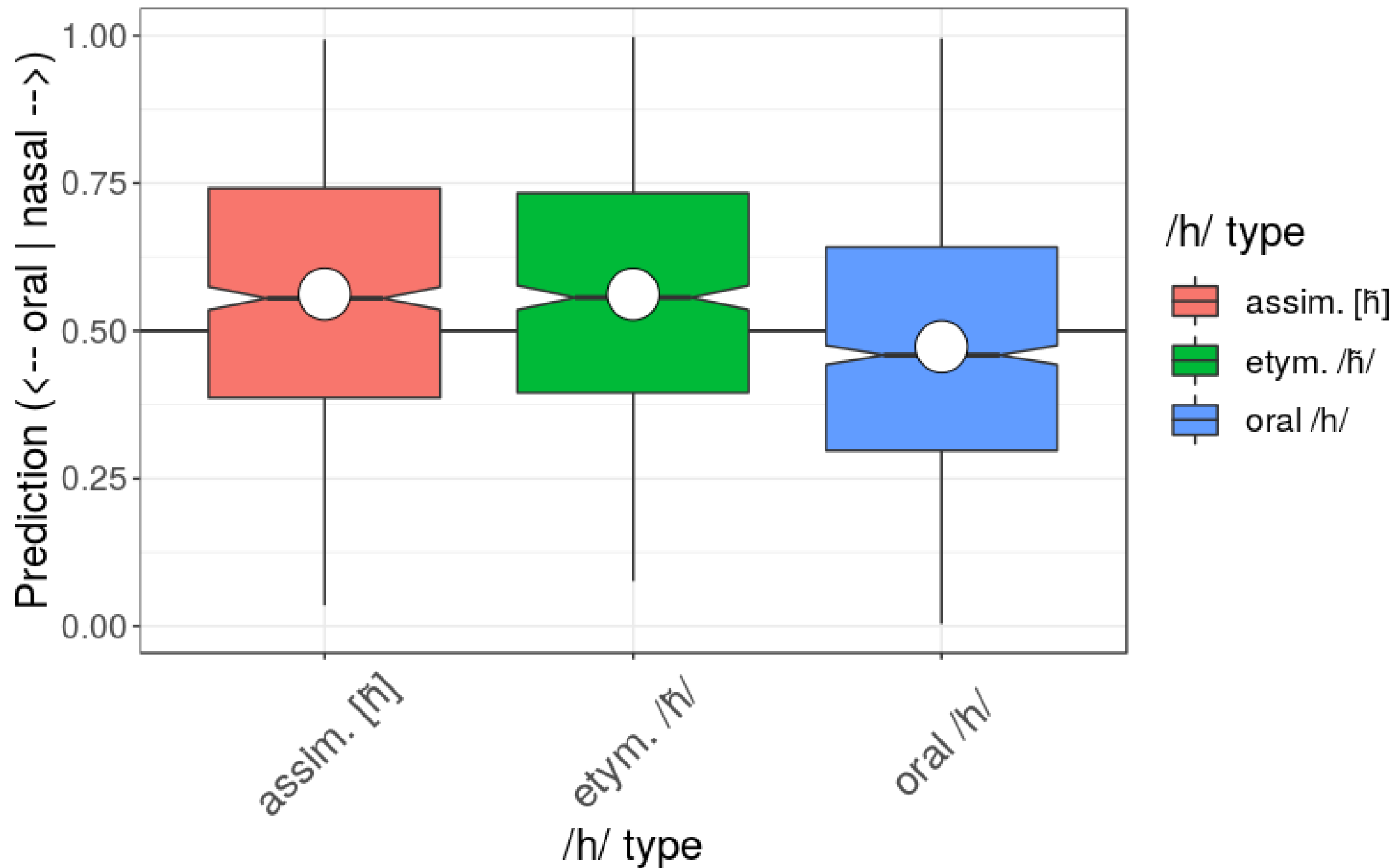


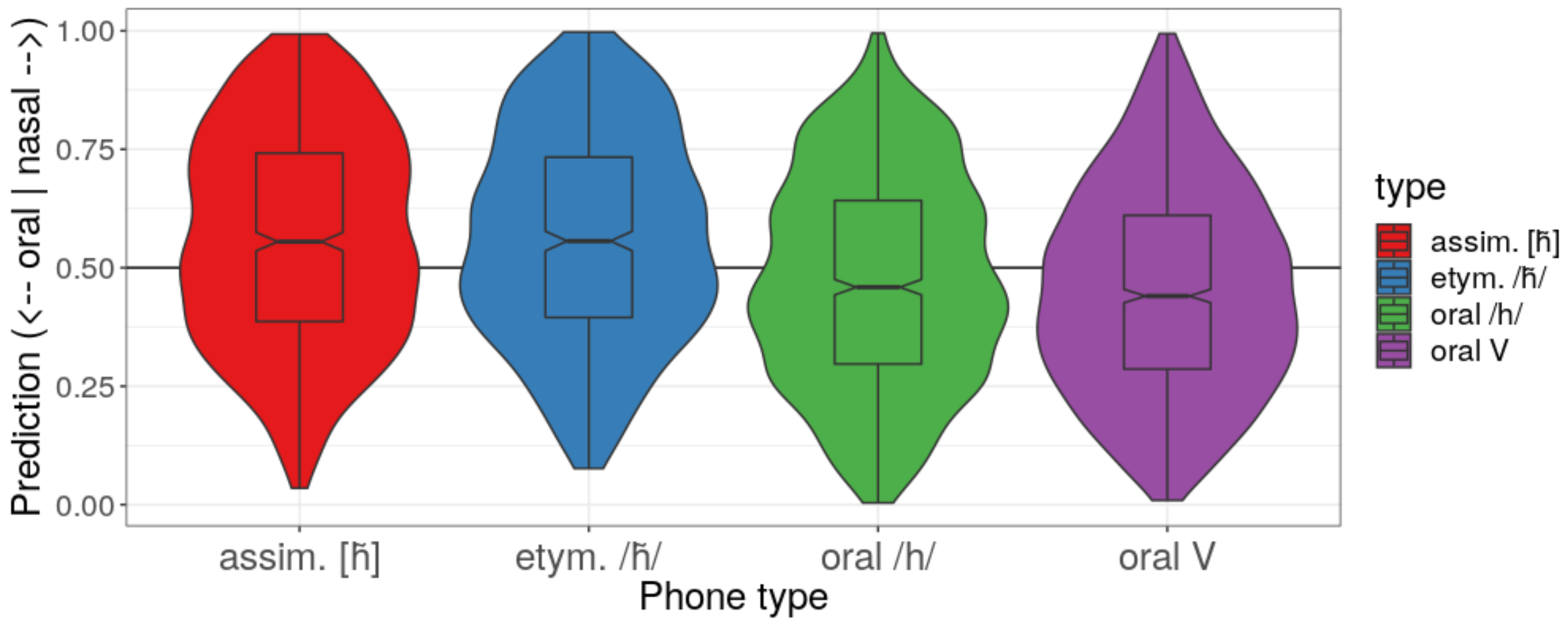
i h i z i n

- Nasality is the only feature that differentiates /h/ and /h̃/.
- No reliable way of measuring acoustic nasality.
- New model to detect nasality.

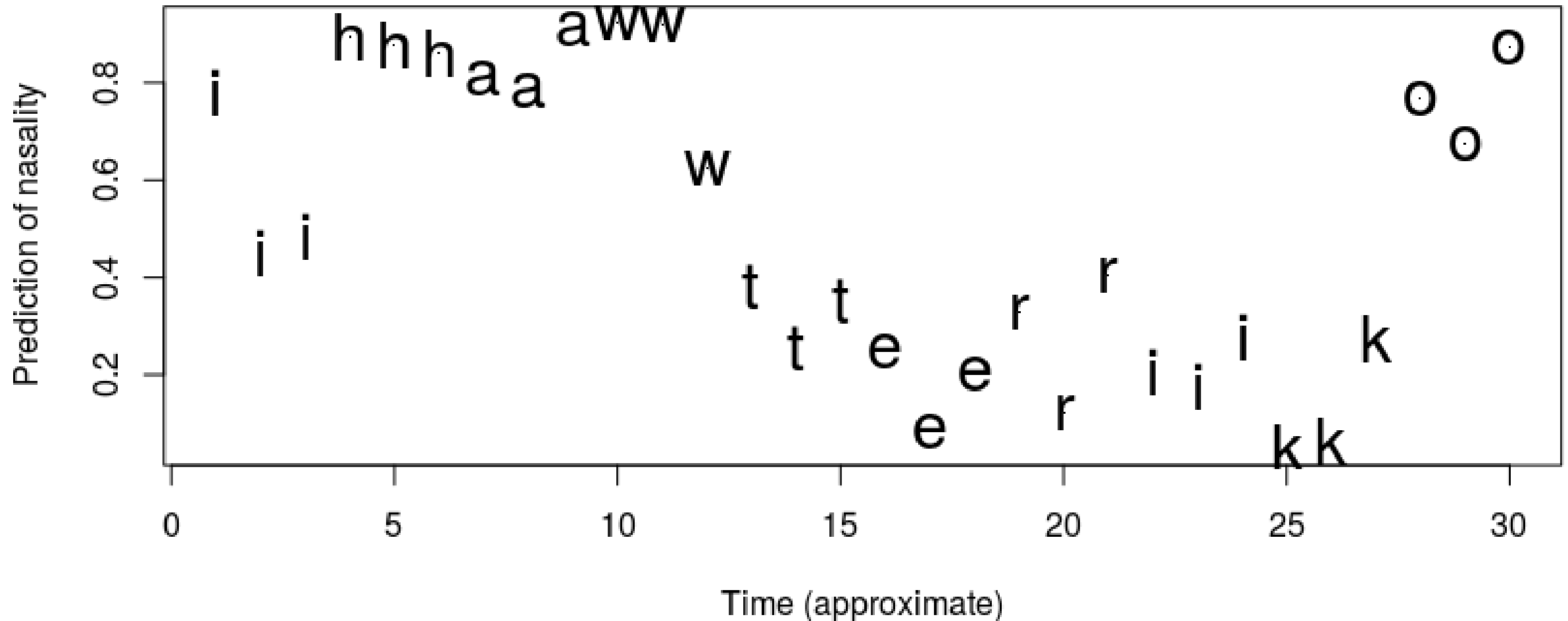
ihizin 'in the hunting'







Nasalized / $\tilde{h}$ / in a word: *ihauteriko* ‘of the carnival’



Eastern accentual system

Basque prosodic systems

The diversity of prosodic systems is wide in modern Basque (cf. Hualde 1997)

- **Central:** word-level stress in the peninitial syllable, marked words +1.
- **Eastern:** word-level stress in the penultimate syllable, marked words –1.
- **Goizueta:** 2 different pitch accents that can fall in either of the first two syllables.
- **Western:** only some words are lexically accented, phrasal pitch accent.

Eastern accentuation

- Unmarked stress: penultimate syllable of the word
 - *gízun* ‘man’, *gizúna* ‘the man’, *néska* ‘girl’, *alhába* ‘daughter’...
- Marked stress: last syllable
 - *alhabá* ‘the daughter’, *neská* ‘the girl’, *ardű* ‘wine’, *añái* ‘ram’
 - monosyllabic suffixes (cf. *-ñí*, *-xkót*, *-liár*, *-tiár*, *-(t)ár*, *-kór* or the borrowed *-ús*).
 - compounds with a monosyllable: *gibel-mín* ‘bile, gall’, *giza-txár* ‘bad man’...
- Plural *-e-* also marked: attracts stress
 - *gizunék* ‘the men (erg.)’, *gizunér* ‘to the men’, *gizunéna* ‘of the men’, *gizunéki* ‘with the men’, *gizunéntako* ‘for the men’...

From the Central to the Eastern system

emakume

emakume





emakume

emak*u*me

emakume¹



emak²ume¹



gizona

gizona





gizona

gizona¹





From the Central to the Eastern system

- A reinterpretation of the peninitial stress in trisyllabic words as originating in the penultimate would give rise to the Eastern pattern (Michelena 1961/77).

$$[\sigma \sigma]'([\sigma \textcolor{red}{\sigma} \acute{\sigma}]) > [\sigma \acute{\sigma}]$$

- This is only possible at a time when most words were trisyllabic: after the grammaticalization of the determiner -a < *ha(r).

From the Central to the Eastern system

- After years of Gascon influence, trisyllabic words were ambiguous. This ambiguity was resolved in an innovative way due to Gascon influence.

$$[\sigma \sigma]'([\sigma \textcolor{red}{\sigma} \acute{\sigma}]) > [\sigma \acute{\sigma}]$$

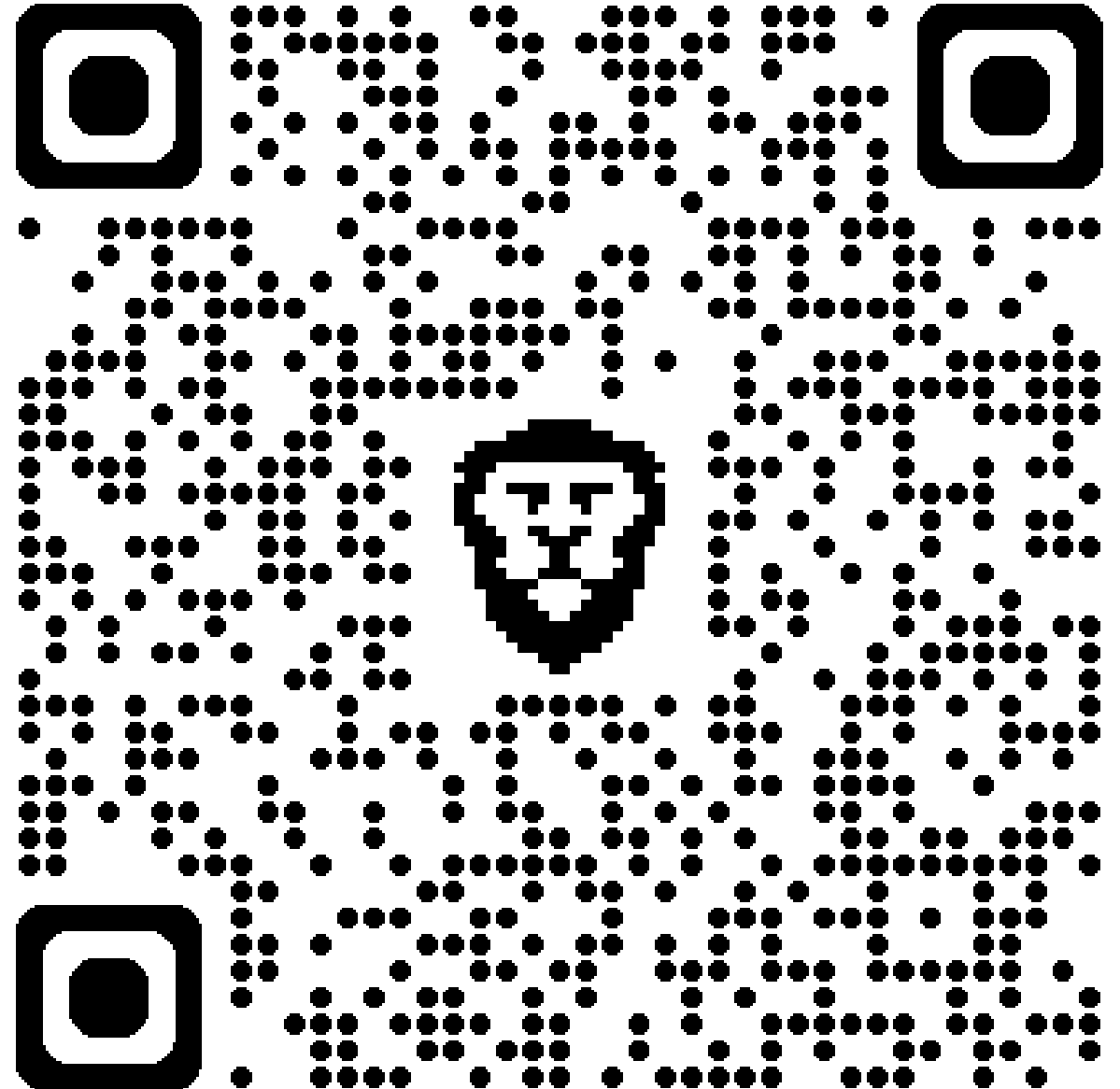
- Bearnese accentuation:
 - Words are either oxytonous or paroxytonous.
 - Accent is assigned based on a number of rules (Mooney 2016).

Today's paper:

Egurtzegi, A. & C. Carignan. 2020. An acoustic description of Mixean Basque. *Journal of the Acoustical Society of America* 147.4, 2791-2802.

Open access here:

<https://doi.org/10.1121/10.0000996>



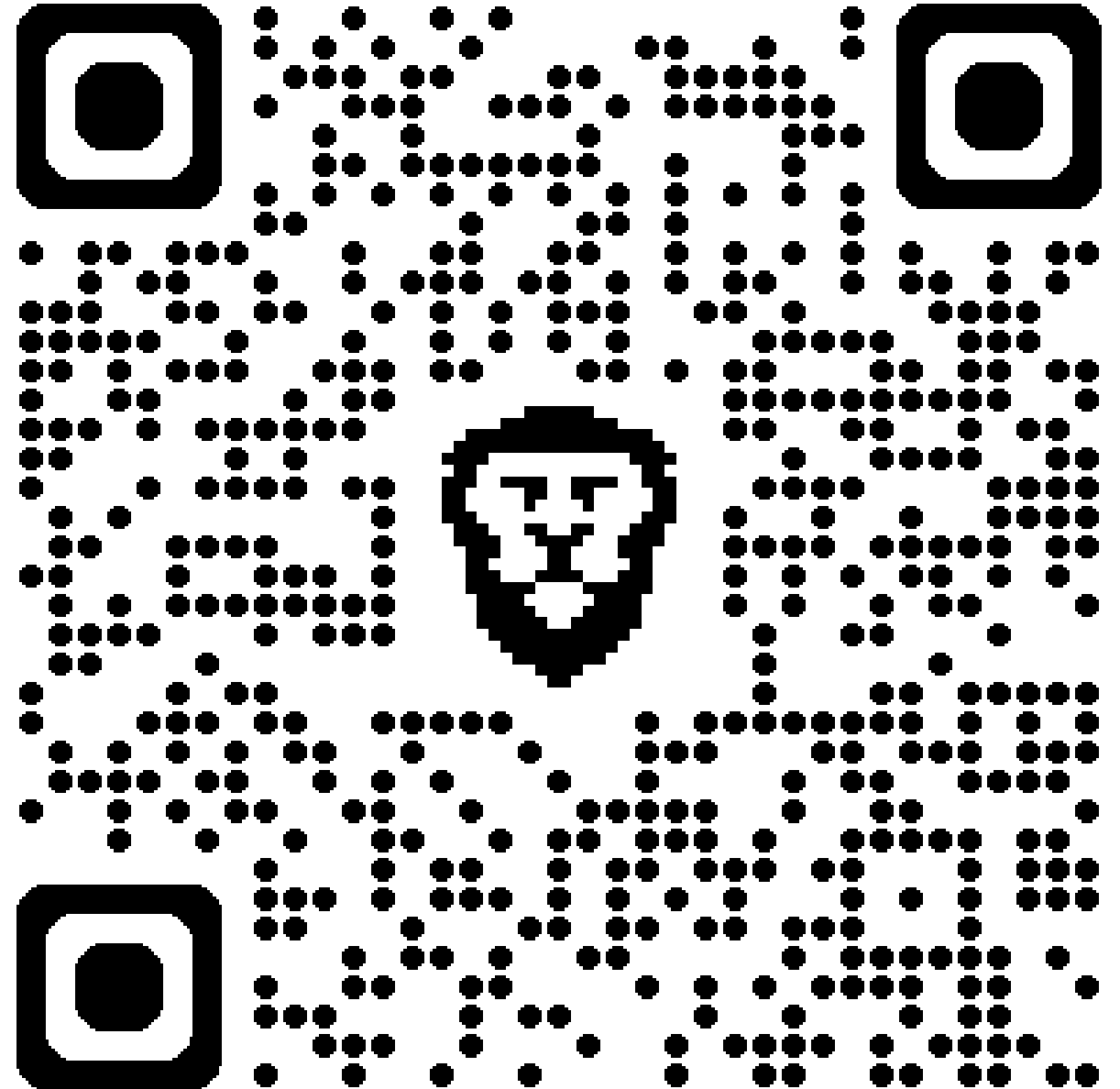
All papers are freely accessible here: egurtzegi.github.io/publications

Next paper:

Egurtzegi, A. 2025. Metathesis. In
Ledgeway et al. (eds.), *The Wiley
Blackwell Companion to Diachronic
Linguistics*. Wiley-Blackwell.

Open access here:

<https://egurtzegi.github.io/papers/MetathesisDiaComEgurtzegi.pdf>



All papers are freely accessible here: egurtzegi.github.io/publications