Perceptual evidence for the representation of English coda voicing

Chelsea Sanker and Robin Karlin

Yale University, UW-Madison

chelsea.sanker@yale.edu; rkarlin@wisc.edu

LSA 2020, New Orleans, LA
January 4, 2020

Yale

WISCONSIN
UNIVERSITY OF WISCONSIN–MADISON

Cues to contrasts

- Listeners use cues other than the "primary" cue for contrasts
- Example: English coda voicing
  - Vowel duration [Raphael 1972]
  - F0 of the preceding vowel [Gruenenfelder & Pisoni 1980]
  - and many others [Lisker 1986]
- Debate over the phonological status of these types of cues—sometimes considered artefacts of phonetic implementation
  - No consensus on articulatory vs. perceptual origin
  - Not a trade-off relation; the existence of actual voicing in the coda does not negatively correlate with vowel duration differences
  - Various secondary cues are not ranked with respect to each other, despite differences in consistency of production [Gruenenfelder & Pisoni 1980, Chen 1970]

Roadmap

- Today: provide additional perceptual evidence that vowel duration for coda voicing is phonological, not just an artefact of implementation
- Roadmap:
  - Part 1: Duration vs. pitch
    - Question A: for English coda voicing
    - Question B: for English vowel quality
  - Part 2: Duration and F1 for English vowel quality
  - Conclusions and future directions

Part 1: Duration vs. F0 as secondary cues

Vowel duration and coda voicing

Production

- Cross-linguistically, vowels are longer before voiced codas (incl. sonorants), though with language-specific variation [Chen 1970, Keating 1979]
- Also language-specific patterns in contexts with voicing neutralization, indicating different degrees of phonologization
  - Vowel length differences preserved in final devoicing in German [Fourakis & Iverson 1984]
  - Neutralized in Dutch final devoicing [Warner et al. 2004]
  - Preserved in English whispering [Sharf 1964]

Perception

- English speakers are more likely to perceive codas as voiced when the preceding vowel is longer, either absolutely [Raphael 1972] or relatively [Denes 1955, Port & Dalby 1982]

F0 and coda voicing

Production

- Strong correlation of F0 with onset voicing: voiceless onsets are followed by higher F0 [House & Fairbanks 1953, Kong, Beckman, & Edwards 2012]
- Weaker evidence with coda voicing: some find parallel effect [Kohler 1982]; others do not [Mohr 1971, Gruenenfelder & Pisoni 1980]

Perception

- Clearer relationship than for production: lower overall F0 [Castleman & Diehl 1996] or falling F0 [Gruenenfelder & Pisoni 1980] in stimuli increases listeners' perception that codas are voiced

Part 1 methodology (two experiments)

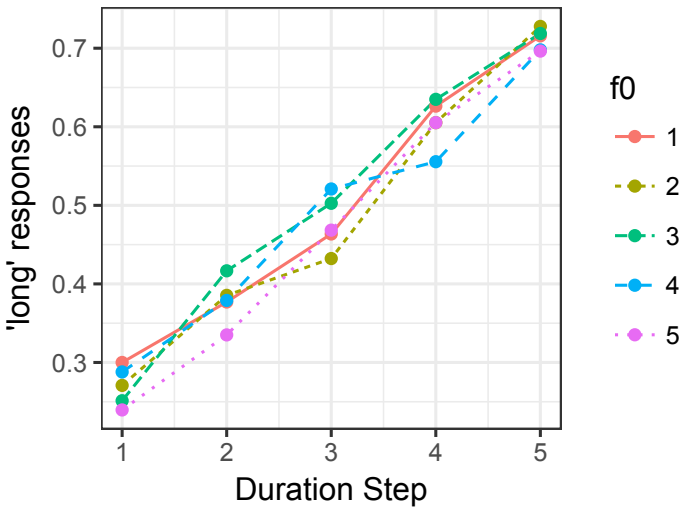
- **Participants**: 24 native speakers of American English in each experiment
- **Stimuli**: Manipulations of CVC forms produced by a female English speaker
  - Voiced (/g/) and voiceless (/k/) codas, which were removed and spliced back in (`Coda`)
  - Vowel duration was manipulated to create a 5-step duration continuum (145 ms to 250 ms) (`DurStep`)
  - F0 was manipulated to create a 5-step continuum (140 Hz to 220 Hz) (`F0Step`)
  - Two vowel environments: /ʌ/, /ɪ/—more on this later!
- **Two experiments in Part 1**:
  - **Duration**: participants rated syllables as 'short' or 'long'
  - **Pitch**: participants rated syllables on pitch, on a scale from 1 (low) to 5 (high)

Statistical analysis

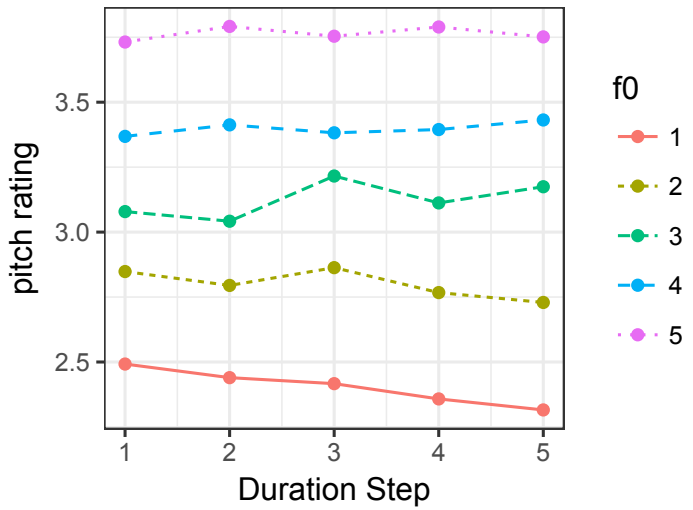
- Responses faster than 250 ms or slower than 5 s were excluded from analysis ($< 1\%$ of the data)
- Regression models for the main analysis — we will illustrate each factor in separate figures for visualization, but significance is demonstrated at the end in the full models

Effects of DurStep and F0Step on duration and pitch

Duration decisions (Exp 1)



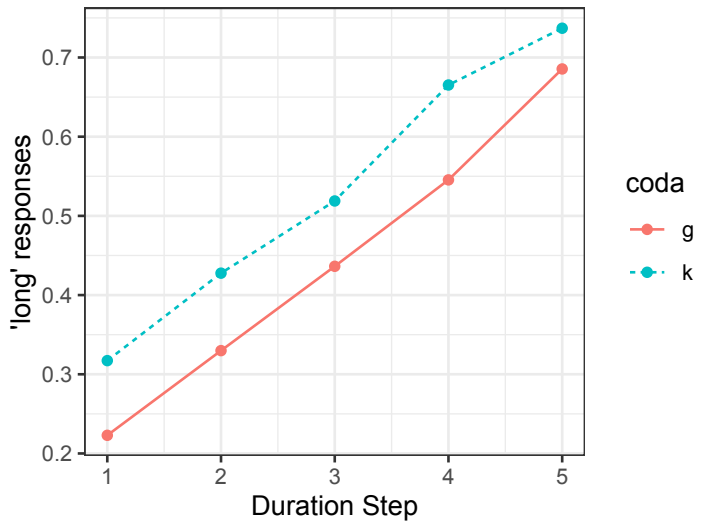
Pitch decisions (Exp 2)



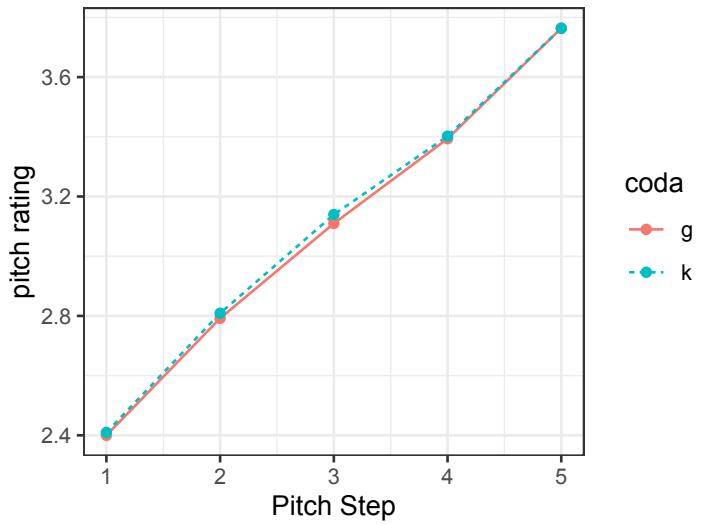
- DurStep and F0Step were major predictors of length and pitch in their respective studies
- Tokens with greater duration were more likely to be rated as long, and tokens with higher F0 were rated as higher pitched

Effects of Coda on duration but not pitch

Duration decisions (Exp 1)



Pitch decisions (Exp 2)



- Coda influenced duration ratings
  - Vowels with voiced codas were rated as shorter than vowels with voiceless codas
  - This suggests that English speakers have the expectation that vowels should be longer before voiced codas
- However, Coda did not influence pitch ratings, indicating that F0 differences are not expected between voiced and voiceless codas

Interim summary: asymmetric behavior of duration and F0

Duration		**F0**	
✓	Consistent production	?	Less consistent production
✓	Readily used for perception	✓	Available, but not as strong of a cue
✓	Vowel duration differences expected in listening—vowels before voiced codas are rated as shorter (this study)	✗	F0 differences not expected in listening; no effect of `F0Step` (this study)

- Suggests that vowel duration is included in the phonology of English coda voicing distinctions, but F0 is not
- Two possibilities:
  - Included in the underlying representation (gestural approach, redundant featural specification, etc.)
  - Included in the phonological grammar—easy enough with this two-way distinction

But wait, there's more...

- **Participants**: 24 native speakers of American English in each experiment
- **Stimuli**: Manipulations of CVC forms produced by a female English speaker
  - Voiced (/g/) and voiceless (/k/) codas, which were removed and spliced back in (`Coda`)
  - Vowel duration was manipulated to create a 5-step duration continuum (145 ms to 250 ms) (`DurStep`)
  - F0 was manipulated to create a 5-step continuum (140 Hz to 220 Hz) (`F0Step`)
  - Two vowel environments: /ʌ/, /ɪ/ (`Vowel`)
- **Two experiments in Part 1**:
  - **Duration**: participants rated syllables as 'short' or 'long'
  - **Pitch**: participants rated syllables on pitch, on a scale from 1 (low) to 5 (high)

Duration and vowel height

Production

- Higher vowels are shorter than lower vowels
  [House & Fairbanks 1953, Solé & Ohala 2010]
- Some evidence that intraspeaker F1 variability within a single vowel quality does **not** correlate with duration
  [Toivonen et al. 2015]

Perception

- High vowels are perceived as longer than lower vowels—compensation effect
  [Gussenhoven 2007, Wang et al. 1976]

F1 and F0

Production

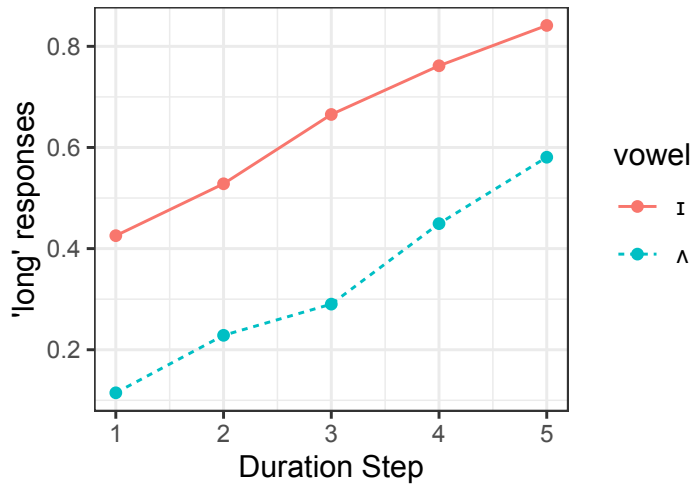
- For a given speaker, higher F0 is associated with lower F1 (higher vowel) [Ohala & Eukel 1987, Whalen & Levitt 1995]

Perception

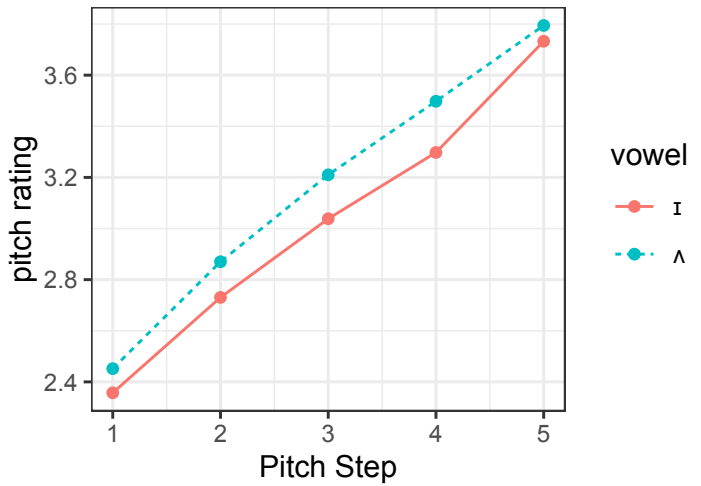
- When F0 is higher, the vowel is perceived as higher [Fujisaki & Kawashima 1968, Traunmüller 1981]

Effects of `Vowel` on both duration and pitch

Duration decisions (Exp 1)



Pitch decisions (Exp 2)



- /ʌ/ was less likely to be identified as long than /ɪ/
  - Consistent with compensation for expected duration of each vowel height
- Smaller but still significant effect on pitch ratings—/ʌ/ was rated as higher pitched than /ɪ/
  - Consistent with compensation for expected F0 of each vowel height

Model for 'long' responses in (Experiment 1)

	β	SE	z-value	p-value
(Intercept)	-1.1	0.17	-6.8	< 0.001
DurationStep	0.29	0.013	22.4	< 0.001
PitchStep	-0.028	0.024	-1.17	0.24
OrigCoda-vclss	-0.02	0.067	-0.31	0.76
Coda-vclss	0.47	0.067	7.0	< 0.001
Vowel-/ʌ/	-1.6	0.07	-22.5	< 0.001

Table: Predictors of 'long' responses (binomial). *Intercept: OrigCoda = Voiced; Coda = Voiced; Vowel = /ɪ/*

Model for pitch ratings (Experiment 2)

	β	SE	t-value	p-value
(Intercept)	2.1	0.07	29.5	< 0.001
DurationStep	-0.0028	0.0048	-0.58	0.56
PitchStep	0.33	0.0096	34.6	< 0.001
OrigCoda-vclss	-0.0081	0.027	-0.3	0.77
Coda-vclss	0.014	0.027	0.5	0.62
Vowel-/ʌ/	0.13	0.027	4.9	< 0.001

Table: Predictors of pitch ratings. *Intercept: OrigCoda = Voiced; Coda = Voiced; Vowel = /ɪ/*

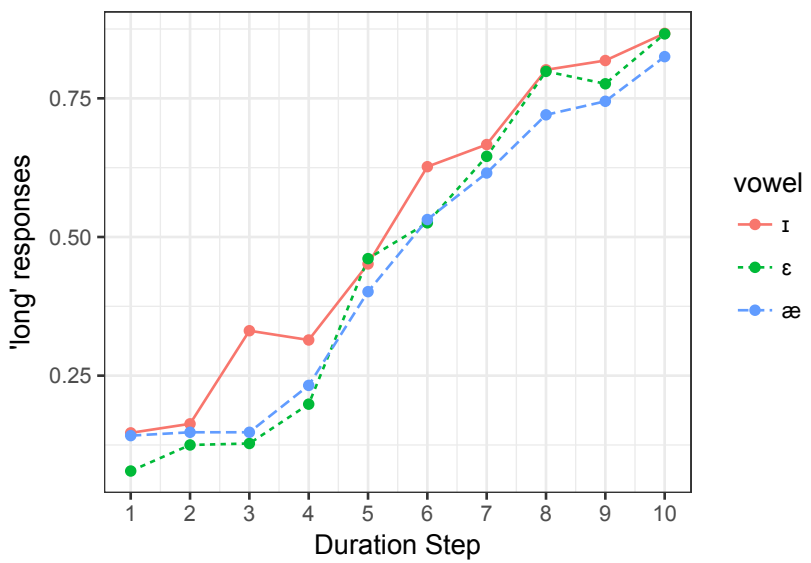
Part 2: More complications of adding vowel duration to the phonology

Experiment 3: Vowel quality effects on perceived duration

- **Participants**: 24 native speakers of American English
- **Stimuli**: Manipulations of CVC forms produced by a female English speaker
  - Onsets and codas were removed in stimulus preparation (only the vowel was presented)
  - Vowel duration was manipulated to create a 10-step duration continuum (130 ms to 252 ms) (`DurStep`)
  - Three vowel qualities: /æ, ɛ, ɪ/ (`Quality`)
  - F1 manipulations within each vowel quality: raised, natural, lowered – 60 Hz +/- 6% of the original (`F1`)
- **Task**: Participants rated syllables as 'short' or 'long'

Effects of Quality

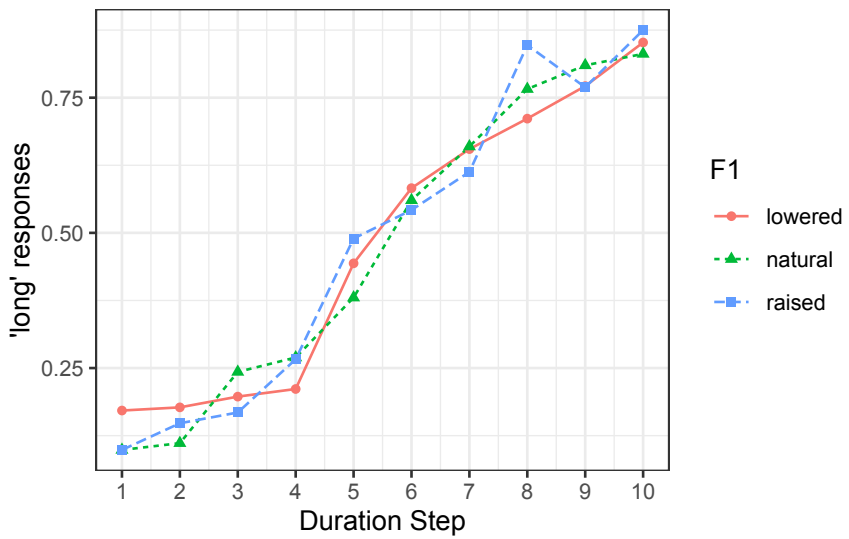
- Consistent with previous results, /ɪ/ is more likely to be identified as long than /æ/ and /ɛ/ (though the difference between /æ/ and /ɛ/ was not significant)



'long' identifications, by duration step and vowel

Effects of F1

- Within vowel categories, there is no clear effect of F1 on perceived duration



'long' identifications, by duration step and within-category F1

Model for 'long' responses in Experiment 3

	β	SE	z value	p value
(Intercept)	-2.95	0.213	-13.8	< 0.001***
Duration Step	0.505	0.0162	31.3	< 0.001***
Vowel /ɪ/	0.359	0.0936	3.84	< 0.001***
Vowel /æ/	-0.061	0.0934	-0.652	0.514
F1	0.0134	0.0467	0.288	0.774
OrigCoda Voiceless	-0.17	0.0763	-2.23	0.026*

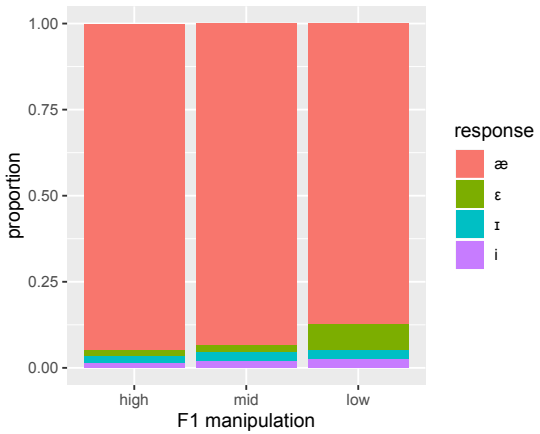
Table: Predictors of 'long' responses (binomial). *Intercept: Vowel = /ɛ/; OrigCoda = Voiced*

Experiment 4: Perceived vowel quality

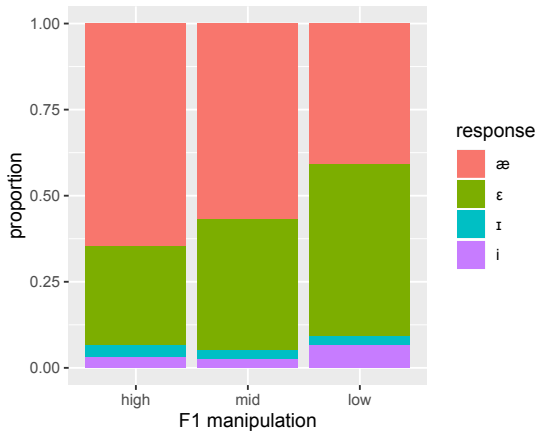
- **Participants**: 25 native speakers of American English
- **Stimuli**: Manipulations of CVC forms produced by a female English speaker (same stimuli from 3)
  - Onsets and codas were removed in stimulus preparation (only the vowel was presented)
  - Three vowel durations: 130 ms, 197 ms, 252 ms (`DurStep`)
  - Three vowel qualities: /æ, ɛ, ɪ/ (`Quality`)
  - F1 manipulations within each vowel quality: raised, natural, lowered (`F1`)
- **Task**: Participants identified each vowel quality by identifying the matching vowel in an array of monosyllabic English words (*beat*, *bit*, *bet*, *bat*)

Effects of F1 on perceived vowel quality

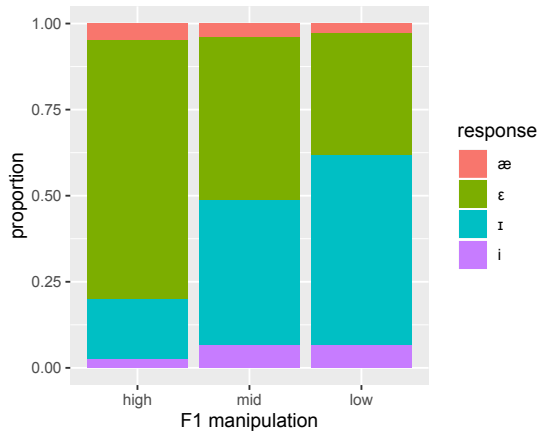
/æ/ items



/ɛ/ items



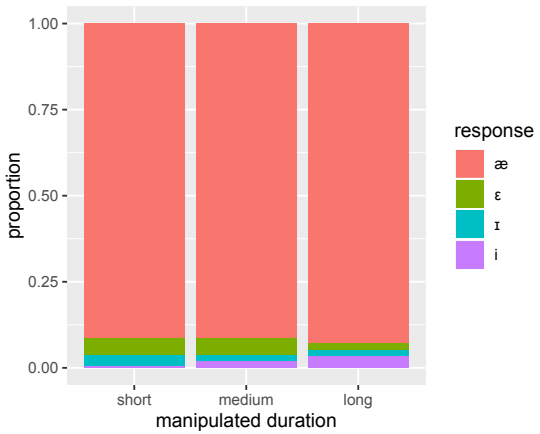
/ɪ/ items



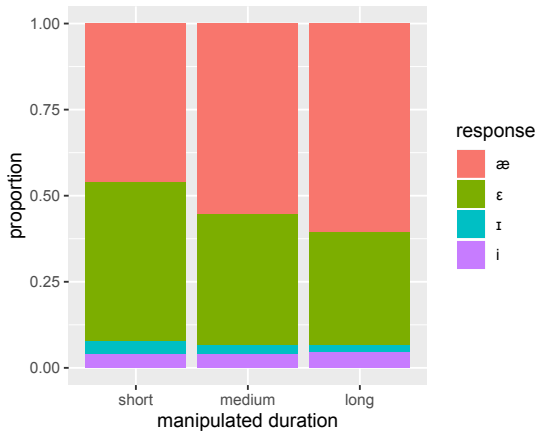
- /æ/ stimuli were consistently identified as /æ/
- /ɛ/ responses for /ɛ/ stimuli increased with lower F1: χ^2 (6, N = 450) = 20.5, p = 0.0022)
- /ɪ/ responses for /ɪ/ stimuli increased with lower F1: χ^2 (6, N = 450) = 57.0, p < 0.001)

Effects of Duration on perceived vowel quality

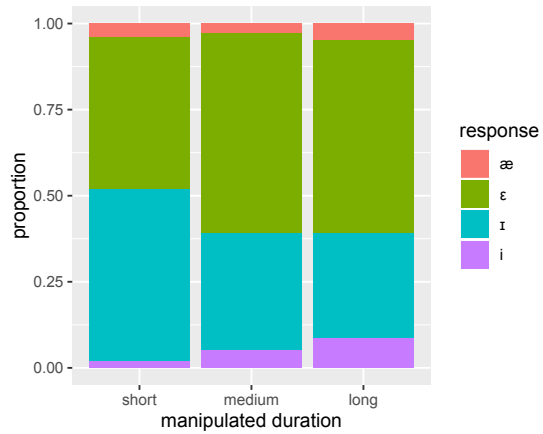
/æ/ items



/ɛ/ items



/ɪ/ items



- /æ/ stimuli were consistently identified as /æ/
- /ɛ/ stimuli exhibited a trend towards more /æ/ identifications at longer durations: χ^2 (6, N = 450) = 7.7, p = 0.26
- /ɪ/ stimuli were more likely to be identified as /ɛ/ at longer durations: χ^2 (6, N = 450) = 18.7, p = 0.0047

Conclusions

How do we decide what cues are phonologized?

- **Criteria for inclusion**:
  - Does the cue exist in *production*?
  - Is the cue used in *perception* for the main contrast?
  - Does the main contrast influence *perception of the secondary cue*?

How do we decide what cues are phonologized?

- **Criteria for inclusion**:
  - Does the cue exist in *production*?
  - Is the cue used in *perception* for the main contrast?
  - Does the main contrast influence *perception of the secondary cue*? —this is the new evidence and strategy we provided today

Our results tease apart different secondary cues

- There is distinct perceptual behavior of two cues for coda voicing in English:
  - Vowel duration is expected and compensated for;
  - Mean F0 is not
- Compensation suggests language-specific cues and thus phonological status, while mere cue usage could stem from weak exemplar memories or a range of perceptual biases
- These differences in behavior indicate the need to distinguish between the phonological status of secondary correlates of contrasts

Further refinement/future questions

Where in the phonology are these secondary cues?

- In the phonological representation?
  - Redundant features (multiply-specified contrast)
  - Gestural representation
- In the phonological grammar?
  - i.e. the output of a process conditioned by the primary contrast

Future directions

- Comparison across languages that have different behaviors of these characteristics in production (e.g. Dutch vs. German)
- Correlations across individuals between compensation and perceptual use of cues

References

References

Castleman, W.A., and Diehl, R.L. (1996). Effects of fundamental frequency on medial and final [voice] judgements. *Journal of Phonetics* 24, 383–398.

Chen, M. (1970). Vowel length variation as a function of the voicing of the consonant environment. *Phonetica* 22, 129–159.

Fourakis, M., and Iverson, G. (1984). On the 'incomplete neutralization' of German final obstruents. *Phonetica* 41(3), 140–149.

Fujisaki, H. and Kawashima, T. (1968). The roles of pitch and higher formants in the perception of vowels. *IEEE TRans. Audio Electroacoust* AU-16(1), 73–77.

Gruenenfelder, T., and Pisoni, D. (1980). Fundamental frequency as a cue to postvocalic consonantal voicing: Some data from speech perception and production. *Perception and Psychophysics* 28(6), 514–520.

Gussenhoven, C. (2007). A vowel height split explained: Compensatory listening and speaker control. In Cole, J. and Hualde, J. I., editors, *Laboratory Phonology 9*, pages 145–172. Mouton de Gruyter.

House, A.S. and Fairbanks, G. 1953. The influence of consonant environment upon the secondary acoustical characteristics of vowels. *Journal of the Acoustical Society of America* 25(1), 105–113.

Keating, P.A. (1979). *A Phonetic Study of Voicing Contrast in Polish*. PhD thesis, Brown University.

Kohler, K.J. 1982. F0 in the production of lenis and fortis plosives. *Phonetica* 39, 199–218.

Kong, E.J., Beckman, M.E., and Edwards, J. (2012). Voice onset time is necessary but not always sufficient to describe acquisition of voiced stops: The cases of Greek and Japanese. *Journal of Phonetics* 40, 725–744.

Lisker, L. (1986). 'Voicing' in English: A catalogue of acoustic features signaling /b/ versus /p/ in trochees. *Lang Speech* 29, 2-11.

Ohala, J.J., and Eukel, B.W. (1987). Explaining the intrinsic pitch of vowels. *In honor of Ilse Lehiste*, pages 207–215.

Port, R.F., and Dalby, J. (1982). Consonant/vowel ratio as a cue for voicing in English. *Perception & Psychophysics* 32, 141–152.

Raphael, L.J. (1972). Preceding vowel duration as a cue to the perception of the voicing characteristic of word-final consonants in American English. *JASA* 51(4), 1296–1303.

Sharf, D.J. (1964). Vowel duration in normal and whispered speech. *Language and Speech* 7, 89–97.

Solé, M.-J. and Ohala, J.J. (2010). What is and what is not under the control of the speaker: Intrinsic vowel duration. In Fougeron, C., Kuehnert, B., Imperio, M., and Vallee, N., eds, *Laboratory Phonology 10*, pp. 607–655. Mouton de Gruyter.

Toivonen, I., Blumenfeld, L., Gormley, A., Hoiting, L., Logan, J., Ramlakhan, N., and Stone, A. (2015). Vowel height and duration. In Steindl, U., Borer, T., Fang, H., Pardo, A.G., Guekguezian, P., Hsu, B., O'Hara, C., and Ouyang, I.C., editors, *Proceedings of the 32nd WCCFL*, pages 64–71.

Traunmüller, H. (1981). Perceptual dimension of openness in vowels. *JASA* 69, 1465–1475.

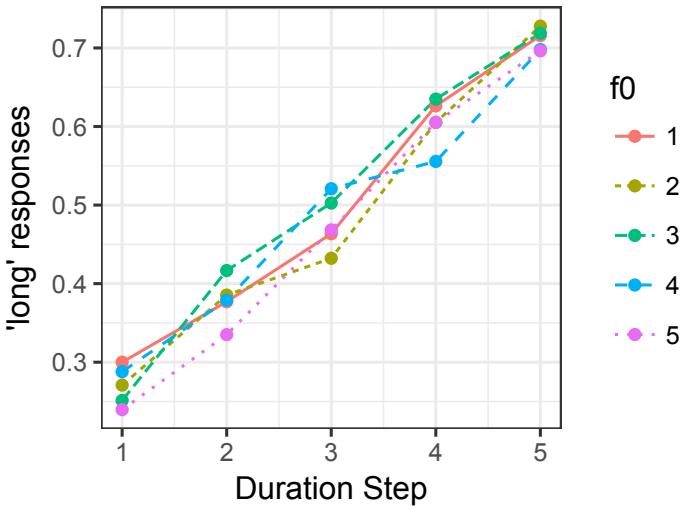
Wang, W. S.-Y., Lehiste, I., Chuang, C.-K., & Darnovsky, N. (1976). Perception of vowel duration. *JASA* 60(S1), S92.

Warner, N., Jongman, A., Sereno, J., and Kemps, R. (2004). Incomplete neutralization and other sub-phonemic duration differences in production and perception: Evidence from Dutch. *Journal of Phonetics* 32, 251–276.

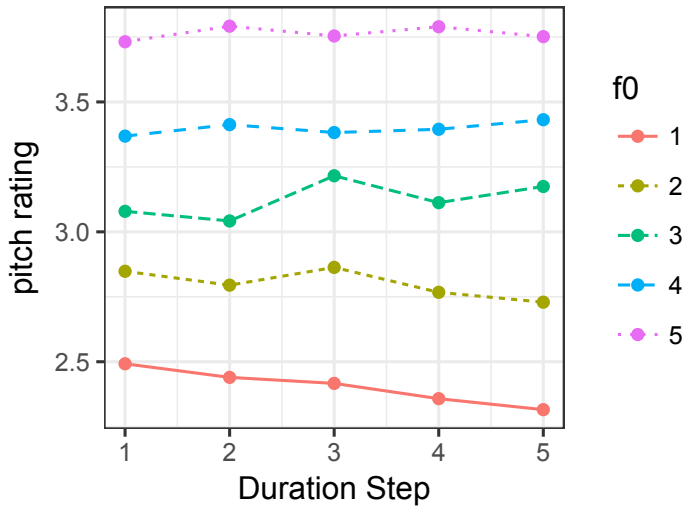
Whalen, D.H., and Levitt, A.G. (1995). The universality of intrinsic F0 of vowels. *Journal of Phonetics* 23, 349–366.

Duration and F0 as predictors

Duration decisions (Study 1)



Pitch decisions (Study 2)

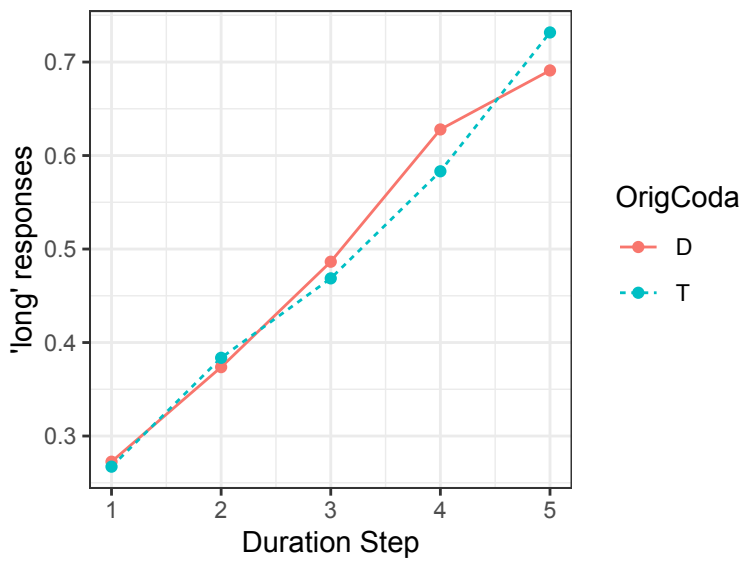


`DurStep` doesn't predict pitch responses and `F0Step` doesn't predict duration responses

This contrasts with some previous work in which higher F0 increases perceived duration (Yu 2010; Gussenhoven & Zhou 2013; Rosen 1977), as does pitch in non-linguistic stimuli (Brigner 1988)

Effects of the original coda

When a coda is present, vowels before voiceless codas are more likely to be identified as long.
However, vowels that were produced with voiced codas are, if anything, more likely to be identified as long.

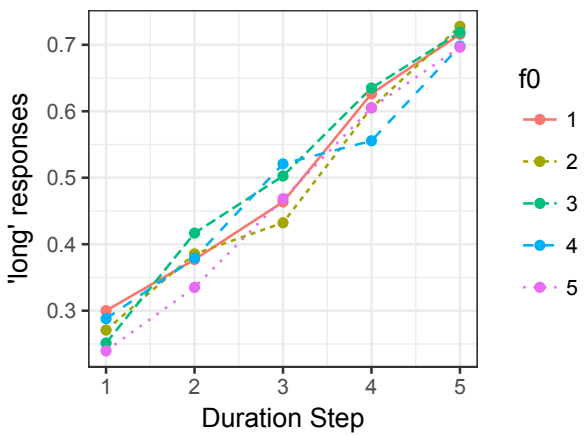


'long' identifications, by duration step and original coda

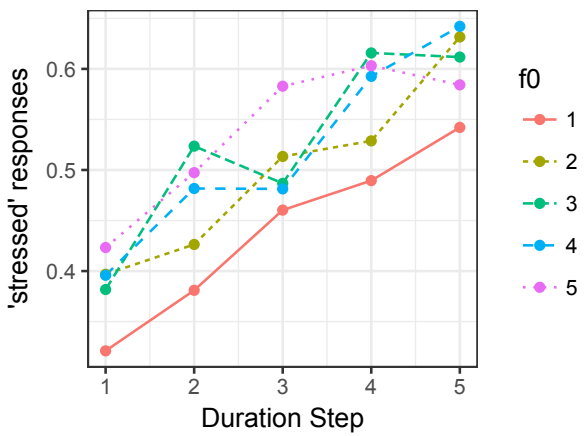
Stress vs. Duration

Listeners' responses differ when asked to identify stress or vowel duration; for example, f0 has a large effect on stress identifications, but a smaller effect on vowel duration identifications.

Duration



Stress



Listeners identifications of the same vowels as long/short or stressed/unstressed, by duration step and f0.