3aSC30

ACOUSTIC MEASURES OF FALSETTO VOICE

Patricia A. Keating
Phonetics Lab
Department of Linguistics
UCLA

May 7, 2014 Providence

Introduction

Falsetto voice is a vocal register above modal voice, in which vibration of stiff vocal fold edges produces a higher-pitched source with a steeper spectral slope.

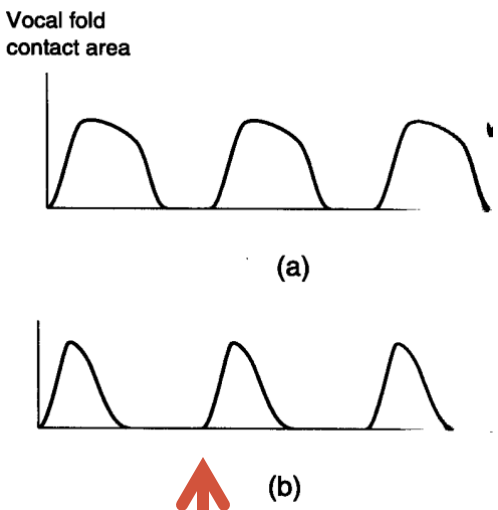
Falsetto voice is common in singing, and in speech it has many social meanings, either expressive or stereotypical, including excitement, deference, mocking, and quotation (e.g. Podesva 2007, Stross 2013). Here falsetto is used in reading a story, to enact female characters.

Physiology of falsetto

- Crico-thyroid and interarytenoids are active, thyroartenoids release
- Vocal fold cover/ligament is stretched thin, decreasing its cross-section area, and increasing its stiffness and frequency of vibration
- Vibrational amplitude and sound intensity are low, but Open Quotient is large
- See figures on next 2 slides

(Titze 2000, van den Berg 1968)

From Titze (2000)



EGG waveform shapes, modal (top) vs. falsetto (bottom)

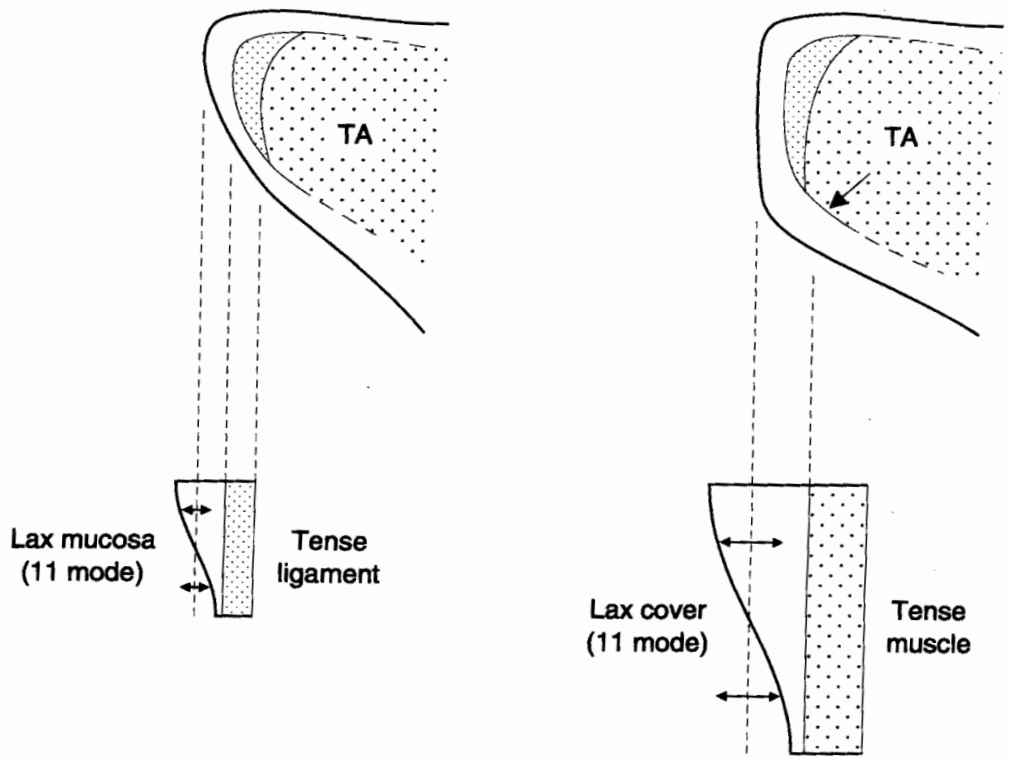
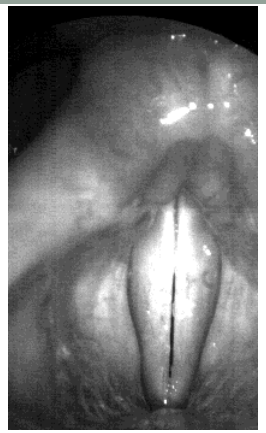


FIGURE 10.5. Vocal fold configuration in coronal section for falsetto register and modal register. In the transition from falsetto to modal register, the thyroarytenoid (TA) muscle contracts, which bulges out the inferior portion of the fold. Vibration in the mucosa and ligament changes to vibration in the cover and body.

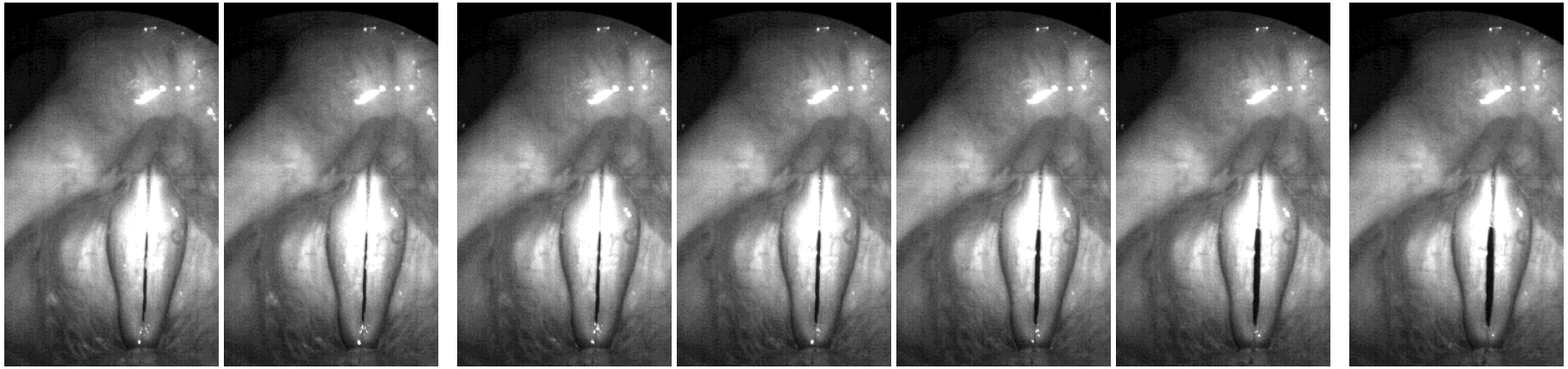
Vocal fold configurations in falsetto (left) vs. modal (right)

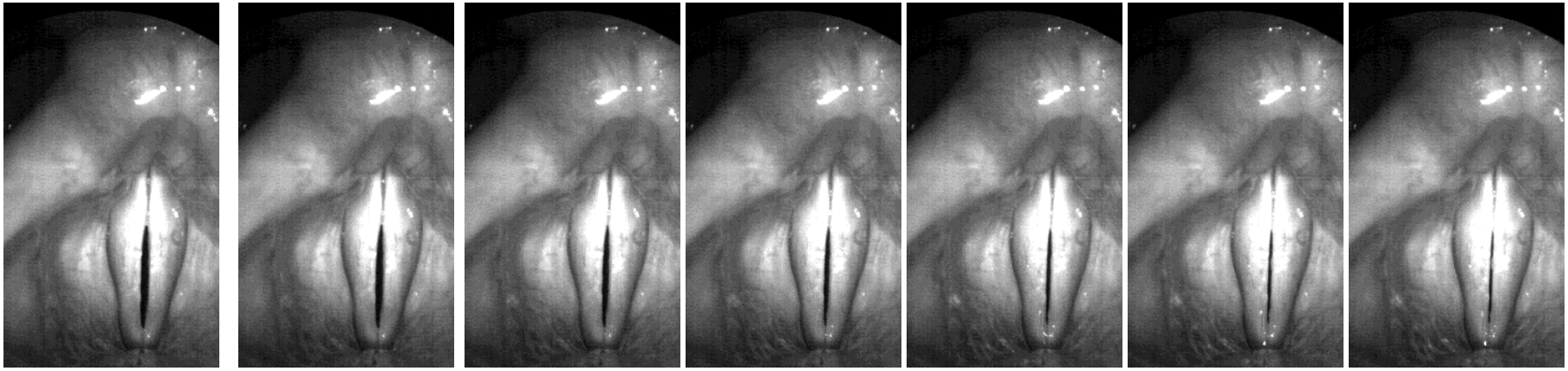


Falsetto glottal cycle

F0 = 800 Hz, frame rate = 10k Hz

(see poster 3aSC2, Chen et al.)

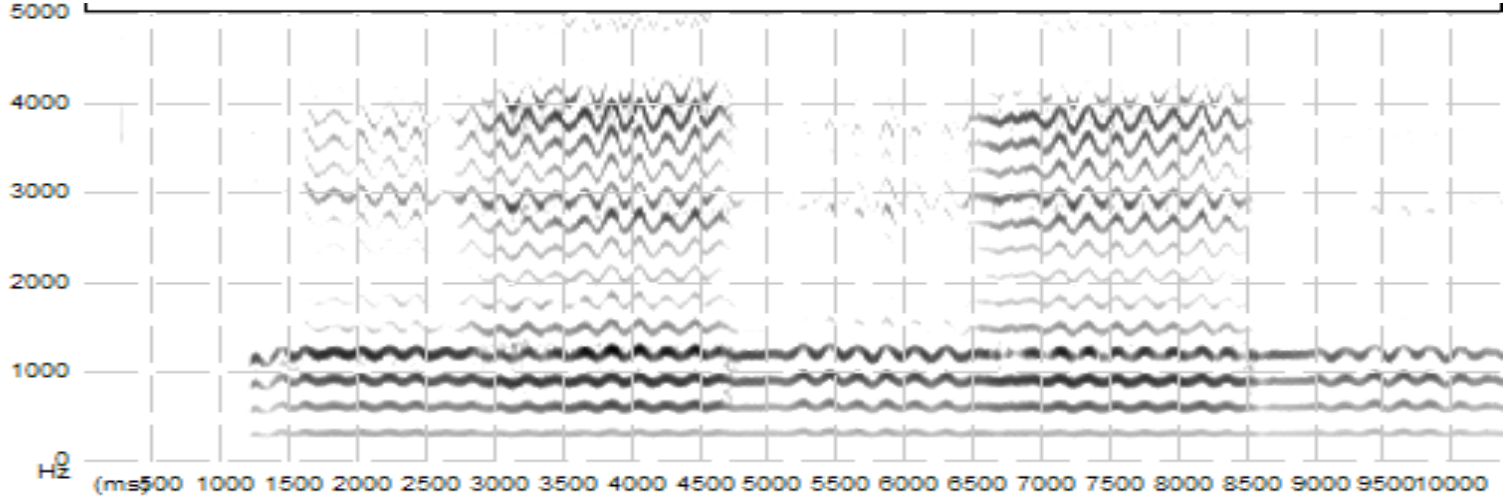




F0 of falsetto

- F0 is typically higher in falsetto than in modal
  - Men's voice break is around 275 Hz, women's around 450 Hz
- But F0 ranges can overlap, with F0s around the break-point possible in either voice register

Singer toggling between head and chest voice (Nair 1999):

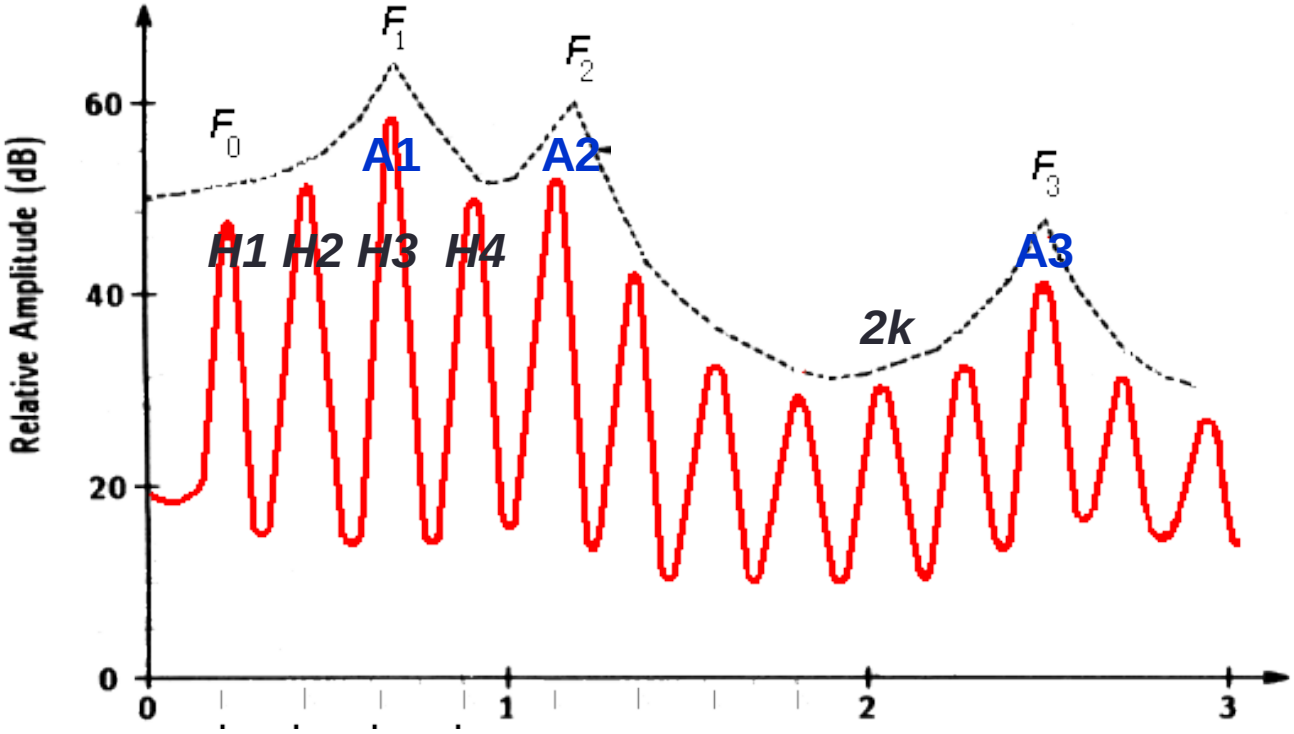


Acoustics of falsetto

- Falsetto shows more energy in the fundamental, less energy in higher harmonics
- Spectral tilt of about 20 dB/octave, including first few harmonics (related to loudness)
- At matched F0, in falsetto, H1 < H2, while in modal H1 > H2
- These are perceptual cues for voice register

(Colton 1973; Hammarberg et al. 1980; Keidar et al. 1987; Neiman et al. 1997; Sundberg 1973; Hanson 1997; Sundberg and Högset 2001)

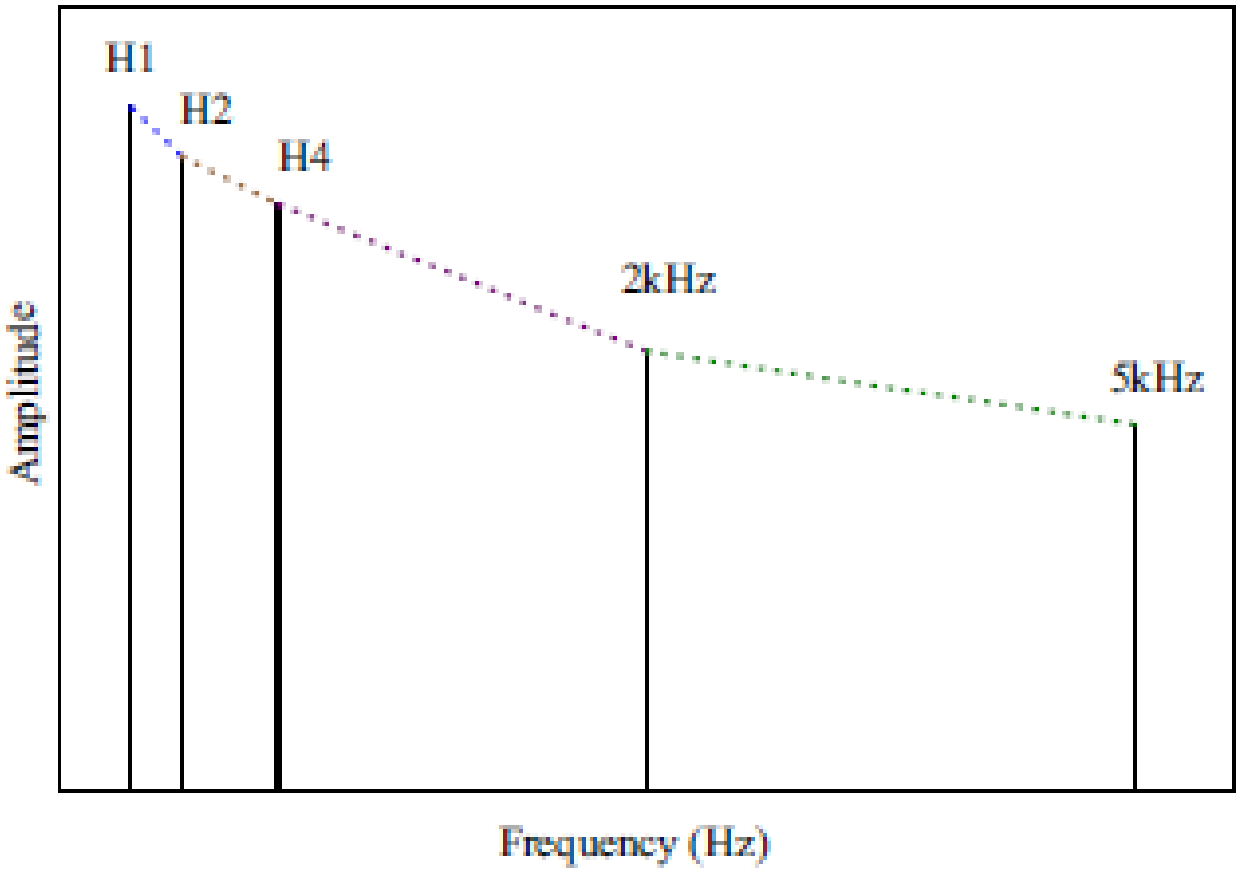
Harmonics in spectrum



Research question

- When a larger set of acoustic measures of the speech spectrum is considered, which ones distinguish modal from falsetto voice?
- Especially, do higher-frequency components (above H1-H2) play a role, as in the Kreiman et al. 4-component model of the voice source spectrum?

Kreiman et al. model of source spectrum (from Garellek et al. 2013)



Methods in present study

- Speakers
  - 11 UCLA students/staff (5 F, 6 M)
- Speech materials
  - Little Red Riding Hood story
  - Read twice: first neutrally, then acting out the character voices (matched samples)
- Recordings made as part of the UCLA project “Production and perception of linguistic voice quality”, freely available online (http://www.phonetics.ucla.edu/voiceproject/voice.html)

Falsetto labeling

- In the character-voice readings, clear falsetto intervals were identified by listening and labeled in Praat textgrids
- Corresponding segments in neutral readings were found and labeled; creaky-voice tokens were excluded
- Utterances with non-nasalized low-vowel tokens with matched falsetto and modal readings were extracted
- Total corpus: 30 pairs of files (matched falsetto and modal tokens)

Vowel tokens

1-6 different words per speaker:

- Grandmother: “Pull the string and the latch will come up.”
- Red Riding Hood:
  - "I am going to my grandmother‘s to take her some nice blackberry wine, for she is quite sick."
  - "Just outside the wood. You can see her cottage through the trees.“
  - "If you please, grandmother, mother has sent me with some blackberry wine.“
  - "And, grandmother, what large ears you have.“

Acoustic analysis

- Middle-fifth of each vowel token analyzed
- VoiceSauce (Shue 2010)
  - F0 from STRAIGHT (max set to 900 Hz)
  - Formants from Praat
- H1*, H2*, H4*, 2k* (harmonic nearest 2kHz)

* Indicates formant-corrected harmonic amplitudes

- H1*-H2*, H2*-H4*, H1*-H4*, H4*-2k*
- Cepstral Peak Prominence, Harmonic-Noise Ratios over several frequency bands; Subharmonic-Harmonic Ratio
- Significant differences (from *lmer*) shown in slides

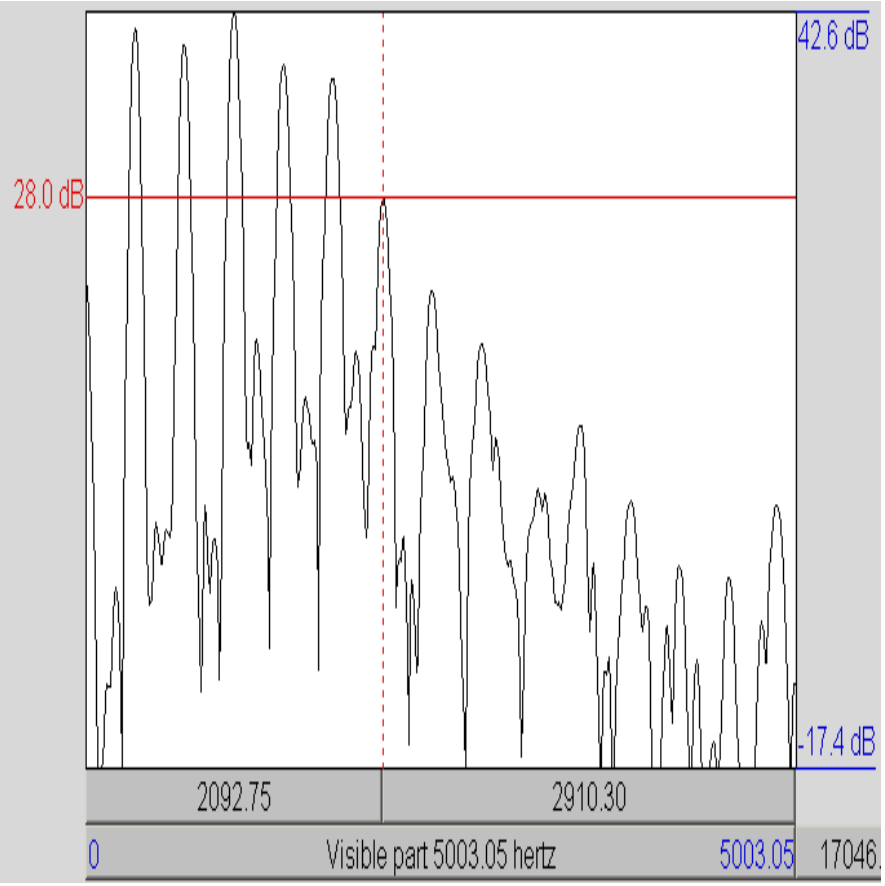
See Shue et al. 2011 for more on parameters

Sample spectra, /æ/ in *latch*

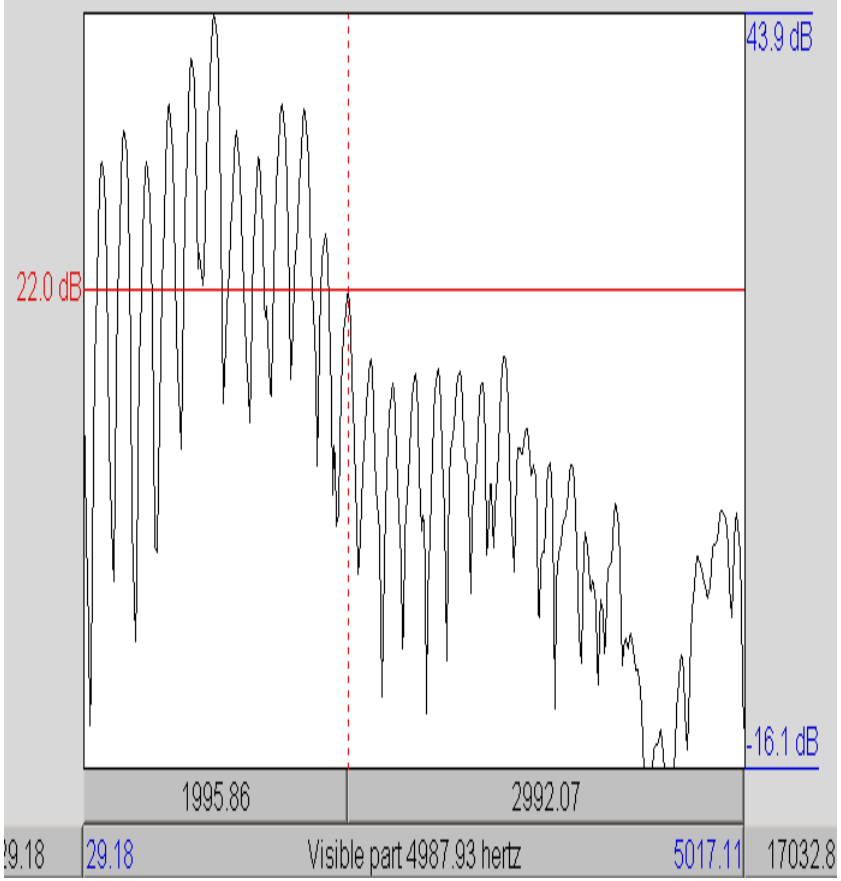
25-year-old female speaker F1 ~1000 Hz

left = imitating Grandma (F0 ~340 Hz), right = neutral (F0 ~170 Hz)

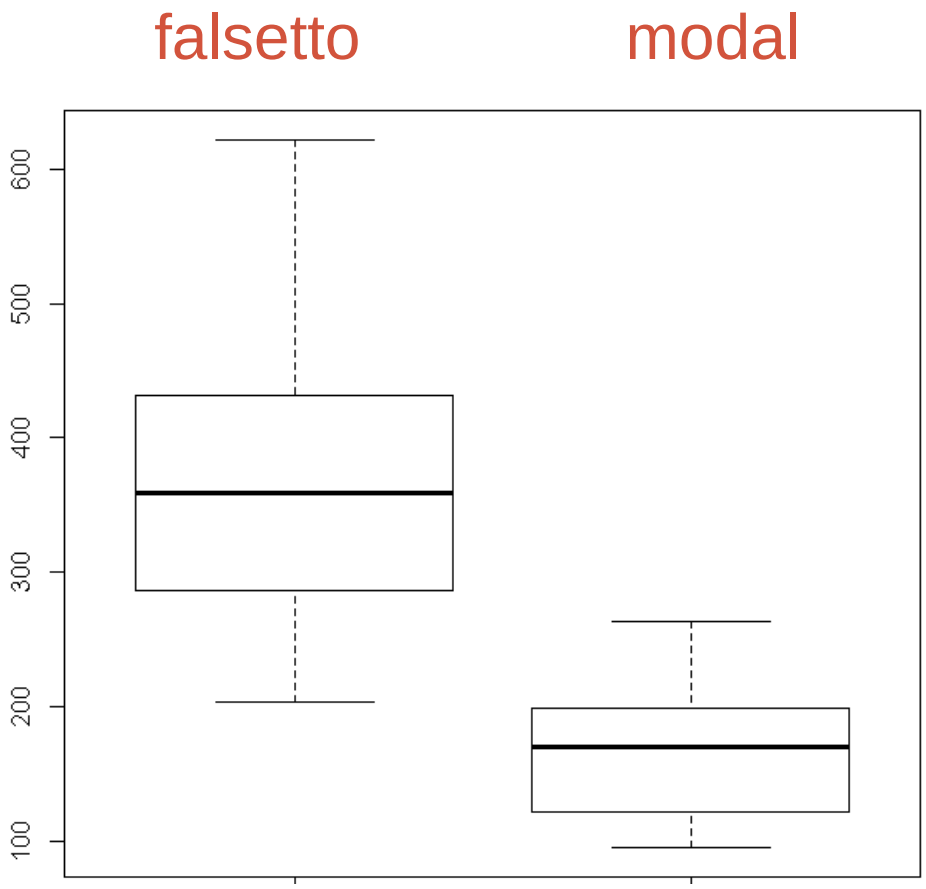
falsetto



modal

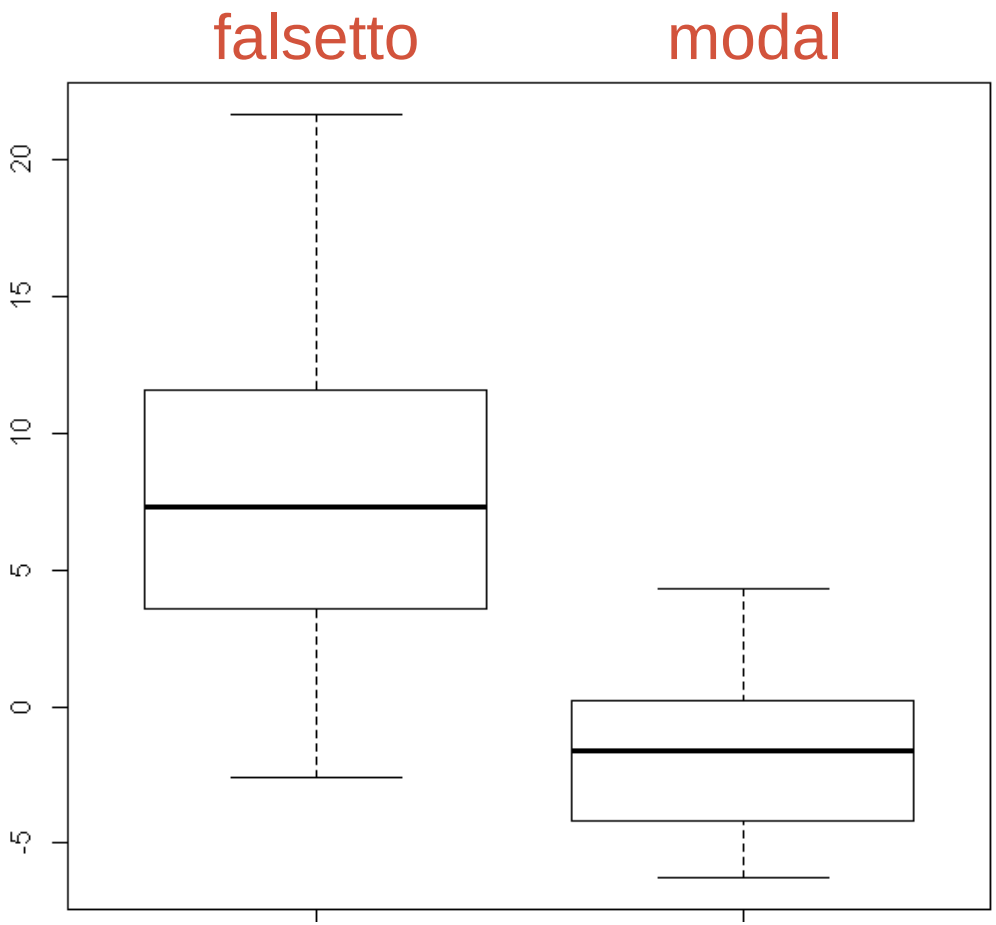


Results: Mid-point F0s



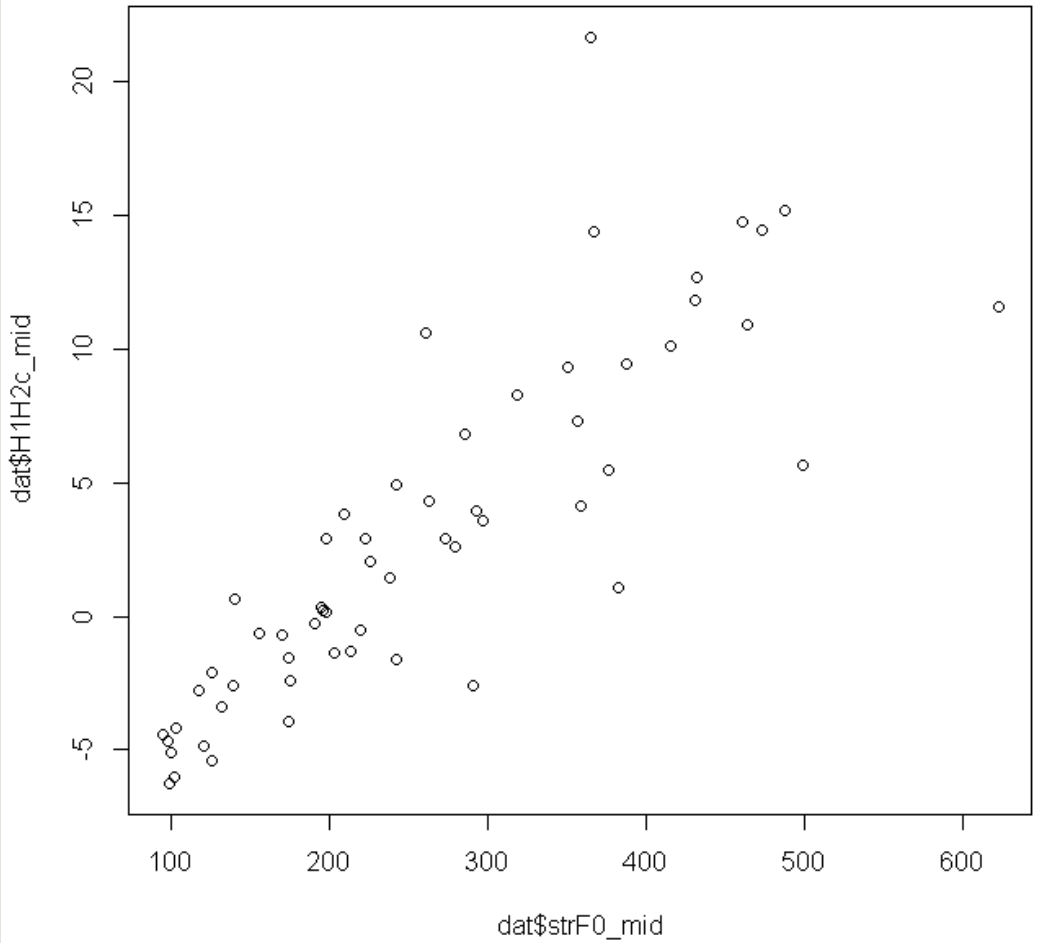
Because F0 differs between the registers, it was included as a regressor in statistical analyses of other voice measures: do these other measures differ by voice register above and beyond the fact that F0 differs?

Results: H1*-H2*



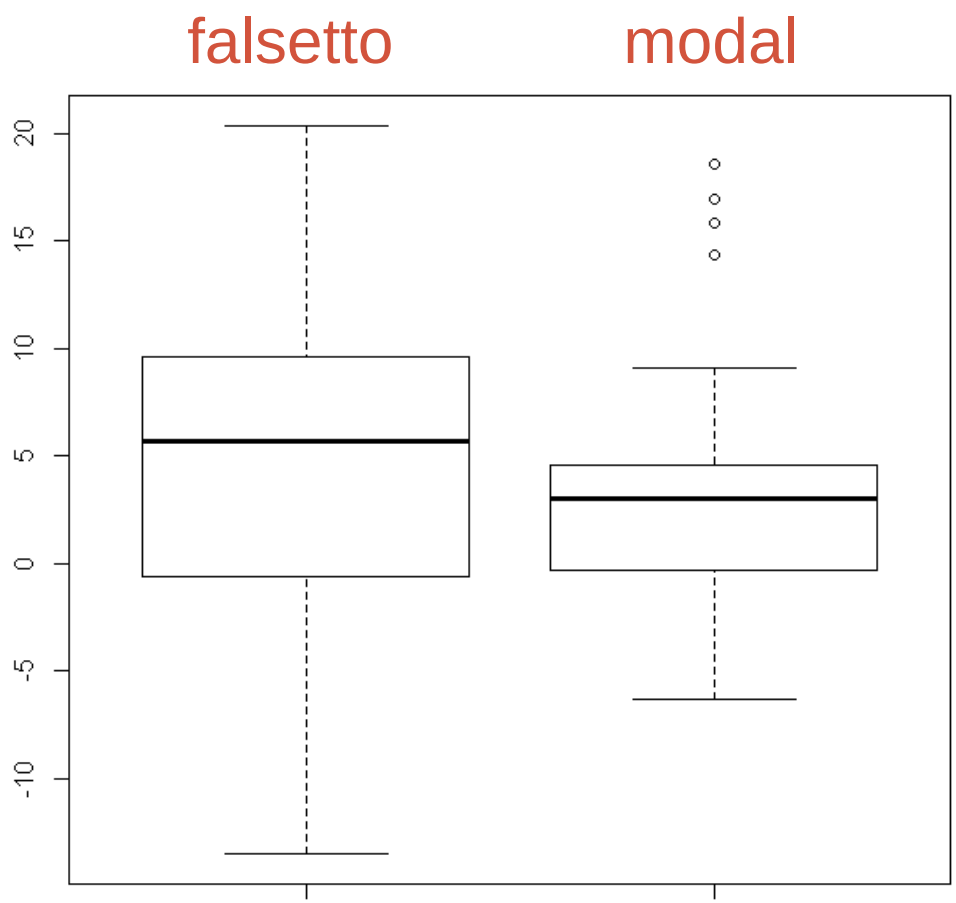
This difference is significant (p<.01) only when F0 is not taken into account; the difference is due to the strong correlation between H1*-H2* and F0 (next slide); gender effects on all voice measures also disappear given F0

H1*-H2* as a function of F0



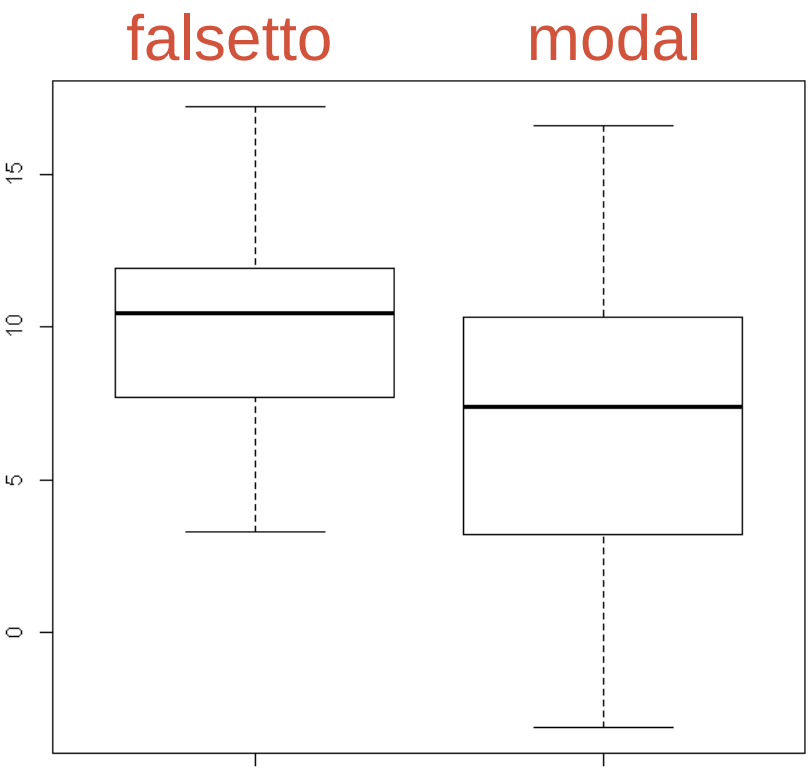
This linear relation is very different from the inverted-v-shaped relation found by Iseli et al. (2007) for normal speaking F0s

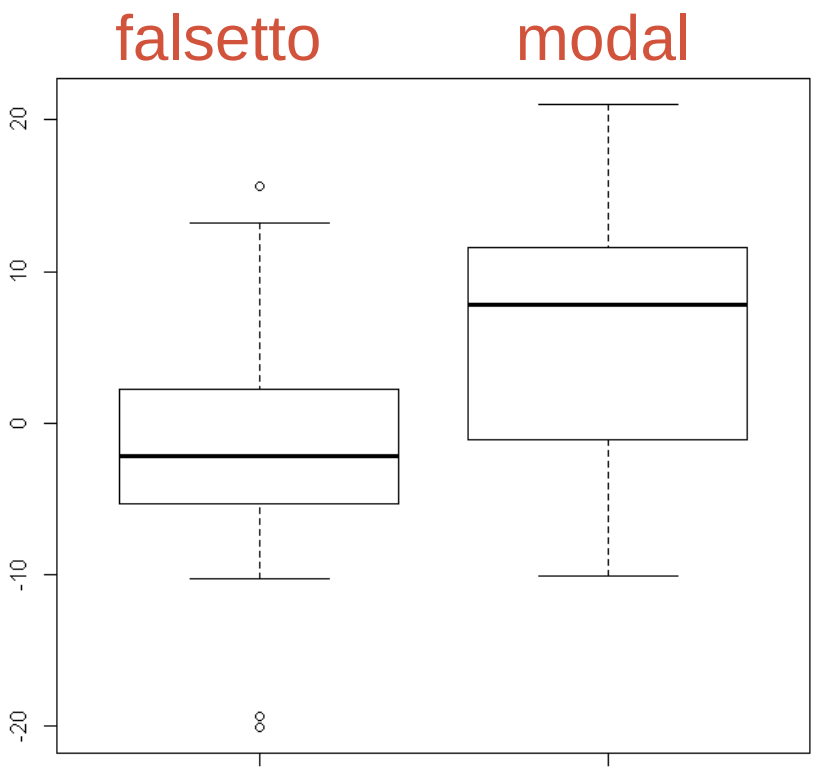
Results: H2*-H4*



This difference is significant *only* when F0 is taken into account (given the F0, H2*-H4* differs between the registers); H2*-H4* is only weakly (and negatively) correlated with F0; is mainly due to H4* (next slide)

Results: H1*, H4*

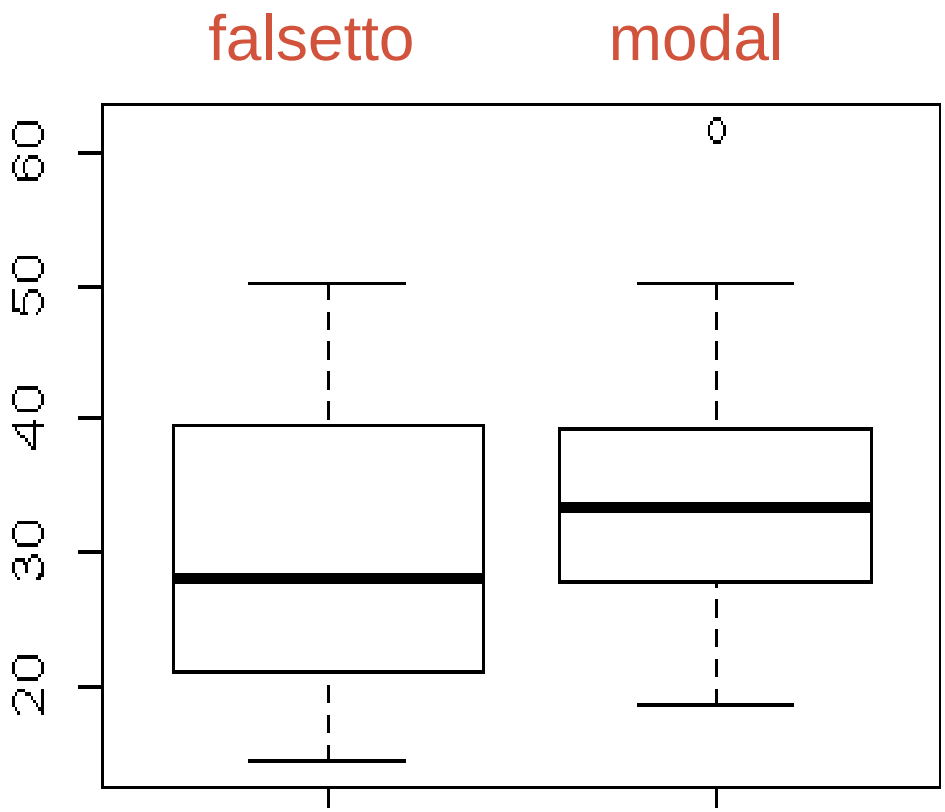




BUT: H2* does not differ significantly between the 2 registers

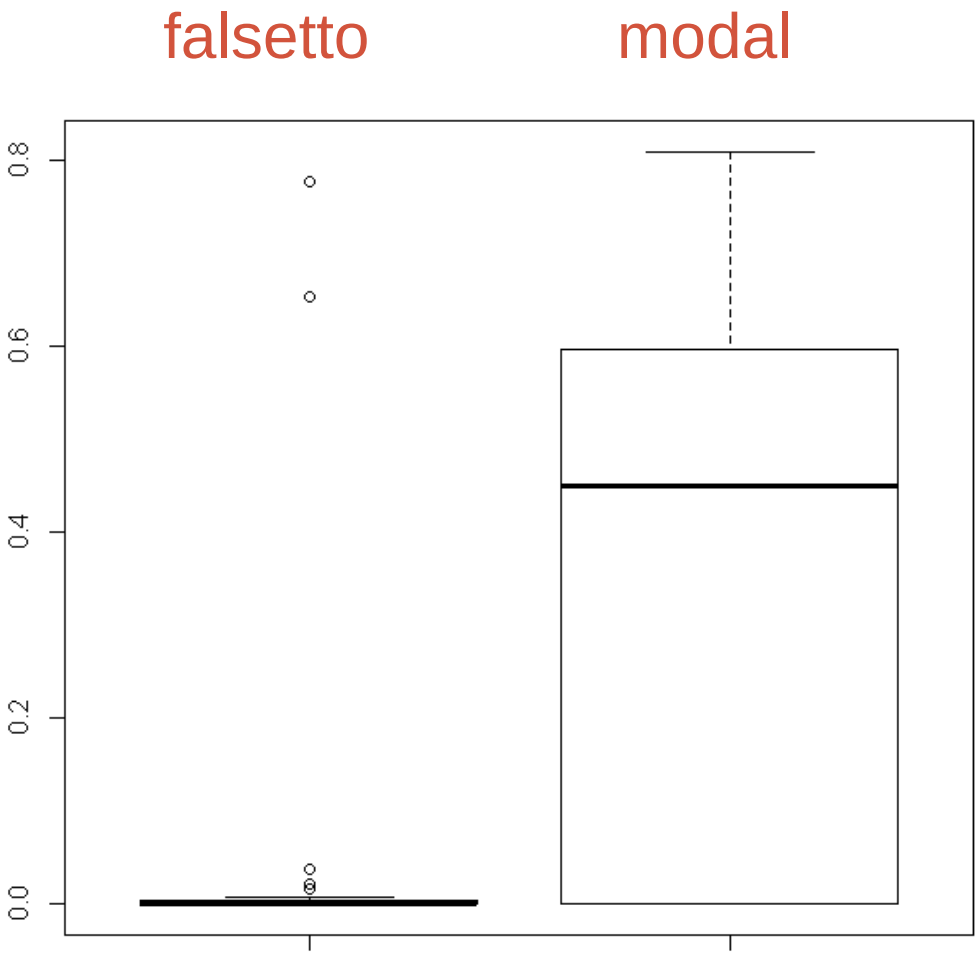
Results: Harm-Noise Ratio

(here, 0-2500 Hz)



All HNRs (over all frequency ranges) are significantly different between the voice registers: modal voice has more harmonic energy in a given frequency range

Results: Subh-Harm Ratio



Falsetto voice has fewer (no) sub-harmonics between harmonics of F0

Summary: Differences

- H1*-H2* differs between falsetto and modal only *because* F0 differs
- H2*-H4* differs only *when* F0 is taken into account
- H2*-H4* difference seems mainly due to H4*: less energy in H4* → larger H2*-H4*
- Absence of subharmonic structure in falsetto

Conclusions

- H1*-H2* is NOT a good measure of falsetto, because its variation with register is due to co-variation with F0
- H2*-H4* (H4*) seems to differ between the registers independently of F0
- Subharmonic-to-Harmonic Ratio also differs between the registers independently of F0, supporting Sun (2002)’s very different use for voice roughness analysis

Acknowledgements

- NSF grants:
  - IIS-1018863 to A. Alwan et al.
  - BCS-0720304 to P. Keating et al.
- Caitlin Smith for recording speakers

References

- Baken & Orlikoff (2000), *Clinical measurement of speech and voice*. Singular.
- Colton (1973), Some acoustic parameters related to the perception of modal-falsetto voice quality. *Folia phoniat*. 25, 302-311
- Garellek et al. (2013), Perceptual sensitivity to a model of the source spectrum. *POMA* 19, 060157
- Hammarberg et al. (1980), Perceptual and acoustic correlates of abnormal voice qualities. *Acta Otol*. 90, 441-451
- Hanson (1997), Glottal characteristics of female speakers: acoustic correlates. *JASA* 101, 466-481

- Iseli et al (2007), Age, sex, and vowel dependencies of acoustic measures related to the voice source. *JASA* 121, 2283-2295
- Keidar et al. (1987), The perceptual nature of vocal register change. *J. Voice* 1.3, 223-233
- Nair (1999), *Voice tradition and technology: A state-of-the-art studio.* Singular.
- Neiman et al. (1997), Acoustic examination of naturalistic modal and falsetto voice registers. *Log. Phon. Vocol.* 22, 135-138
- Podesva (2007), Phonation type as a stylistic variable: The use of falsetto in constructing a persona. *J. Sociolinguistics* 11, 478-504.
- Shue (2010), *The voice source in speech production: Data, analysis and models*. UCLA dissertation.
- Shue et al. (2011), VoiceSauce: A program for voice analysis. *Proc. ICPhS XVII*, 1846-1849
- Stross (2013), Falsetto voice and observational logic: Motivated meanings. *Language in Society* 42, 139-162.
- Sun (2002), Pitch determination and voice quality analysis using Subharmonic-to-Harmonic Ratio. *Proc. ICASSP-2002*, Orlando, 333-336
- Sundberg (1973), The source spectrum in professional singing. *Folia phoniat.* 25, 71–90
- Sundberg & Högset (2001), Voice source differences between falsetto and modal registers in counter tenors, tenors and baritones*. Log. Phon. Vocol*. 26, 26-36.
- Titze (2000), *Principles of Voice Production*. Denver: National Center for Voice and Speech.
- van den Berg (1968), Mechanisms of the larynx and the laryngeal vibrations. In Manual of Phonetics (ed. Malmberg), pp. 278-308, Amsterdam: North Holland Publishing.