Acoustic classification of velar fricatives in Assamese

Charles Redmon
Department of Linguistics, University of Kansas
171st Meeting of the Acoustical Society of America: Salt Lake City, 2016

Research Questions

1. To what extent does allophonic variation affect acoustic discrimination of obstruents in Assamese?
2. Is the behavior of this contrast system consistent across gender and dialect sub-groups?

Background

- Fricative identification can be modeled robustly in acoustics and perception at a global level (in the absence of within-category place and manner variation)[1,2]
- But allophony is expected to interact with category discrimination; e.g. in German the realization of /x/ as [ç] after /i/ influences the contrast with /ʃ/[3]
- Posterior obstruents in Assamese: /k, kʰ, g, gʱ, x, h/
- Optional word-initial fortition of /h/ ($h \rightarrow x$) and post-vocalic lenition of /kʰ/ ($k^h \rightarrow x$) may reduce acoustic discriminability of /x/
  - These processes are reported to apply with varying frequency depending on gender and dialect[4,5]

Participants

- 8 speakers from Jorhat District: 4 female (JF), 4 male (JM)
- 4 speakers from Nalbari District: all male (NM)
- Gender contrast: $JF \longleftrightarrow JM$
- Dialect contrast: $NM \longleftrightarrow JM$

Materials

- Near-minimal pairs in triggering contexts for allophonic processes on /h/ and /kʰ/

CV	VCV	VC
xa-, ha-	-ɔxɔ-, -ɔhɔ-, -ɔkʰɔ-	-ax, -akʰ

- 3 words exhibiting each target sequence were recorded in sentence frames and repeated in 3 randomized blocks

Acoustic Measures

- RMS amplitude (RMS)
- Low- and high-frequency spectral tilt (LFT, HFT)[6]
- Spectral moments (M1, M2, M3, M4)[7,8]
- *All measurements taken at 30 ms windows at consonant onset, midpoint, and offset (C_i, C_m, C_f), transitions (VC/CV), and vowel offset/onset (V_1/V_2)*

Models

- Speaker information was partialled out of raw parameter values, with residuals then used in principal components analysis (PCA) to reduce the parameter size to 3
- Cross-validated logistic regression on /x/-/h/ and /x/-/kʰ/ contrasts was run stepwise by window

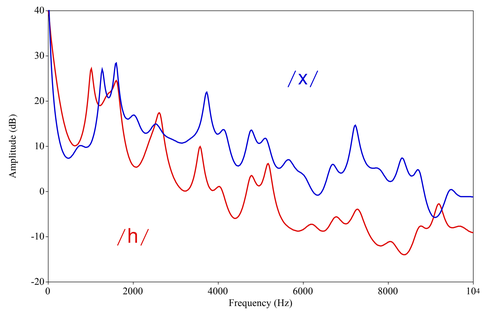

Figure 1: Sample velar and glottal fricative spectra at consonant midpoint (C_m) from tokens /xat/ 'seven' and /hat/ 'market' produced by a male speaker from Jorhat.

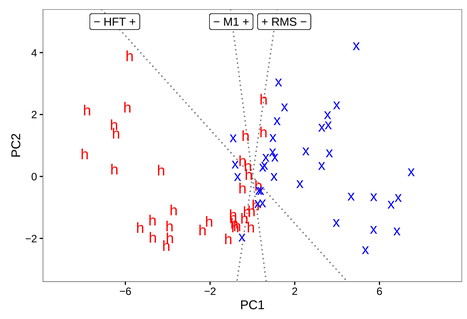

Figure 2: Sample PCA reduction (JM /xa/-/ha/ contrast at C_m). Dotted lines mark boundaries perpendicular to parameter eigenvectors (+/− signs indicate directionality).

Key Classification Patterns

Word-initial (CV) and word-final (VC) contexts consistently reduce discriminability of /x/, but the Jorhat Female group shows greater influence of /h/ fortition word-initially and Nalbari Male speakers show greater influence of /kʰ/ lenition post-vocalically.

/x/–/h/

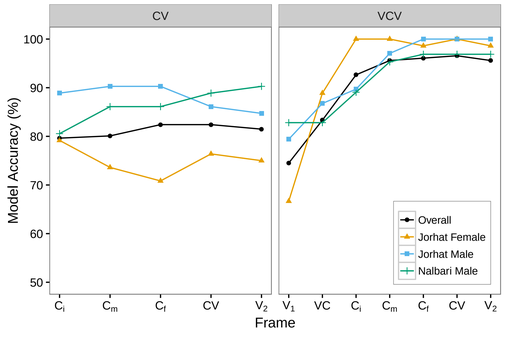

Figure 3: Classification accuracies for models of the /x/–/h/ contrast in word-initial (CV) and intervocalic (VCV) positions.

- CV < VCV in /x/–/h/ accuracy ($p < 0.001$)
- JF < JM = NM in CV /x/–/h/ accuracy ($p < 0.05$)
- Spectral mean and RMS amplitude are the most critical parameters in distinguishing /x/ and /h/ (based on the size of AIC reduction from inclusion in the model)

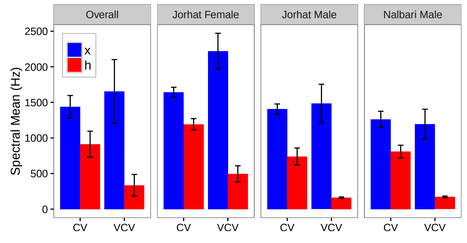

Figure 4: Spectral means of /x/ and /h/ at consonant midpoint (error bars represent ± 1 standard error about the mean).

/x/–/kʰ/

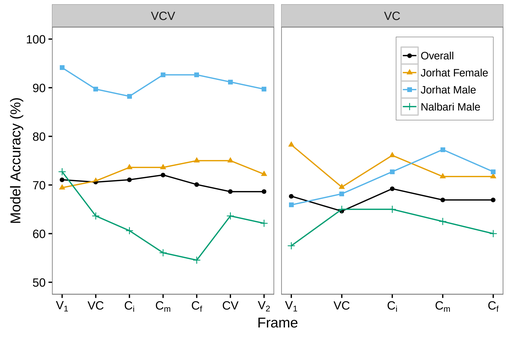

Figure 5: Classification accuracies for models of the /x/–/kʰ/ contrast in intervocalic (VCV) and word-final (VC) positions.

- VC = VCV ($p = 0.975$) in overall /x/–/kʰ/ accuracy, but group differences were marginal word-finally, suggesting coda position is a stronger trigger for lenition
- NM < JF < JM in VCV /x/–/kʰ/ accuracy ($p < 0.01$)
- No parameters consistently distinguish /x/ and /kʰ/

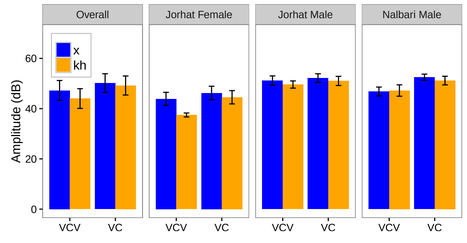

Figure 6: RMS amplitudes of /x/ and /kʰ/ at consonant midpoint (error bars represent ± 1 standard error about the mean).

Phonotactic Frequency

- Incorporation of lexical information should bias classification toward the velar fricative: Relative proportions of /xa/- and /ha/-initial words in a written corpus of Assamese[9] were 0.64 and 0.36, respectively
- Word-final aspirated stops are generally uncommon in Assamese, as was revealed for /kʰ/ in the present corpus, which had a type frequency of 39 as compared with 359 for /x/
  - Only four items formed minimal pairs, however, suggesting /x/-/kʰ/ discrimination may be less affected by lenition than predicted in naïve acoustic models

Conclusions

- Classification of velar fricatives in Assamese significantly depends on allophonic variation in acoustically similar posterior obstruents
- The relative impact of processes such as /h/ fortition and /kʰ/ lenition was not constant across speaker groups, and particularly affected Jorhat Female and Nalbari Male classification models
- Posterior obstruent contrasts may be better modeled as separate systems according to gender and dialect
- The extent to which listeners can cope with these differences in perception has yet to be tested

Future Directions

- Predictions from acoustic models need to be validated on perception data from Assamese listeners
- Manipulation of lexical and phonotactic frequencies in perception stimuli will help determine the degree to which listeners can compensate for acoustic ambiguity in maintaining posterior obstruent contrasts

References

1 McMurray, B., & Jongman, A. (2011). What information is necessary for speech categorization? Harnessing variability in the speech signal by integrating cues computed relative to expectations. *Psychological Review, 118(2)*, 219-246.
2 Silbert, N., de Jong, K., Regier, K., & Albin, A. (2015). The mapping between phonological categories and acoustic cues in the production of English obstruents. *J. Acoust. Soc. Am., 138(6)*, 3834-3845.
3 Jannedy, S., Weirich, M. & Helmeke, L. (2015). Acoustic analyses of differences in [ɛ] and [ʃ] productions in Hood German. In *Proceedings of the 18th International Congress of Phonetic Sciences:* Glasgow, UK.
4 Sarma, M. K. (2006). *A phonological study of syllables in Assamese and Assamese English.* Unpublished M.Phil. Dissertation, EFLU, Hyderabad.
5 Mahanta, S. (2012). Assamese. *Journal of the International Phonetic Association, 42(02)*, 217-224.
6 Narayanan, S., & Alwan, A. (2000). Noise source models for fricative consonants. *IEEE Transactions on Speech and Audio Processing, 8(3)*, 328-344.
7 Forrest, K., Weismer, G., Milenkovic, P., & Dougall, R. N. (1988). Statistical analysis of word-initial voiceless obstruents: Preliminary data. *J. Acoust. Soc. Am., 84(1)*, 115-123.
8 Jongman, A., Wayland, R., & Wong, S. (2000). Acoustic characteristics of English fricatives. *J. Acoust. Soc. Am., 108(3)*, 1252-1263.
9 Baker, P., Hardie, A., McEnery, T., Cunningham, H., & Gaizauskas, R. J. (2002). EMILLE, A 67-million word corpus of Indic languages. In *LREC*. http://www.emille.lancs.ac.uk/

Acknowledgements

Thanks to Dr. Priyankoo Sarmah, Dr. M. Prasanna, and Pamir Gogoi for assisting with data collection at IIT Guwahati, and to Drs. Indranil Dutta, Allard Jongman, Joan Sereno, Annie Tremblay, and 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
W: http://redmonc.github.io