Background on Hmong

- Hmong-Mien language family
- Spoken in China, Laos, Vietnam, Thailand
- Also large populations in California, Minnesota, and Wisconsin
- Several varieties (White, Green, Black, etc.) with various levels of mutual intelligibility
- ~3-4 million speakers of all dialects, 200,000 Hmong in US (most of whom speak White or Green varieties)

White Hmong tones

Tone (Esposito, to appear)	Orthographic tone symbol	Example in IPA	Example in White Hmong orthography
High-rising (45)	-b	[pɔ^{45}]	pob ‘ball’
Mid (33)	∅	[pɔ^{33}]	po ‘spleen’
Low (22)	-s	[pɔ^{22}]	pos ‘thorn’
High-falling (52)	-j	[pɔ^{52}]	poj ‘female’
Mid-rising (24)	-v	[pɔ^{24}]	pov ‘to throw’
Low-falling creaky (21)	-m	[pɔ̰^{21}]	pom ‘to see’
High-falling breathy (42, 52)	-g	[pɔ̤$^{52/42}$]	pog ‘grandmother’

- There is also an eighth (-d) tone, which is a syntactic variety of the –m tone

Importance of phonation

- The relative importance of phonation cues to tone identification in White Hmong is unclear
- Breathy (52/42) tone and modal (52) tone both have similar falling pitch  
- Creaky (21) tone and modal (22) tone are less similar in pitch
- Low modal (22) tone is significantly longer in duration than creaky (21) tone[1]  

1. Esposito (to appear)

Tone and phonation contrasts

- Languages may contrast phonation → 'register' languages (e.g. Chong)[1]
- Many others contrast tones (e.g. Thai)[2]
- Some languages cross-classify tones and phonation types (e.g. Jalapa Mazatec)[3]
- In some tone languages, phonation changes are associated with certain tones (e.g. Hmong)[4]

1. DiCanio (2009); 2. Tingsabadh & Abramson (1993); 3. Garellek & Keating (2011); 4. Esposito (to appear)

Phonation in tone systems

- Phonation can be used as an independent dimension
  → PHONATION AND PITCH INDEPENDENCE

- Non-modal phonation (types of creaky/laryngealized voice) can accompany pitch height due to physiological interdependencies[1]
  - Vocal fold tenseness is common at very high F0
  - Creak or vocal fry is common at very low F0
  - Creaky phonation can be used to reach pitch target, but breathy phonation can occur at any pitch height as an additional contrast[2]
  - Conversely, certain voice registers (e.g. faucalized voice) are accompanied by changes in pitch[3]

  → PHONATION AND PITCH INTERDEPENDENCE

- Both of these possible for role of phonation in White Hmong

1. Sundberg (1987); 2. Kuang (2012); 3. Edmondson & Esling (2006)

Hmong tone perception

- Little is known about White Hmong tonal perception
- Andruski (2006) found better identification of natural tokens of breathy/creaky tones than modal ones in White Hmong/Green Mong
- Possible that improved identification is facilitated by non-modal phonation of the breathy (52) and creaky (21) tones
- But *relative* importance of phonation compared to other cues (F0, duration) is still unknown

Insight from other tone languages

- Studies on other tone languages show that non-modal phonation helps in the identification of certain lexical tones
  - Cantonese[1]
  - Karen[2]
  - Mandarin[3]
  - Vietnamese[4]
- These studies point to the advantage of *creaky voice* in particular
- Ongoing work on Black Miao (Black Hmong) reveals perceptual advantage of breathiness [5]

1. Yu & Lam (2011); 2. Brunelle & Finkeldey (2011); 3. Belotel-Grenié & Grenié (1997); 4. Brunelle (2009); 5. Kuang (to appear)

Goals of this study

- Determine how phonation cues are used in White Hmong, where non-modal phonation is associated with certain tones
- Determine the relative importance of phonation cues in tonal recognition in Hmong
- Better understand the relationship between pitch and voice quality in tone

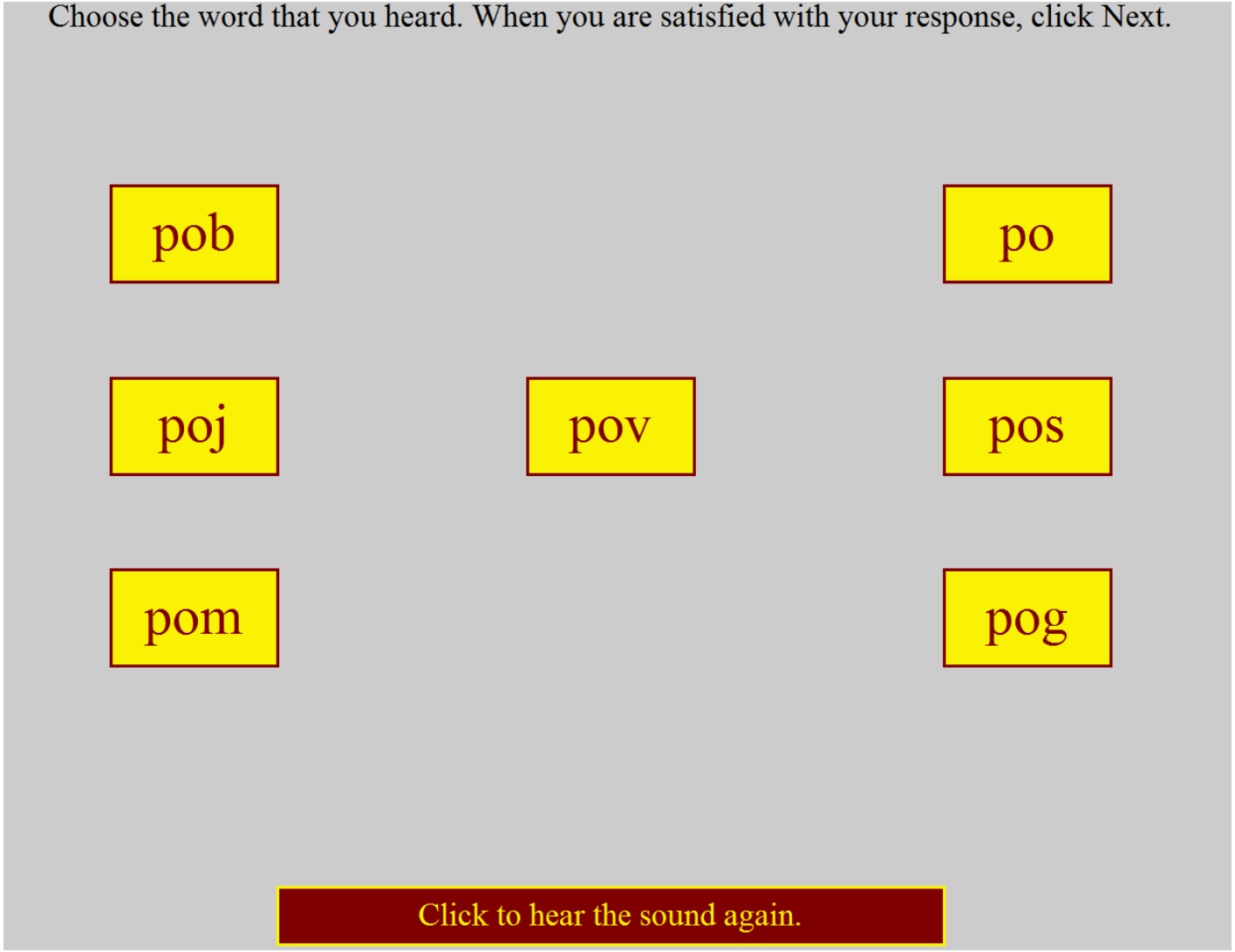
Present study

- 7-alternative forced-choice task, implemented in Praat
  - chose which word they heard (7 tones → 7 alternatives)
- 15 participants (8 female, 7 male), all native speakers of White Hmong
- Study was conducted at the Hmong-American Partnership in St. Paul, MN

Experimental setup

- Participants chose which word they heard
- They could hear the stimulus as many times as they chose
- They could change their response before moving to the next stimulus

Experimental setup



Stimuli

- Sample tokens of /pɔ/ with 6 of 7 tones were recorded by female native speaker (tone 33 resynthesized from 22)
- F0 was resynthesized in Praat using PSOLA, which preserves voice quality (i.e. spectral, noise) characteristics[1]
- Post-hoc acoustic analysis indeed revealed no change in voice quality as a function of pitch manipulation

1. Moulines & Charpentier (1990)

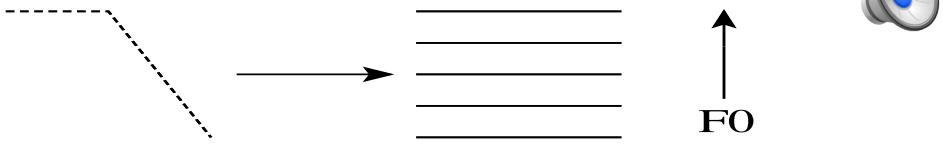
F0 and duration manipulations

- The F0 manipulations are used to determine the extent to which phonation can be relied on when the F0 of a stimulus does not match the expected pitch contour
  - E.g. if stimulus is breathy but has a very different pitch contour than 52/42, will it still be heard as breathy?

- Because the modal (22) and creaky (21) also differ in duration
  - Will a longer stimulus with creak be heard as the creaky (21) tone?
  - Will a shorter stimulus with a 21 contour but no creak be heard as the modal (22) tone?

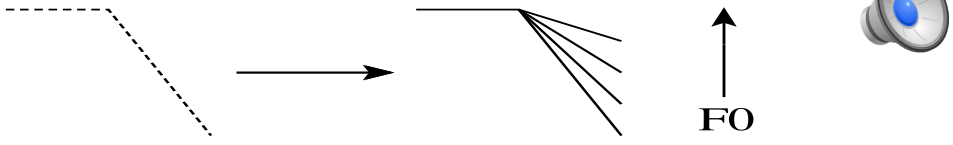
Breathy tone manipulations

Manipulation 1: Flat F0 at different levels



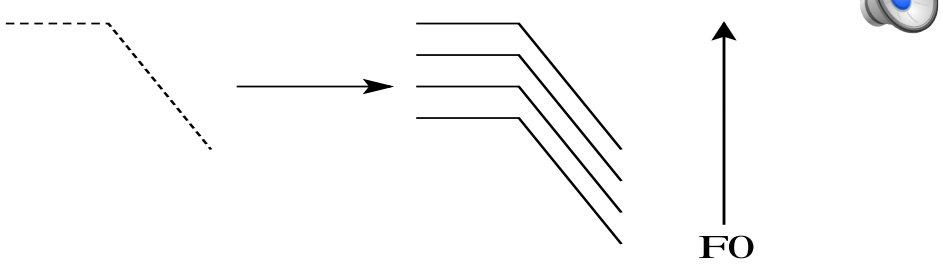
Original breathy contour → Flattened contour lowered at 10 Hz increments

Manipulation 2: Varying end of F0 fall



Original breathy contour → End of F0 fall raised by 10 Hz increments

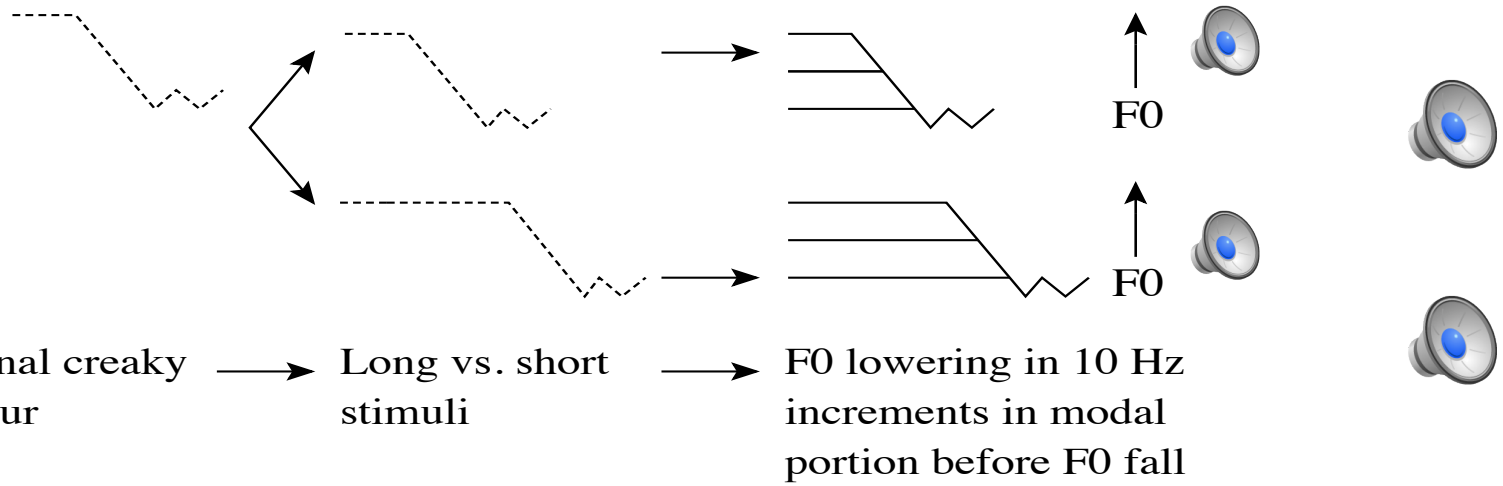
Manipulation 3: F0 shift of entire contour



Original breathy contour → Entire contour shifted by 10 Hz increments

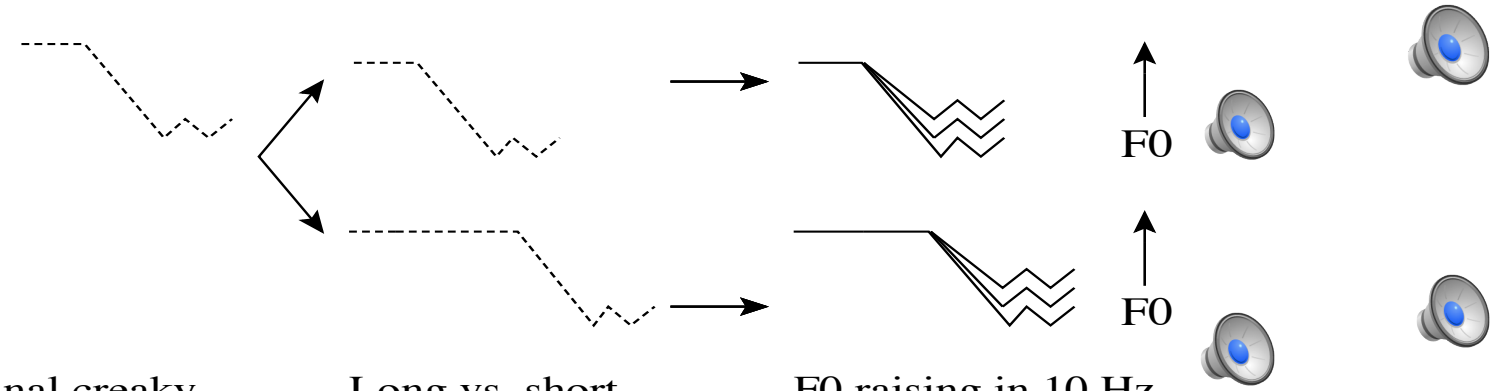
Creaky tone manipulations

Manipulation 1: Varying F0 in modal portion of creaky tone's F0 contour



Original creaky contour → Long vs. short stimuli → F0 lowering in 10 Hz increments in modal portion before F0 fall

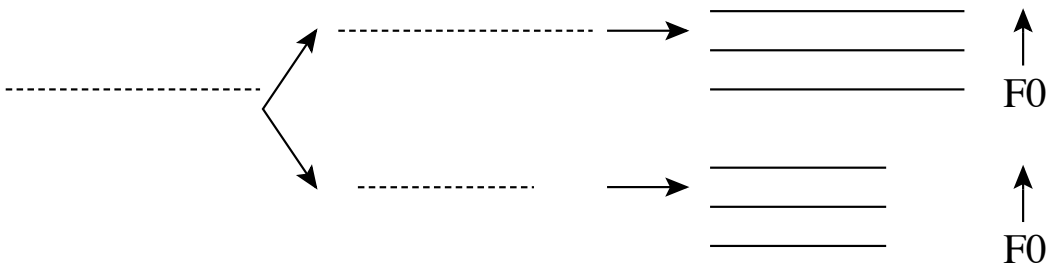
Manipulation 2: Varying F0 in creaky portion of creaky tone's F0 contour



Original creaky contour → Long vs. short stimuli → F0 raising in 10 Hz increments in creaky portion at F0 fall

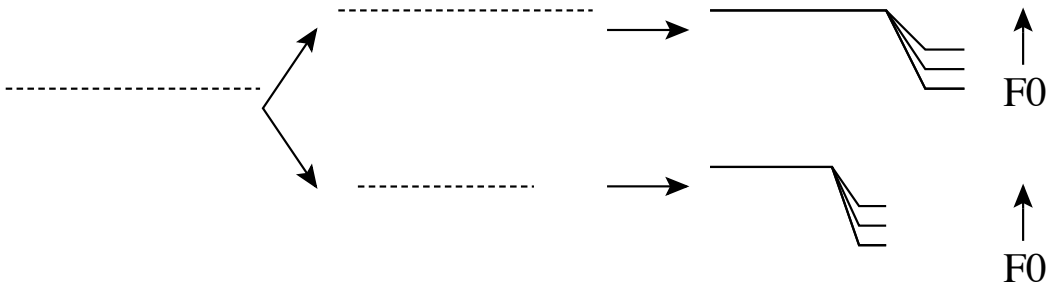
Low modal tone manipulations

Manipulation 1: Varying F0 level for the low modal tone



Original low modal contour → Long vs. short stimuli → F0 raised by 10 Hz increments

Manipulation 2: Falling F0 contour for the low modal tone



Original low modal contour | Long vs. short stimuli | F0 falls by 10 Hz increments

Other tone manipulations

- Tokens of /pɔ/ with 3 other modal tones (45, 52, 24) were included
- These also had F0 manipulated: whole F0 contour was raised or lowered in 10 Hz increments
- 127 stimuli were created, each presented twice
  - 25 from breathy (52) tone
  - 30 from creaky (21) tone
  - 24 from low-modal (22) tone
  - 38 from other modal tones

Results: analysis

- Two logistic mixed-effects models were used to determine what factors were significant in predicting
  - breathy-tone response (*pog*) vs. a modal-tone response (*pob, poj, pov, po, pos*)
  - creaky-tone response (*pom*) vs. a modal-tone response

Models' fixed effects	Models' random effect
Original phonation (B, M, C)	Participant
Start F0	
End F0	
Mean F0	
Pitch contour (flat/dynamic)	
Vowel length (long/short, for creaky model)	

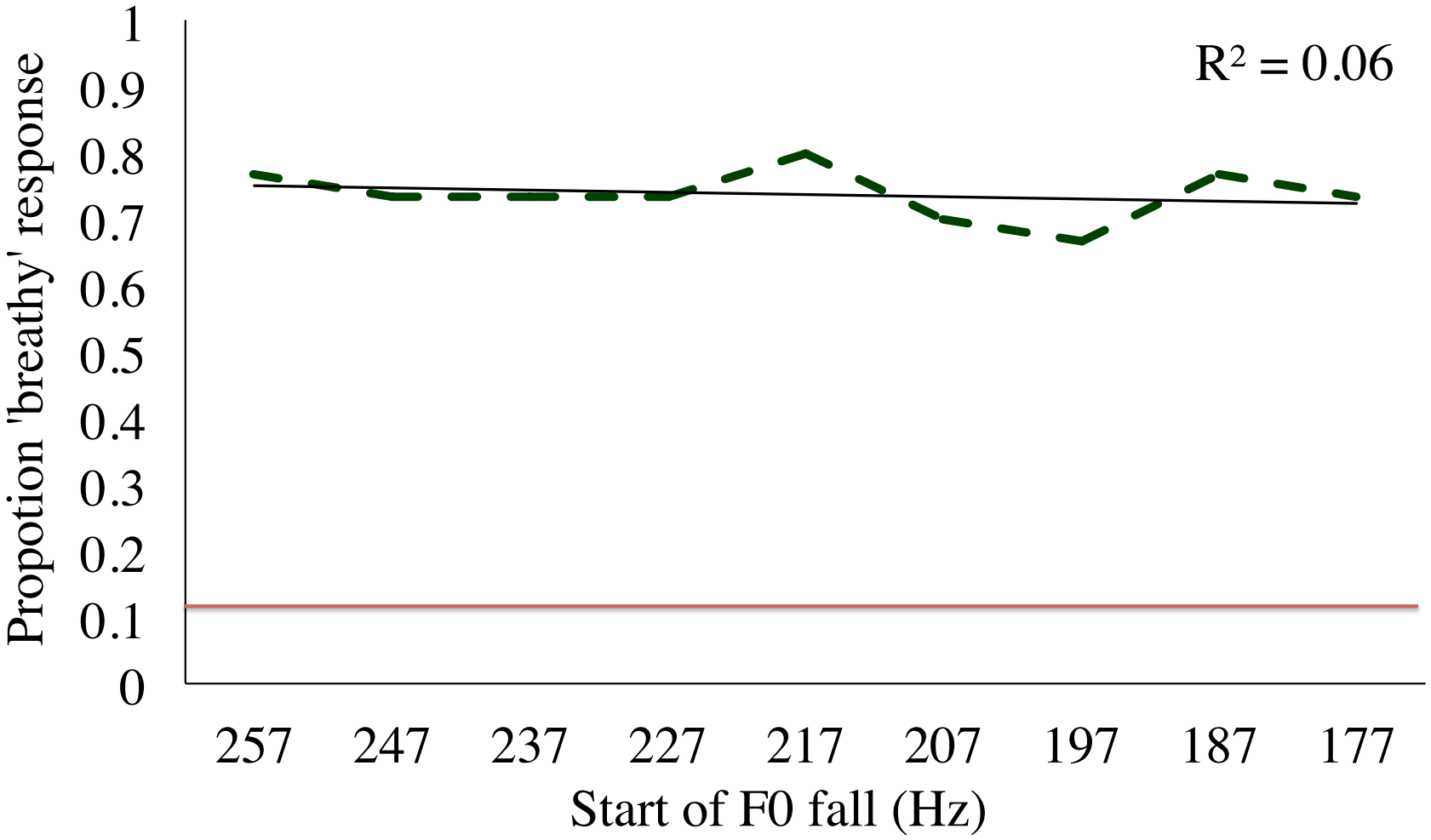
Results: breathy-tone responses

- Originally breathy stimuli increased likelihood of obtaining a ‘breathy tone’ response

- No other factor was significant

	estimate	*SE*	$Z-score$	$p-value$
Intercept	−2.48	0.38	−6.58	$< 0.0001 * **$
Orig. tone=breathy	3.98	0.18	21.57	$< 0.0001 * **$
Mean F0	−0.0008	0.01	−0.09	0.93
F0 in 1st ninth	−0.01	0.01	−1.91	0.06
F0 in final ninth	0.01	0.01	1.02	0.31
F0 slope - flat	−0.04	0.16	−0.27	0.79

Results: breathy-tone responses



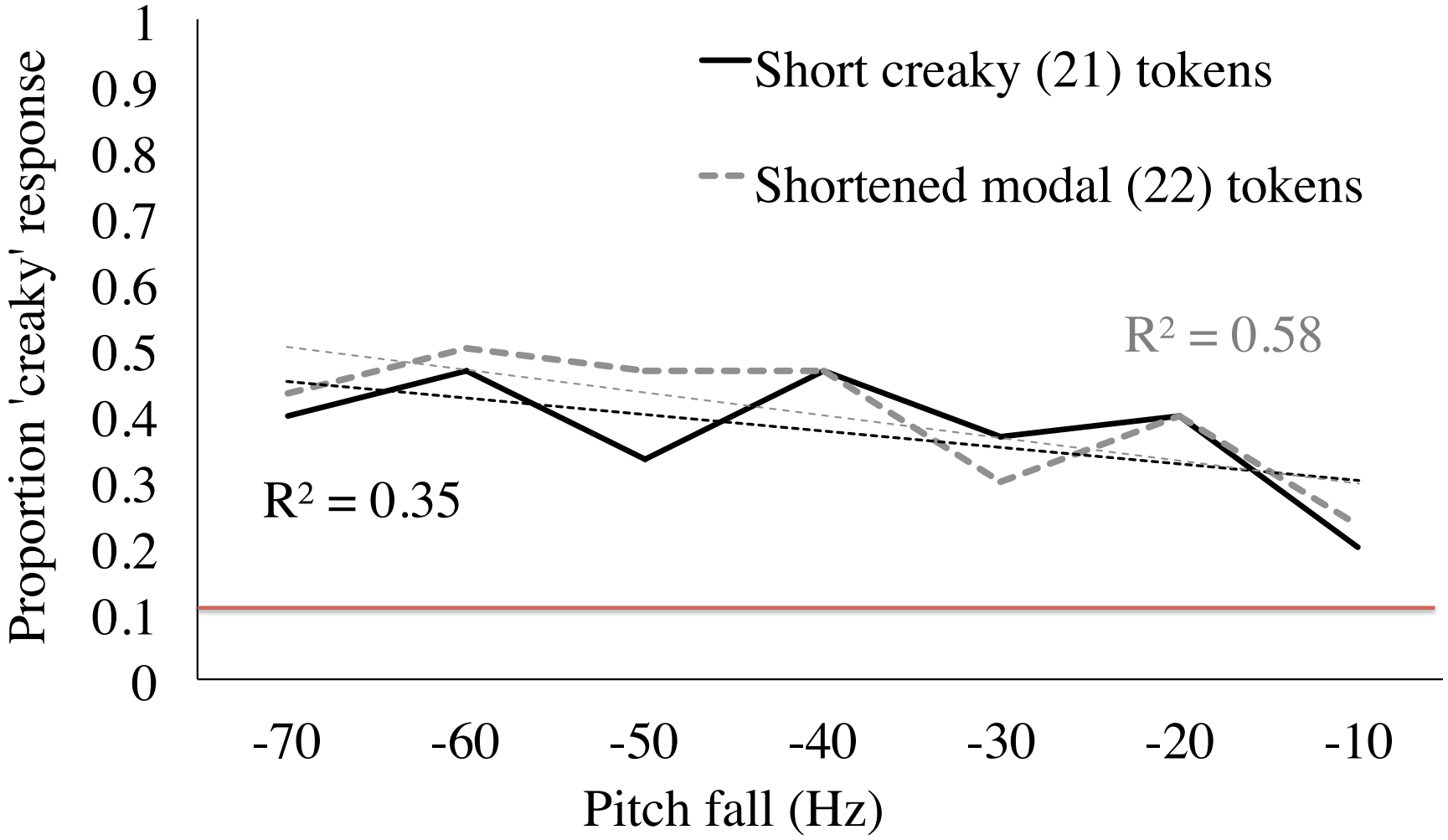
Results: creaky-tone responses

- Originally creaky stimuli did *not* increase likelihood of obtaining a 'creaky tone' response

- Falling F0 and shorter duration were significant

	estimate	*SE*	$Z-score$	$p-value$
Intercept	1.30	0.37	3.55	$< 0.001 **$
Orig. tone=creaky	0.09	0.14	0.61	0.54
Mean F0	−0.005	0.01	−0.94	0.35
F0 in 1st ninth	−0.001	0.004	−0.34	0.73
F0 in final ninth	−0.02	0.004	−3.45	$< 0.001 **$
F0 slope - flat	−1.10	0.18	−6.18	$< 0.0001 ***$
Length - short	1.11	0.13	8.69	$< 0.0001 ***$

Results: creaky-tone responses



Summary of results

- Originally-breathy stimuli were significantly more likely to be chosen as breathy-toned, *regardless of F0*
- Originally-creaky stimuli were *not* significantly more likely to be chosen as creaky-toned
- Creaky-toned responses best predicted by F0 fall and duration
  - Bigger F0 drop → more likely to be creaky-toned
  - Short duration → more likely to be creaky-toned

Discussion

- Creaky phonation in White Hmong enhances the low-falling creaky tone's pitch dynamics, but is not a primary cue in tonal recognition

- The creaky tone is sometimes called 'checked' ([V?])[1] which appears to be appropriate given our findings
  - Also, acoustic evidence showing that creakiness is predominately at end of vowel[2, 3]

1. Ratliff (1992); 2. Garellek (2012); 3. Esposito (to appear)

Discussion

- Breathy voice is the most important factor in predicting breathy-toned responses
  - Likely due to similarity in pitch between breathy –g and modal –j tones
  - In Karen & N. Vietnamese, breathiness was not significant factor in identification of tones with breathy voice quality
  - But breathiness also significant in Black Miao (Hmong) tone recognition
- Unlike previous work on Cantonese, Karen, Mandarin, and Vietnamese
  - Creaky voice in Hmong does not aid significantly in tone identification

Discussion

- Phonation and pitch can be independent…
  - Breathy phonation in Hmong → categorical shift in voice quality from modal tones, not affected by F0
- …but they are also interdependent
  - Creaky phonation in Hmong → Enhances the low pitch of the low-falling tone, not necessary for tone recognition

Discussion

- Listeners might still hear F0 falls as creaky!

- A sharp drop in F0 can be perceptually equivalent to glottalization/creaky voice, even when no other irregularity is found[1,2]
  - Listeners might hear F0 drop as creaky and thus still listening for 'creak'

1. Hillenbrand & Houde (1996); 2. Gerfen & Baker (2005)

Conclusions

- Non-modal phonation can play a primary and secondary role in tone identification *within the same language*
  - Cue weighting depends on the contrast
  - In Hmong, breathy phonation is the primary cue to the identification of the breathy (52) tone, unlike breathy phonation in other languages
  - Creaky phonation, insofar as not tied to pitch dynamics, is not a major cue to the creaky (21) tone

Ua tsaug [u^{3}ɔ^{3} tʂʌ̤^{5}ṳ^{2}] Thank you

Thanks also to Jody Kreiman, Susan Yang, the Hmong-American Partnership, and members of the UCLA Phonetics Lab. Work supported by NSF grant BCS-0720304.

References

- Andruski, J. E. (2006). “Tone clarity in mixed pitch/phonation-type tones,” *JPhon* **34**, 388–404.
- Belotel-Grenié, A. & Grenié, M. (1997). “Types de phonation et tons en chinois standard,” *Cahiers de linguistique - Asie orientale* **26**, 249–279.
- Brunelle, M. (2009). “Tone perception in Northern and Southern Vietnamese,” *JPhon* **37**, 79–96.
- Brunelle, M. &Finkeldey, J. (2011). “Tone perception in Sgaw Karen”, in *Proceedings of ICPhS 17*, 372–375.
- DiCanio, C. T. “The phonetics of register in Takhian Thong Chong,” *JIPA* **39**, 162–188.
- Edmondson, J. E., & Esling, J. H. (2006). The valves of the throat and their functioning in tone, vocal register, and stress: laryngoscopic case studies. *Phon* **23**, 157,191.
- Esposito, C. M. (to appear). “An acoustic and electroglottographic study of White Hmong phonation,” *JPhon*.
- Garellek, M. (2012). “The timing and sequencing of coarticulated non-modal phonation in English and White Hmong,” *JPhon* **40**, 152–161.
- Garellek, M. & Keating, P. (2011). “The acoustic consequences of phonation and tone interactions in Jalapa Mazatec,” *JIPA* **41**, 185–205.
- Gerfen, C. & Baker, K. (2005). “The production and perception of laryngealized vowels in Coatzospan Mixtec,” *JPhon* **33**, 311–334.
- Hillenbrand, J. M. & Houde, R. A. (1996). “Role of F0 and amplitude in the perception of glottal stops,” *JSHR* **39**, 1182–1190.
- Kuang, J. (to appear). “How is a 5 level tone contrast possible?” Talk to be presented at *Acoustics 2012*, Hong Kong, China.
- Moulines, E. & Charpentier, F. (1990). “Pitch-synchronous waveform processing techniques for text-to-speech synthesis using diphones,” *Speech Communication* **9**, 453–467.
- Ratliff, M. (1992). “*Meaningful Tone: A study of tonal morphology in compounds, form classes and expressive phrases in White Hmong*,” Monograph Series on Southeast Asia (Northern Illinois University, Center for Southeast Asian Studies, DeKalb, IL).
- Sundberg, J. (1987). “*The science of the singing voice*,” Northern Illinois University Press, DeKalb IL.
- Tingsabadh, M.R. K. & Abramson, A. S. (1993). “Thai (Illustrations of the IPA),” *JIPA* **23**, 24-28.
- Yu, K. M. & Lam, H. W. (2011). “The role of creaky voice in Cantonese tonal perception,” in *Proceedings of ICPhS 17*, 2240–2243.