Source properties of dorsal fricatives

Charles Redmon and Allard Jongman

Department of Linguistics, University of Kansas

The 5th Joint Meeting of the Acoustical Society of America and the Acoustical Society of Japan: Honolulu, 2016

Background

- Models of voiceless fricative acoustics rely on the assumption that the sound source at the supralaryngeal constriction is aperiodic[1]
  - This simplifying assumption accounts for a wide range of fricative places of articulation[1–6]
- However, for voiceless **velar** /x/ and **uvular** /χ/ fricatives (jointly referred to as **dorsal fricatives, X**) the potential for high-velocity airflow to induce vibration of the uvula means such sounds may in some instances be excited by a **mixed source**
  - Data from previous studies support this expectation[1,7–11]

Research Questions

- How common are mixed-source productions of voiceless dorsal fricatives and what is their relative positional and vocalic distribution?
- What are the acoustic consequences of uvular vibration when present?

Participants

- 4 native speakers (2F) each of Arabic, Persian, and Spanish
- All 12 participants grew up in Saudi Arabia, Iran, or Spain, respectively

Materials

- 2 fricatives /s, X/ in 3 vowel contexts /i, a, u/ in 2 positions (CV, VC) formed the target set
- 3 words containing each target sequence were recorded in sentence frames in 4 repetitions
- The same list was recorded twice:
  1. Acoustic data was recorded with a head-mounted microphone in an anechoic chamber
  2. Oral and nasal airflow was recorded with the Scicon R&D (2015) system
- Acoustic signals were digitized at 22.05 kHz, and airflow signals were digitized at 11 kHz and low-pass filtered at 200 Hz

Source Estimation

- Frequency of oscillation in airflow was estimated through the power spectral density (PSD) of the autocorrelation function $R(\tau)$ over 60 ms windows sampled in 10 ms steps
- Reference pitch floors by speaker were used to ensure any periodicity was not due to voicing

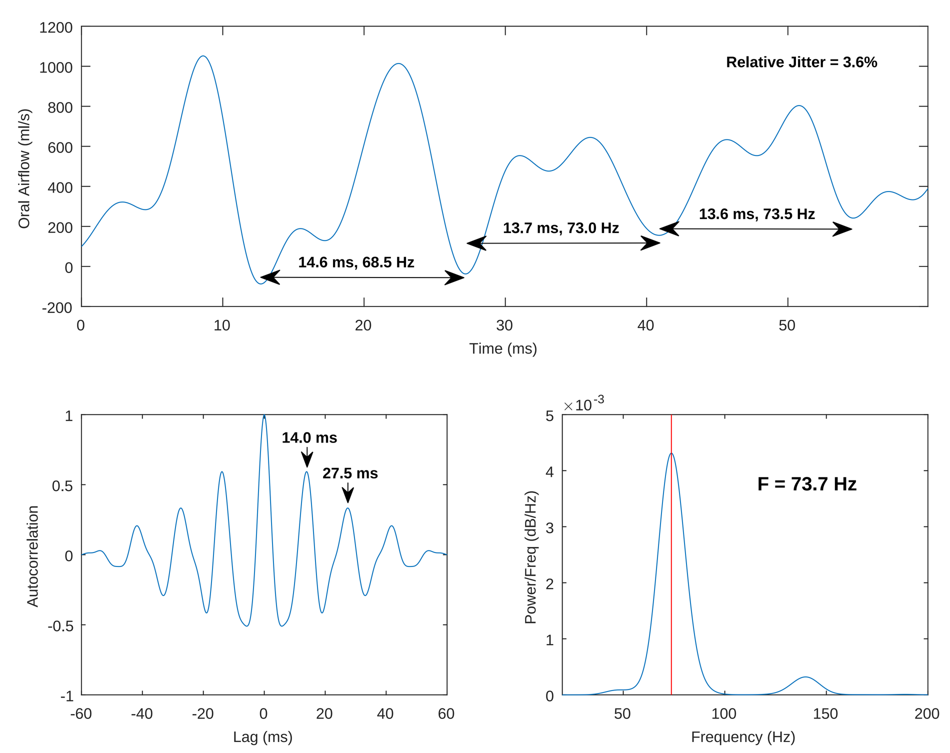

Figure 1: Measurement illustration and definitions.

Aerodynamic Data

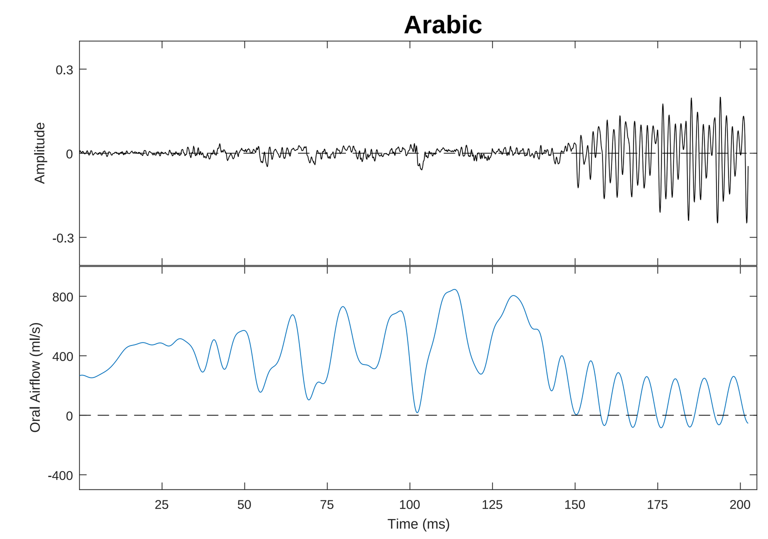

Figure 2: Sample production of the Arabic word /Xa:ru:f/ by speaker AM02. The estimated oscillation frequency in this token is 61.8 Hz.

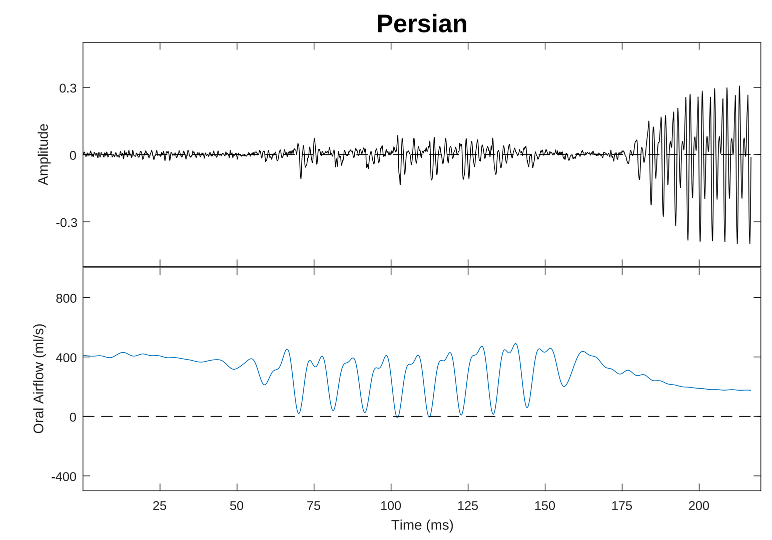

Figure 3: Sample production of the Persian word /Xa:k/ by speaker PF02. The estimated oscillation frequency in this token is 93.5 Hz.

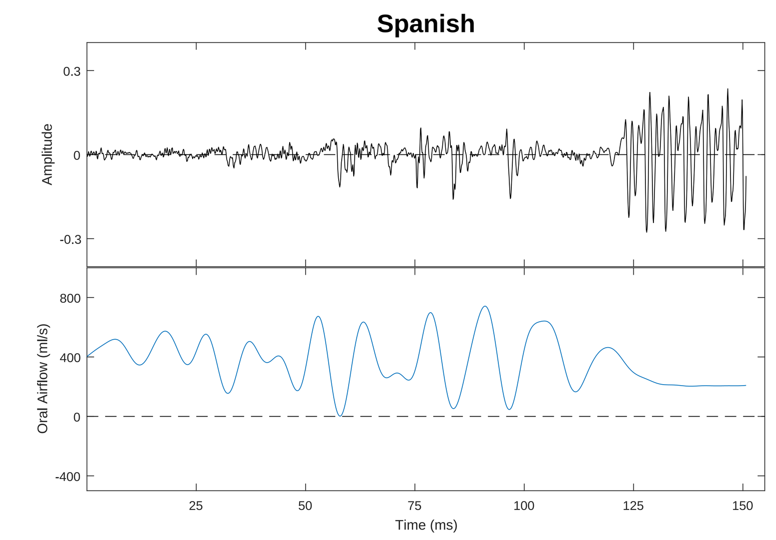

Figure 4: Sample production of the Spanish word /Xade/ by speaker SF01. The estimated oscillation frequency in this token is 77.4 Hz.

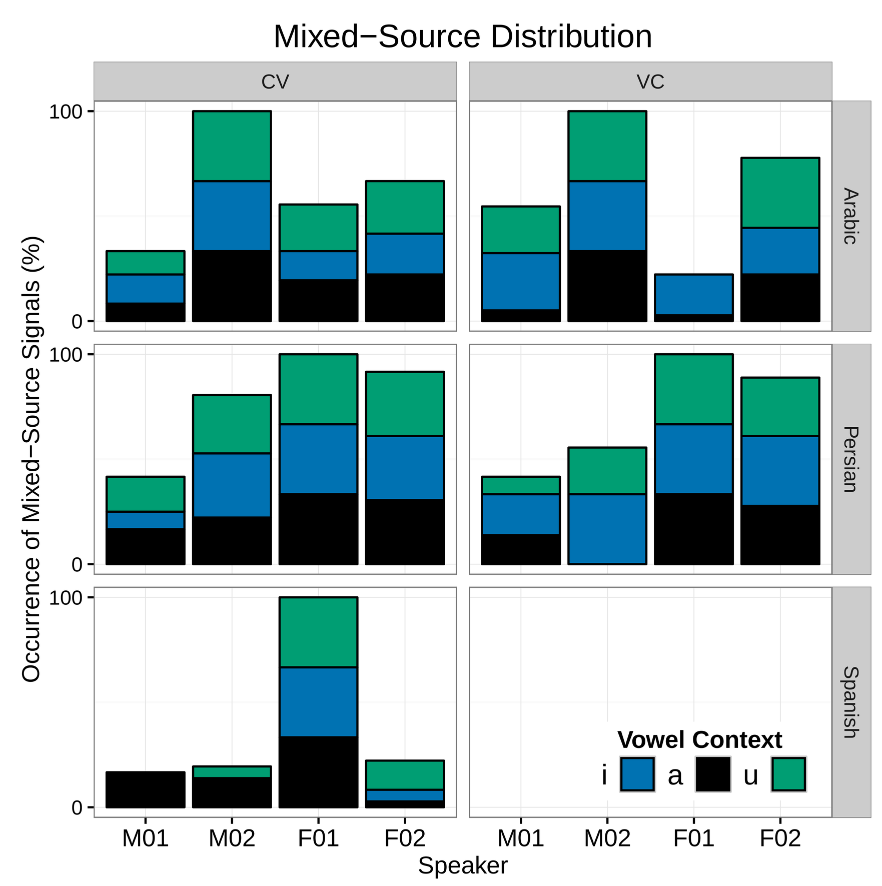

Figure 5: Distribution of mixed-source /X/ tokens by Language, Position, and Vowel. Speaker IDs are truncated (e.g. M01 for PM01) for space.

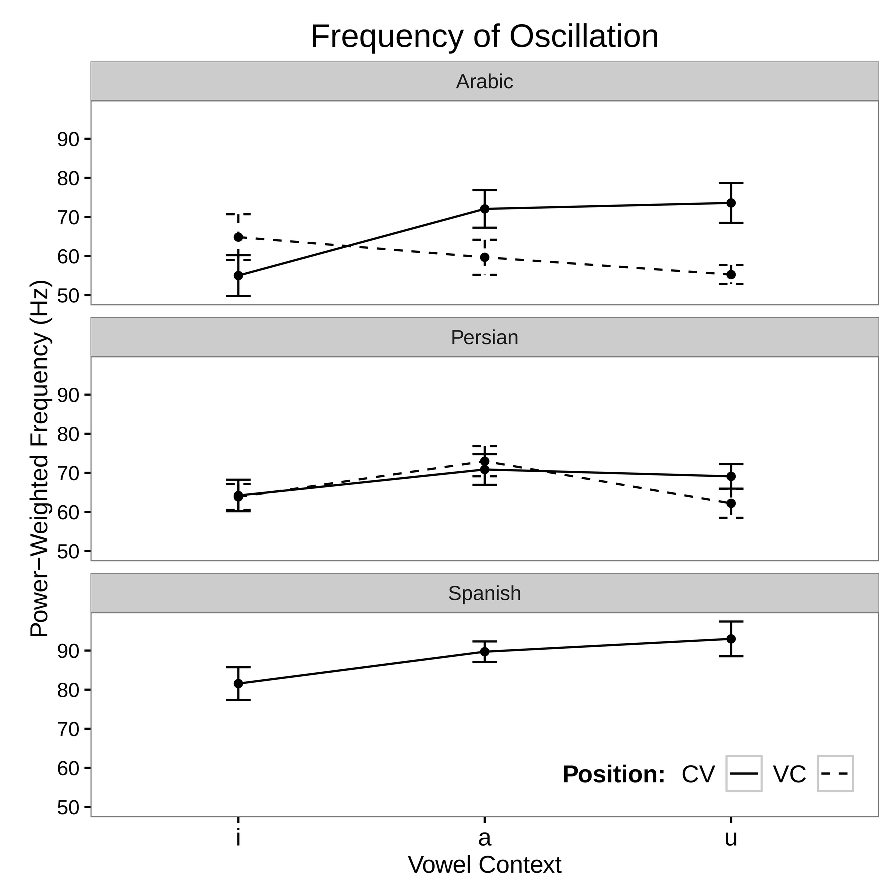

Figure 6: Estimated frequencies of oscillation by Language, Position, and Vowel Context. Error bars represent ± 1 standard error.

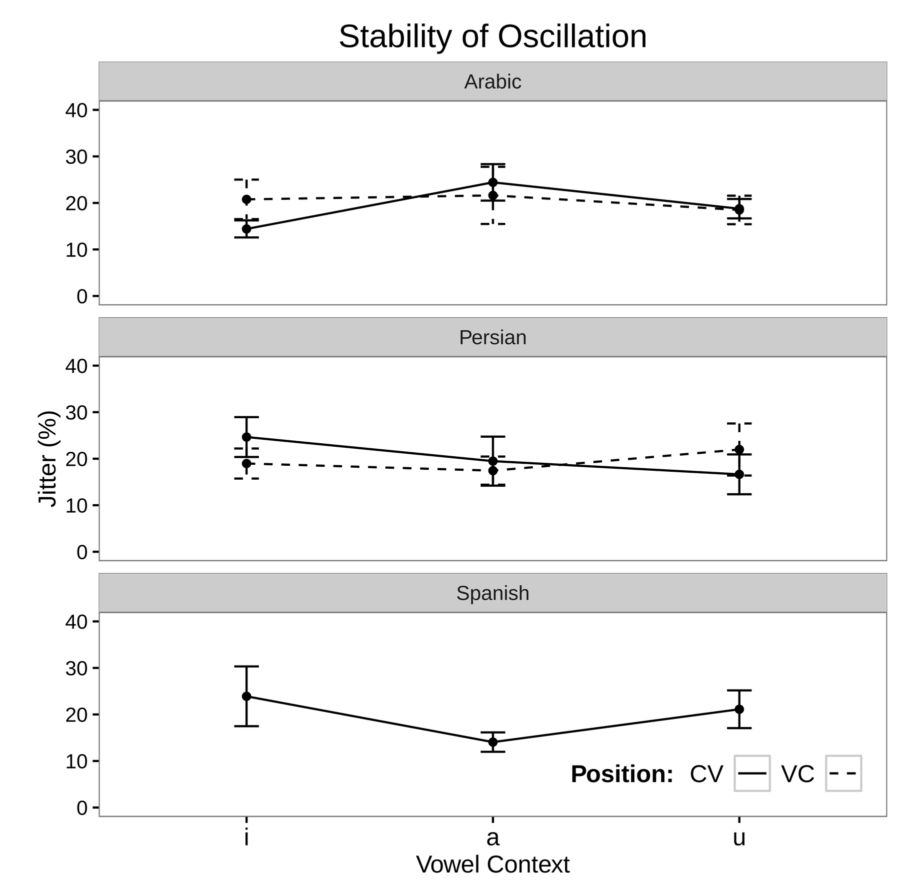

Figure 7: Relative Jitter in the base frequency of a representative sample (n = 148) of mixed-source tokens across Language, Position, and Vowel.

- Arabic and Persian, relative to Spanish, show more consistent productions of the dorsal fricative with some degree of uvular vibration
- Oscillation frequency in word-initial position is reliably lower in the context of the high front vowel /i/ than before the low back vowel /a/
- Oscillation frequency is much less stable (higher percentage jitter) than would be expected for voiced fricatives or trills, which is consistent with the characterization of the process as a random vibration. No clear patterns emerged in Position or Vowel Context comparisons across languages, however.

Acoustic Data

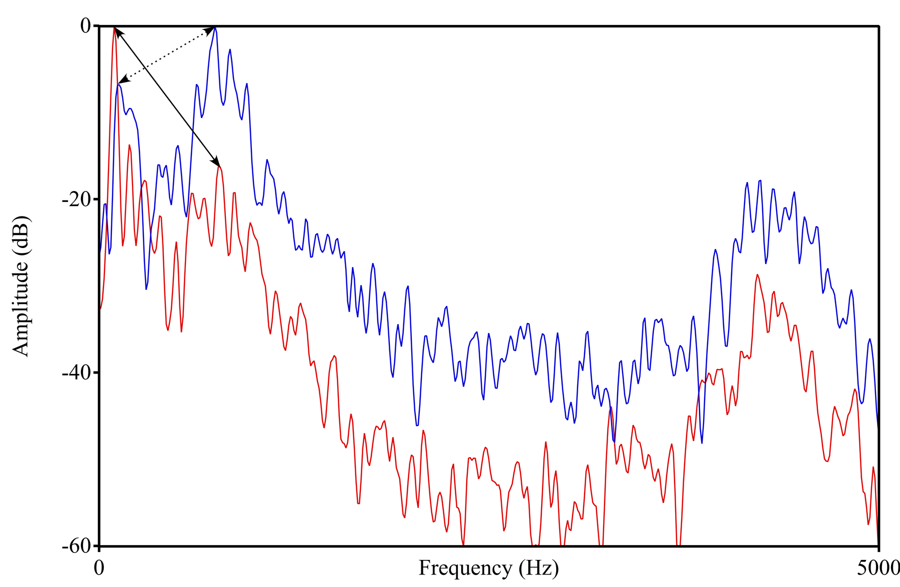

Figure 8: Spectra from the middle 60 ms of /X/ in two repetitions of the word /Xutˤba/ produced by speaker AF02, one with uvular vibration (red) and one without (blue). An index of the presence of vibration, A_s–A_f, the source component amplitude minus the filter component (spectral peak) amplitude, is shown with the solid and dashed arrows, respectively. In Figures 9 and 10 we have used this measure to classify dorsal fricative tokens (Acc.: 98%, Spec.: 100%) into Aperiodic and Mixed source types.

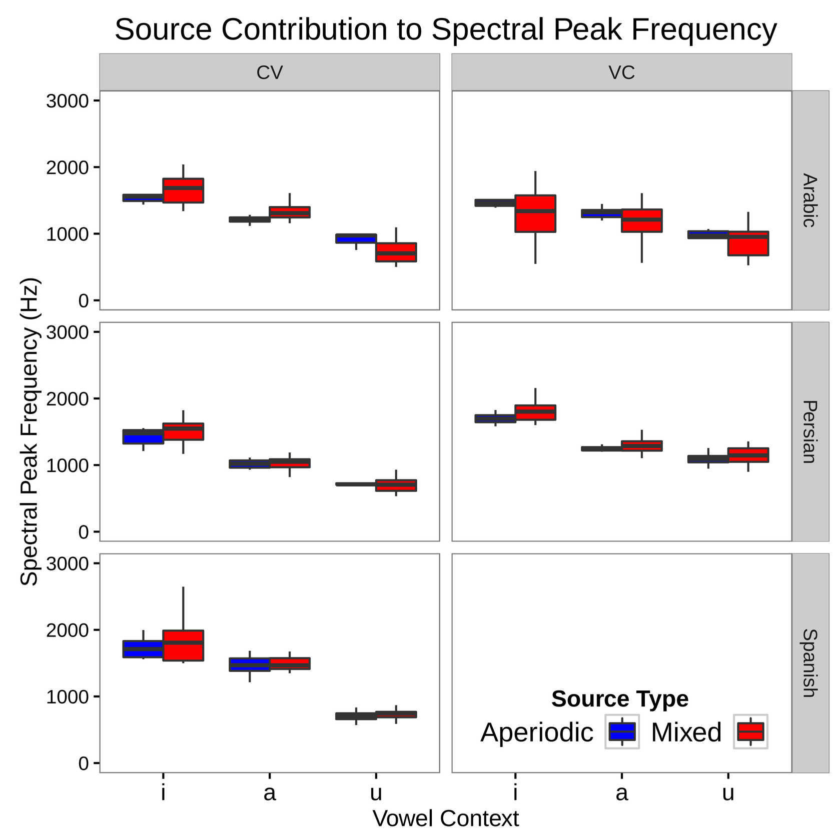

Figure 9: Spectral peak frequencies by source type.

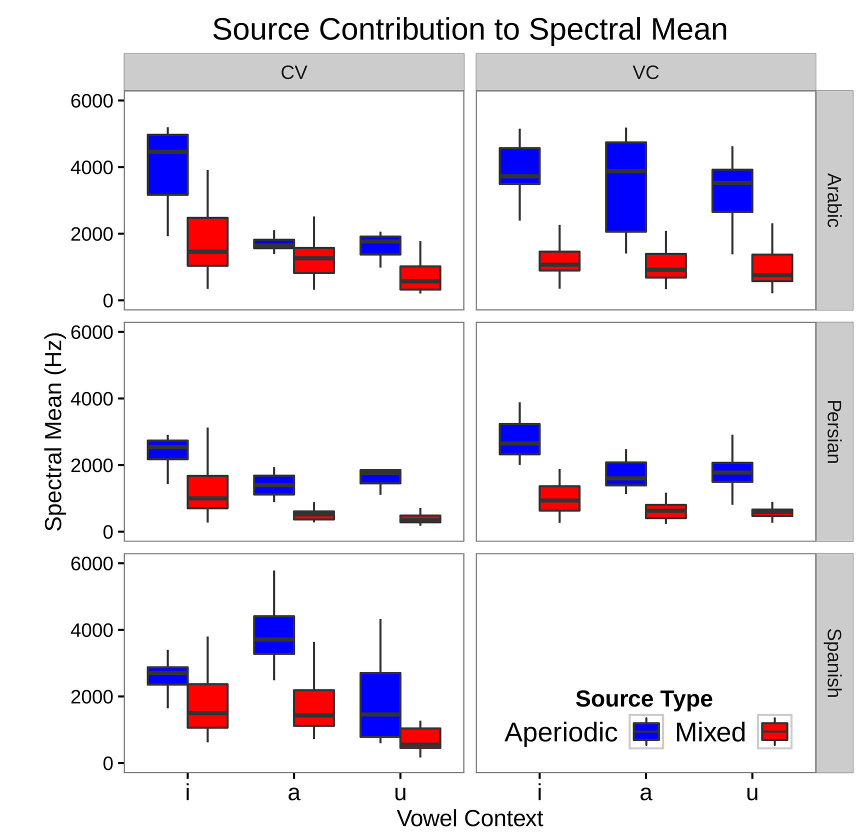

Figure 10: Spectral means by source type.

Synthesis

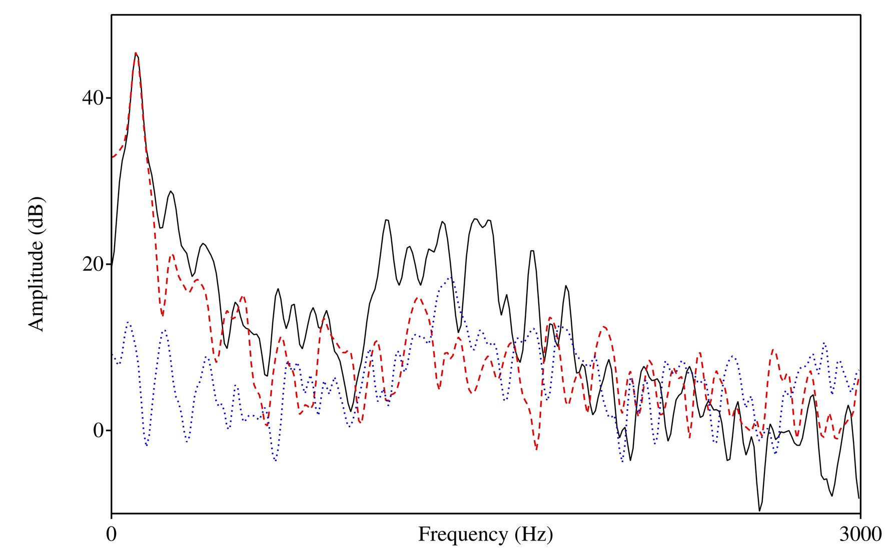

Figure 11: Real (solid line) and synthesized (dashed/dotted lines) spectra of the middle 60 ms of /X/ in the Persian word /ʃaX/ produced by speaker PF01. The dotted blue line is from an LPC filter of the real signal excited by an aperiodic source composed of Gaussian noise. The dashed red line is from excitation of the same filter by a mixed source composed of the same Gaussian noise overlaid on a 100 Hz sine wave.

Conclusions

- Uvular vibration is pervasive in Arabic and Persian productions of voiceless dorsal fricatives
- The presence of uvular vibration has robust acoustic effects and critically affects common spectral measures such as the first moment
- Effects of vowel context and position on the nature of this vibration vary by language
  - One pattern, /i/ < /a/ in oscillating frequency, does hold across languages
  - This pattern may be aerodynamically motivated, though without imaging data it is unclear whether this is due to effects of posterior cavity size or constriction size

Future Directions

- **Modeling the vibration process:** Pharyngeal pressure measurements are necessary to provide a complete aerodynamic model, with imaging data an important supplement
- **Perception:** Is the periodic component we observed in the acoustics perceptually salient? Moreover, is it critical to the maintenance of contrasts with posterior fricatives like /ħ, h/ in Arabic and /h/ in Persian?
- **Synthesis:** At present the sine wave component to the source does not reflect frequency and amplitude perturbations common in real speech

References

1 Fant, G. (1960). *Acoustic Theory of Speech Production*; 2 Stevens, K. (1971). *JASA, 50(4)*; 3 Shadle, C. (1985). Ph.D. Diss., MIT; 4 Alwan, A. (1986). M.S. Diss., MIT; 5 Narayanan, S., & Alwan, A. (2000). *IEEE SAP, 8(2)*; 6 Jongman, A., *et al.* (2000). *JASA, 108(3)*, 7 Zeroual, C. (2003). *ICPhS Proc.*; 8 Jesus, L., & Shadle, C. (2005). *JIPA, 35*; 9 Shosted, R., & Chikovani, V. (2006). *JIPA, 36*; 10 Shosted, R. (2008). *Acoustics '08 Proc.*; 11 Yeou, M., & Maeda, S. (2011). In *Instrumental Studies in Arabic Phonetics*

Acknowledgements

We would like to thank Ahmad Razi, Reema Al-Mutair, and Maria Martinez Garcia for their help with stimulus design, and Joan Sereno, Annie Tremblay, Jie Zhang, and the members of the Experimental Research Seminar at KU for their helpful feedback on earlier versions of this work.

Contact Information

E: redmon@ku.edu, jongman@ku.edu

W: redmonc.github.io, kuppl.ku.edu/allard-jongman