Perceptual sensitivity to a model of the source spectrum

Marc Garellek[a,b], Robin A. Samlan[b], Jody Kreiman[b], & Bruce Gerratt[b]

[a]Department of Linguistics, UCLA; [b]Department of Head and Neck Surgery, UCLA

marcgarellek@ucla.edu

Introduction

- The source spectrum is often schematized with harmonic amplitudes falling off at 12 dB/octave (Ní Chasaide & Gobl 1997).
  - But, there is much variation across voices and voice qualities (Fig.1).

Fig 1

- *How can we model the source spectrum to capture differences across speakers & voice quality types?*
- Kreiman et al. (2011) model the source spectrum using 4 component slopes:
  - H1-H2, H2-H4, H4-2kHz (H4-2K), 2kHz-5kHz (2K-5K)

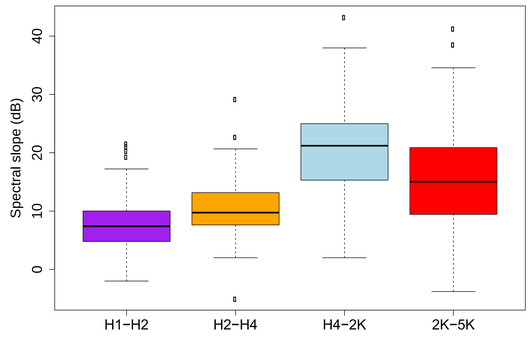
Research questions:

- Part of a broader effort to understand which aspects of vocal fold motion determine voice quality via the relationship between quality and acoustics, and acoustics and production
- What is the extent to which speakers vary along these 4 component slopes?
- To what extent might the slope of one component depend on the slopes of others and on F0/noise?
- How sensitive are listeners to each of these component slopes?

Cross-speaker variability

- 144 voices (79 F, 65 M) – normal and disordered
- Voices were inverse filtered and copy-synthesized using custom software (Kreiman et al., 2010).
- Spectral slopes, F0, and noise (NHR) measured from synthetic copy.

Fig 2: Range of values for the 4 slopes



- Linear regression used to predict one component slope as a function of other components' slopes, F0, and NHR

RESULTS:

- Higher values of a component's slope → <u>lower</u> values in adjacent components' slopes

Table 1: Adjacent slope relations

	H1-H2	H2-H4	H4-2K	2K-5K	F0	NHR
H1-H2		−0.11			+0.32	
H2-H4	−0.11		−0.37		+0.42	
H4-2K		−0.37		−0.21	−0.55	
2K-5K			−0.21			−0.18

Listener sensitivity

- *How sensitive are listeners to each of these component slopes?*
- Just-noticeable difference (JND) task –
  - 1 up, 2 down paradigm
- 2 voices – 1 Female (F0=200 Hz), 1 Male (F0=104 Hz)
- H1-H2, H2-H4, H4-2K: varied in 0.5 dB steps within a range of naturally occurring values
- 2K-5K: varied in 1 dB steps
- 18 listeners (11 F, 7 M) with normal hearing

RESULTS:

Fig 3: JND by component slope and voice

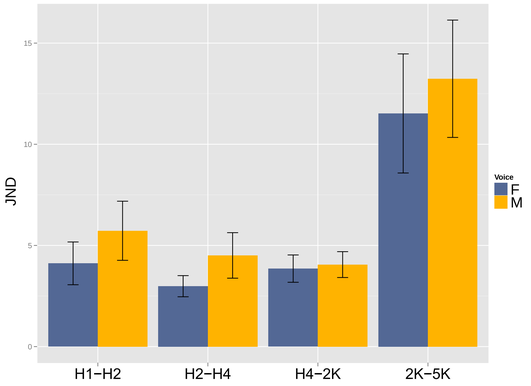


Table 2: mean JND / range in 144 voices

	Female	Male
H1-H2	0.17	0.24
H2-H4	0.09	0.13
H4-2K	0.09	0.09
2K-5K	0.26	0.29

Discussion

- Component slopes "trade off" with slopes of adjacent components:
  - As one goes up, the other goes down.
  - General flattening of spectrum towards higher frequencies (12 dB/octave expectation) is not the norm.
- The source spectrum is more variable above H4 (H4-2K, 2K-5K) than below H4 (H1-H2, H2-H4).
- More variability ⇏ less sensitivity:
  - Listeners are generally good at hearing small differences in harmonic amplitudes at frequencies below 2K.
- Difference in JND between the male and female voices could be due to speaker characteristics or F0.
- JNDs for H1-H2 are roughly consistent with those reported previously (cf. Kreiman & Gerratt, 2010).
  - Differences across listeners and/or voices?
- More work is needed:
  - To see how sensitivity to one component slope varies as a function of the others.
  - To tie vocal physiology to acoustic production and the perception of quality it evokes.

Acknowledgments

Thanks to Norma Antoñanzas-Barroso and Sherrie Rastifar. This work is supported by NIH/NIDCD grant DC01797.

References

1. Kreiman, J., & Gerratt, B.R. (2010). Perceptual sensitivity to the first harmonic amplitude in the voice source. *JASA* 128, 2085-2089.
2. Kreiman, J., Antoñanzas-Barroso, N., & Gerratt, B.R. (2010). Integrated software for analysis and synthesis of voice quality. *Behavior Research Methods* 42, 1030-1041.
3. Kreiman, J., Garellek, M., & Esposito, C. (2011). Perceptual importance of the voice source spectrum from H2 to 2 kHz. *JASA* 130, 2570.
4. Ní Chasaide, A. & Gobl, C. (1997). Voice source variation. In *The Handbook of the Phonetic Sciences*, Oxford: Blackwell.