Acoustic correlates of stress and their use in diagnosing syllable fusion in Tongan

James White & Marc Garellek
UCLA

Introduction

- Goals:
  - To determine the acoustic correlates of primary and secondary stress in Tongan.
  - To use these acoustic cues in determining whether syllable fusion occurs.

Background: Tongan basics

- Five vowels: /i, e, a, o, u/.
- Stress assignment (Churchward 1953; Feldman 1978) is the following:
  - Primary stress is moraic trochee at right edge.
  - Secondary stress depends on morphology, but in our words will be left-edge trochees.
  - E.g., *maemaeni* → [ˌmaema'eni] 'somewhat withered-DEM'

    E.g., *pēpeeni* → [ˌpeepe'eni] 'baby-DEM'

Background: Syllable fusion

- Vowel sequences (and "long" vowels) have recently been analyzed as disyllabic (Taumoefelau 2002; Anderson & Otsuka 2006):
  - *Sio* [ˈsi.o] 'to see'

- Certain vowel sequences, notably those for which the second vowel is higher than the first (e.g. *ai, ei*) may become one syllable, resulting in a diphthong. Identical vowels are also said to fuse into a long vowel (Churchward 1953; Feldman 1978; Poser 1985; Schütz 2001).
  - **E.g.,** *hū* → [ˈhuː] *[ˈhu.u] 'to go in',
    kai → [ˈkai] *[ˈka.i] 'to eat'

Phonological implications of syllable fusion: Why look at it?

- Poser (1985) claimed that syllable fusion is part of an ordering paradox with the definitive accent:

- Definitive accent (DA): addition of mora that is a copy of word-final vowel:
  - e.g., *ika* [ˈi.ka] 'fish' → *iká* [i.ˈkaː] 'fish+DA'.
- DA is phrasal in nature
  - Like English genitive *'s*, applies to whole NPs.
  - E.g., Ko e [me'alele kulokulá].

    PRED REF car red.DA

    'It's the red car.'

    *Ko e me'alelé kulokula (Anderson & Otsuka, 2006)

Phonological implications of syllable fusion: Why look at it?

- Syllable fusion must be lexical.
  - It cannot be post-lexical because it does not apply across all morpheme boundaries:
    - *Fakaafā* 'to bring a hurricane' /faka+afaa/ [fa.ka.a.ˈfaː], *[fa.kaː.ˈfaː]
    - Syllable Fusion = lexical, DA = phrasal, thus SF precedes DA
- BUT: DA must occur before syllable fusion:
  - *pō*+DA 'night (def.)' → [po.ˈoː], *[ˈpoː.o]
  - So, DA precedes SF

Poser's Paradox

Phonological implications of syllable fusion: Why look at it?

- More generally, this can be thought of a look-ahead problem involving stress:

- Certain vowel sequences make up a single syllable:
  - *hū* → [ˈhuː],
  - *hū+fi+a* → [huː.ˈfi.a],
  - *kai* → [ˈkai]

- But this syllabification is blocked just in case the second vowel of the sequence would carry primary stress ("breaking"):
  - *hū+fi* → [hu.ˈu.fi], *[ˈhuːfi]
  - *kai+ni* → [ka.ˈi.ni], *[ˈkai.ni]

- Syllabification needs to look ahead (anticipate stress placement) to determine whether to syllabify the vowels as a single syllable or as two syllables.

Phonological implications of syllable fusion: Why look at it?

- With Optimality Theory (Prince & Smolensky 1993/2004), there is no look-ahead problem:

	EDGEMOST	ONSET
☞ hu.(ú.fi)		*
(húu).fi	*!	

- But theories such as harmonic serialism (McCarthy 2008) also have to deal with the look-ahead problem.

Disagreement about syllable fusion

	ˈmai	maˈini	ˈmia	miˈani
Churchward (1953)	✗	✓	✗	✗
Feldman (1978) Poser (1985) Schütz (2001)	✓	✗	✗	✗
Taumoefolau (2002)	✗	✗	✗	✗

Does fusion occur?

- Tongan's role in the argument against serial grammars crucially relies on the syllable fusion process. But the data gathered so far have been impressionistic. There is also disagreement as to whether fusion occurs at all.

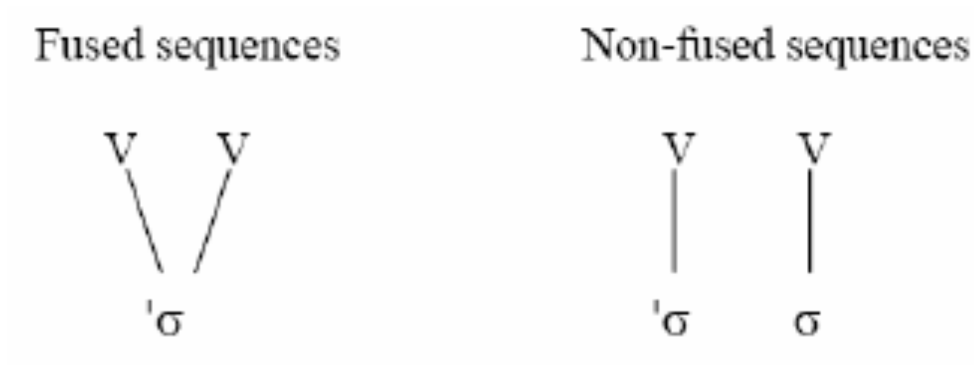
- AND no one says what fusion would even sound like or how you would otherwise know it has occurred.

- This study looks for empirical evidence for syllable fusion in Tongan.

- We will argue that "fusing" sequences have a different phonetic realization than non-fusing sequences, but there is little evidence for a phonological rule.

Stress and fusion

- If syllable fusion occurs, fused sequences and non-fused sequences should have a different syllabification. Stress should be associated with both vowels in a fused sequence:



Stress cues should be found throughout the sequence in fusing sequences.

Background: Stress correlates

- Common acoustic correlates of stressed syllables are greater pitch and intensity, longer duration, vowel quality differences (e.g., Gordon & Applebaum 2010).

- Greater spectral tilt (i.e., difference in voice quality) has also been found to be a cue to stress (Sluijter & van Heuven 1996).

- Correlates of secondary stress may differ from those of primary stress (e.g., Adisasmito-Smith & Cohn 1996).

Expt 1: Correlates of primary stress

- Method
  - Recorded 4 female native speakers of Tongan
  - CV'CVCV words where V1 = V2; compared V1 (unstressed) to V2 (stressed)
  - Carrier sentence: *Angimui ʻae foʻilea ko e ____ kiateau.* (“Repeat the word ____ for me.”)
  - 10 words used for each vowel, 3 tokens each, for a total of 1248 tokens
  - Tokens were labeled in Praat, and the acoustic measures were obtained using VoiceSauce (Shue et al 2009).
    - VoiceSauce available from http://www.ee.ucla.edu/~spapl/voicesauce/index.html

What do these measures indicate?

Measure	**Correlate of:**
F0	Pitch
Duration	Length
F1	Height
F2	Frontness
RMS Energy	Loudness
H1-H2	Voice quality
CPP	Periodicity

Statistical analysis

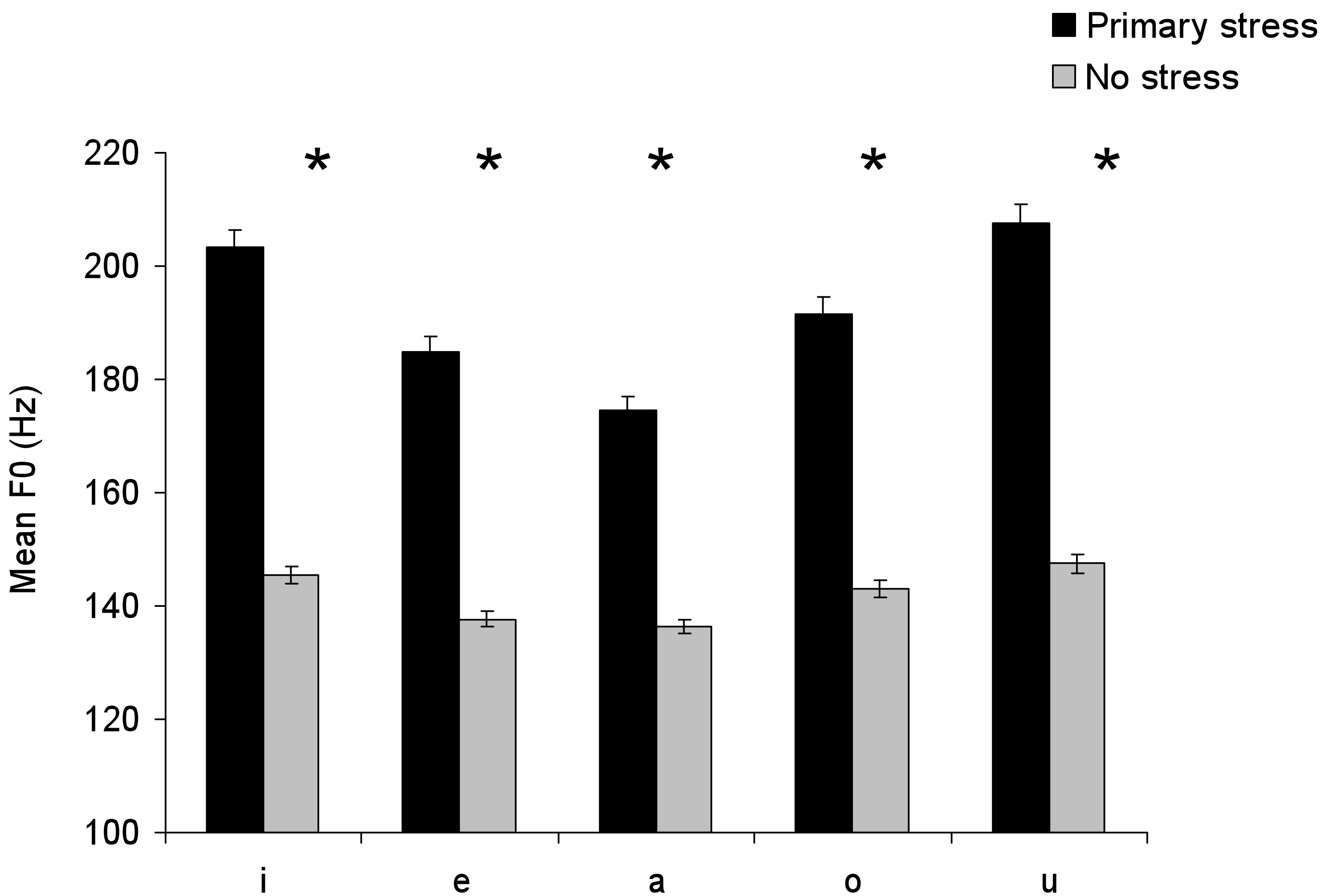
- We ran a linear mixed-effects model for each measure, with stress (primary vs. none) as a fixed effect and random effects for speaker, word, repetition and vowel.

- Post-hoc by-vowel analysis were run for the significant measures using the same LME model, but over each vowel individually.

Results for primary stress

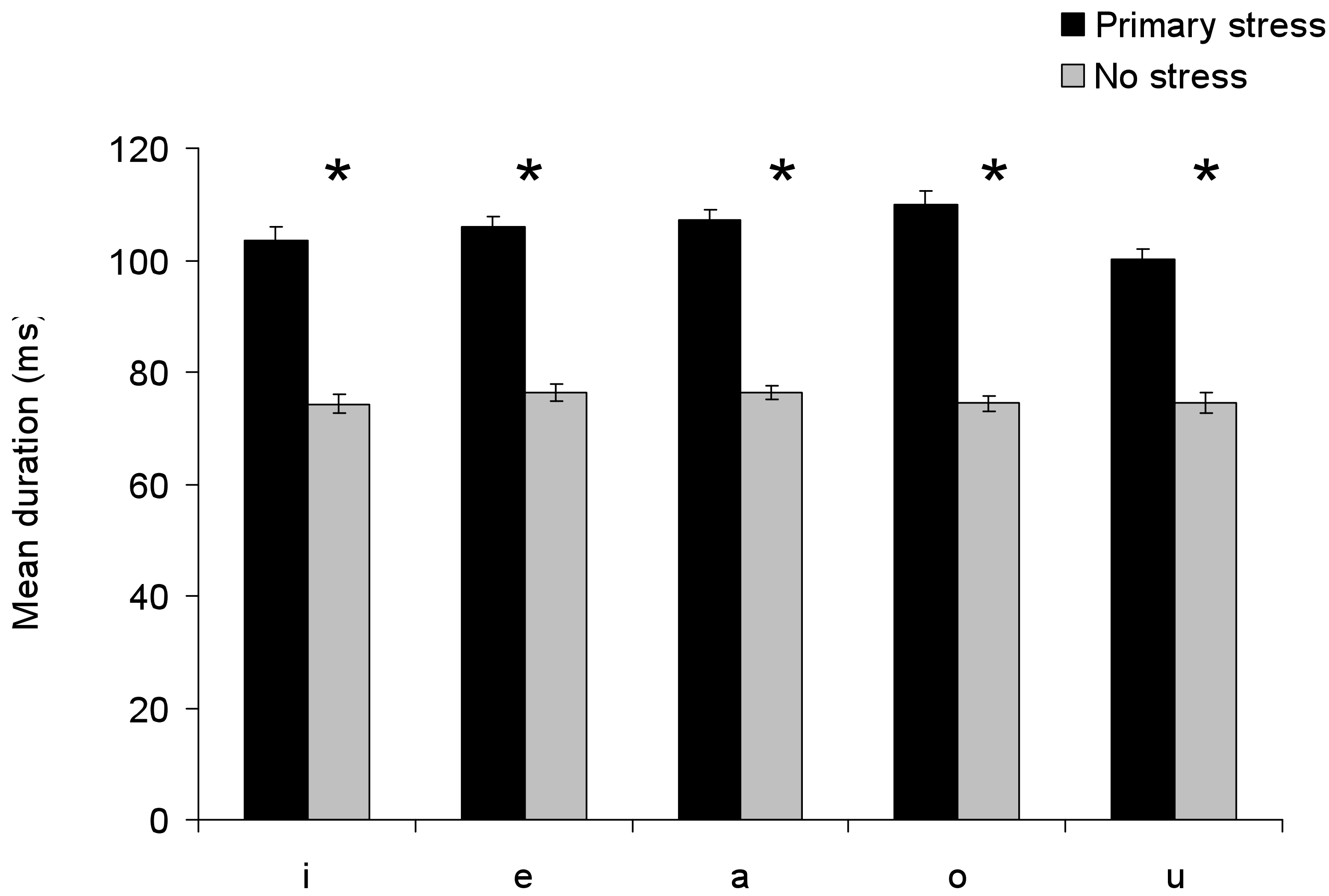
- All measures except for F2 show main effects for primary stress!

F0 in primary stress



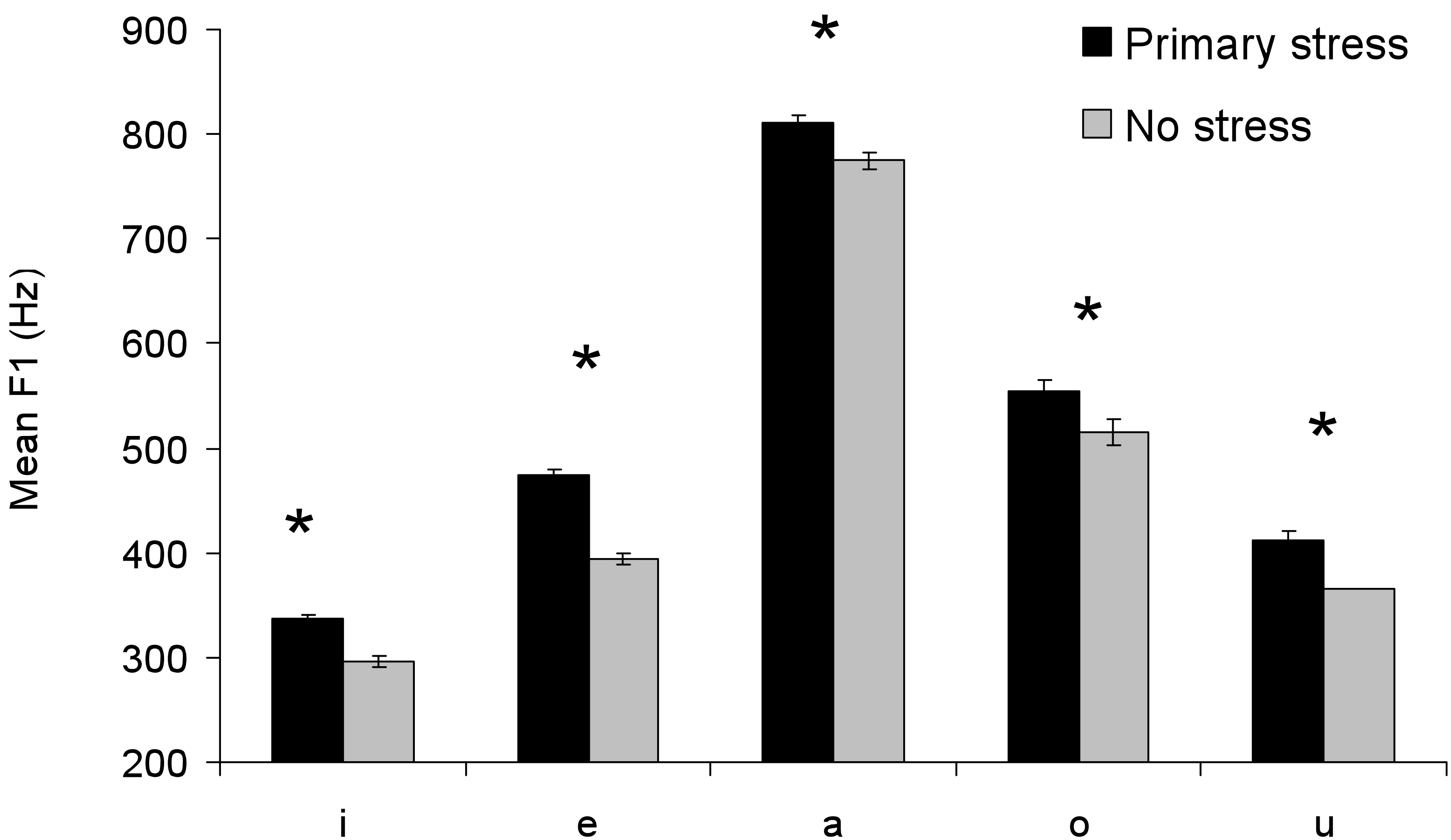
Vowels of all qualities have **higher** mean F0 when stressed by about 60 Hz, $p < 0.0001$.

Duration in primary stress



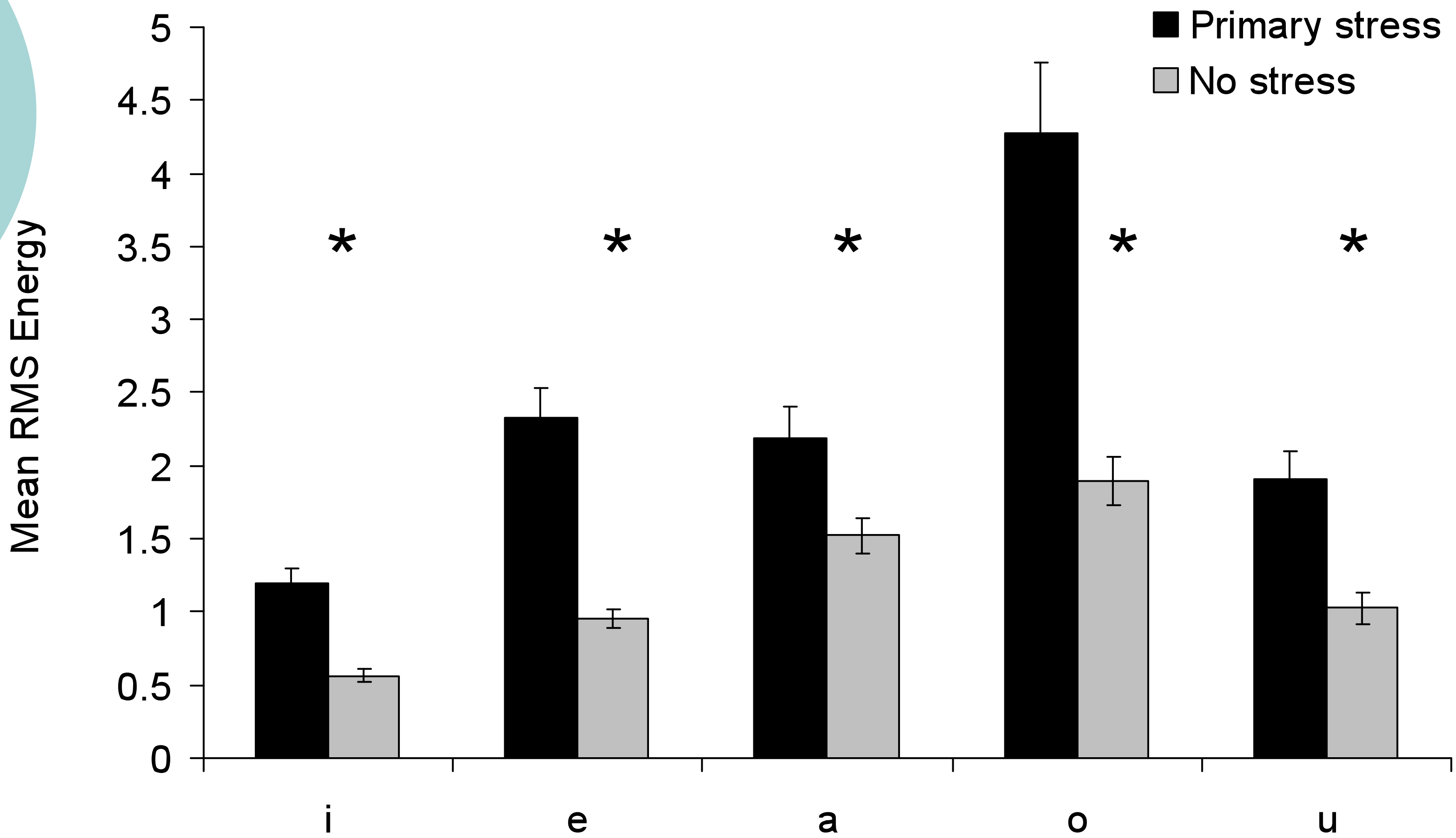
Vowels of all qualities are **longer** when stressed by around 27 ms, $p < 0.0001$.

Vowel quality in primary stress: F1



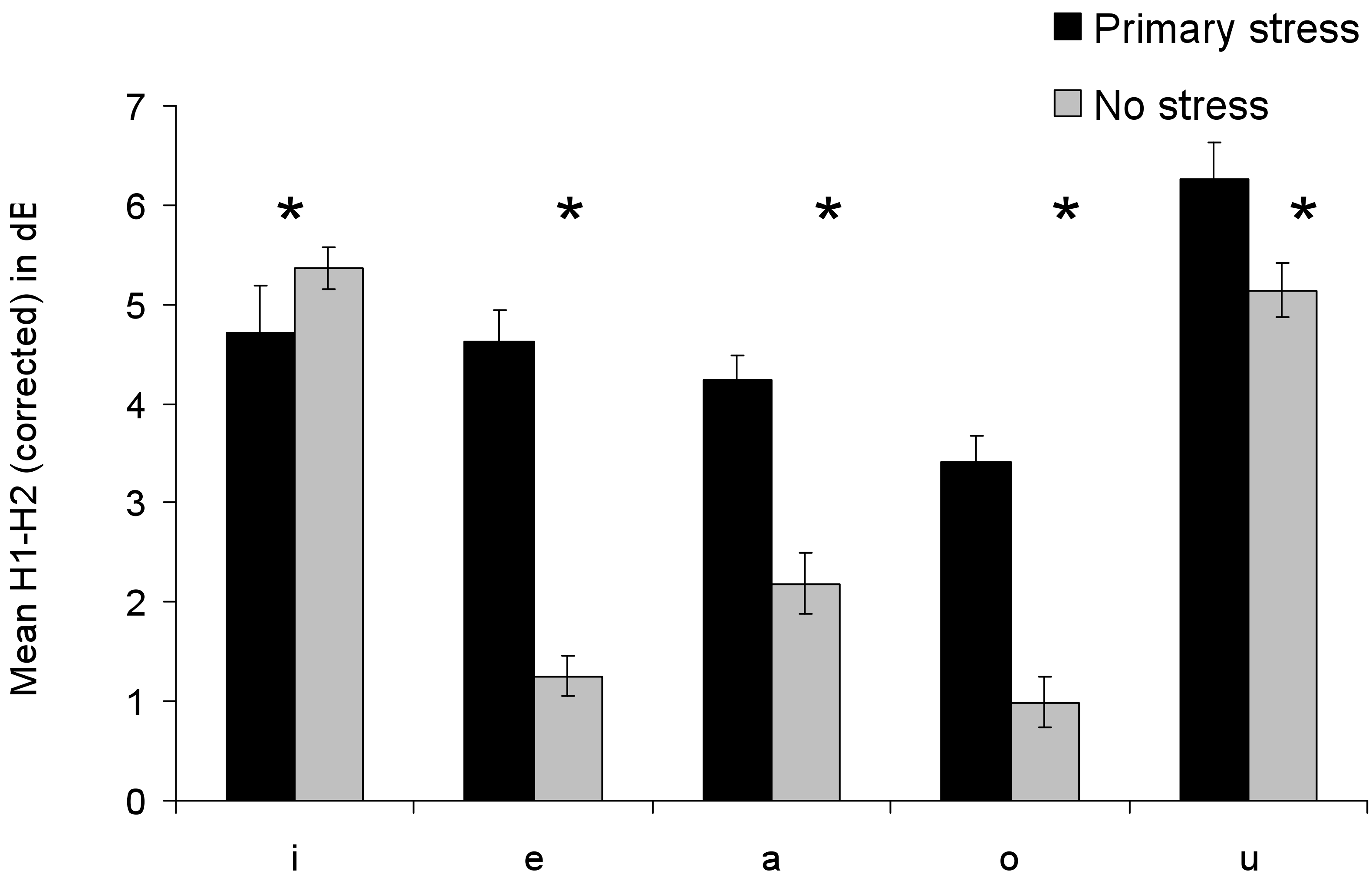
Vowels are **lower** in quality (higher F1) by about 55 Hz when stressed, p< 0.0001.

Energy in primary stress



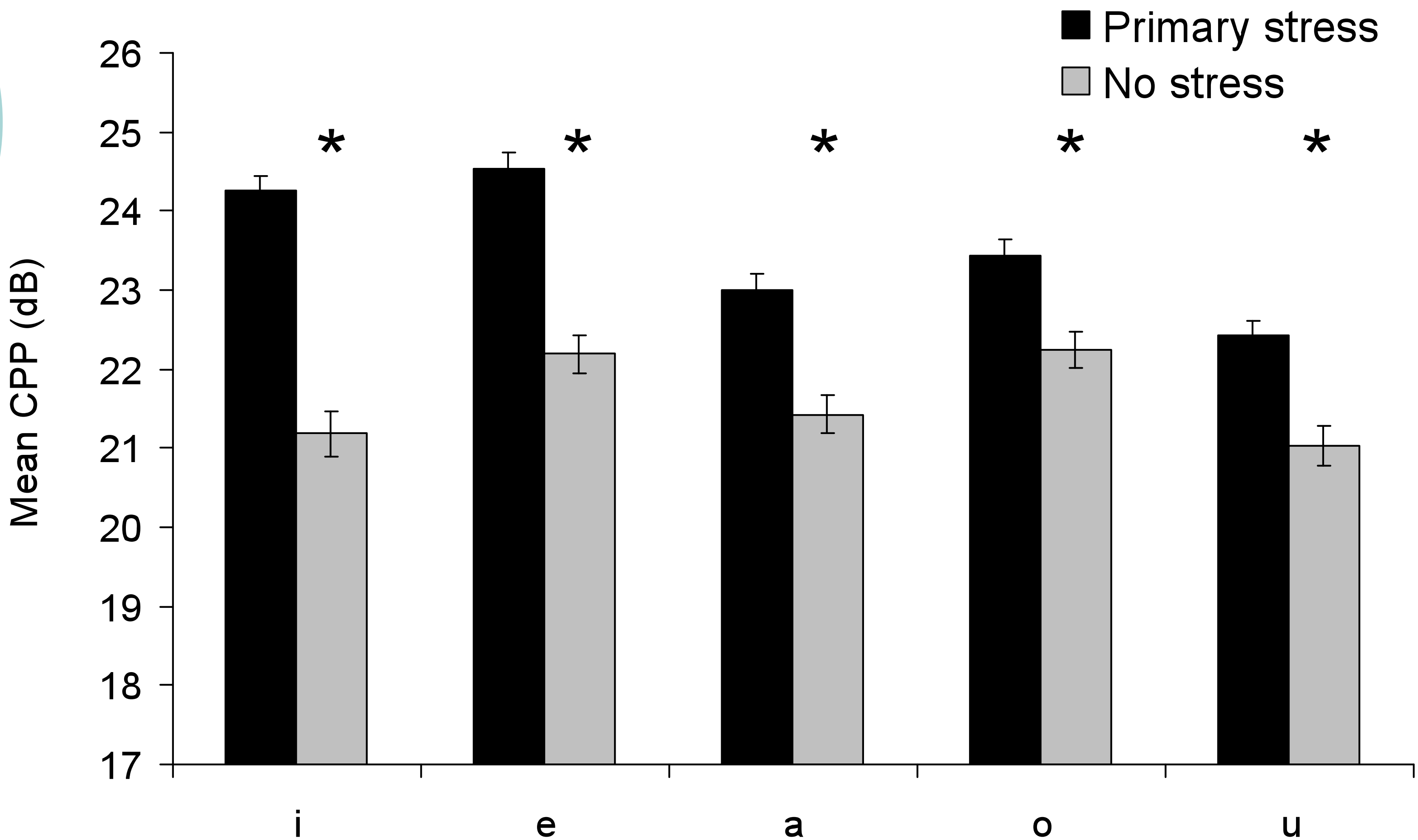
Vowels are louder when stressed by about 1.31u, $p < 0.001$.

Voice quality in primary stress



Stressed vowels have higher H1-H2 by about 2 dB, $p < 0.001$
(but cf. /i/)

Vowel periodicity in primary stress



Stressed vowels have higher CPP (more periodic) by about 2 dB, $p < 0.001$.

Summary of primary stress correlates

- Stressed vowels:
  - Are higher pitched
  - Are longer
  - Are lower
  - Are louder
  - Are more periodic

What are the best predictors of primary stress?

- Logistic mixed-effects model was run, with all the acoustic measures as fixed effects, and repetition, vowel quality, word, and speaker as random effects.

- F0 was best predictor ($p < 0.0001$), followed only by CPP ($p = 0.03$)

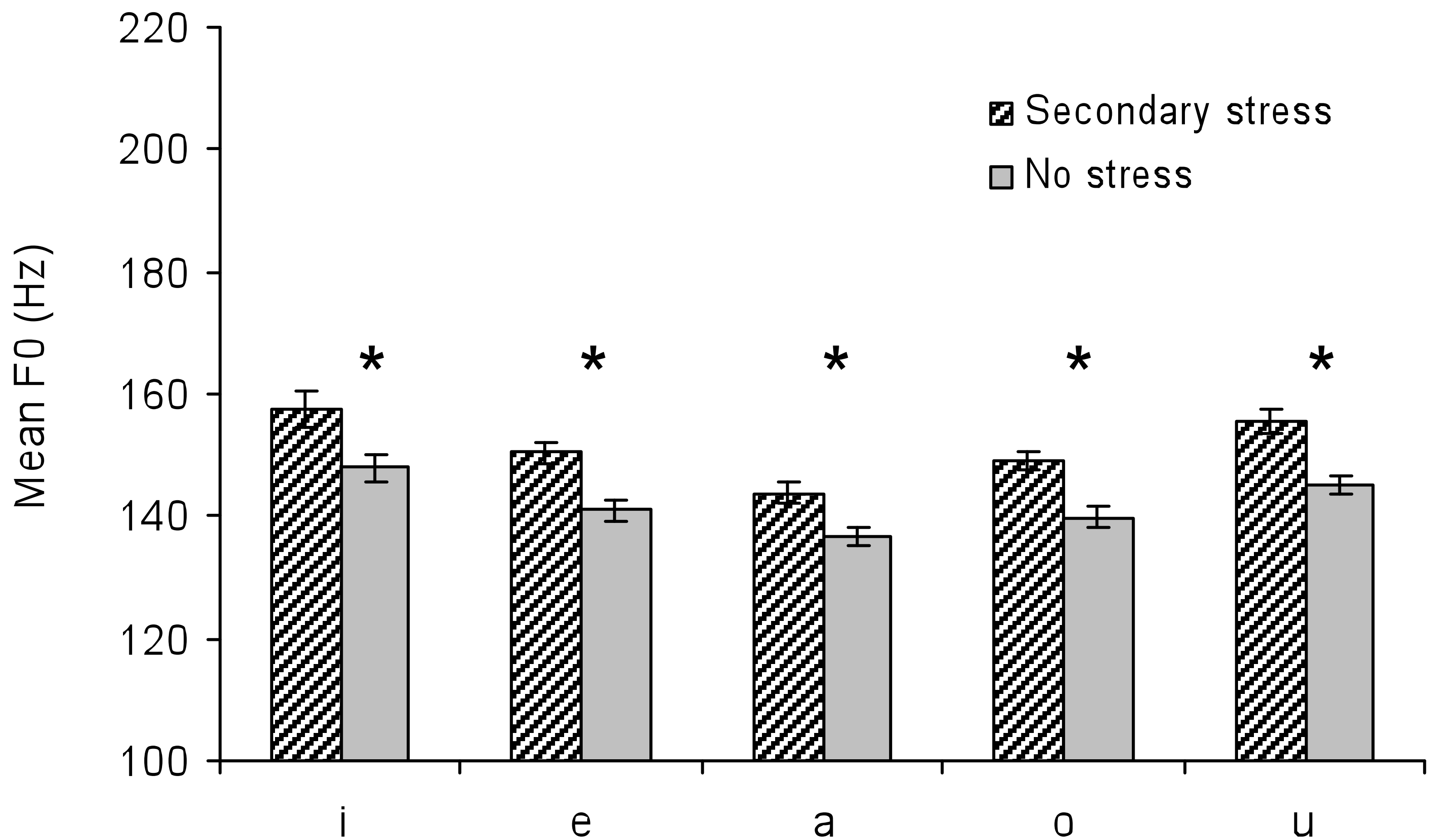
Expt 2: Correlates of secondary stress

- Method
  - Same four speakers
  - ˌCVCVˈCVCV words where V1 = V2; compared V1 (secondary-stressed) to V2 (unstressed).
  - Target words were same as for experiment 1 (primary stress correlates), but with a CV suffix, usually *–ni*, attached. These suffixed words form a prosodic word (Kuo & Vicenik 2010).
  - Same carrier as in Expt 1.
  - 10 words used for each vowel, 3 tokens each, for a total of 1326 tokens.
  - Same labeling and analysis method as in Expt 1.

Secondary stress results

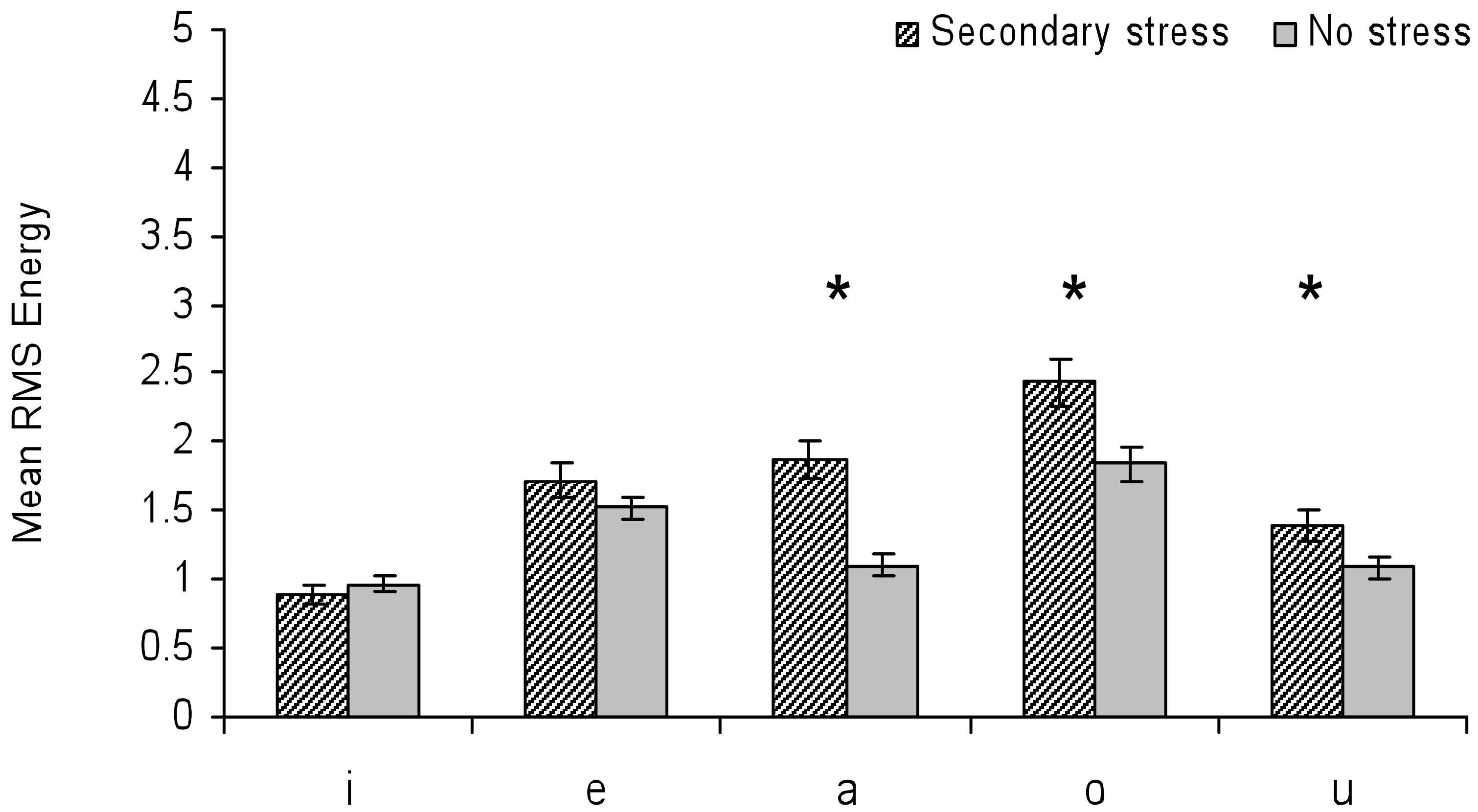
- Significant main effects for F0, energy, and duration.
- Duration was longer for unstressed vowels than for vowels with secondary stress
- **BUT** secondary-stressed vowels were always word-initial, so potential confound
  - However, in some languages, stressed vowels have been found to be shorter (Gordon & Applebaum 2010)

F0 in secondary stress



Vowels of all qualities have **higher** mean F0 when stressed by about 9 Hz, $p < 0.0001$.

Energy in secondary stress



/a, o, u/ are louder under secondary stress by about 0.2u, $p < 0.0001$.

What are the best predictors of secondary stress?

- Logistic mixed-effects model was run, with F0 and energy, and repetition, vowel quality, word, and speaker as random effects.

- F0 was best predictor ($p < 0.0001$), followed only by Energy ($p = 0.08$)

Primary/secondary correlates comparison

	Primary Stress	Secondary Stress
Pitch	Higher	Higher
Duration	Longer	Shorter
Vowel quality	Lower	
Intensity	Louder	Louder
Voice quality	More periodic	

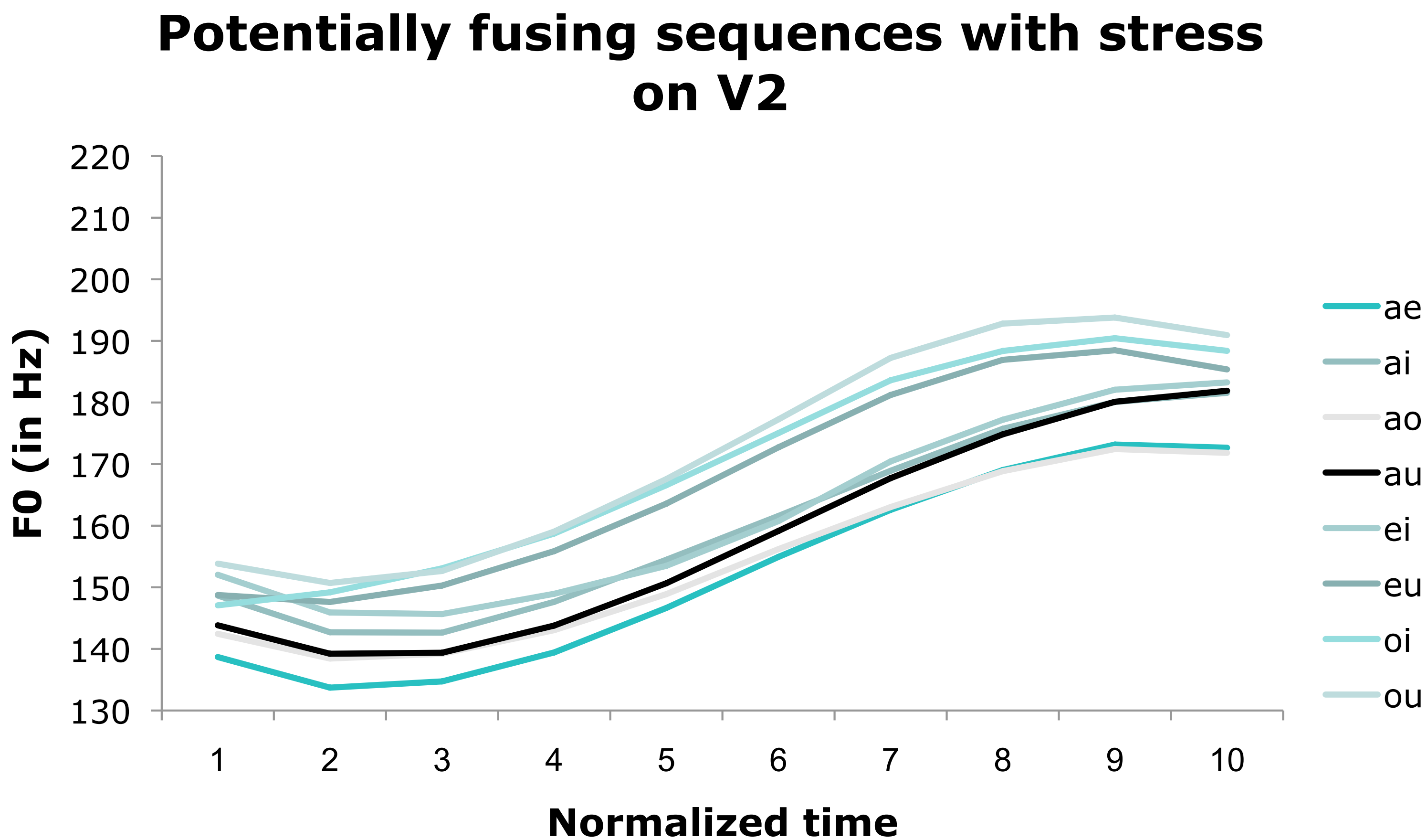
Expt 3: Syllable fusion

- Goal: To find acoustic evidence for/against syllable fusion
- If syllable fusion occurs, we expect lower-to-higher VV sequences should be realized differently than corresponding higher-to-lower sequences, e.g.:
  - Have a later pitch peak (corresponding to one stress assigned throughout) rather than a pitch target on the first vowel only
    - pitch peak in Tongan is at the end of the syllable (Kuo & Vicenik, 2010).
  - The second vowel of a fusing sequence should show cues of stress.

Expt 3: Syllable fusion

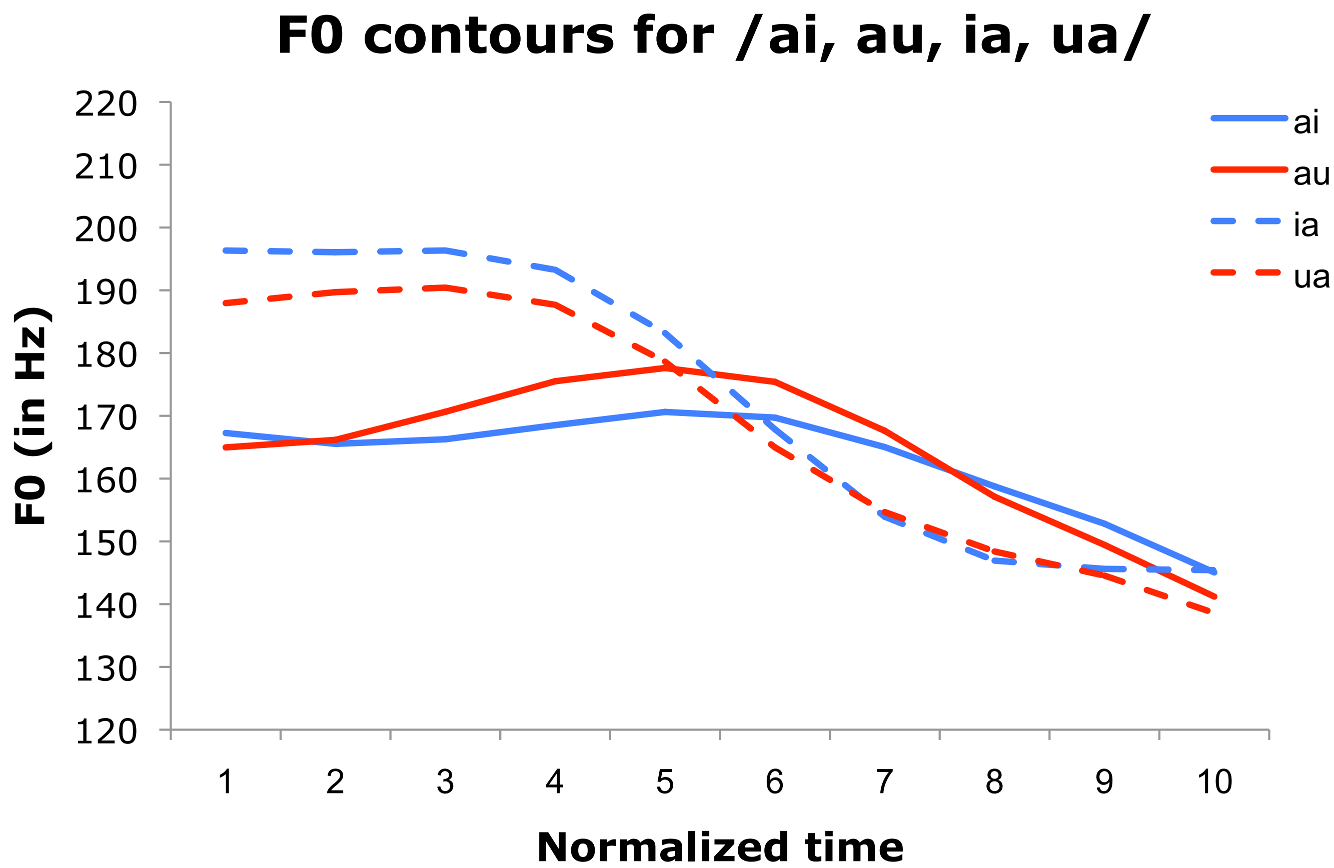
- Method
  - Data from the same speakers as in Expts 1 and 2.
  - For primary stress: 'CVV words compared with ˌCVVCV'VCV words
  - We compared falling sequences (e.g. *ai*) and corresponding rising ones (e.g. *ia*) across words.
  - Same carrier as before.
  - Sequences: *ai (ia), au (ua), ae (ea), ao (oa), ei (ie), eu (ue), oi (io), ou (uo)*
  - Approx. 3 words used for each sequence examined, 3 reps each, for a total of ~550 tokens.

What happens when primary stress falls on V2 of a sequence?



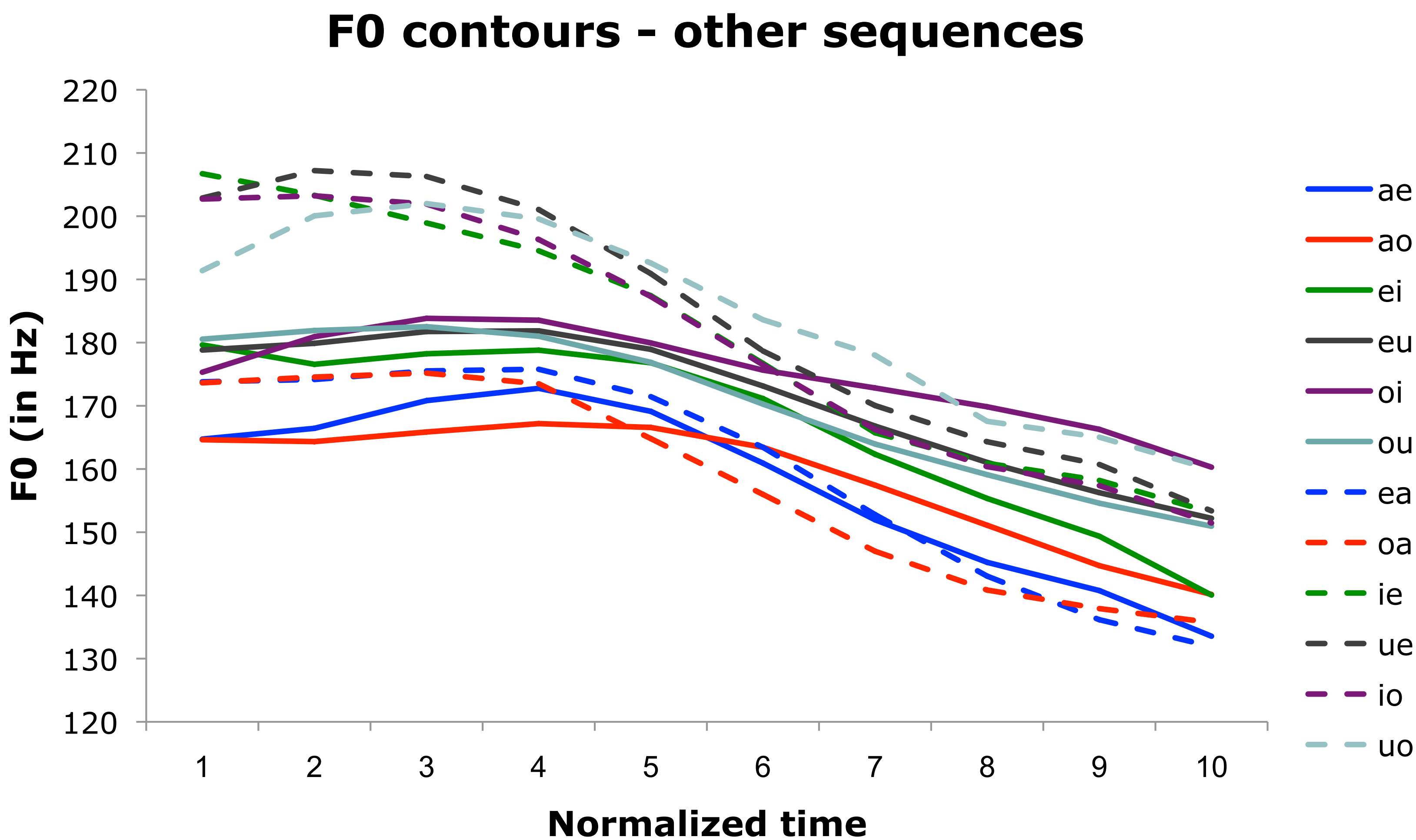
F0 peaks occur towards end of sequence → No syllable fusion

Pitch contours



- Later peak for /ai, au/ than for /ia, ua/

Pitch contours



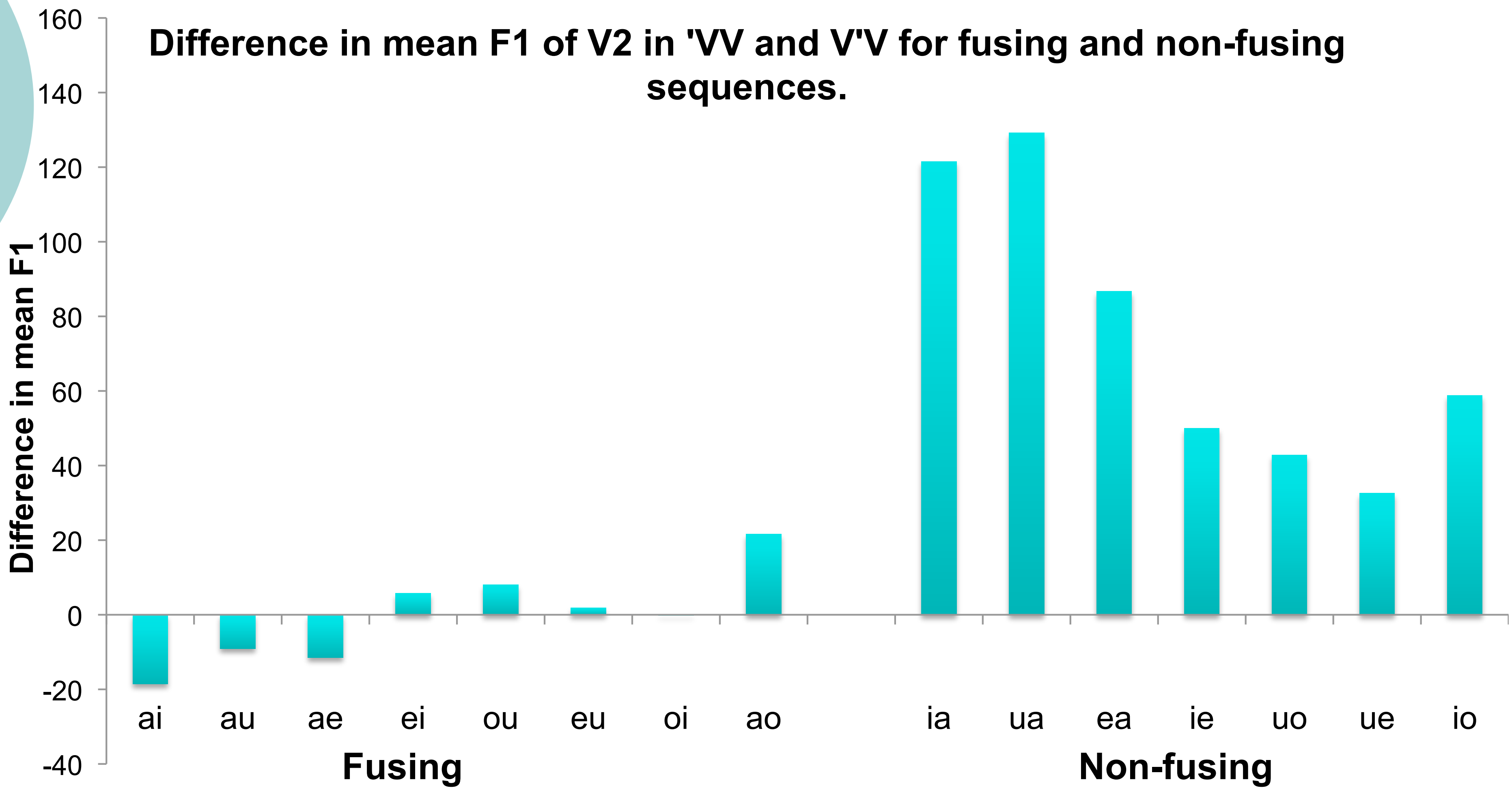
Other potentially fusing sequences do not seem to have later peaks.

ˈVV and VˈV comparisons

If fusion occurs, then there should be less difference in the vowels underlined in the left column than those in the right column:

	Fusing Sequence	Non-fusing Sequence
Stressed	[ˌkaikaˈini]	[ˌkiakiˈani]
Unstressed	[ˈkai]	[ˈkia]

'VV and V'V comparisons (F1)



V2 of fusing sequences looks like a stressed vowel, but not for non-fusing sequences.

Other differences?

- We did similar analyses (contours, V2 comparisons) for our other stress correlates
  - little to no difference between potentially fusing and non-fusing sequences

- Also, no difference in duration for, e.g., *ai* and *ia.*

Discussion

- Various scholars (Churchward, Feldman, Poser, Schütz, us) hear a difference between fusing and non-fusing sequences.
- What are we hearing?
- There are acoustic differences between fusing and non-fusing sequences:
  - F0 contours suggest yes (for /ai, au/)
  - Other measures (mainly F1) suggest all lower-to-higher sequences are different.

Discussion

- **SO,** there is phonetic evidence that fusing sequences are realized differently in Tongan.
- It remains unclear whether syllable fusion is robust enough to be an across-the-board phonological rule.
- Based on evidence so far, fusion seems more like a rule of phonetic implementation, occurring (perhaps optionally) when stress falls on the first V of a lower-to-higher sequence.

Discussion

- This makes a prediction: If fusion is only phonetic, then it is likely to occur more frequently at faster speech rates
  - a phonological rule should be independent of speech rate

- If fusion is actually phonological, ideally we'd want to find:
  - Phonological evidence
  - Experimental evidence that speakers treat those sequences differently

Conclusions

- Our study finds that Tongan stress is manifested by multiple acoustic cues
- Fusion results in phonetic differences in some sequences
- But there does not appear to be strong evidence for syllable fusion as a phonological rule in the language
- So – while syllable fusion has been regarded as a problem for serial-based theories of phonology (e.g., Poser's paradox), under this analysis, it is not a problem after all.

Mālō ‘aupito!

- Thanks to:
  - Our consultants
  - Kie Zuraw and Hilda Koopman
  - The 2010 UCLA field methods class participants
  - Audience at the UCLA phonology seminar

References

- Adisasmito-Smith, N. & Cohn, A. C. (1996). Phonetic correlates of primary and secondary stress in Indonesian: A preliminary study. *Working papers of the Cornell phonetics laboratory* 11*,* Cornell.
- Anderson, V. & Otsuka, Y. (2006). The phonetics and phonology of "Definitive Accent" in Tongan. *OL* 45(1)*,* 25-42.
- Churchward, C. M. (1953). *Tongan grammar*. Oxford: Oxford University Press.
- Feldman, H. (1978). Some notes on Tongan Phonology. *OL* 17, 133-139.
- Gordon, M. & Applebaum, A. (2010). Acoustic correlates of stress in Turkish Kabardian. *JIPA* 40, 35-58.
- Kuo, G. & Vicenik, C. (2010). Intonation of Tongan. Poster presented at the Acoustical Society of America, Cancun, MX.
- McCarthy, J. (2008). The gradual path to cluster simplification. *Phonology, 25,* 271-319.
- Poser, W. J. (1985). Cliticization to NP and Lexical Phonology. In J. Goldberg et al. (eds.), *Proceedings of West Coast Conference in Formal Linguistics* 4, Stanford Linguistics Association, CSLI, 262–272
- Schütz, A. J. (2001). Tongan accent. *OL* 40, 307-323.
- Shue, Y.-L., Keating, P. A., & Vicenik, C. 2009. VoiceSauce: A program for voice analysis. Poster presented at the Acoustical Society of America, San Antonio, TX.
- Sluijter, A. M. C. & van Heuven, V. J. (1996). Spectral balance as an acoustic correlate of linguistic stress. *JASA* 100(4), 2471-2485.
- Taumoefolau, M. (2002). Stress in Tongan. *MIT Working Papers in Linguistics, 44,* MIT.