Subglottal coupling as a quantal basis for the feature [back]

Morgan Sonderegger[1]   Xuemin Chi[2]

[1] Department of Computer Science, University of Chicago

[2] Speech Communication Group, RLE, MIT

ASA 156th Meeting
November 12, 2008

Outline

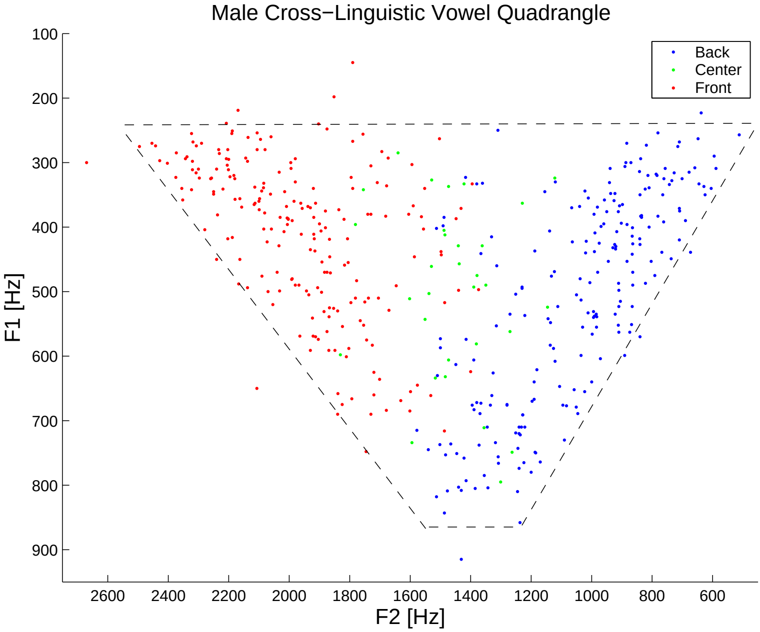
- Introduction
- Proposed quantal relation
- Acoustic model, predictions
- Diphthong data
- Monophthong data
- Discussion

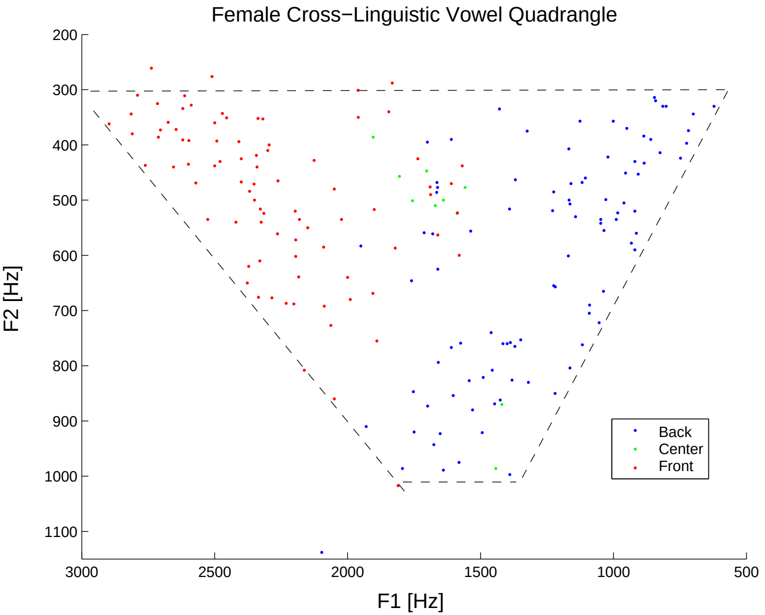
Introduction

- Broad question: what carves up vowel space?
- Quantal theory [Stevens, 1989] predicts phonetic features arise from acoustic-articulatory nonlinearities: quantal relations.
- How to test a proposed quantal relation?
- For a proposed QR for [feat]:
  1. Modeling of proposed QR.
  2. Production: Check whether QR occurs robustly in speech, matches predictions of model.
  3. Perception
- **Proposed QR for [back]**: Oral/subglottal coupling effects near SubF2, predicted interaction with breathiness.

Motivation: [back] realization

- From phonology, vowel backness has $\approx$ 3 natural classes: front ([-back]), back ([+back]), central (unspecified).
- How clear a divide phonetically, cross-linguistically?
- Published F1/F2 data from 46 languages (male), 19 languages (female), $\geq$ 3 speakers/study.

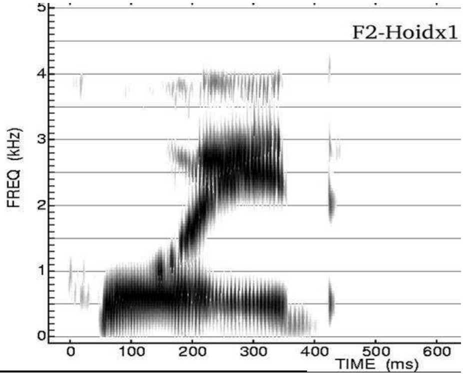

Male Cross−Linguistic Vowel Quadrangle
Back
Center
Front
F1 [Hz]
100
200
300
400
500
600
700
800
900
F2 [Hz]
2600
2400
2200
2000
1800
1600
1400
1200
1000
800
600


Female Cross−Linguistic Vowel Quadrangle
200
300
400
500
600
700
800
900
1000
1100
F2 [Hz]
3000
2500
2000
1500
1000
500
F1 [Hz]
Back
Center
Front

- Fairly sharp front/back division
- Dividing line at (depends on fitting method) 1385 Hz (male), 1592 Hz (female).
- Why there, and why sharp?

Proposed QR for [back]

- [Stevens, 1998]: Front/back divide near second subglottal resonance (SubF2)?
- SubF2: M 1408 ± 129 Hz, F 1566 ± 121 Hz.[1]
- Why? Oral/subglottal coupling non-negligible when F2 near SubF2. Sometimes see A2 attenuation, F2 discontinuity near SubF2.
- Clearest in back/front diphthongs:

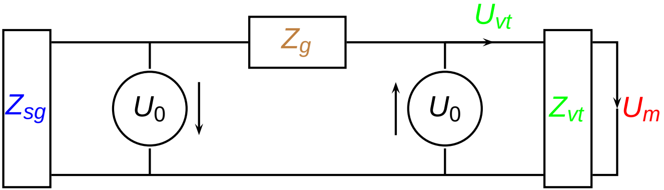


---

[1] From [Lulich 2006] survey

Acoustic model I

- Equivalent circuit model for oral/subglottal interaction:



- Approximate Z_{sg} as lossy tube, linear $Z_g = R_g + j\omega M_g$:

$$R_g = 12\mu \frac{h_g}{l_g {d_g}^3} + K\rho U_g \frac{1}{(l_g d_g)^2}, \quad M_g = \rho \frac{h_g}{l_g d_g}$$

- $|Z_g|$ inversely proportional to glottal area ($l_g d_g$).

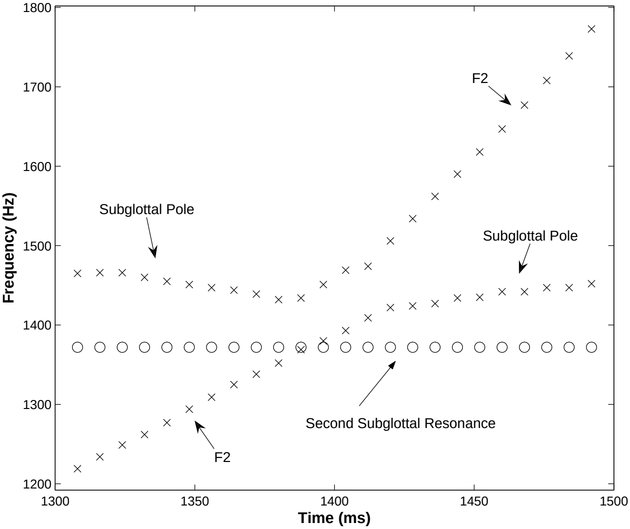
Acoustic Model II

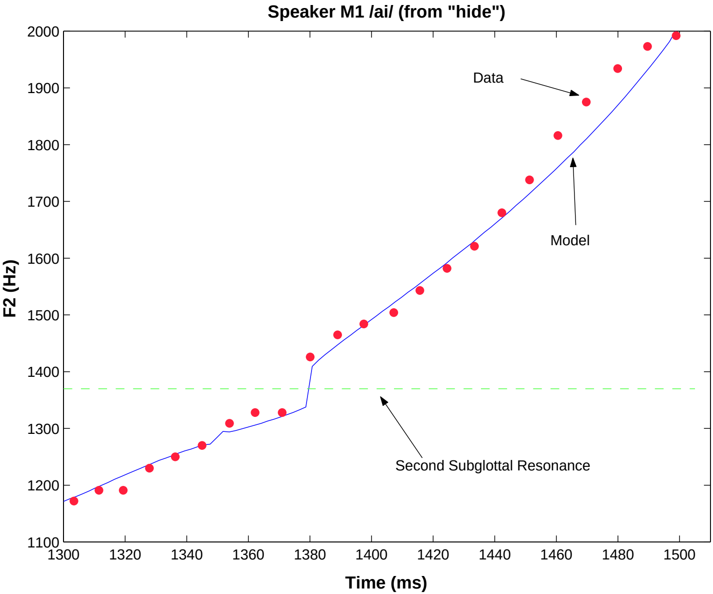
- Transfer function is $\frac{U_m}{U_{vt}} \frac{U_{vt}}{U_0}$, input to VT is

$$\frac{U_{vt}}{U_0} = \frac{Z_g}{Z_g + Z_{vt} + Z_{sg}}$$

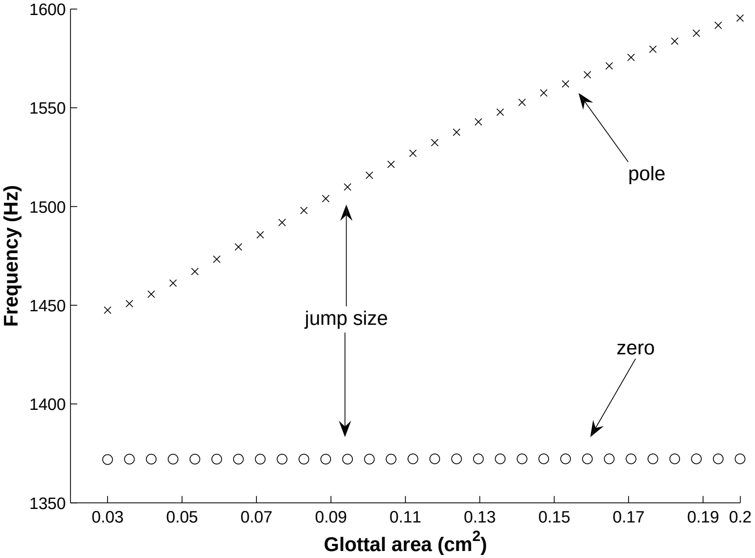
- $= 1$ for large $|Z_g|$ (no coupling).
- With coupling, extra pole-zero pair in transfer function, zero at max Z_{sg} (i.e. SubF2).
- $|Z_g|$ inversely prop to glottal area $\implies$ coupling $\propto$ glottal area.

Predicts jumps in F2 near SubF2, size $\approx$ new pole/zero separation. Back-front diphthong:

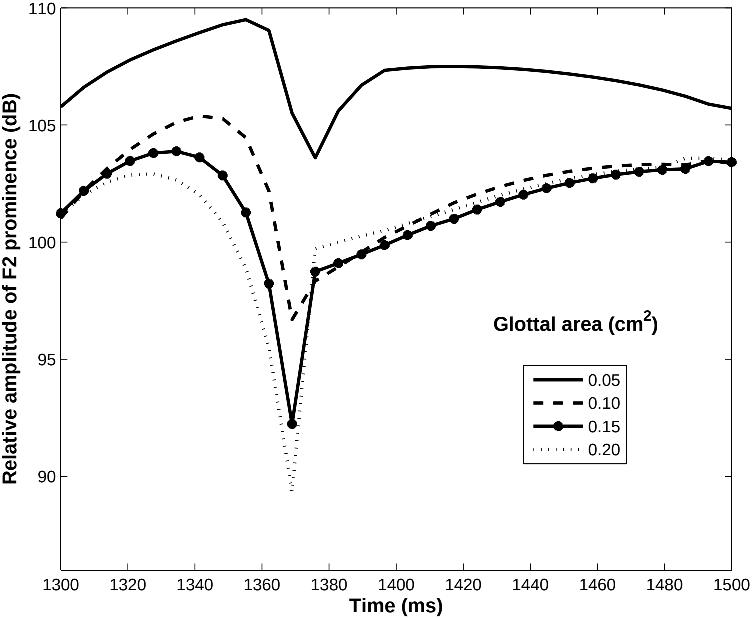



Speaker M1 /ai/ (from "hide")
F2 (Hz)
2000
1900
1800
1700
1600
1500
1400
1300
1200
1100
1300
1320
1340
1360
1380
1400
1420
1440
1460
1480
1500
Time (ms)
Data
Model
Second Subglottal Resonance

Predicts F2 jump size grows with glottal area:



Predicts A2 attenuation near SubF2 grows with glottal area:

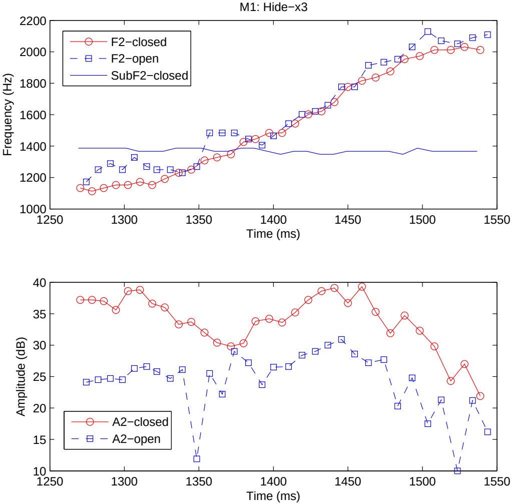


Production data: diphthongs

- Model predicts F2 jumps, A2 attenuation occur near SubF2
- To check if occur, examine back-front diphthongs.
- 3 male, 3 female American English speakers recorded "hVd, say hVd again." for V=/oi/, /ai/, x5.
- Subglottal pressure measured using small external accelerometer attached above sternal notch [Cheyne 2002].
- Small number of tokens because all measurements by hand – formant trackers can miss F2 jump.

Open vs. closed phase

- Open-phase and closed-phase F2 and A2 tracks taken.
- Coupling effects occur in closed-phase tracks, but more robustly in open-phase (expected: larger glottal area):



Diphthong results II

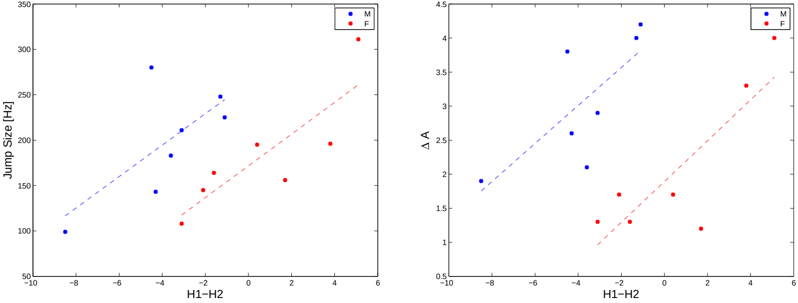
- Without controlling for open vs. closed phase (i.e. as in spectrogram), measuring by pitch period, $\Delta A2$ always occurs, jumps usually.
- Open-phase measurements, average over 10 diphthongs/speaker:

Speaker	Jump in $F2$(Hz)	Falling/rising $\Delta A2$ at jump(dB)
M1	299	-8.3, +7.5
M2	348	-8.0, +6.5
M3	237	-5.6, +6.5
F1	294	-8.5, +9.4
F2	252	-8.4, +6.2
F3	201	-6.2, +8.3

- But: no correlation between $\Delta A2$ and F2 jump size.
- Coupling effects occur robustly, $\Delta A2$ more robust.

Breathiness $\propto$ coupling effects?

- Over larger set of speakers (7 M, 7 F): breathiness correlated with coupling effect size?
- $F2$ jump, $\Delta A2$ calculated for 3–4 /oi/ tokens/speaker, averaged.
- H1-H2 ($\approx$ breathiness [Hanson, 1997]) measured in /a/ productions (high F1), averaged.



- Across speakers:

	Jump/H1-H2	$\Delta A2$/H1-H2
M	r=0.69	r=0.74
F	r=0.84	r=0.83
All	r=0.49	r=0.24

($p < 0.05$, $p < 0.1$, $p > 0.25$)

- Coupling effect size pos. corr. with speaker breathiness.

[+back]/[-back] monophthongs

- Coupling effects in back-front dipthongs near SubF2 robust $\Longrightarrow$ SubF2 could be used as front/back divide: is it?
- SubF2-F2 for vowels should be $<$ 0 for front, $>$ 0 for back.
- Examined 10 V tokens of each English monophthong for 14 speakers, same carrier sentences.
- Call each set of 10 “vowel group”: e.g. SubF2-F2 for 10 productions of /uw/ by speaker M1.
- Do [+back] groups and [-back] groups have expected signs?

Results

- All front vowel groups have signif. negative ($p < 0.05$) SubF2-F2.
- 83% back vowel groups signif. positive, 8% n.s. positive.
- But:
    - Coarticulation increases F2 (/hVd/ environment)
    - Most of the n.s. groups are “who’d” (fronting for many AE speakers), “hud” (back or central?).
- Still, indicates for AE vowels, SubF2 as [back] dividing line plausible.

Discussion

- Have shown modeling, production steps to testing proposed QR for back.
- Coupling effects occur robustly (attentuation $>$ jumps), effect size correlated with breathiness.
- Last step: show used in perception [Lulich 2006, Lulich et al. 2007].
- Quantal theory predicts front/back boundary and that F2 region near boundary unstable: possible explanation for rarity of (phonological) central vowels?
- Implications for theories of vowel space.

- Details: Sonderegger 2004, Chi & Sonderegger 2007
- Thanks to Ken Stevens, Steven Lulich, Speech Group!

References

- Cheyne, H. (2002). *Estimating glottal voicing source characteristics by measuring and modeling the acceleration of the skin on the neck*. Ph.D. thesis, MIT.
- Chi, X. and Sonderegger, M. (2007). Subglottal coupling and its influence on vowel formants. *JASA*, 122:1735–1745.
- Hanson, H. (1997). Glottal characteristics of female speakers: Acoustic correlates. *JASA*, 101: 466-481.
- Lulich, S. (2006). *The Role of Lower Airway Resonances in Defining Vowel Feature Contrasts*. Ph.D. thesis, MIT.
- Lulich, S.M., Bachrach, A., and Malyska, N. (2007). A role for the second subglottal resonance in lexical access. *JASA*, 122:2320-2327.
- Sonderegger, M. (2004). Subglottal coupling and vowel space. B.S. thesis, MIT.
- Stevens, K. (1989). On the quantal nature of speech. *J. Phonetics*, 17(1):3–45.
- Stevens, K. (1998) *Acoustic Phonetics*. MIT Press, Cambridge, MA.