



The production of phonation type in San Martín Peras Mixtec

Ben Eischens and Jahnavi Narkar
University of California, Los Angeles



Acknowledgements

This is joint work with Jahnvi Narkar (UCLA), supported by undergraduate Research Assistant Saíd Bracamontes

We are extremely grateful to members of the San Martín Peras community for generously sharing their language with us. Without their knowledge and effort, none of this work would have been possible.

We especially thank Angie Mendoza Rivera, Angel Mendoza Rivera, Gerardo Espinosa Gracida, Juan Carlos Gracida, Juan Gracida Ortiz, Javier Mendoza Rivera, Lorena Espinosa Rivera, Reina Mendoza Rivera, and Vicenta Espinosa Rivera



Introduction

The **phonology** of [ʔ] in Mixtec languages is a topic of frequent debate. It is variably analyzed as...

- A consonant (Pike and Cowan 1967; North and Shields 1977; Peters 2018; Belmar 2024)
- The realization of glottalized phonation on the vowel, root, or foot (Josserand 1983; Macaulay and Salmons 1995; Gerfen 1996; Macken and Salmons 1997; Penner 2019)

But remarkably little is known about the **phonetics** of [ʔ] in Mixtec languages (c.f., Gerfen & Baker 2005; Cortés et al 2023)



Introduction

This presentation: Acoustic studies of [ʔ] and [h] in San Martín Peras Mixtec (SMP Mixtec), replicating the methodology of previous studies.

High interspeaker **consistency**, as opposed to previously reported interspeaker **variation** (Gerfen & Baker 2005; Cortés et al 2023)

- [ʔ] and [h] almost categorically realized with **voicelessness**, despite intervocalic position (cf. Davidson 2021; Garellek et al 2023)



Background



Background

SMP Mixtec is spoken by ~11,500 in and around San Martín Peras, Oaxaca, Mexico (INEG 2020)

Also spoken by diaspora communities in California (Mendoza 2020)





Ahuejutla, Oaxaca

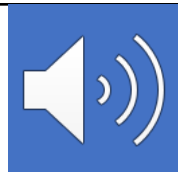


Background: The language

Canonical root shape is bimoraic, with optional onsets and no codas

- Two short vowels or one long vowel

CVCV	VCV	CVV	VV
Ntivi [ⁿ tiβi] ‘beautiful’ ‘bonito’	ìvi [iβi] ‘two’ ‘dos’	Ntsì [ⁿ tsì̃] ‘dead’ ‘muerto’	lin [ĩ] ‘one’ ‘uno’



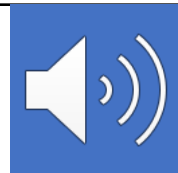


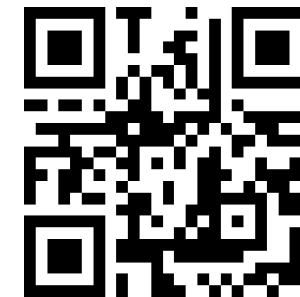
Background: The language

Each mora may host one of (at least) five contrastive tones (Peters 2018)

- H(igh), M(id), L(ow), R(ising), and F(alling)

H-L	M-L	L-L	LH-L	HL-L
Nánà [nǎnǎ] 'mother' 'madre'	Konì [konǐ] 'will see' 'verá'	Ònì [ònǐ] 'three' 'tres'	Xǐyò [ʃijò] 'dress' 'vestido'	Xânù [ʃânù] 'cigarette' 'cigarillo'





Background: The language

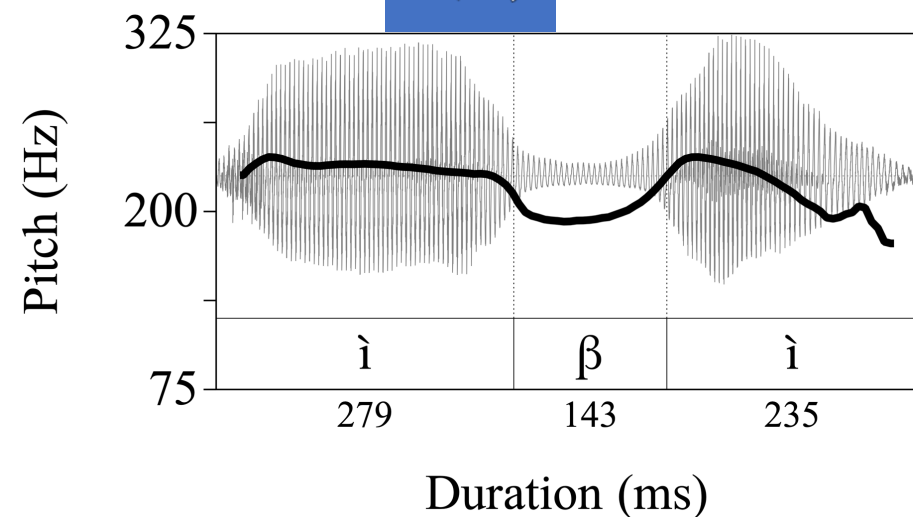
Contrast between [V] (1) and [Vʔ] (2), largely irrespective of tone (Eischens and Hedding to appear)

1. ìvì

[ìβì]

'two'

'dos'

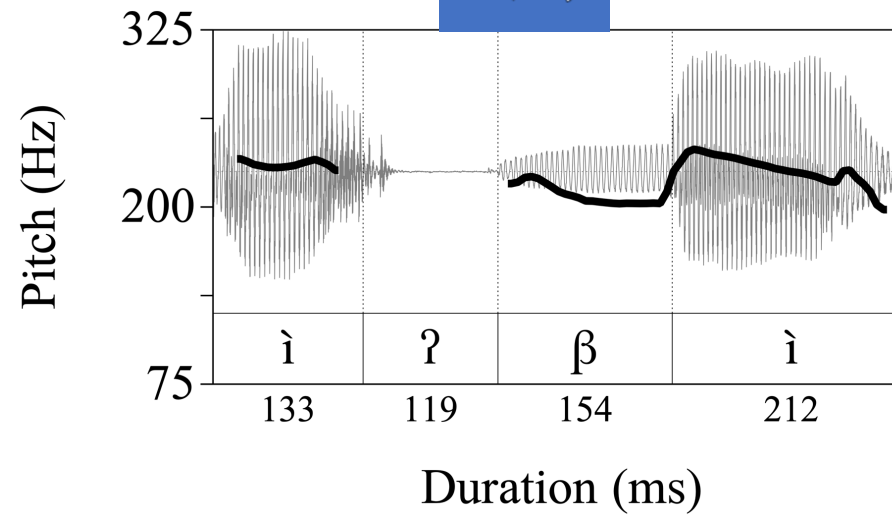
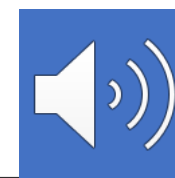


2. ì'vì

[ìʔβì]

'pain'

'dolor'





Background: The language

Contrast between $[V^ʔ]$ (1) and $[V^h]$ (2) (Eischens and Hedding to appear).

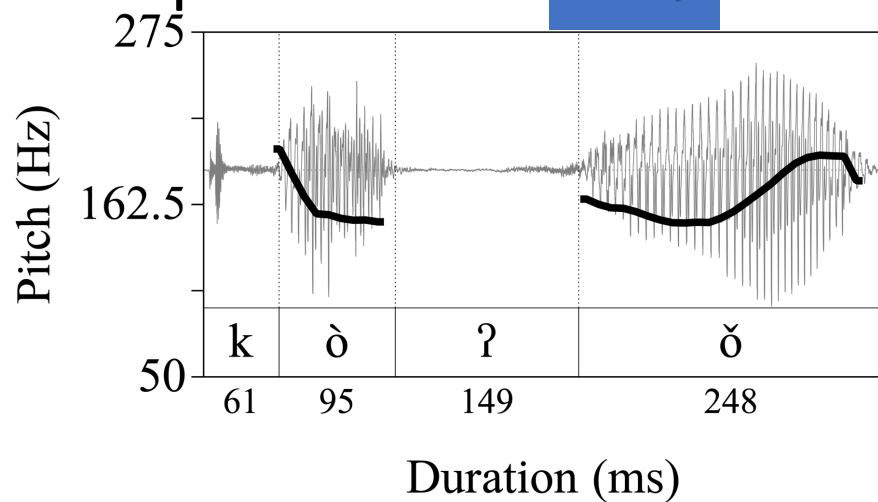
- $[V^h]$ is tonally restricted (Eischens et al 2024)

1. Kò'ǒ

[kòʔǒ]

'plate/bowl'

'plato'

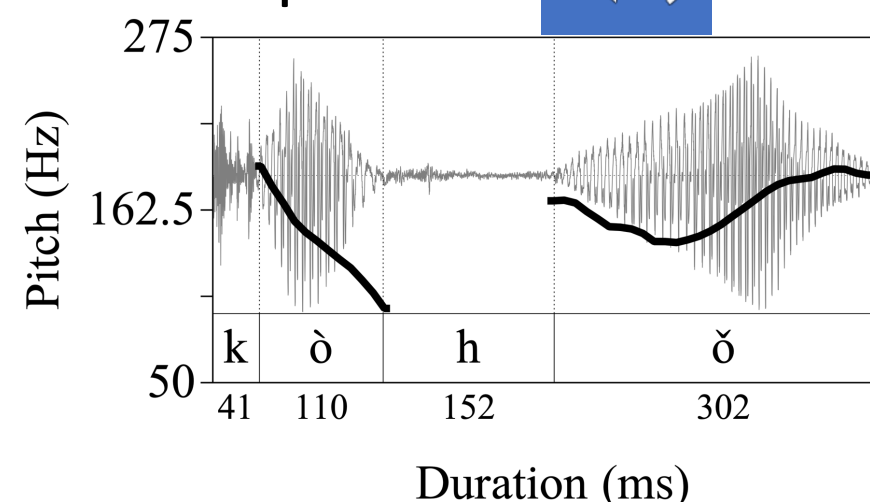


2. Kòjǒ

[kò^hǒ]

'snake'

'serpiente'

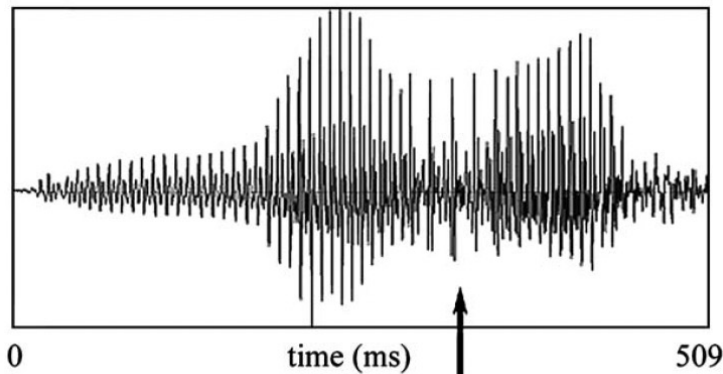




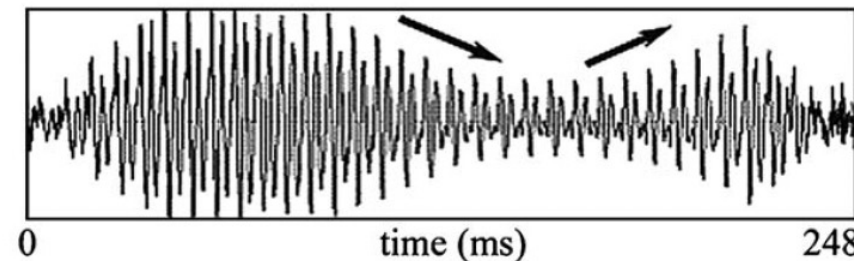
Background: Previous work

Gerfen and Baker (2005) examined the production and perception of laryngealized vowels in Coatzospan Mixtec

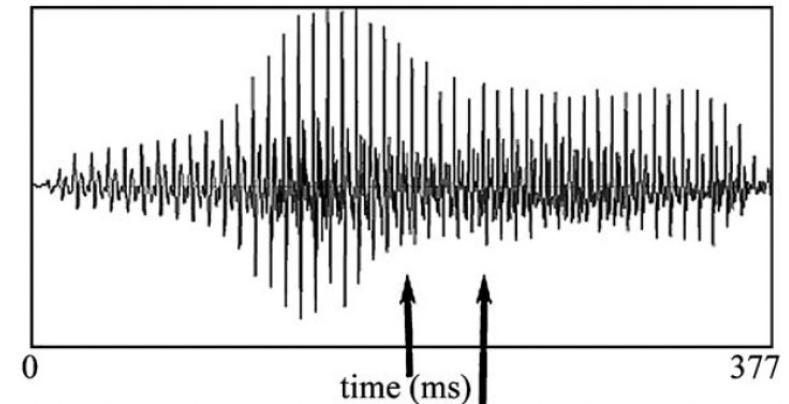
- Production of words in isolation
- Variety of realizations, many very subtle



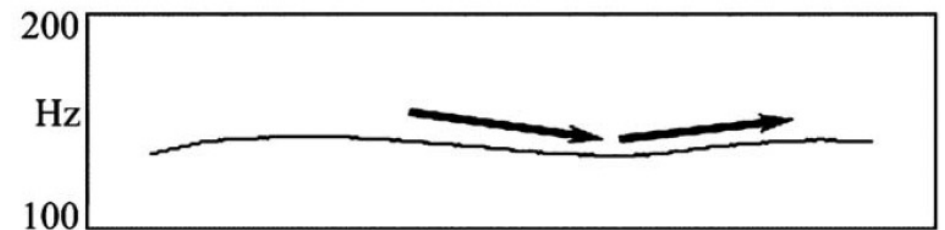
Creaky voicing



Amplitude dip



No clear laryngealization, but still H1-H2 differences



Small pitch dip



Background: Previous work

Cortés et al (2023) investigated the production of [VʔV] sequences in San Sebastián del Monte Mixtec

- Productions of words in a carrier phrase
- Variety of realizations, high inter- and intra-speaker variation



Production study

Production study: Methods and materials

8 participants (3 female, ages 18-75) in Ahuejutla, Oaxaca produced words from a wordlist in isolation

Words presented as pictures, alongside working orthography, Spanish translation, and example pronunciation

- No widely used orthography



Production study: Methods and materials

Wordlist contained 52 words, presented 6 times in a randomized order

- 312 productions per participant

Following Gerfen and Baker (2005), the wordlist contained minimal pairs between [V] and [Vʔ]

- 10 [V] ~ [Vʔ] minimal pairs (20 words total, 38% of items)

- 12 roots with [V^h]

- 12 roots with medial voiceless C

- 4 other [Vʔ] and [V] words not in minimal pairs

Production study: Analysis

All productions hand-annotated in Praat by undergraduate Research Assistant and double-checked by first author

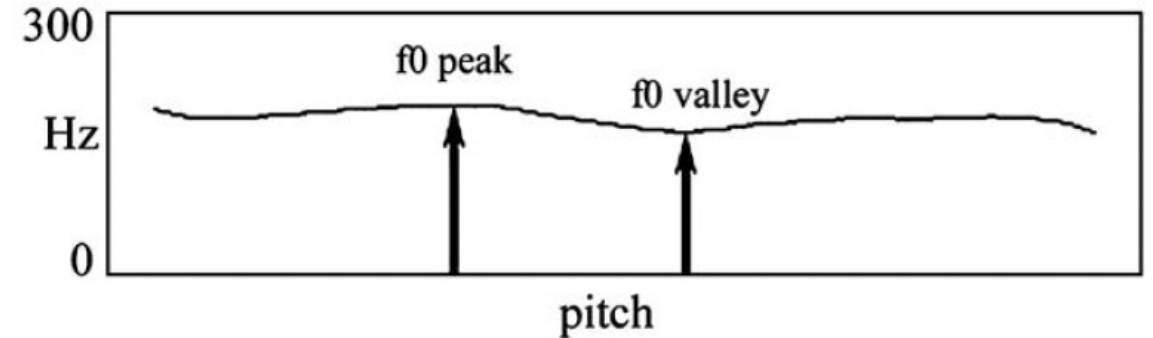
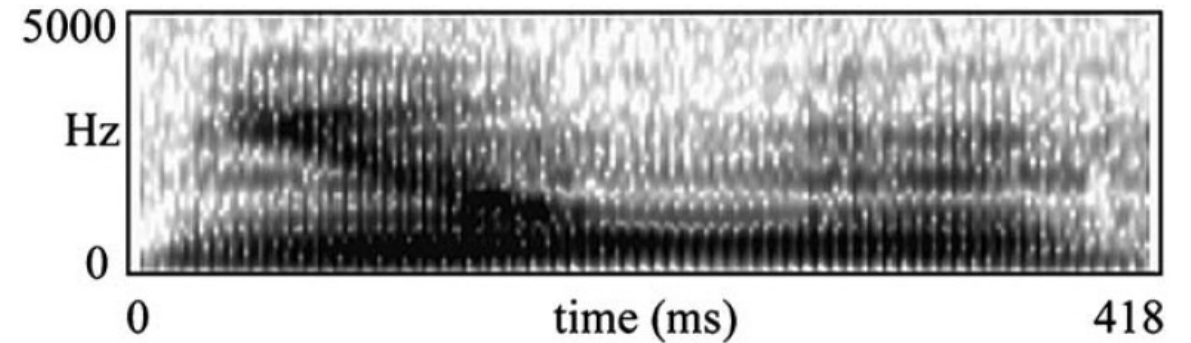
Investigated differences between phonation types in the following measures:

- Duration
- Pitch
- H1-H2
- Voicing

Production study: Analysis

Caveat: Gerfen and Baker (2005) examined these measures through the entire $[V^?]$ or $[V^?V]$ sequence.

- $[V^?]$ and $[V^?V]$ were almost always periodically voiced throughout



Production study: Analysis

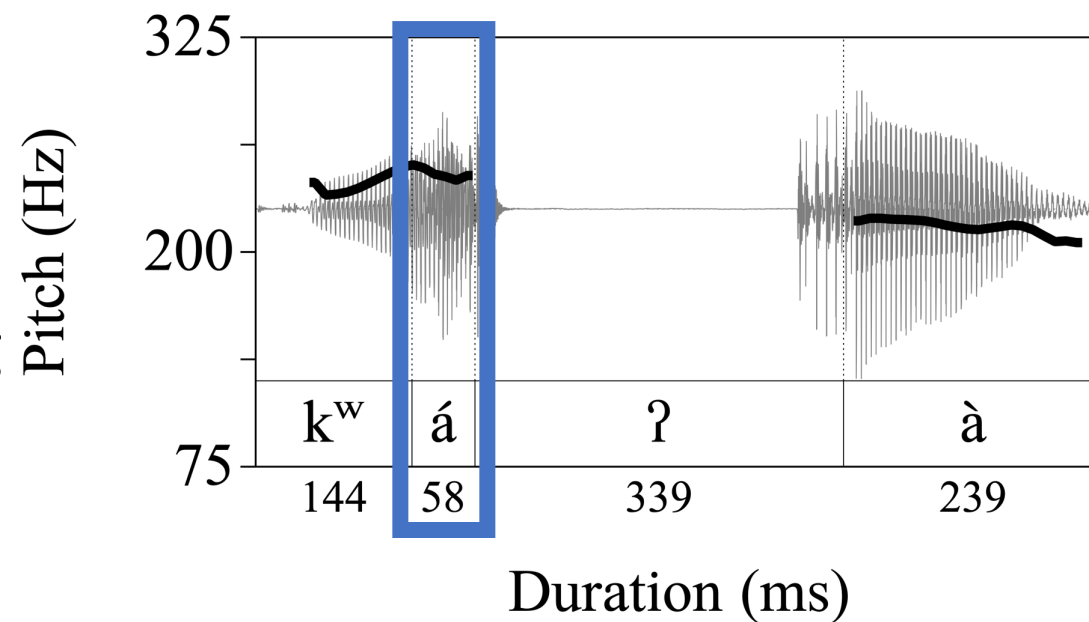
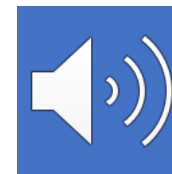
But our $[V^ʔ]$ and $[V^h]$ almost always contained voicelessness

- Pitch, H1-H2 are not comparable across $[CVV] \sim [CV^ʔV] \sim [CV^hV]$

Compared measures from the **periodically-voiced portion** preceding $[ʔ]$ and $[h]$

- We'll call this **V1**

1. Kuá'à
 $[k^w á^ʔ à]$
 'red'
 'rojo'



Production study: Analysis

We compared these to modal vowels followed by a medial consonant

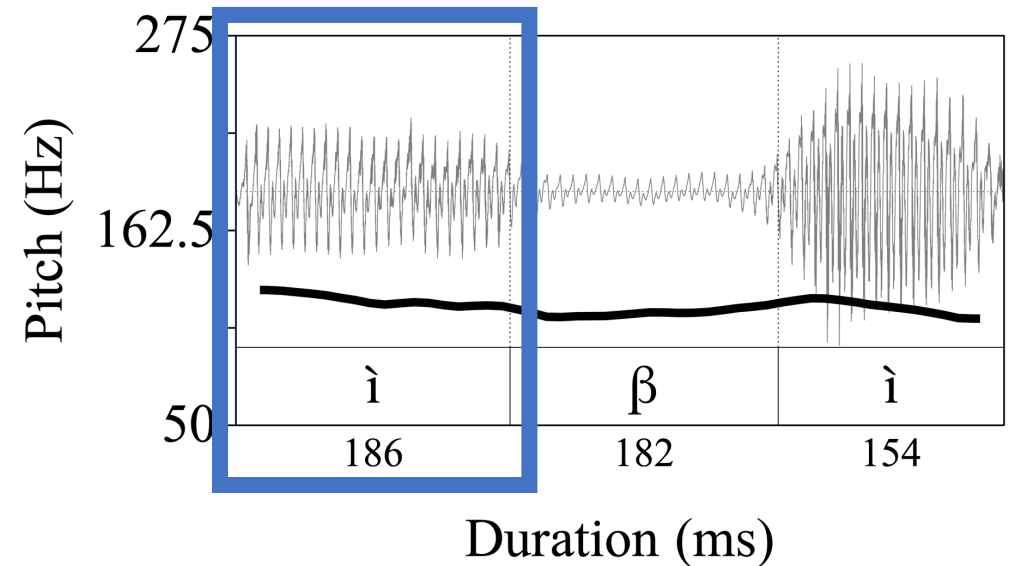
- Comparable sequential structure

1. ìvì

[ìβì]

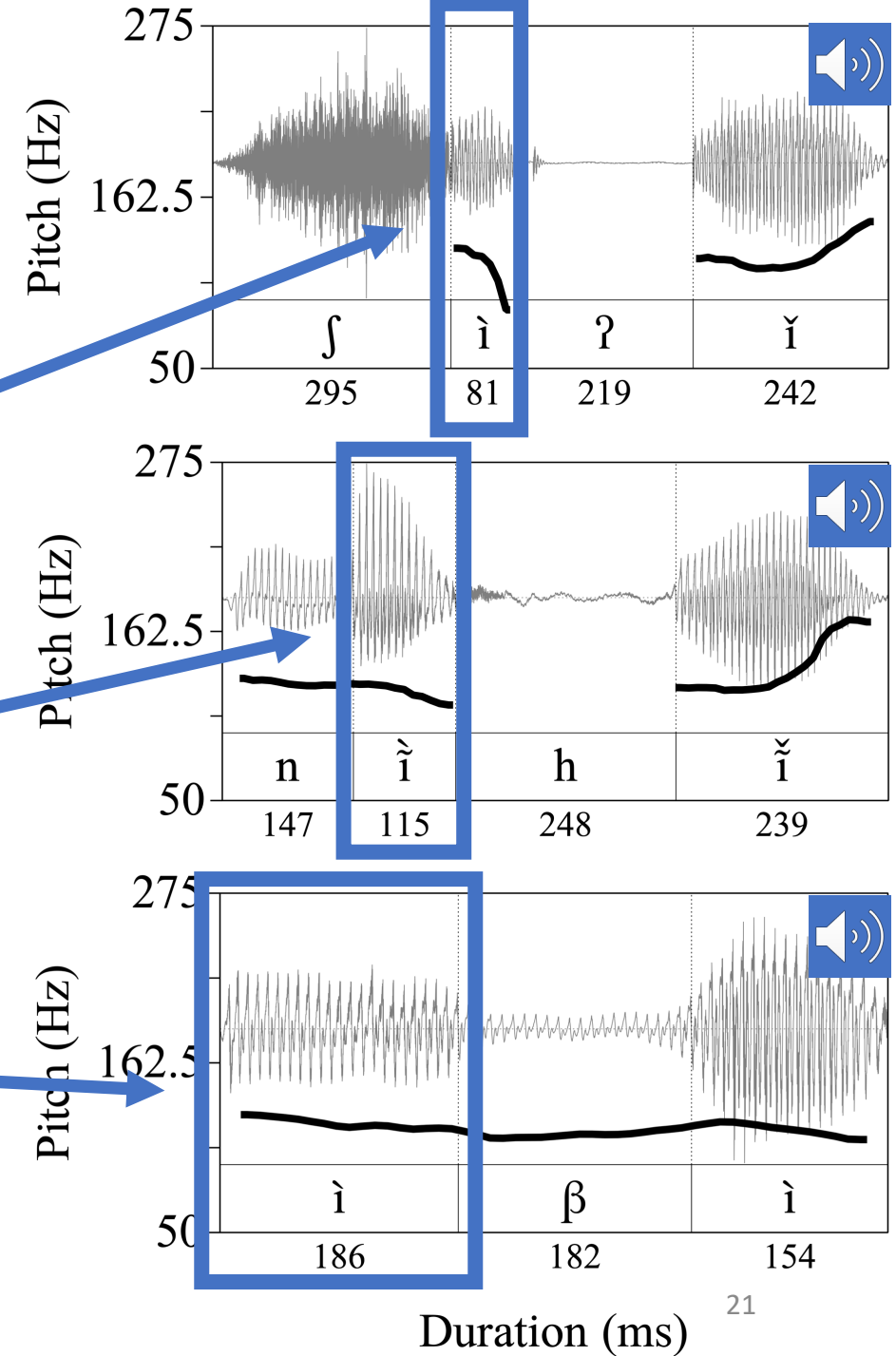
'two'

'dos'



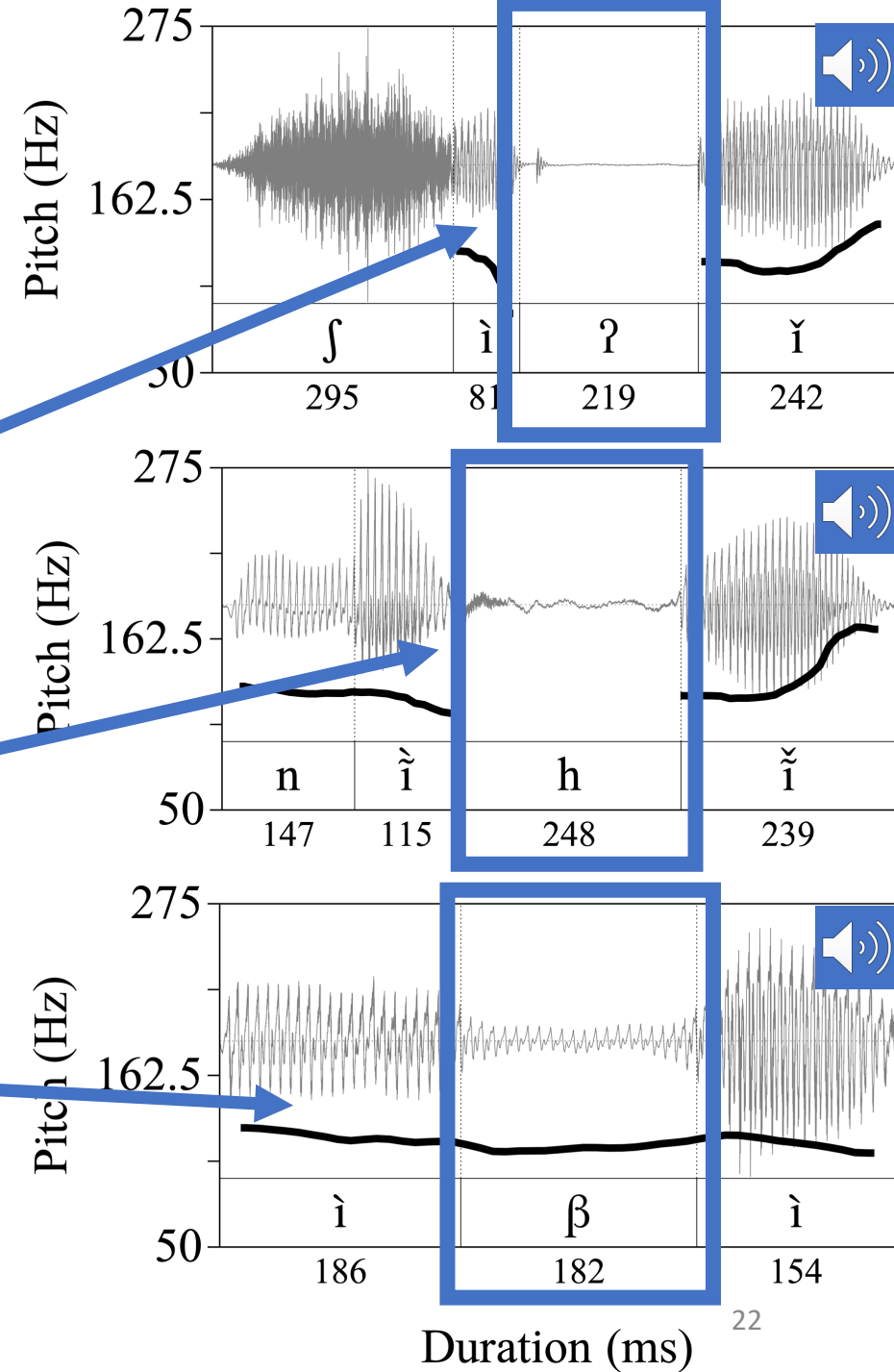
Production study: Analysis

	C	V ₁	X	V ₂
'mushroom'	ʃ	ì	ʔ	ĩ
'blood'	n	ĩ	h	ĩ
'two'	--	ì	β	ì



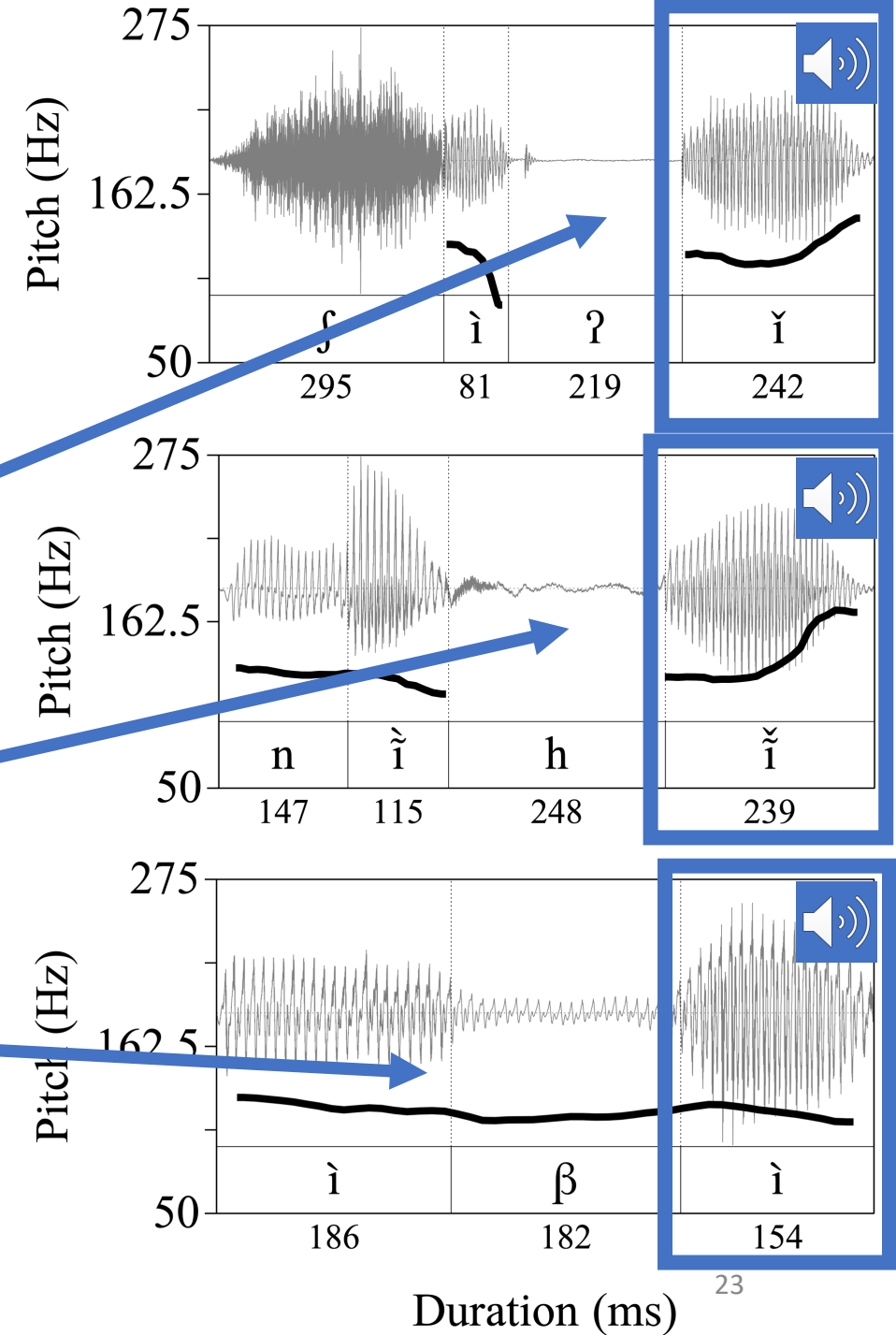
Production study: Analysis

	C	V ₁	X	V ₂
'mushroom'	ʃ	ì	ʔ	ĩ
'blood'	n	ĩ	h	ĩ
'two'	--	ì	β	ì



Production study: Analysis

	C	V ₁	X	V ₂
'mushroom'	ʃ	ì	ʔ	ĩ
'blood'	n	ĩ̃	h	ĩ̃
'two'	--	ì	β	ì



Production study: Analysis

Bayesian mixed effects model (Bürkner 2017) in R

- Different dependent variables, independent variable of phonation
- Random intercepts for item, random slopes for participants
- Weakly informative priors

Output:

- Effect relative to baseline
- Credible interval
 - If it does contains 0, the effect is not credible
 - If it does not contain 0, the effect is credible

Production study: Results

Duration

Compared V_1 duration of $[V_1^?]$ and $[V_1^h]$ to modal $[V_1C]$

V1

$[V_1^?]$	$[V_1^h]$	Modal $[V_1C]$
[s <u>i</u> [?] ʃ]	[n <u>i</u> ^h ʃ]	[<u>i</u> β i]

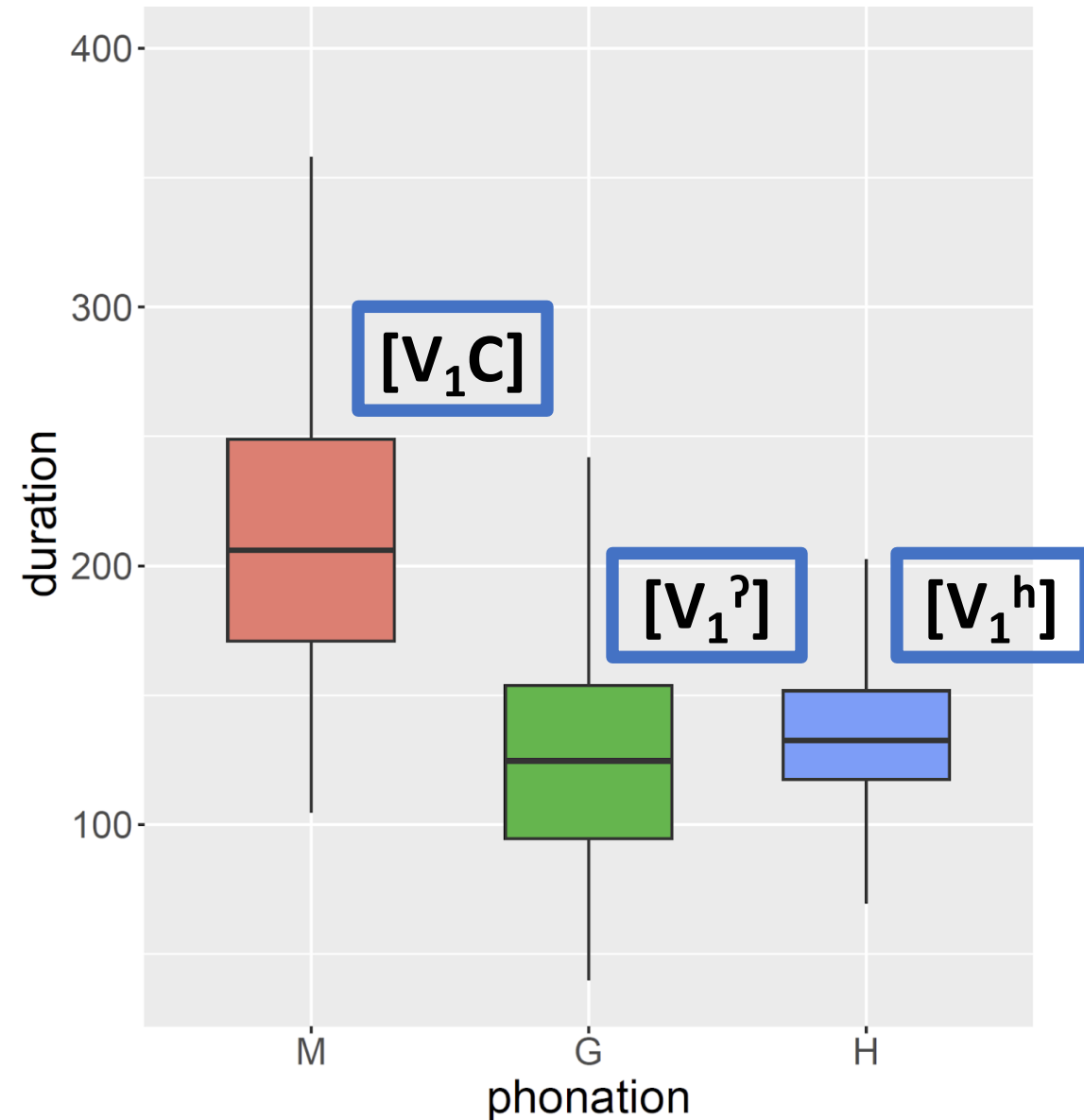
Production study: Results

Duration

V_1 in $[V_1^?]$ and $[V_1^h]$ credibly shorter than modal $[V_1C]$ baseline

$[V_1^?]$ effect: -84.5 (CI = [-118, -49])

$[V_1^h]$ effect: -74.58 (CI = [-104, -45])



Duration of V_1 (ms) by phonation type across all speakers

Production study: Results

Finding:

- V_1 of $[V_1^?]$ and V_1 of $[V_1^h]$ are reliably shorter than modal $[V_1C]$

$[V_1^?]$	$[V_1^h]$	Modal $[V_1C]$
Shorter V_1 duration		Longer V_1 duration

Production study: Results

Pitch

Due to tonotactic restrictions, we only compared initial L tones across vowel types

- In L-L and L-LH melodies

V1

$[V_1^?]$	$[V_1^h]$	Modal $[V_1C]$
$[s \underset{\square}{i}^? \check{y}]$	$[n \underset{\square}{i}^h \check{y}]$	$[\underset{\square}{i} \beta \underset{\square}{i}]$

Production study: Results

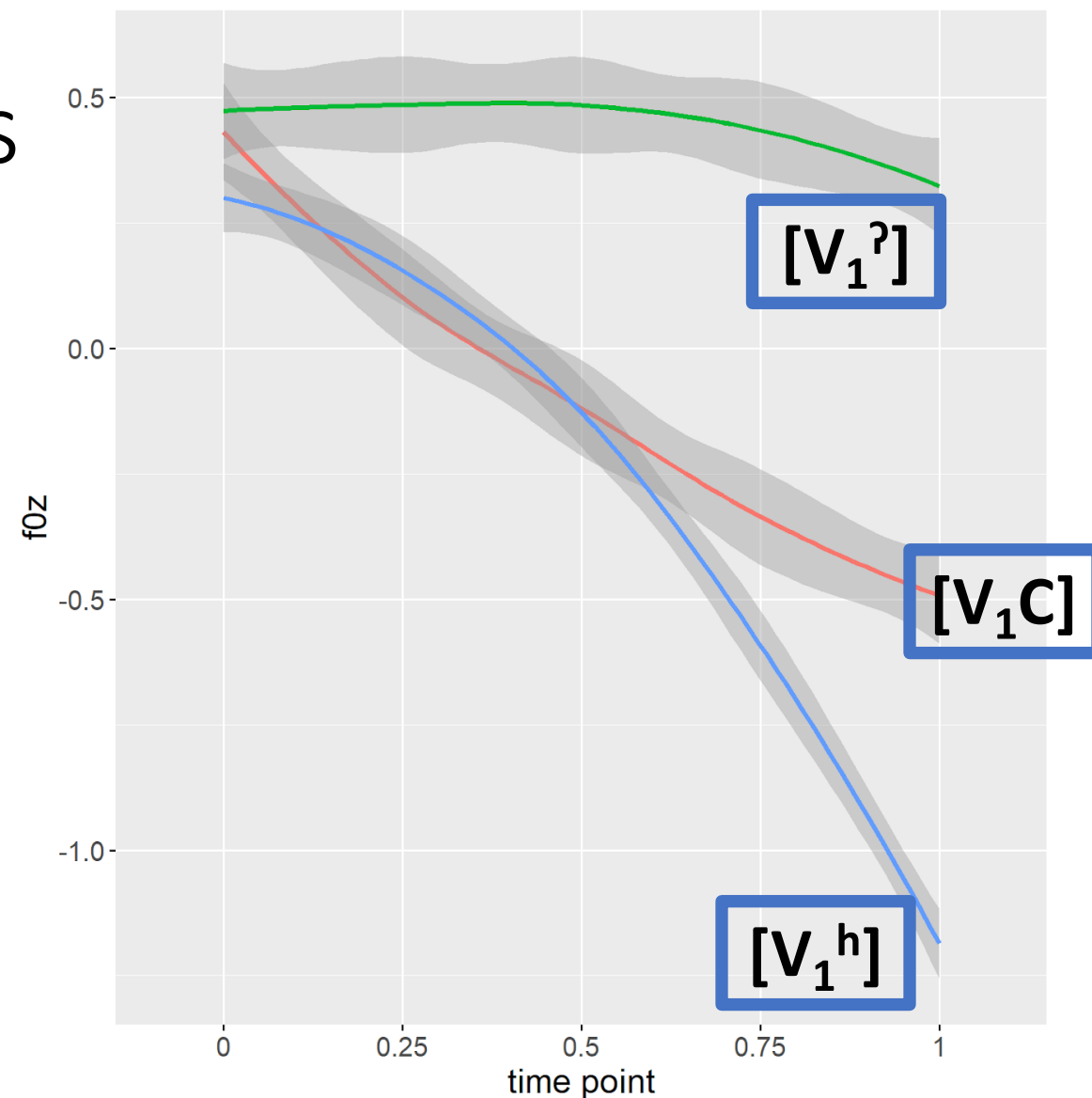
Pitch fell in all conditions

Pitch on modal $[V_1C]$ fell from onset to offset

- Effect: -0.91 (CI = $[-1.06, -0.76]$)

Pitch on $[V_1^?]$ lowers, but not as much as modal $[V_1C]$

- Effect: $-0.91 + 0.76 = -0.15$
(CI = $[0.58, 0.94]$)

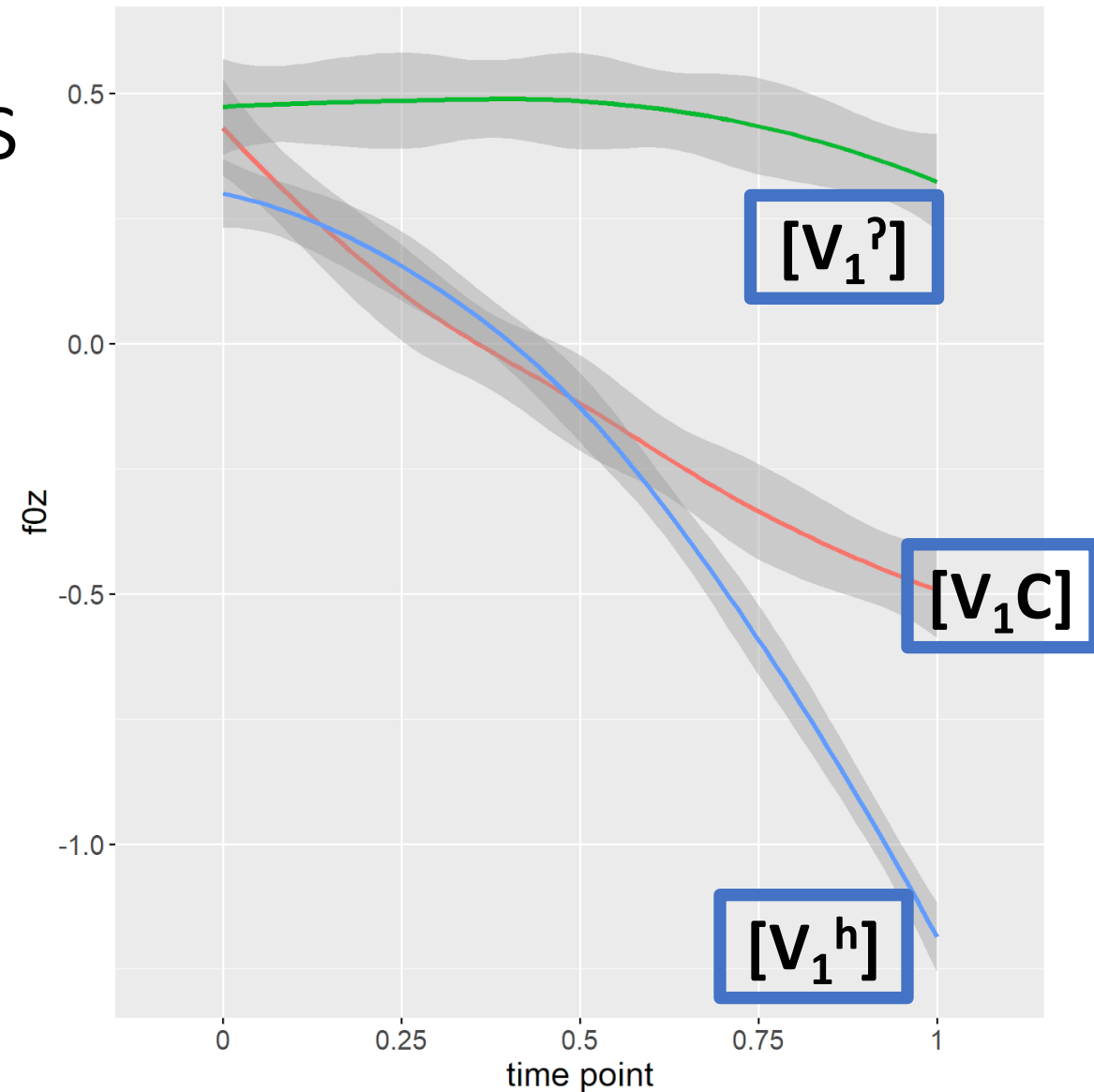


Production study: Results

Pitch fell in all conditions

Pitch on $[V_1^h]$ lowers more than on modal $[V_1C]$

- Effect = $-0.91 + -0.58 = -1.49$
(CI = $[-0.75, -0.4]$)



Normalized pitch on V1 across phonation types

Production study: Results

Finding:

- Pitch of all L tones falls, but fall is smaller in $[V_1^?]$ and larger in $[V_1^h]$

$[V_1^?]$	$[V_1^h]$	Modal $[V_1C]$
Shorter V_1 duration		Longer V_1 duration
Small pitch fall	Steep pitch fall	Moderate pitch fall

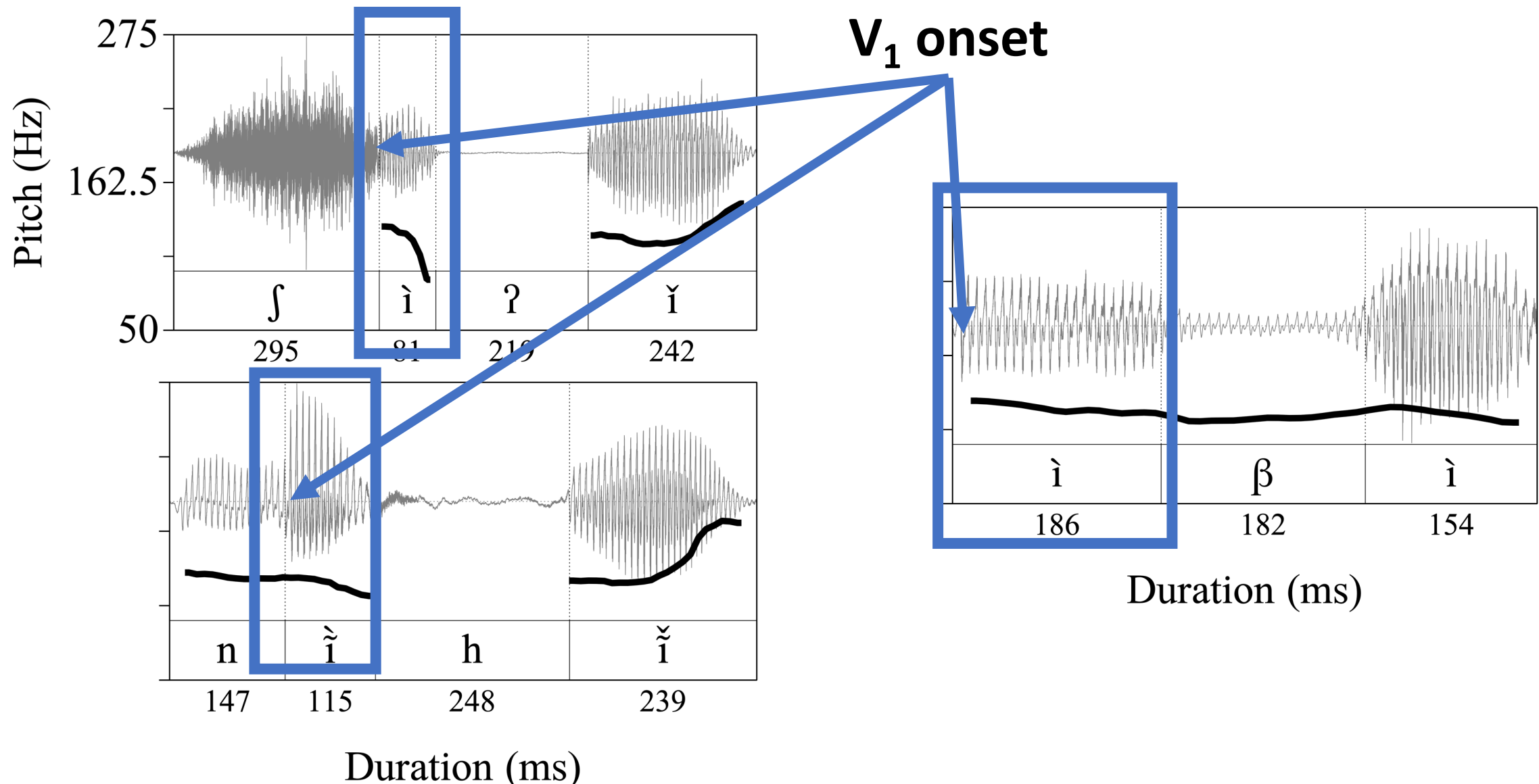
Production study: Results

H1-H2

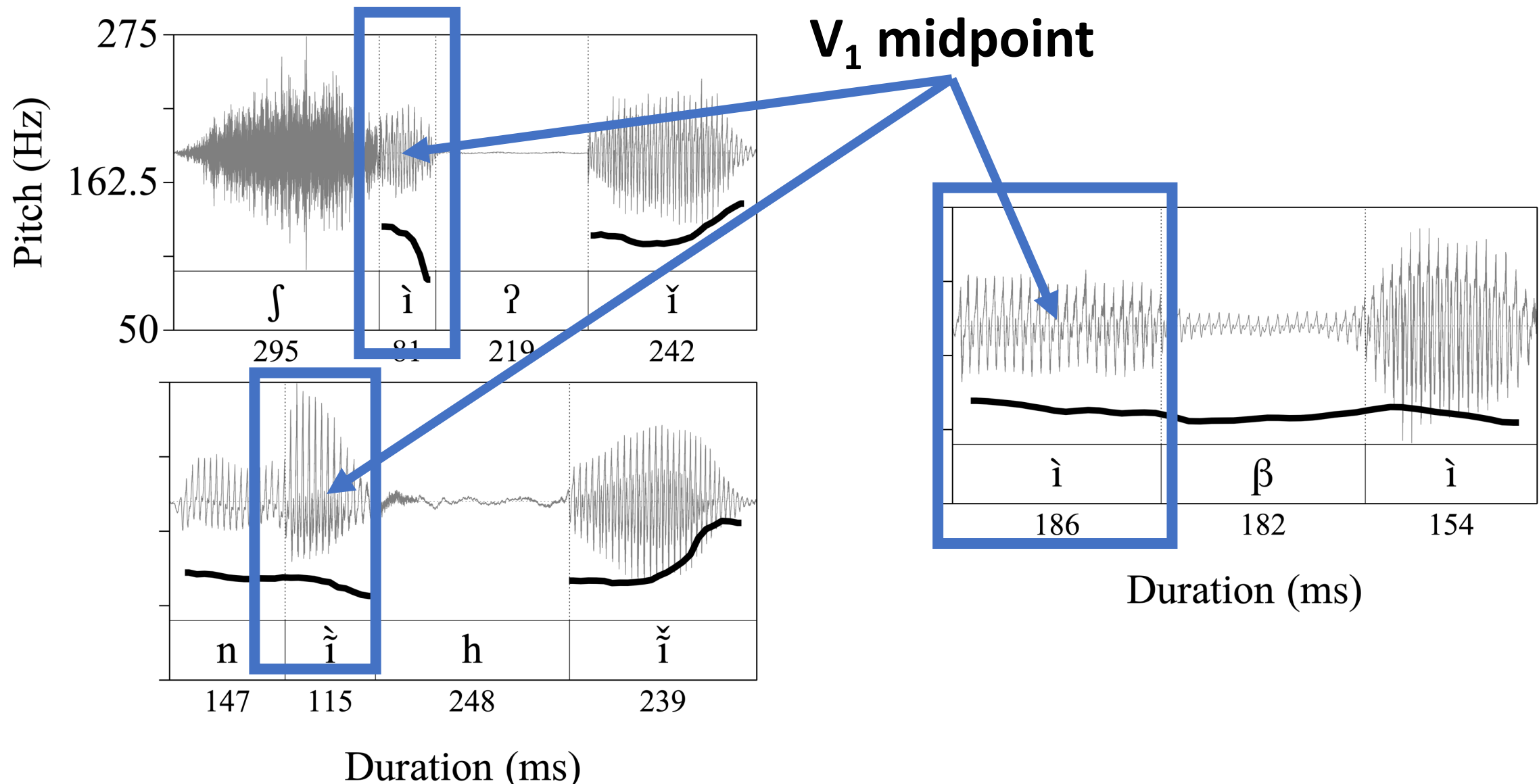
Usually lower in creaky voice, higher in breathy voice (Kreiman et al 2010; Keating et al 2015)

Compared formant-corrected *H1-H2 on V1 at three time points

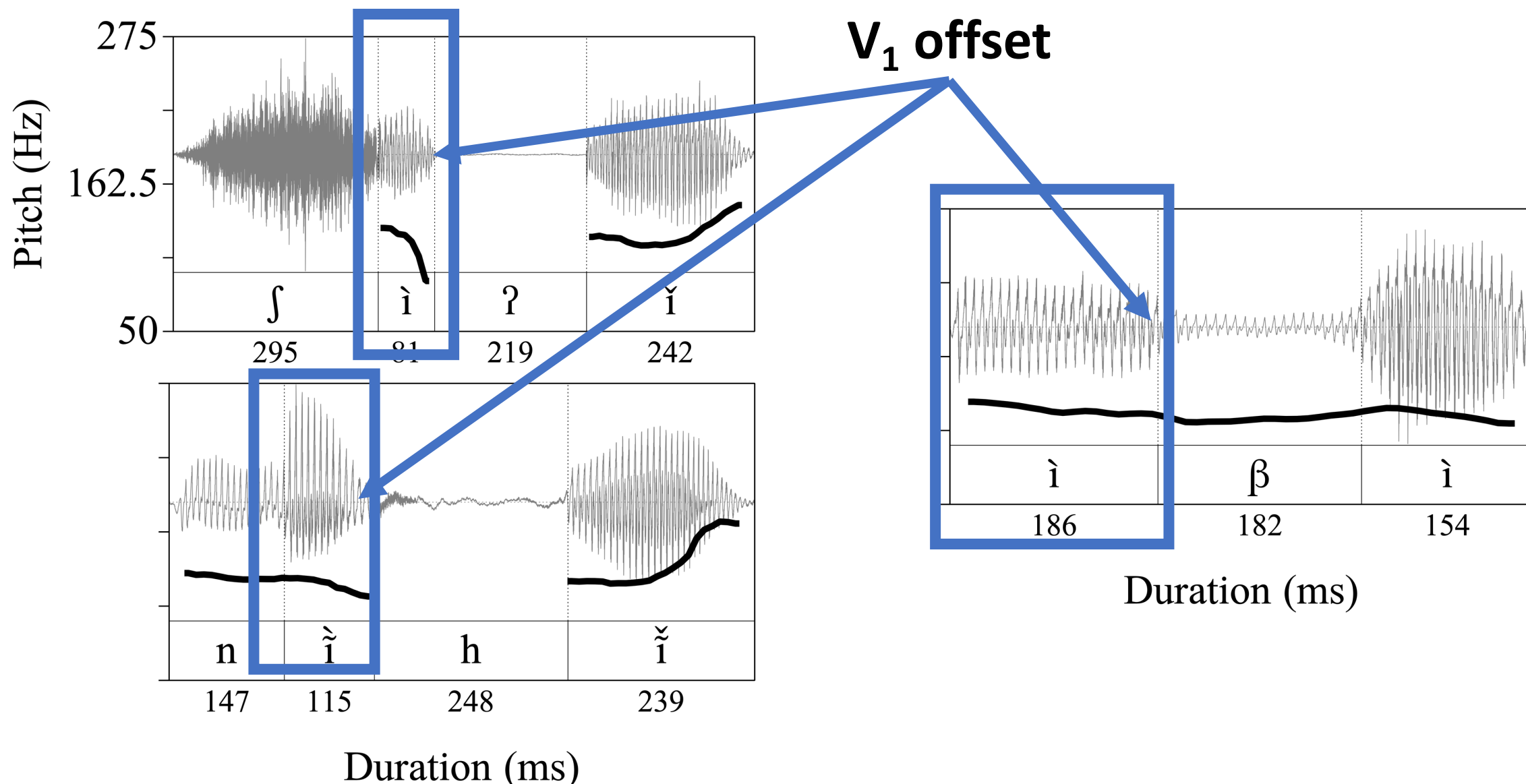
Production study: Results



Production study: Results



Production study: Results



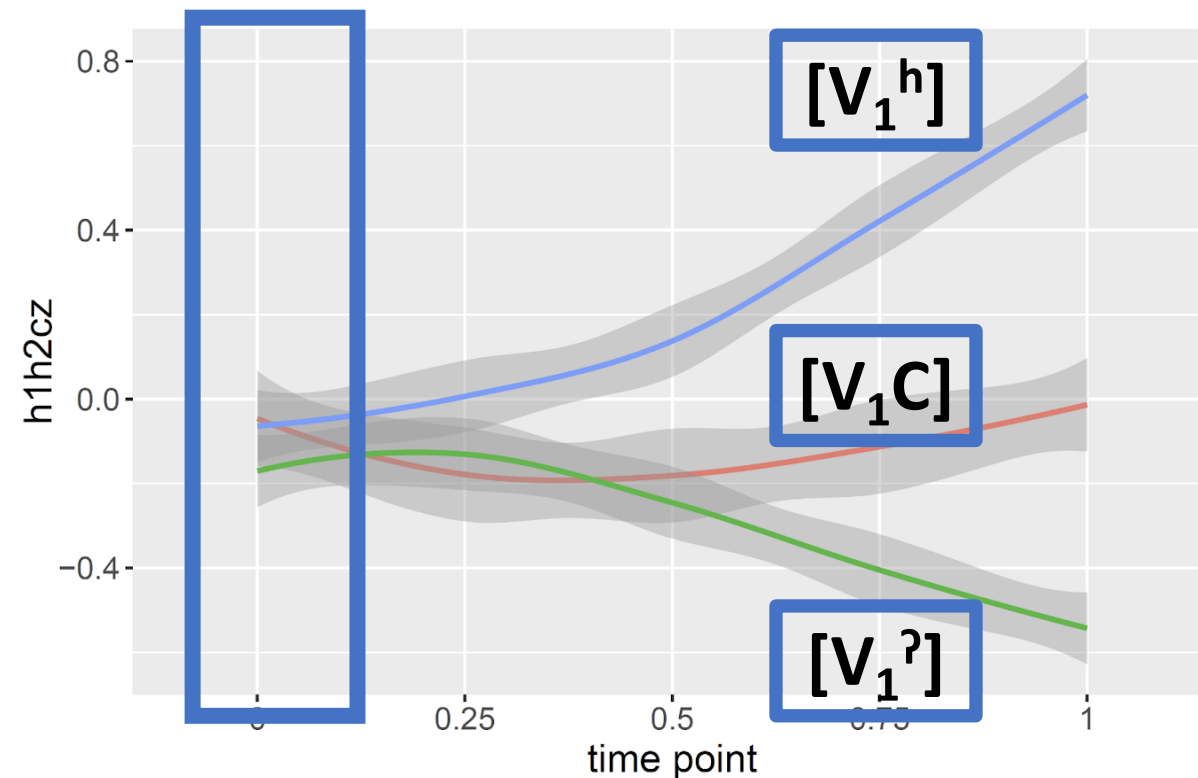
Production study: Results

H1-H2:

At vowel onset, no difference between categories

$[V_1^?]$ effect: = -0.13 (CI: [-0.35, 0.09])

$[V_1^h]$ effect = -0.02 (CI: [-0.24, 0.18])



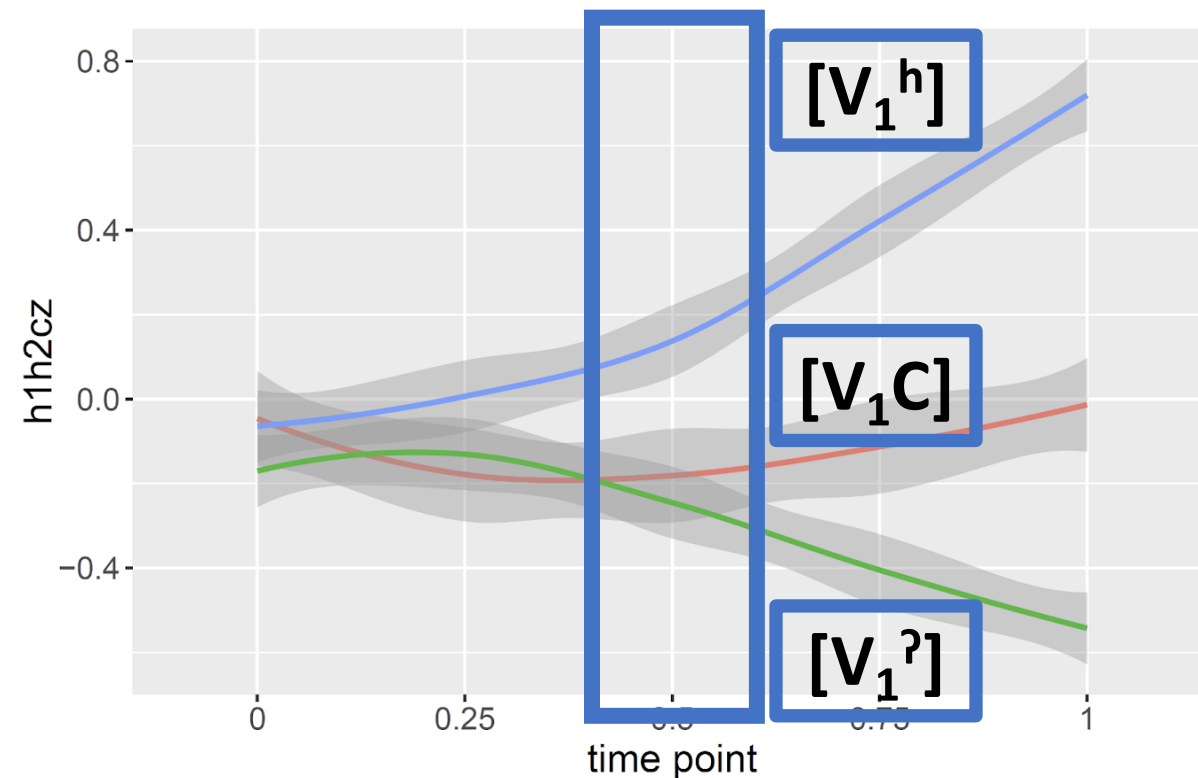
Production study: Results

H1-H2:

At vowel midpoint, $[V_1^h]$ is higher than modal $[V_1C]$, but $[V_1^?]$ is not different from modal $[V_1C]$

$[V_1^?]$ effect = 0.11 (CI: [-0.22, 0.49])

$[V_1^h]$ effect = 0.47 (CI: [0.14, 0.82])



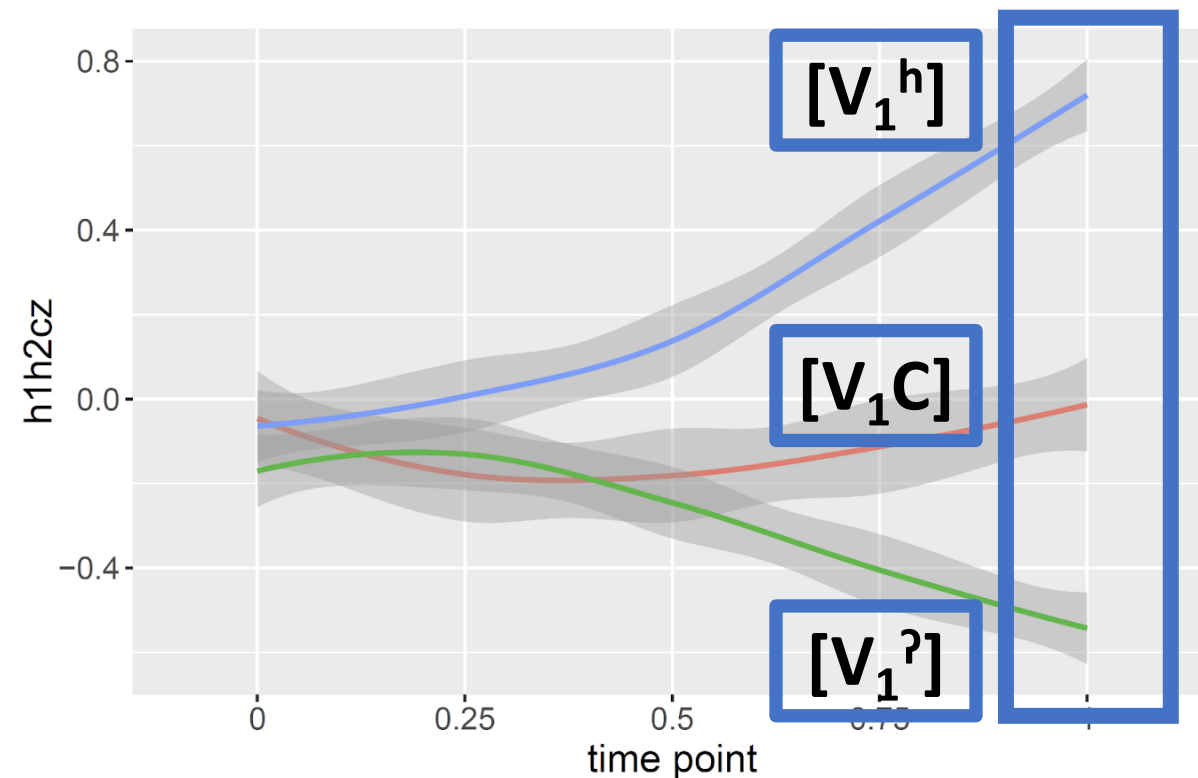
Production study: Results

H1-H2:

At vowel offset, $[V_1^h]$ is higher than modal $[V_1C]$, and $[V_1^?]$ is lower than modal $[V_1C]$

$[V_1^?]$ effect = -0.66 (CI: [-1.12, -0.24])

$[V_1^h]$ effect = 0.79 (CI: [0.38, 1.2])



Production study: Results

Finding:

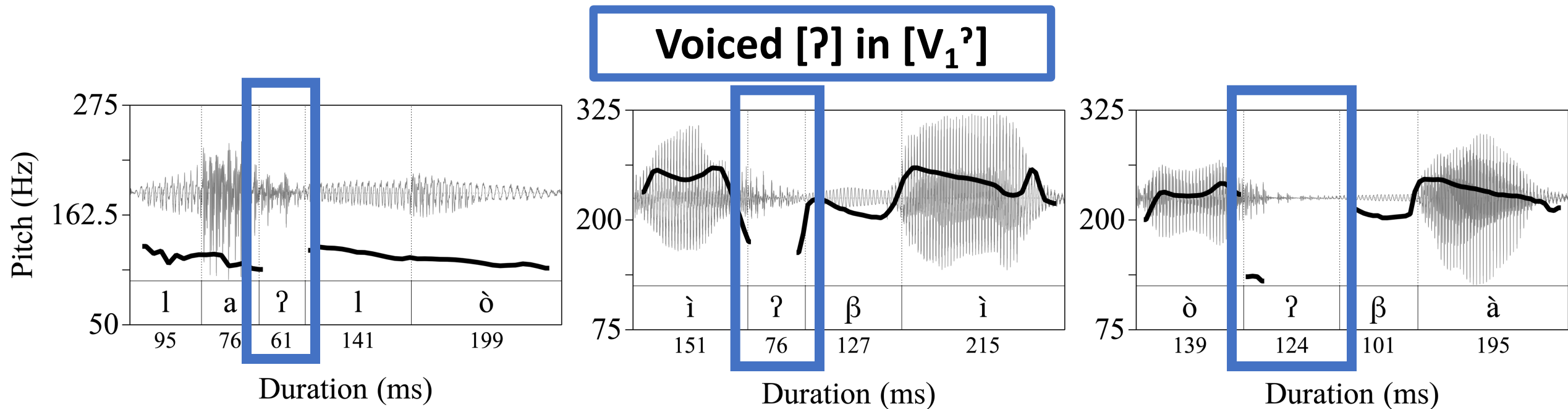
- *H1-H2 falls at end of $[V_1^?]$, rises throughout $[V_1^h]$

$[V_1^?]$	$[V_1^h]$	Modal $[V_1^C]$
Shorter V_1 duration		Longer V_1 duration
Small pitch fall	Steep pitch fall	Moderate pitch fall
Late *H1-H2 fall	Gradual *H1-H2 rise	Steady *H1-H2

Production study: Results

Voicing

Classified tokens as **voiced** and **voiceless**

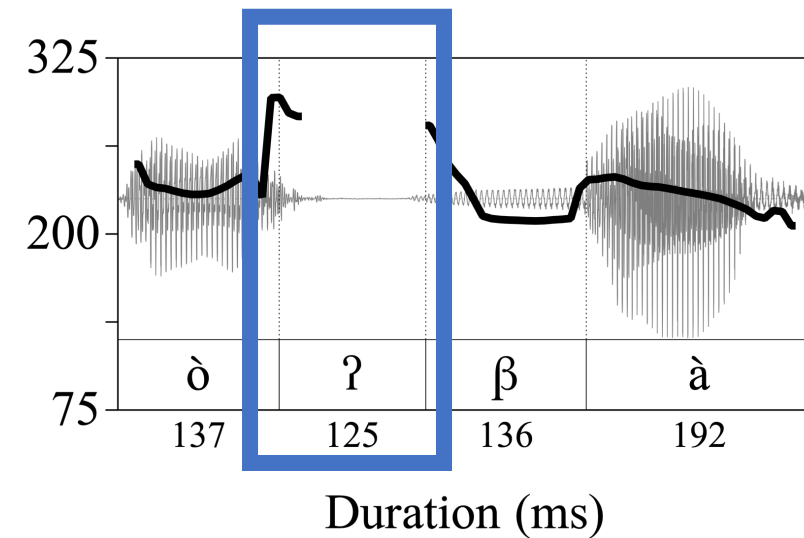
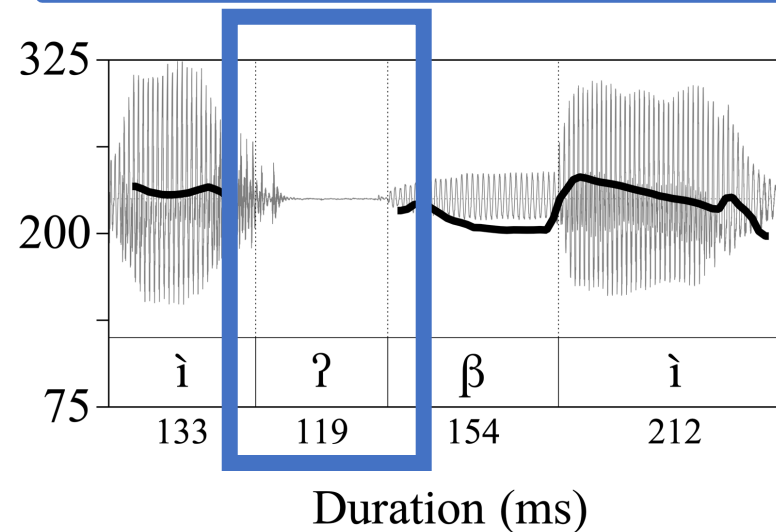
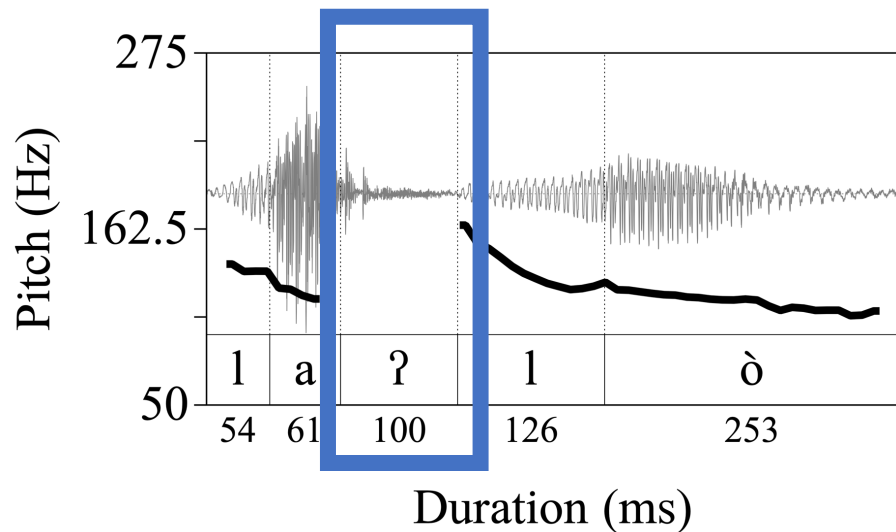


Production study: Results

Voicing

Classified tokens as **voiced** and **voiceless**

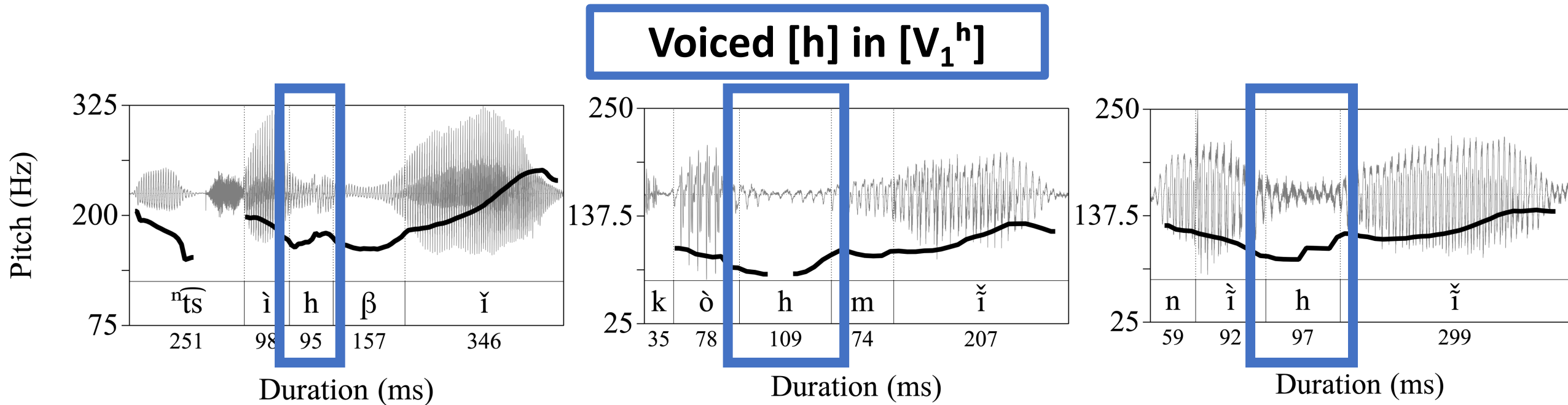
Voiceless [ʔ] in [V₁ʔ]



Production study: Results

Voicing

Classified tokens as **voiced** and **voiceless**

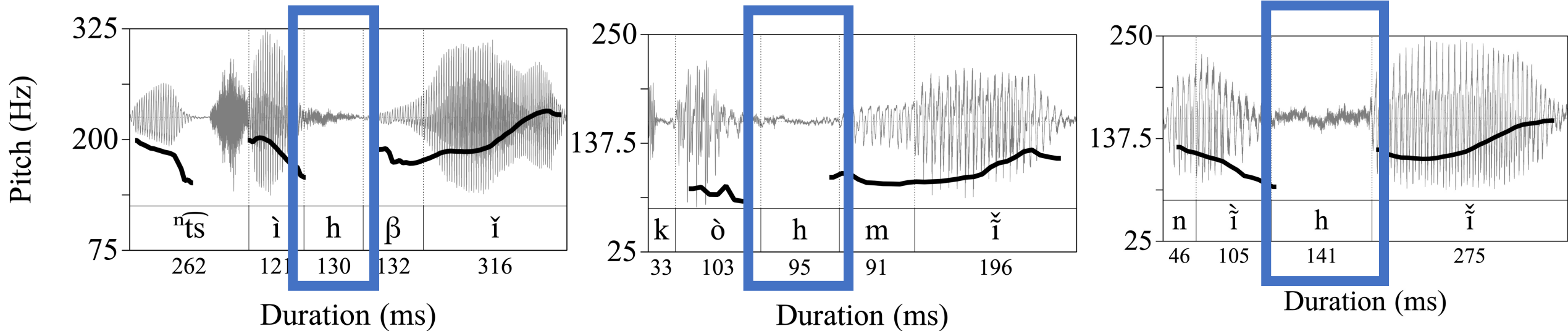


Production study: Results

Voicing

Classified tokens as **voiced** and **voiceless**

Voiceless [h] in [V₁^h]



Production study: Results

Voicing

The [ʔ] in [V₁^ʔ] was almost categorically voiceless

AMR1	AMR2	GEG	JCE	JGO	JMR	LER	RMR
67/70 (96%)	63/63 (100%)	56/69 (81%)	28/54 (52%)	62/62 (100%)	69/72 (96%)	69/71 (97%)	72/72 (100%)

Number of voiceless [ʔ] productions over all [ʔ] productions

Production study: Results

Voicing

The [ʔ] in [V₁^ʔ] was almost categorically voiceless

AMR1	AMR2	GEG	JCE	JGO	JMR	LER	RMR
67/70 (96%)	63/63 (100%)	56/69 (81%)	28/54 (52%)	62/62 (100%)	69/72 (96%)	69/71 (97%)	72/72 (100%)

Number of voiceless [ʔ] productions over all [ʔ] productions

Production study: Results

Voicing

The [h] in $[V_1^h]$ was almost categorically voiceless

AMR1	AMR2	GEG	JCE	JGO	JMR	LER	RMR
72/72 (100%)	71/71 (100%)	61/71 (86%)	58/62 (94%)	62/62 (100%)	72/72 (100%)	70/73 (96%)	71/72 (99%)

Number of voiceless [h] productions over all [h] productions

Production study: Results

Voicing

The [h] in $[V_1^h]$ was almost categorically voiceless

AMR1	AMR2	GEG	JCE	JGO	JMR	LER	RMR
72/72 (100%)	71/71 (100%)	61/71 (86%)	58/62 (94%)	62/62 (100%)	72/72 (100%)	70/73 (96%)	71/72 (99%)

Number of voiceless [h] productions over all [h] productions

Production study: Results

Finding:

- [ʔ] and [h] in [V₁^ʔ] and [V₁^h] almost always voiceless

[V ₁ ^ʔ]	[V ₁ ^h]	Modal [V ₁ C]
Shorter V ₁ duration		Longer V ₁ duration
Small pitch fall	Steep pitch fall	Moderate pitch fall
Late *H1-H2 fall	Gradual *H1-H2 rise	Steady *H1-H2
Almost always voiceless		Always voiced

Discussion

Discussion

Two acoustic dimensions differentiate **modal** $[V_1C]$ from $[V_1^?]/[V_1^h]$

$[V_1^?]$	$[V_1^h]$	Modal $[V_1C]$
Shorter V_1 duration		Longer duration
Period of voicelessness		Always voiced

Discussion

Two other acoustic dimensions differentiate **between** $[V_1^?]$ and $[V_1^h]$

$[V_1^?]$	$[V_1^h]$	$[V_1^C]$
Small pitch fall	Steep pitch fall	Moderate pitch fall
Late *H1-H2 fall	Gradual *H1-H2 rise	Steady *H1-H2

$[V_1^h]$ appears to be more gradually phased, while $[V_1^?]$ is more abrupt (cf. DiCanio 2012)

Discussion

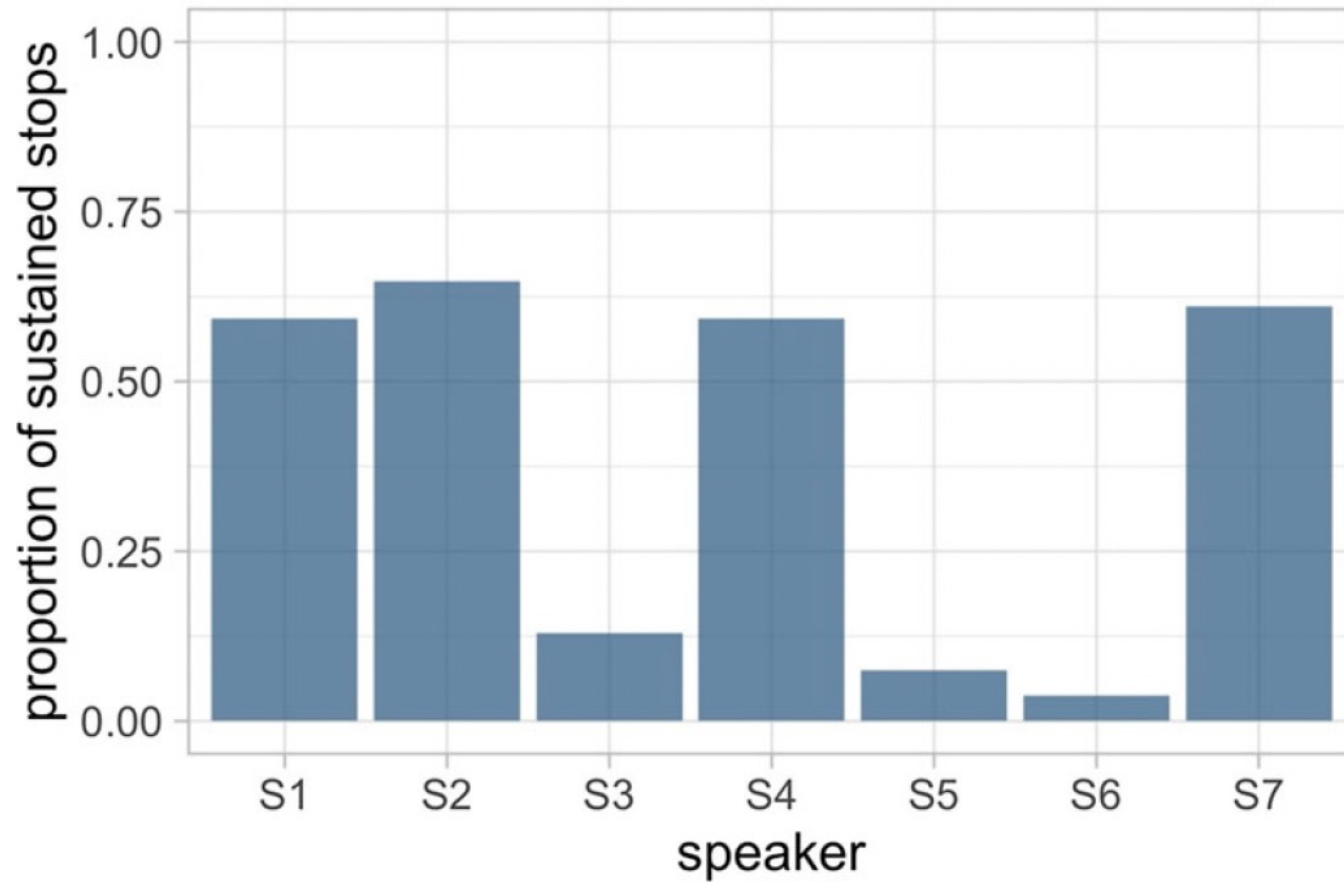
One surprising result is the **prevalence of voicelessness** in $[V_1^?]$ and $[V_1^h]$

Gerfen and Baker (2005:319): “eight tokens [out of 144] were realized with heavy creak, such that Praat’s pitch tracker failed to track an f0 contour throughout the laryngealized portion of the vowel”

- 136/144 (94%) of tokens had periodic voicing throughout [?]

Discussion

One surprising result is the **prevalence of voicelessness** in [ʔ] and [h].



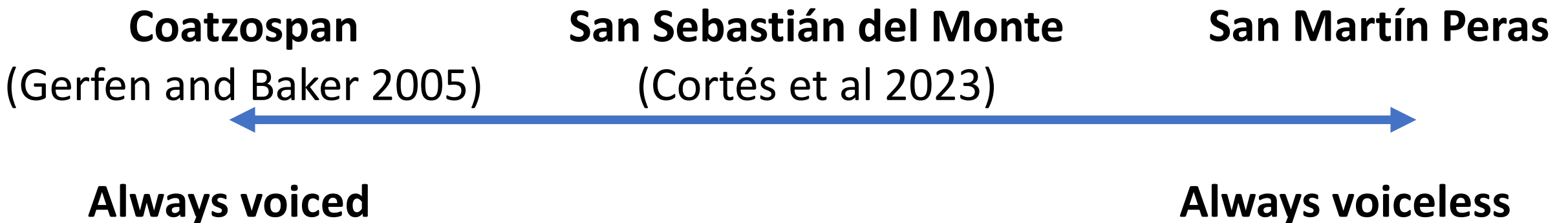
Cortés et al (2023): 4/7 speakers produced [VʔV] sequences with a sustained glottal stop 60-70% of the time.

3 others almost never did

Discussion

Different prevalences of voicelessness across the three languages

- Coatzospan: Almost always voiced, wide variation
- San Sebastián del Monte: Mixed of voiced and voiceless, wide variation
- San Martín Peras: Almost always voiceless, little variation



Discussion

Especially surprising given half of all $[V_1^ʔ]$ and $[V_1^h]$ had intervocalic $[ʔ]$ and $[h]$.

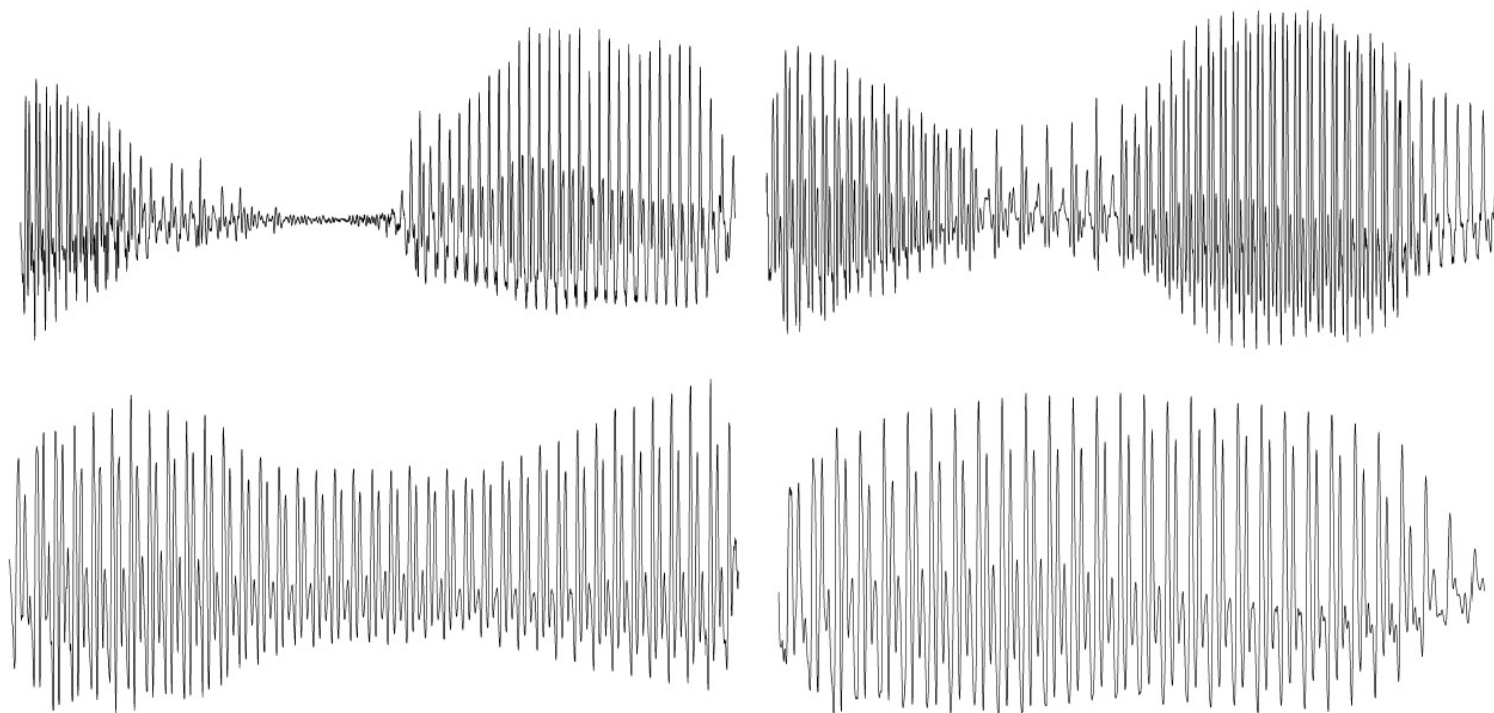
- $[V_1^ʔV_2]$ and $[V_1^hV_2]$

These were just as likely to be voiceless as those followed by a consonant

- $[ʔ]$ and $[h]$ are susceptible to intervocalic voicing (Davidson 2021; Garellek et al 2023)

Discussion

Also surprising because previous work has shown that [V^ʔV] sequences in SMP Mixtec are highly variable, and often periodically voiced in connected speech (Eischens to appear)



Four productions of the vocalic portion of the word [loʔo] ('small') by the same speaker in the same context

Discussion

One participant who produced 100% of [ʔ] and [h] as voiceless regularly produces them as voiced in a narrative

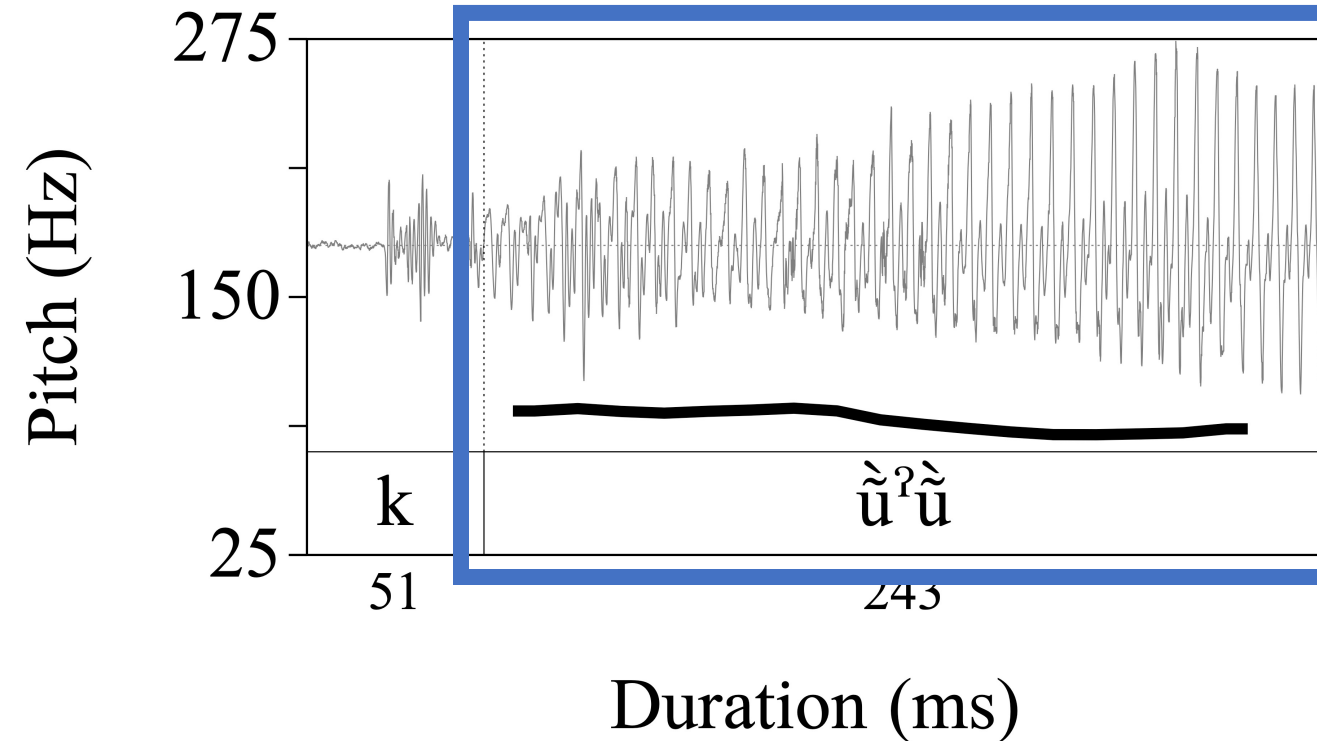
1. Tsia ná kù'ùn ntó ...

[tsʲa ná kũʔũ ˈntó]

then MOD go.POT 2PL

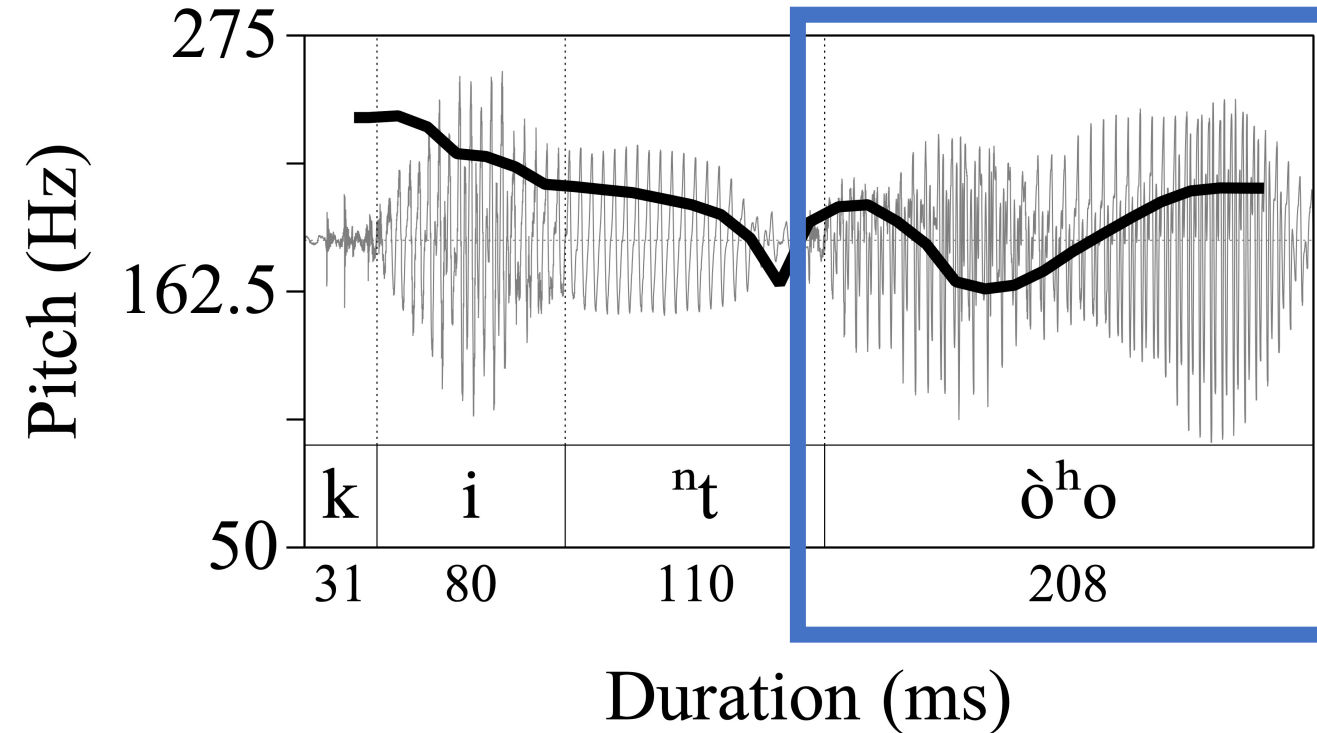
‘Then why don’t you go and...’

‘Y porque ustedes no van y...’



Discussion

One participant who produced 100% of [ʔ] and [h] as voiceless regularly produces them as voiced in a narrative



1. Tá ná kintòjo và'a ún ra ...

[tá ná kiⁿtò^ho
for MOD stay.POT

'So that you end up well'

'Para que quedas bien'

βà^ʔ=ũ
well=2SG

ra]
and



Discussion

One potential reason for high rates of voicelessness is **prosodic prominence**.

- Tokens were produced in isolation, which is often associated with utterance-final prosody
- When a [V^ʔV] sequence is utterance-final in SMP Mixtec, its laryngealization is stronger than elsewhere (Eischens to appear)

Discussion

Prosodic domain edges are often associated with **increased gestural magnitude** (Fougeron and Keating 1997)

- In SMP Mixtec, the entire utterance-final root may be strengthened, leading to devoicing of [ʔ] and [h]
- Creates testable prediction: Increased gestural magnitude should hold of consonants, too

Conclusion

Conclusion

Production study with 8 speakers of SMP Mixtec

Multiple acoustic measures distinguish $[V_1^ʔ]$, $[V_1^h]$, and modal $[V_1C]$

- Duration
- Pitch
- H1-H2
- Voicing

Conclusion

Gerfen and Baker (2005) and Cortés et al (2023) found high inter- and intra-speaker **variation** in the production of [ʔ]

We found remarkable inter- and intra-speaker **consistency** in the production of [ʔ] and [h].

- Most often realized with voicelessness during [ʔ] and [h]

Prevalence of voicelessness may be explained by prosodic prominence

Thank you!

Appendix A: Duration

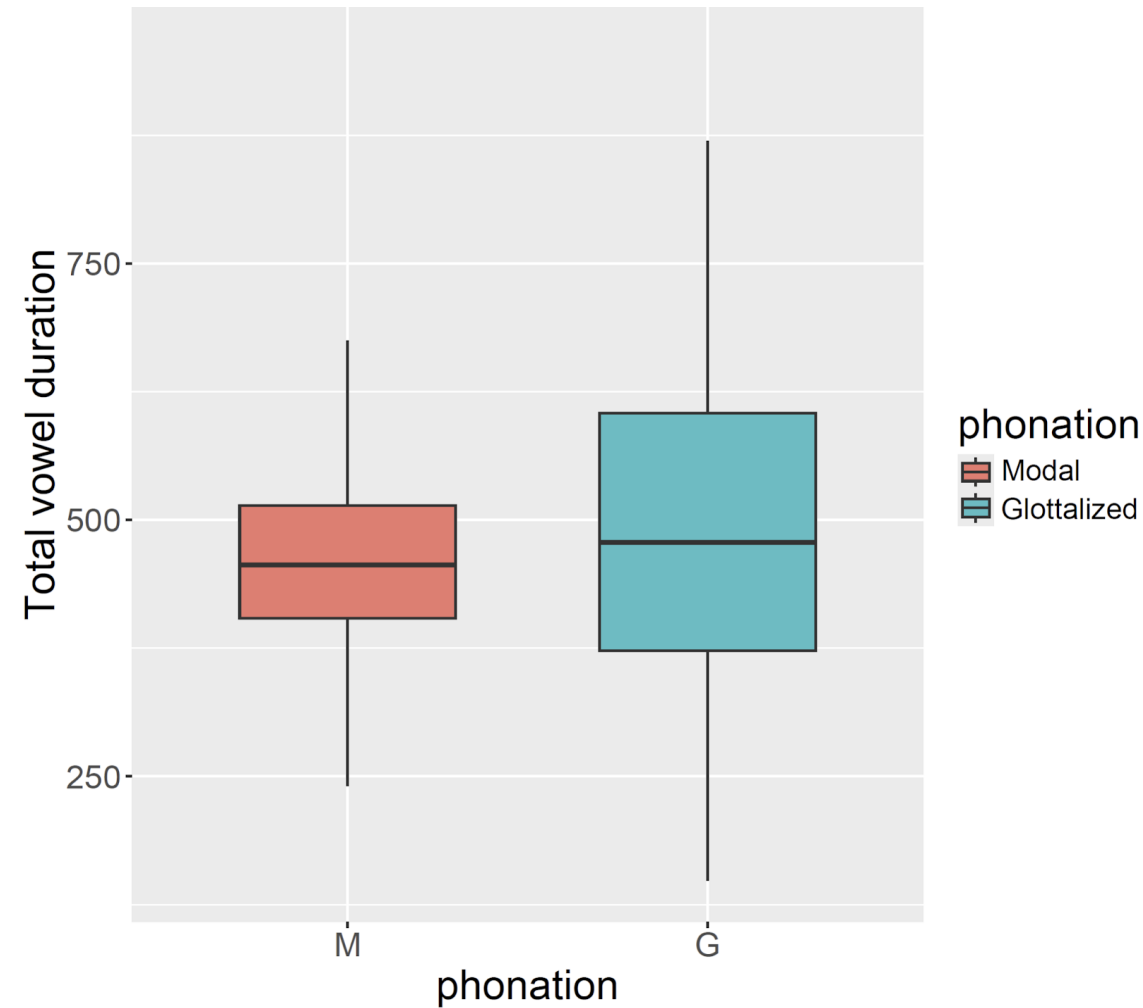
Duration

Compared duration of [VʔV] to [VV]

Compared to modal baseline...

[VʔV] effect = +32.36

- **Not credible:** CI = [-25.72, 92.52]



Appendix B: Task differences

Why do we see this large difference?

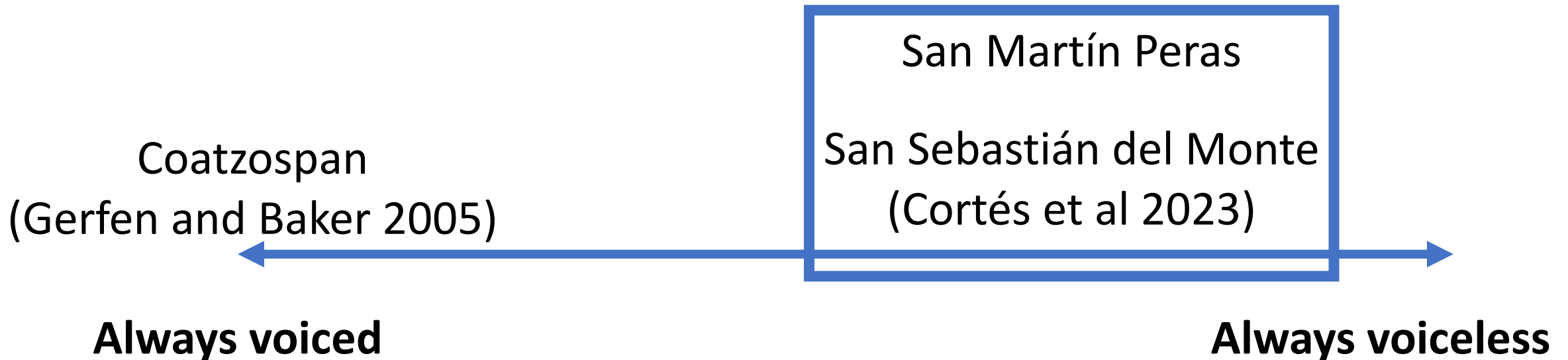
- Differences in the tasks could have created an artificial difference between San Sebastián del Monte Mixtec and SMP Mixtec

San Sebastián del Monte Cortés et al (2023)	San Martín Peras
<p>Words in a carrier phrase</p> <p>(25) Tiáái ___ viti. /tʃáái ___ viti/ Tiáá=i ___ viti. write.POT = 1SG ___ now ‘I am going to write ___ now.’</p>	<p>Words in isolation</p> <p>Words in isolation are likely to have strong glottalization (Eischens to appear)</p>

Appendix B: Task differences

Why do we see this large difference?

- Perhaps the continuum looks more like this



Appendix B: Task differences

To test this, we carried out a follow-up study

- The same as study 1, but words were embedded into a carrier phrase

1. Só'o káchi é ___ tũ'un ntá'vi

[só'o ká^htʃ=ě ___ tũ'ũ n^htá'βi]

like.this say.CONT=1PL.INCL ___ word poor/Indigenous

'This is how you say ___ in Mixtec'

'Así se dice ___ en mixteco'

Appendix B: Task differences

Four participants, three of whom participated in Study 1 one year prior

Data not fully processed yet, but voicing results are the same

AMR1	LER	RMR	VER
64/71 (90%)	71/71 (100%)	68/70 (97%)	61/64 (95%)

Number of voiceless [ʔ] productions over all [ʔ] productions

Appendix B: Task differences

Four participants, three of whom participated in Study 1 one year prior

Data not fully processed yet, but voicing results are the same

AMR1	LER	RMR	VER
72/72 (100%)	74/74 (100%)	72/72 (100%)	68/68 (100%)

Number of voiceless [h] productions over all [h] productions

Appendix B: Task differences

Difference in prevalence of voiceless [ʔ] and [h] between San Sebastián del Monte Mixtec and SMP Mixtec not due to presence/lack of carrier phrase.

Appendix C: Intervocalic vs non-intervocalic

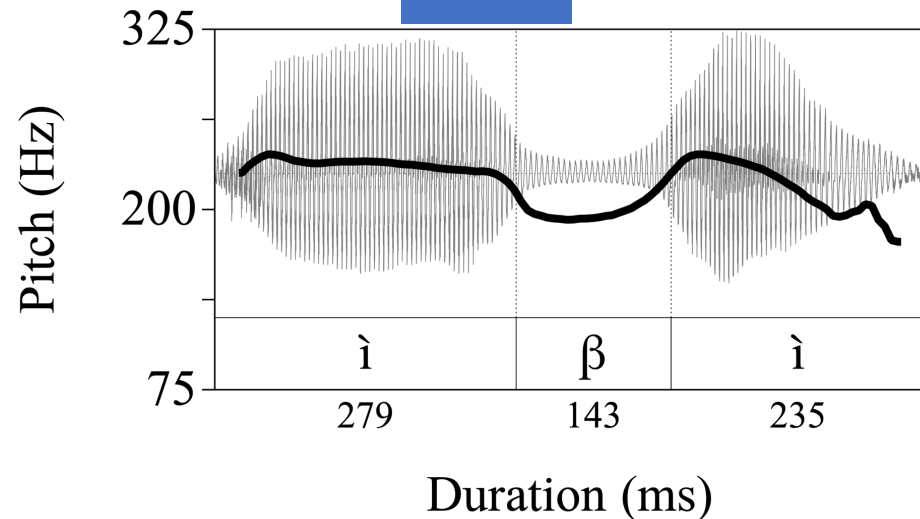
Some items had medial C (about half of $[V_1^?]$ and $[V_1^h]$)

1. ìvì

$[i\beta i]$

'Two'

'Dos'

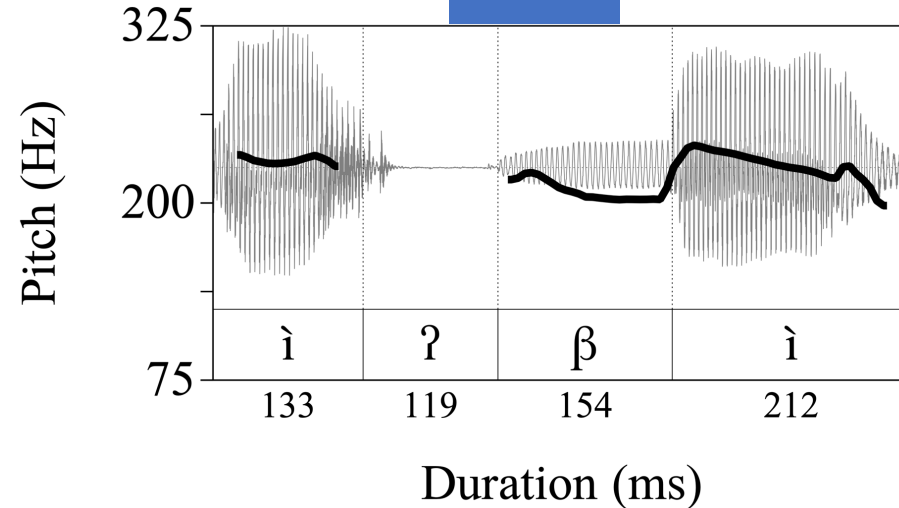
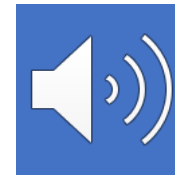


2. ì'vì

$[i^? \beta i]$

'Pain'

'Dolor'



Appendix C: Intervocalic vs non-intervocalic

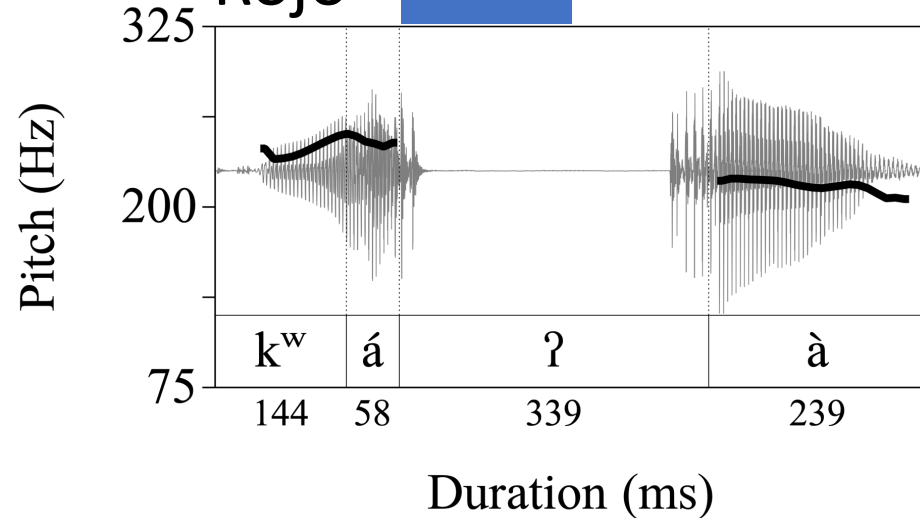
Others had no medial C (about half of $[V_1^ʔ]$ and $[V_1^h]$)

1. Kuá'à

$[k^wáʔà]$

'Red'

'Rojo'

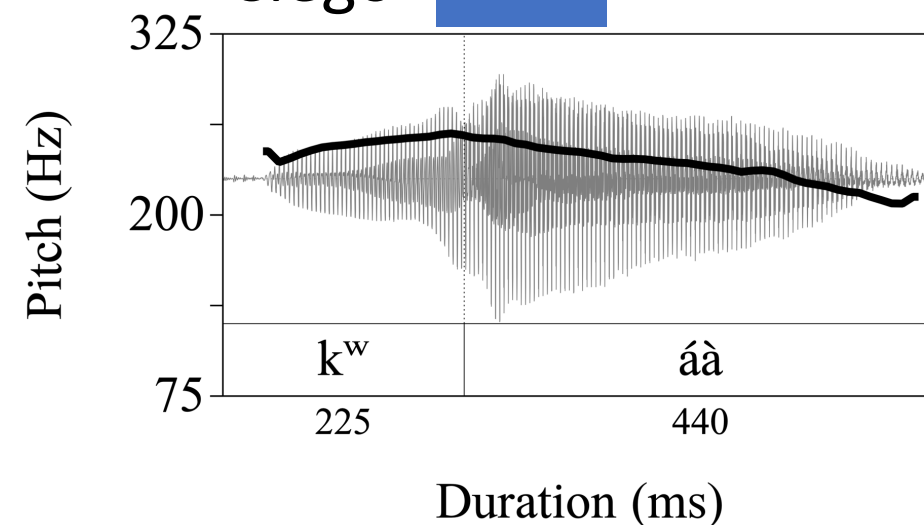
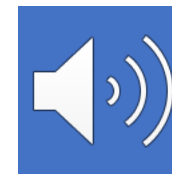


2. Kuáà

$[k^wáà]$

'Blind'

'Ciego'



Discussion

[ʔ] was just as likely to be voiceless when intervocalic then when pre-consonantal

AMR1	AMR2	GEG	JCE	JGO	JMR	LER	RMR
40/41	38/38	40/41	19/35	38/38	42/42	41/42	42/42
(98%)	(100%)	(98%)	(54%)	(100%)	(100%)	(98%)	(100%)

Number of voiceless [VʔV] productions over all [VʔV] productions

[ʔ] is very often (at least partially) voiced intervocalically (Davidson 2021; Garellek et al 2023)

Discussion

[h] was also just as likely to be voiceless when intervocalic as when pre-consonantal

AMR1	AMR2	GEG	JCE	JGO	JMR	LER	RMR
37/37	36/36	36/36	29/32	28/28	36/36	36/36	36/36
(100%)	(100%)	(100%)	(91%)	(100%)	(100%)	(100%)	(100%)

Number of voiceless [V^hV] productions over all [V^hV] productions

[h] is also very often voiced intervocalically (Garellek et al 2023)

Appendix D: Preaspiration

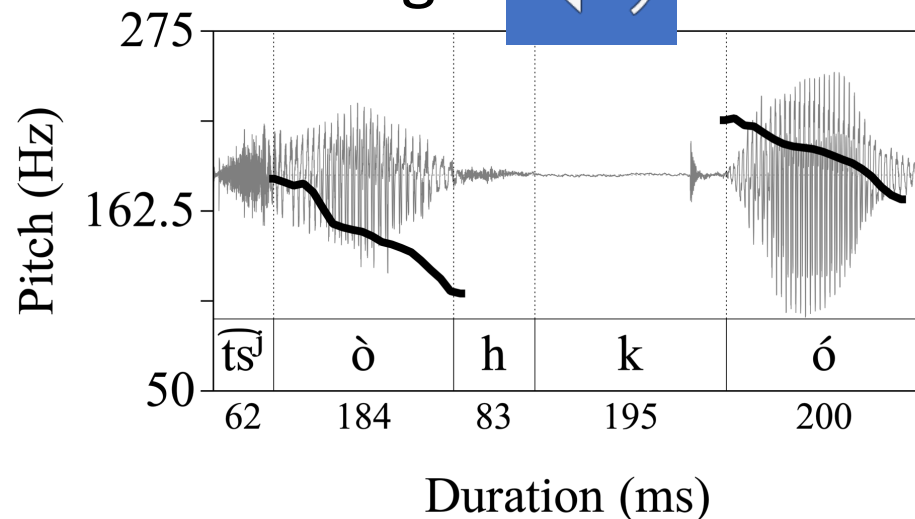
Root-medial voiceless consonants are predictably pre-aspirated
(Eischens and Hedding to appear)

1. Tsiòkó

[ts^jò^hkó]

'Ant'

'Hormiga'



2. Usù

[u^hsù]

'Deer'

'Venado'

