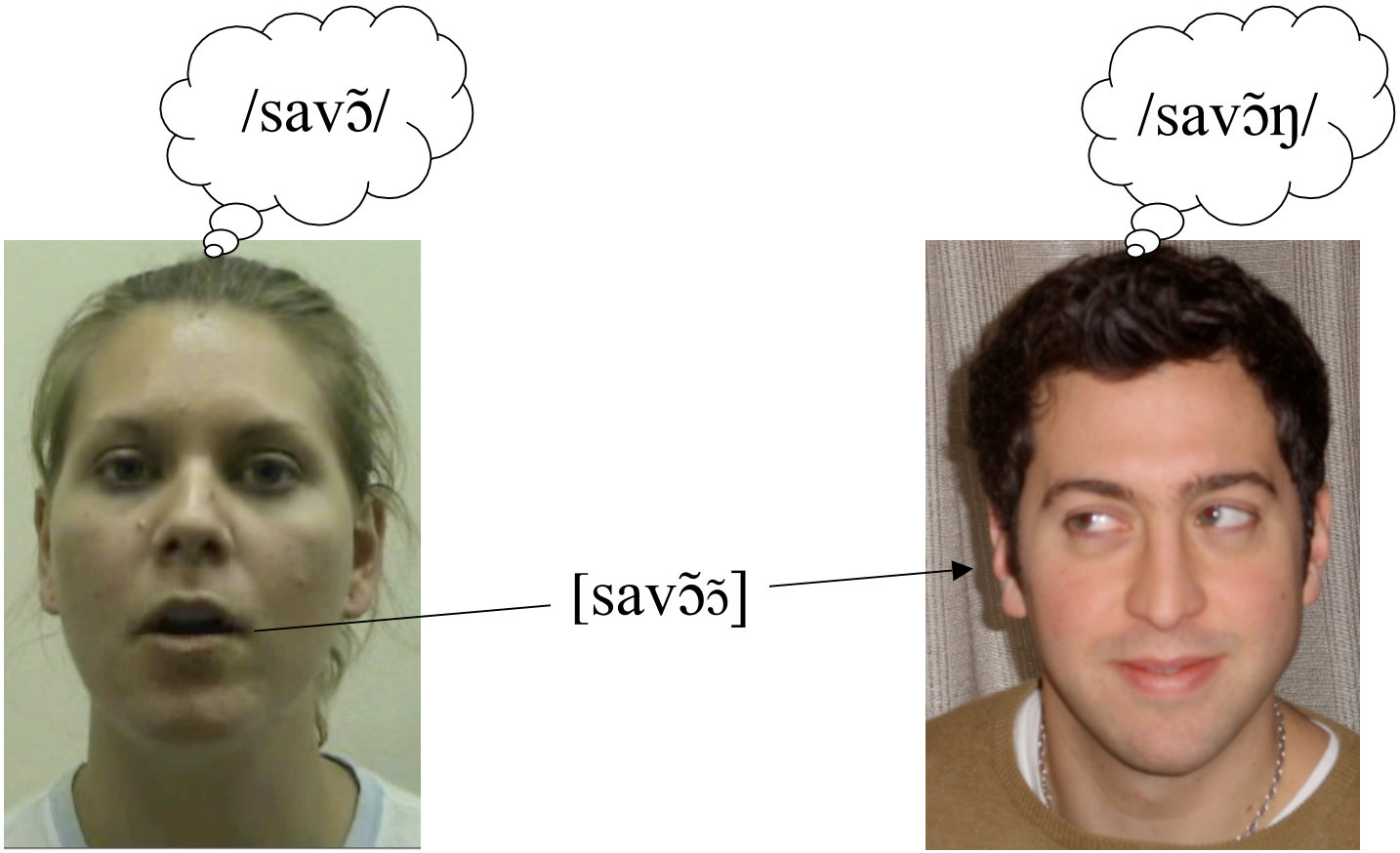
The Acoustic and Visual Phonetic Basis of Place of Articulation in Excrescent Nasals

Keith Johnson & Christian DiCanio
UC Berkeley

Laurel MacKenzie
University of Pennsylvania

- Ohala (1981) proposed that the listener is the source of sound change, because listeners parse incomplete or ambiguous acoustic signals to determine the linguistic intentions of the speaker, and this parsing process is sometimes incorrect.



Problem:
Why are unmarked coda nasals velar?

Excrescent Nasal Velarity in Midi French

		Midi	*Standard*
savon	"soap"	[savɔ̃ŋ]	[savɔ̃]
cf. *savone*	"to soap up"	[savɔne]	[savɔne]
pain	"bread"	[pɛɛ̃ŋ]	[pɛ̃]

Coda Velars

- Morais-Barbosa (1962, p.692)

  Standard Portuguese [lã] “wool” ~ dialectal variant [lãŋ]

- Wiese (1996) French borrowed into German

  [ʀɛstoʀã] “restaurant” > [ʀɛstoʀaŋ]

- Paradis & Prunet (2000) French borrowings in Fula.

	French	Fula	
(a)	[seʀʒã]	[sarsaŋ]	“sergeant”
(b)	[ljøtnã]	[lijetinaŋ]	“lieutenant”
(c)	[kɔ̃feʀãs]	[kɔŋferas]	“conference”

Coda Velars, cont.

- Trigo (1988)

  English borrowings in Puerto Rican Spanish

  “train” (v.) > [tɾeŋ]

- Dialectal Variation in Spanish

	Standard Spanish	Andalusian Spanish
“bread”	[pan]	[paŋ]
“glutton”	[gloton]	[glotoŋ]

Two Related Phenomena

Excrescent nasal velarity: [ã] > [ãŋ]

Final velarity: [an] > [ãn] > [ã] > [ãŋ]

[am] > [ãm] > [ã] > [ãŋ]

Previous Approaches

- Paradis & Prunet (2000) *Nasal vowel unpacking*

  Nasal vowels are underlyingly biphonemic, which accounts for the surfacing of excrescent nasals.

- Rice (1996) *Default Variability Hypothesis*

  Failure to fill in an unmarked phonological feature results in a default phonetic interpretation of velar.

- Howe (2004) *Revised Articulator Theory*

  All vowels are "dorsum articulated." The spread of [dorsal] from the vowel to the coda causes coda velarity.

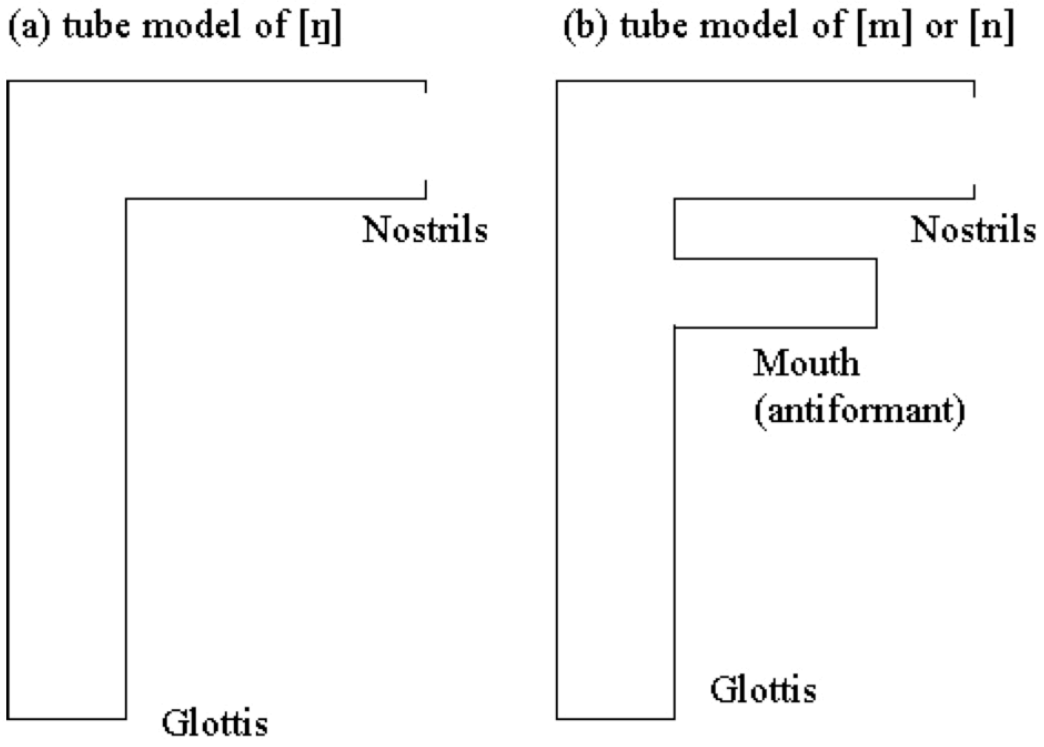
Previous Approaches (cont.)

- Velar nasals tend to alternate with nasalized vowels because they are similar to each other.
- “Sharing a feature” (Howe, 2004) or “default interpretation” (Rice, 1996) are ways to *express* similarity, but do not explain why it exists.
- Ohala's *Listener-based model of sound change*

  A theory describing the historical and cognitive mechanisms by which similarity leads to a language sound pattern.

Hypothesis 1: Acoustic Similarity (Ohala, 1975)

- Nasalized vowels are more acoustically similar to velar nasals than they are to labial or coronal nasals because the velar nasal [ŋ] (like nasalized vowels) has no oral acoustic antiformants while [m] and [n] do.



Hypothesis 2: Visual Similarity

- Nasalized vowels are more *visually* similar to velar nasals than they are to labial or coronal nasals because the velar nasal [ŋ] (like nasalized vowels) has no visible oral closure while [m] and [n] do.

[ŋ] [ɔ̃] [n] [m]

Experiment 1: Nasal Place Identification in audio-visual stimuli

- **Methods**

  Subjects: Fifteen (8 women, 7 men) UCB undergraduates.

  Materials: Audio-only and Audio-visual recordings of 36 CVN words.

  V = [**i, ə, ɔ, eɪ**], N = [**m n ŋ**]

  Video editing: Each movie clip had a still-frame for 0.5 seconds, then the closure and opening movements for C, V, and N, and then a still-frame for 0.5 seconds.

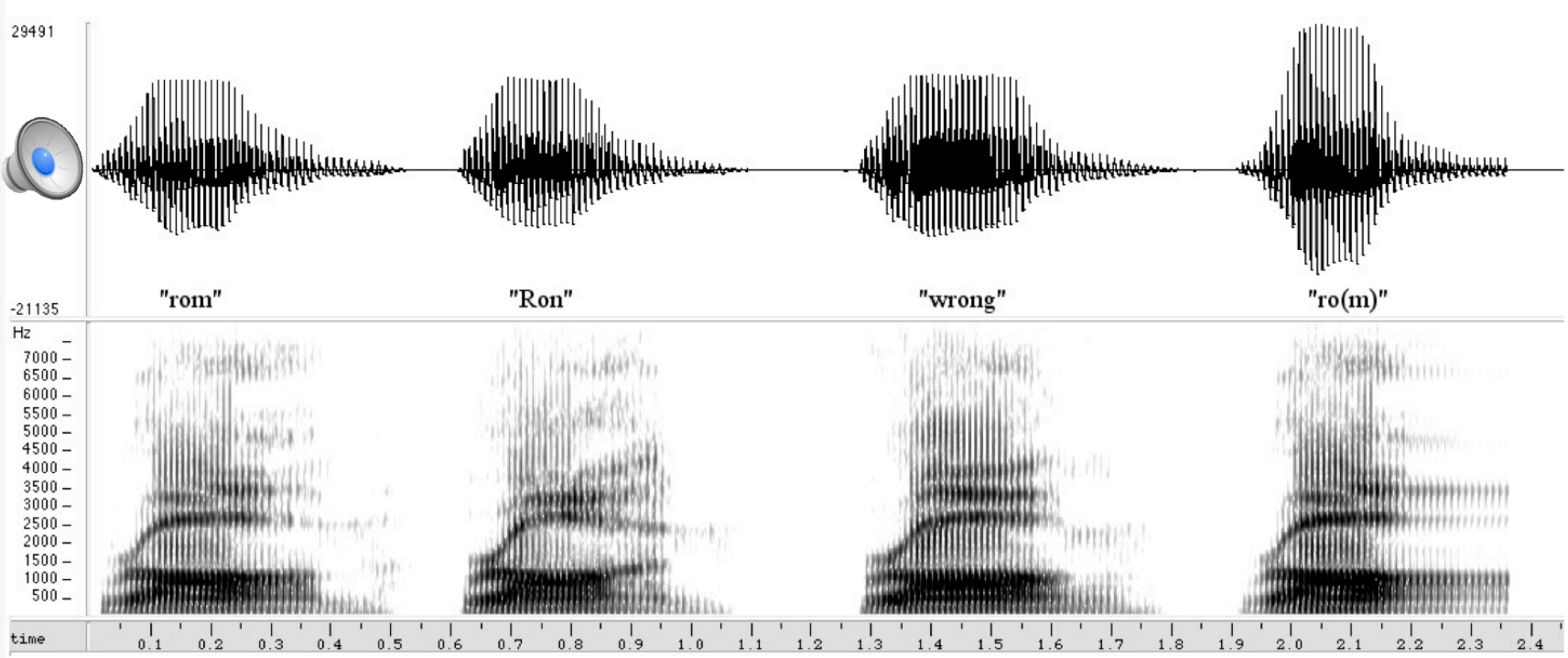
  Audio editing: "Placeless" stimuli constructed from tokens ending in [m] by deleting the last part of V and all of N. The last glottal pulse of the truncated token was repeated so that the token's duration matched the original to within 5 ms and an amplitude envelope which also matched the original [m] token was applied.

Stimuli

Table 1. The thirty-six words used as stimuli in experiments 1 and 2.

[**i**]	beam, bean, bing	[**eɪ**]	dame, dane, dang
	seem, seen, sing		fame, feign, fang
	ream, reen, ring		same, sane, sang
[**ɔ**]	calm, con, kong	[**ə**]	dumb, done, dung
	pom, pawn, pong		rum, run, rung
	rom, ron, wrong		sum, sun, sung

Audio-only stimuli



All these stimuli were embedded in white noise (SNR = 0 dB.)

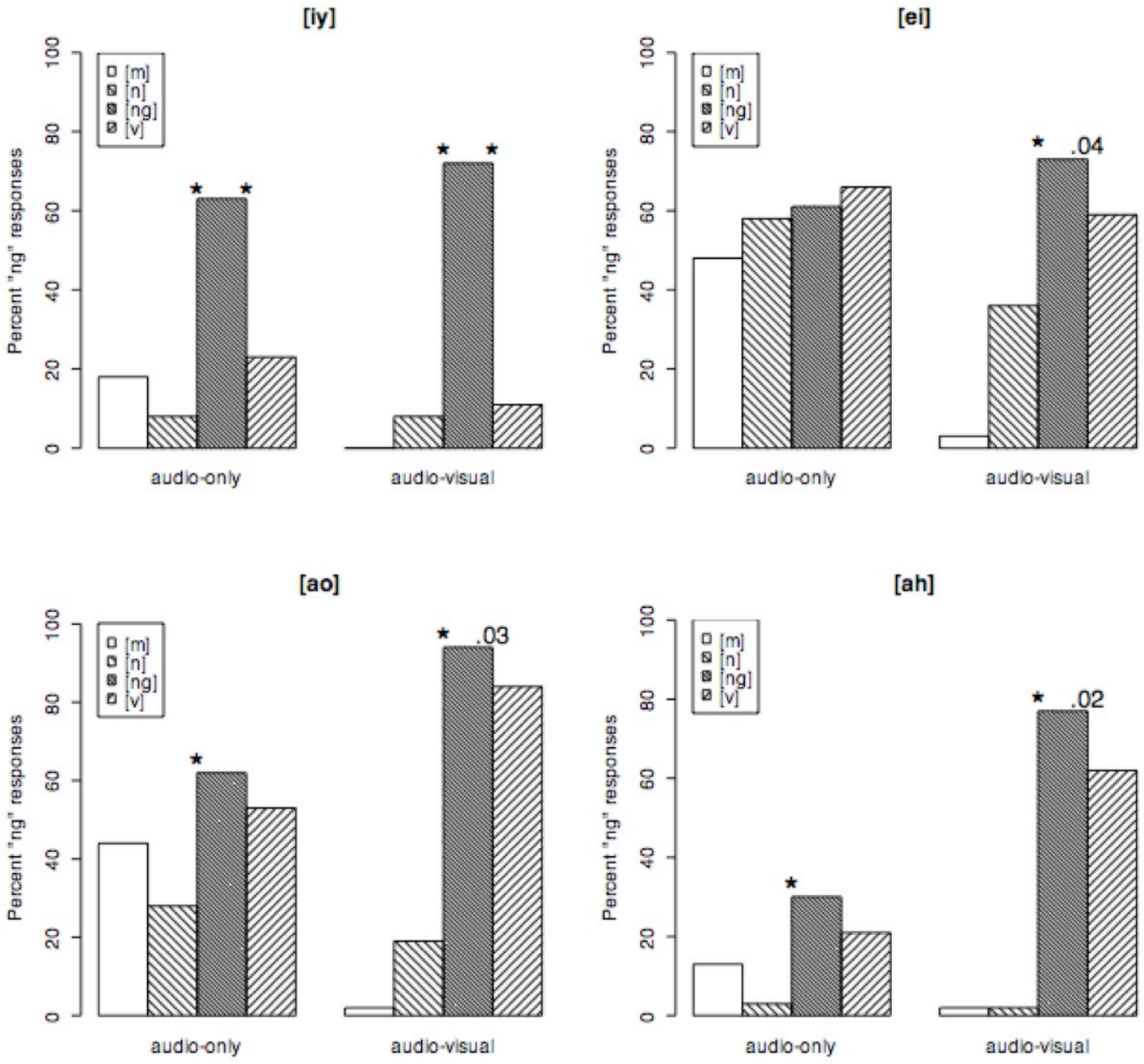
Procedure: 96 audio-only trials

Task: Identify the final consonant as "m", "n", or "ng".

Audio-visual stimuli

Procedure: 144 audio-visual trials

Task: Identify the final consonant as "m", "n", or "ng".


[iy]
[ei]
[ao]
[ah]
[m]
[n]
[ng]
[v]
Percent "ng" responses
0
20
40
60
80
100
audio-only
audio-visual
.04
.03
.02

Results from Experiment 1

(responses to [i] tokens removed)

Percent of "m", "n", or "ng" responses (columns) to audio-only and audio-visual stimuli (rows)

		"m"	"n"	"ng"
audio-only	[m]	33	32	35
	[n]	11	58	30
	[ŋ]	17	31	51
	[x̃]	21	32	46
audio-video	[m]	96	2	3
	[n]	3	78	19
	[ŋ]	2	16	82
	[x̃]	3	28	69

Results from Experiment 1 (cont.)

(responses to [i] tokens removed)

Percent "m", "n", and "ng" responses (columns) to "placeless nasal" audio [x̃] tokens as a function of the video display of the token (rows)

[x̃]

		"m"	"n"	"ng"
video	/m/	92	4	4
	/n/	3	43	53
	/ŋ/	3	28	69

Experiment 2: [ŋ] detection in audio-visual stimuli

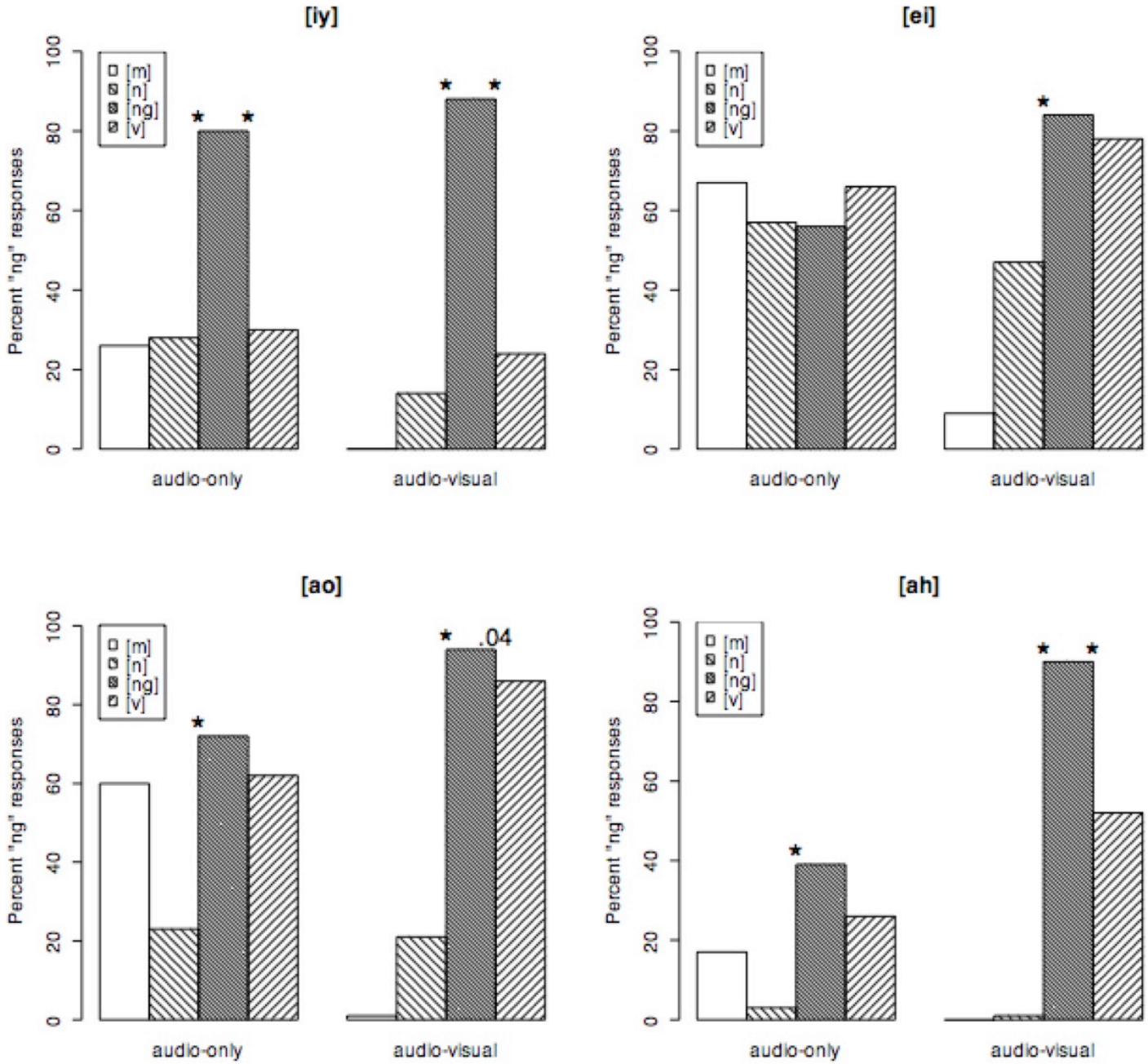
What if listeners did not really think that nasalized vowels sounded like [ŋ]?

Task: [ŋ] detection - "Did the token end with "ng" or not?"

Subjects: Fifteen (8 women, 7 men) UCB undergrads.

Materials: The same as in experiment 1.

Procedure: A block of 96 audio-only trials then a block of 144 AV trials.


[iy]
[ei]
[ao]
[ah]
[m]
[n]
[ng]
[v]
Percent "ng" responses
0
20
40
60
80
100
audio-only
audio-visual
*
.04

Results from Experiments 1 and 2

(responses to [i] tokens removed)

Percent of "m", "n", or "ng" responses (columns) to audio-only and audio-visual stimuli (rows) and percent "ng" responses in experiment 2.

		Exp 1			Exp 2
		"m"	"n"	"ng"	"ng"
audio-only	[m]	33	32	35	48
	[n]	11	58	30	27
	[ŋ]	17	31	51	56
	[x̃]	21	32	46	51
audio-video	[m]	96	2	3	3
	[n]	3	78	19	23
	[ŋ]	2	16	82	89
	[x̃]	3	28	69	72

Results from Experiments 1 and 2 (cont.)

(responses to [i] tokens removed)

Percent "m", "n", and "ng" responses in experiment 1 and percent "ng" detections in experiment 2 (columns) to "placeless nasal" audio [ɴ̃] tokens as a function of the video display of the token (rows)

		Exp 1			Exp 2
		"m"	"n"	"ng"	"ng"
video	/m/	92	4	4	1
	/n/	3	43	53	61
	/ŋ/	3	28	69	72

Experiment 3: Nasal identification in audio-visual stimuli

Fundamentally the same as experiment 1 with two changes:

Naturally produced nasalized vowels

- instead of the "placeless nasals" of experiments 1 and 2
- Tests the visual hypothesis with movies of nasalized vowels

Noise replacement of final nasal

- instead of random noise through-out the token
- Forces greater reliance on visual cues (of nasalized vowels)

Experiment 3

- **Methods**

  Subjects: Nineteen (12 women, 7 men) UCB undergraduates.

  Materials: For each of three vowel environments - [ə, ɔ, eɪ] we selected three sets of words ending in the final nasals [m], [n], and [ŋ]. The speaker, who is a phonetically trained, native speaker of English, and L2 speaker of French, also produced a "word" which consisted of the initial consonant and a nasalized version of the vowel.

  Video editing. Movie clips of each word were produced using the same method as in experiment 1.

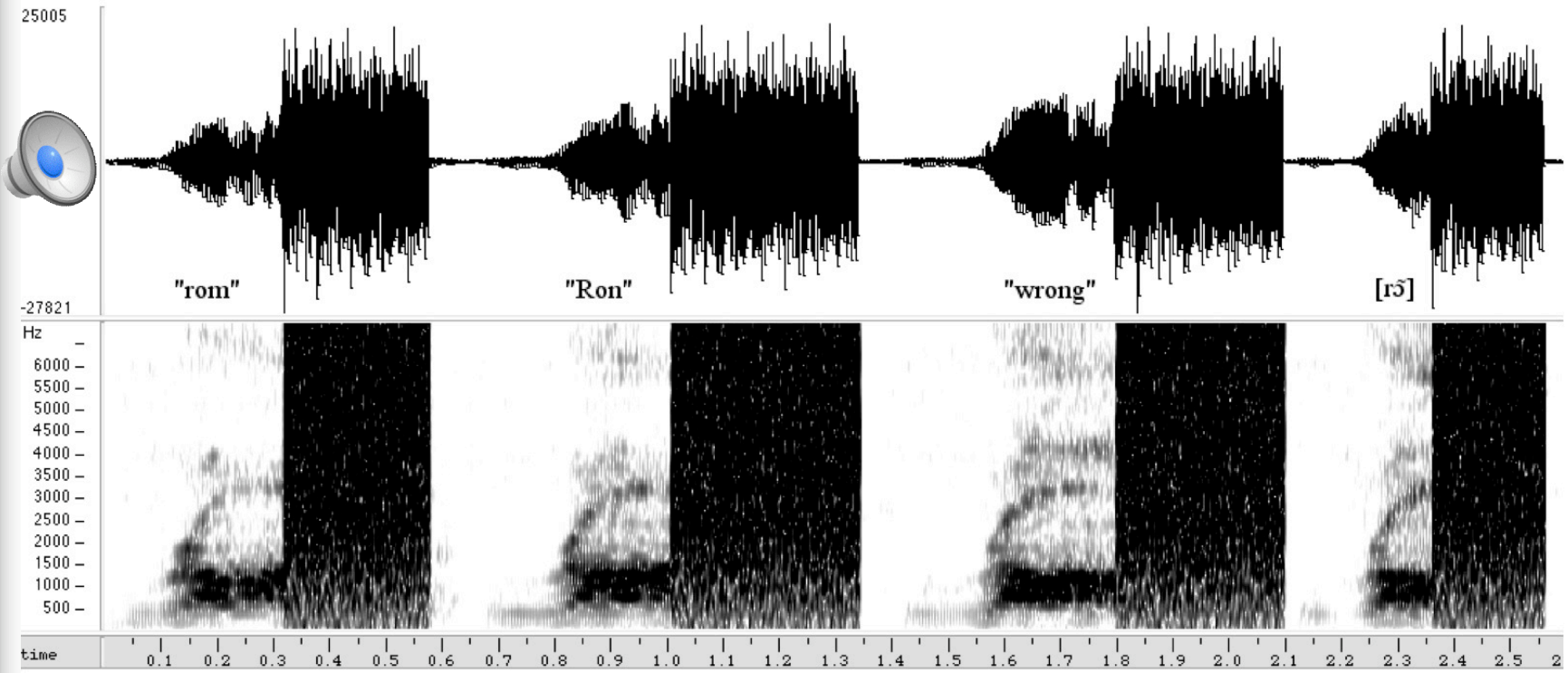
  Audio editing. The audio sound track was edited so that the last half of the vowel and all of the nasal segment were deleted and replaced by a burst of Gaussian white noise of the same duration as the deleted portion.

Stimuli

Table 1. The thirty-six "words" used as stimuli in experiments 3 and 4.

[ə]	dumb, done, dung, [**də̃**]	[eɪ]	dame, dane, dang, [**dẽɪ**]
	rum, run, rung, [**ɹə̃**]		fame, feign, fang, [**fẽɪ**]
	sum, sun, sung, [**sə̃**]		same, sane, sang, [**sẽɪ**]
[ɔ]	calm, con, kong, [**kɔ̃**]		
	pom, pawn, pong, [**pɔ̃**]		
	rom, ron, wrong, [**ɹɔ̃**]		

Audio-only stimuli



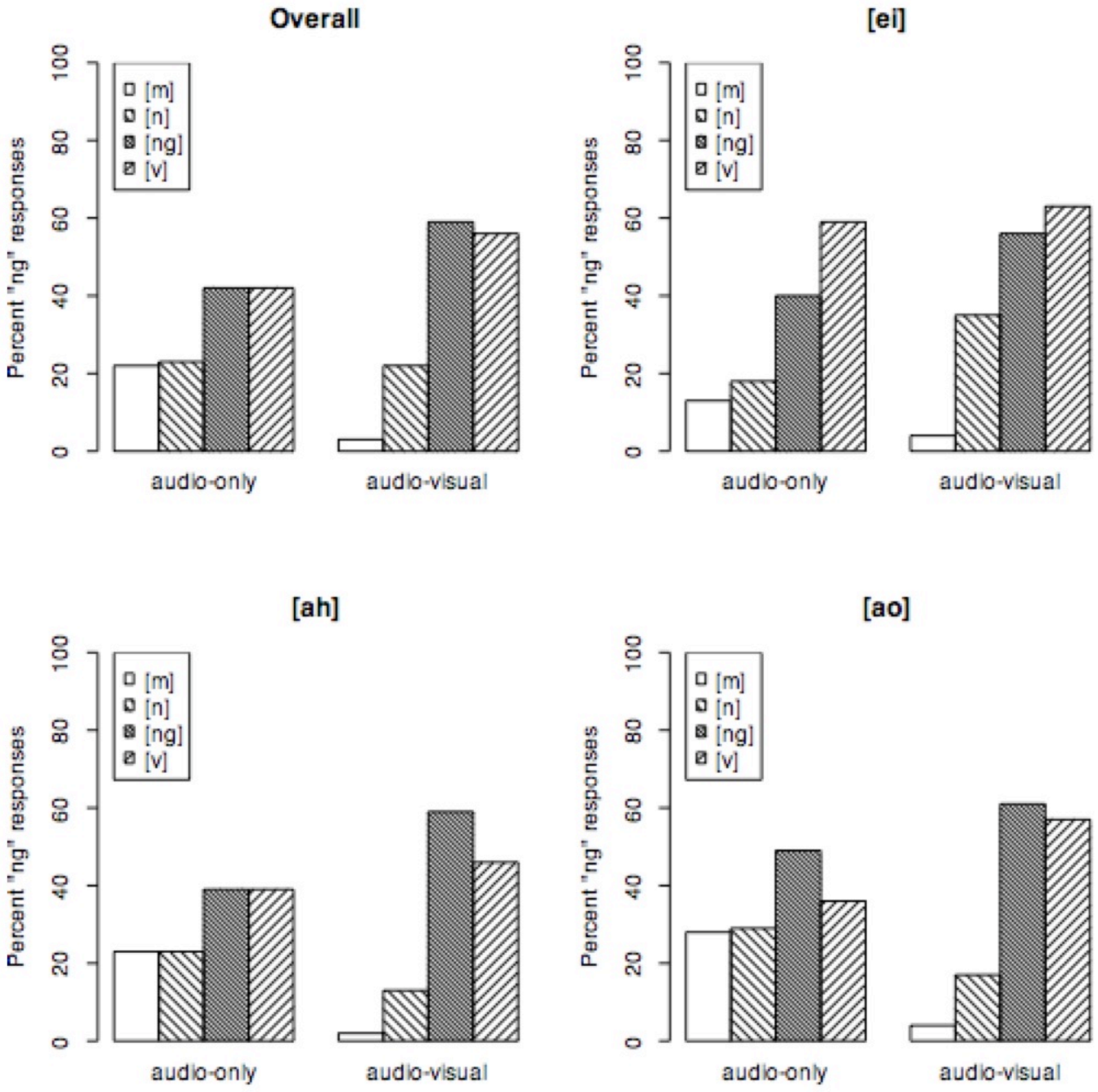
Procedure: 36 audio-only trials

Task: Identify the final consonant as "m", "n", or "ng".

Audio-visual stimuli

Procedure: 36 audio-visual trials

Task: Identify the final consonant as "m", "n", or "ng".


Overall
[ei]
[ah]
[ao]
Percent "ng" responses
0
20
40
60
80
100
[m]
[n]
[ng]
[v]
audio-only
audio-visual

Results from Experiment 3

Percent "m", "n", and "ng" responses (columns) to the audio-only and AV stimuli ending in different nasal consonants or a nasalized vowel (rows)

		"m"	"n"	"ng"
audio-only	[m]	25	55	22
	[n]	21	56	23
	[ŋ]	17	40	42
	[x̃]	28	31	42
audio-visual	[m]	92	5	3
	[n]	4	75	22
	[ŋ]	4	38	59
	[x̃]	7	38	56

Experiment 4: [ŋ] detection with noise obliterated stimuli

- Analogous to experiment 2.

  What if listeners didn't really think that nasalized vowels sound like [ŋ]?

Task: [ŋ] detection - "Did the token end with "ng" or not?"

Subjects:	Fifteen (8 women, 7 men) UCB undergrads.
Materials:	Identical to Experiment 3.
Procedure:	Block of 36 audio-only trials then a block of 36 audio-visual trials.

Results from Experiments 3 and 4

Percent "m" "n" and "ng" responses (columns) to the audio-only and AV stimuli ending in different nasal consonants or a nasalized vowel in experiment 3 (rows), and percent "ng" responses in Experiment 4.

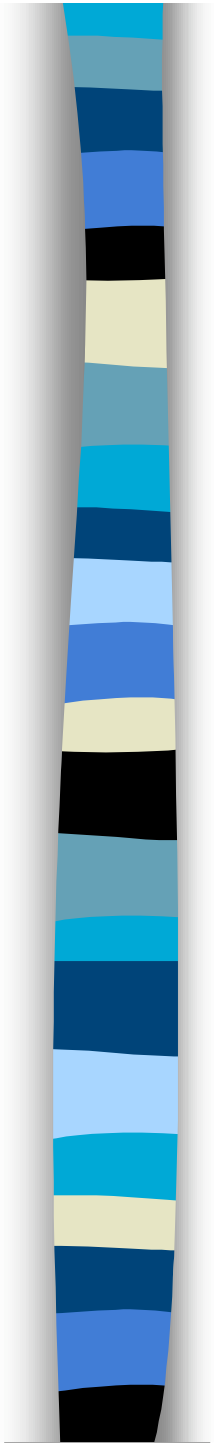
		Exp 3			Exp 4
		"m"	"n"	"ng"	"ng"
audio-only	[m]	25	55	22	33
	[n]	21	56	23	33
	[ŋ]	17	40	42	47
	[x̃]	28	31	42	39
audio-visual	[m]	92	5	3	4
	[n]	4	75	22	45
	[ŋ]	4	38	59	69
	[x̃]	7	38	56	47

Conclusions

1. Both the acoustic and visual similarity hypotheses are supported. Listeners are sensitive to the acoustic and visual similarity of nasalized vowels and velar nasals when deciding the place of articulation of a coda nasal.

2. Innocent misperception of nasalized vowels can lead listeners to posit the existence of a coda velar nasal.

3. A plausible historical phonological explanation of excrescent nasal velarity is that the change originates from the misperception of nasalized vowels.



Slides from this talk are available at:

https://webfiles.berkeley.edu/cdicanio/public_html/talks