Distributional factors in Telugu sibilant production

Charles Redmon, Allard Jongman, and Jie Zhang

Department of Linguistics, University of Kansas; Lawrence, KS, USA

BACKGROUND

- Telugu is a Dravidian language spoken in South India
- Unlike many languages in the region which lost the three-way distinction between alveolar, palatal, and retroflex sibilants present in Sanskrit, Telugu purportedly preserves the contrast[1-4]
- Such dense systems are typologically rare and have been shown (e.g., in Polish and Mandarin) to be acoustically unstable[5-7]

GOAL OF THE STUDY

We seek to characterize the acoustics of the sibilant contrast system in Telugu, information which is largely absent from the literature.

PARTICIPANTS

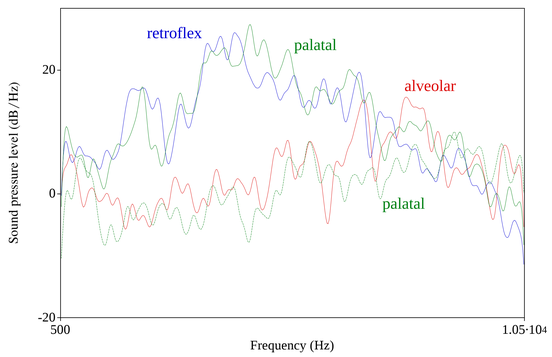
- 16 native speakers of Telugu (8 female, 8 male) recorded in Hyderabad at the English and Foreign Languages University
- 14/16 from Telangana (8 of whom were from Hyderabad)

MATERIALS

- 240 stimuli (120 words × 2 reps)
  - 3 sibilant fricatives (alveolar, retroflex, palatal)
  - 60 word-initial (CV), 60 word-medial/final (VC)
  - Critical vowel contexts: 12 /a/, 2 each of /i, e, o, u/
  - Half of the /a/-context items have 2nd-order neighbors (near-minimal pairs) contrasting in sibilant place; half do not
- **We focus in this presentation on studying the contrast in the /aCa/ context, because (1) it is the most common environment in which all three sibilants occur, and (2) word-initial retroflex sibilants are largely limited to English loanwords**

SIBILANT SPECTRA

The following are sample spectra from Speaker F01, where the dotted palatal line illustrates the occasional alveolar-like realization observed in many speakers' data.

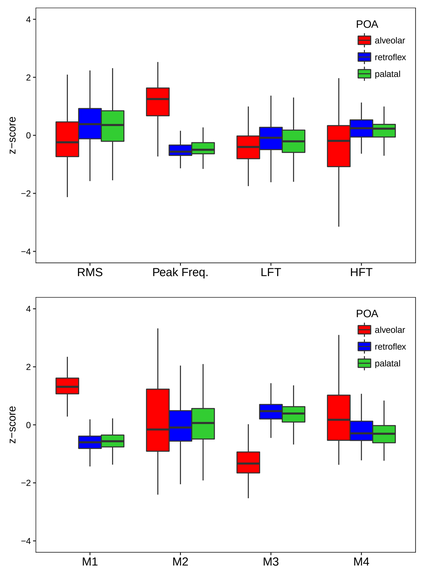


MEASUREMENTS

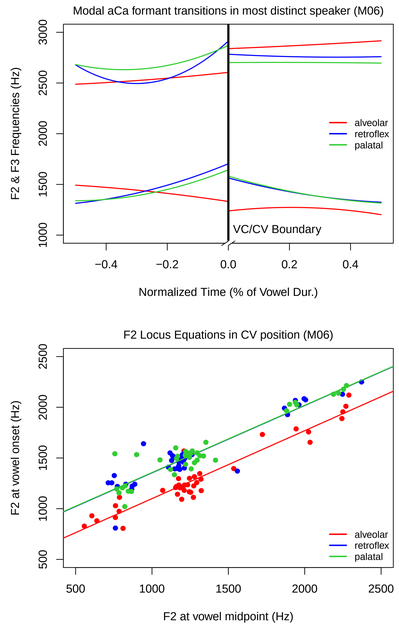
- Noise amplitude (RMS)
- Spectral peak frequency (PeakF)
- Spectral tilt below (LFT) and above PeakF (HFT)
- Spectral moments at consonant midpoint (M1–M4)
- F2 and F3 transitions (modeled with coefficients of quadratic polynomial fits to VC/CV transitions; for simplicity the table in the next panel shows F2/F3 at vowel midpoint and offset/onset)

ACOUSTIC FEATURES

Consonantal Parameters (aCa)



Coarticulatory Information

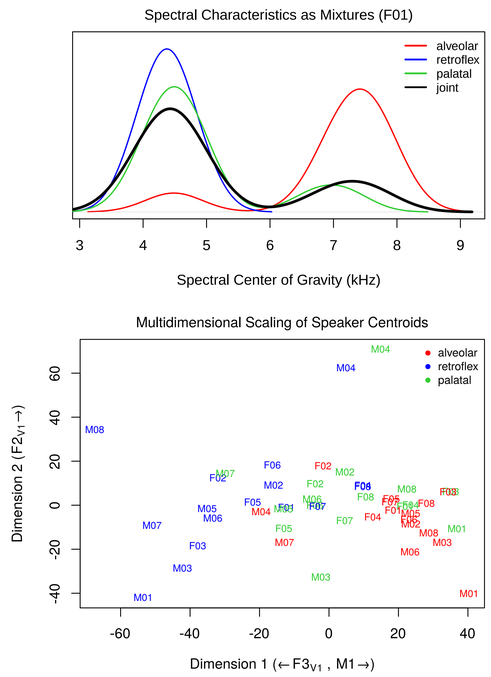


Univariate tests of place effects

		Sibilant Means		
Param.	$\Delta LL_{Sib.}$	**alv.**	**ret.**	**pal.**
M1 (Hz)	477*†	7047	4450	4594
M3	329*†	-1.11	0.81	0.68
PeakF (Hz)	266*	6969	3743	3868
$F2_{CV}$ (Hz)	80*	1513	1702	1678
$F2_{VC}$ (Hz)	80*	1539	1757	1752
HFT	27*	-4e-3	-2e-3	-2e-3
$F2_{V1}$ (Hz)	27*	1405	1332	1416
RMS (dB)	26*	54.5	56.4	56.2
$F3_{V2}$ (Hz)	13*	2877	2759	2791
$F2_{V2}$ (Hz)	9*	1434	1452	1394
$F3_{V1}$ (Hz)	9*	2892	2788	2807
M2 (Hz)	5*	1835	1732	1771
M4	4*	2.96	1.69	1.79
$F3_{VC}$ (Hz)	3*	2981	2920	2926
LFT	2	3e-3	4e-3	4e-3
$F3_{CV}$ (Hz)	2	2923	2888	2874

Model: Linear mixed effects regression with Speaker as random intercept
*significant omnibus effect of sibilant place
†all pairwise differences significant

CONTRAST SEPARATION



PATTERN OF PALATAL SIBILANT MISCLASSIFICATIONS (%) BY SPEAKER IN THE aCa CONTEXT

	F01	F02	F03	F04	F05	F06	F07	F08	M01	M02	M03	M04	M05	M06	M07	M08
Alveolar	18.6	0	12.0	0	1.7	0	0	0	0	0	0	0	0	0	15.3	13.8
Retroflex	31.3	38.3	31.9	22.9	28.5	14.1	17.3	23.5	24.6	49.2	38.1	35.8	29.3	33.5	42.2	23.6

CLASSIFICATION RESULTS

Structure of the classification model:

- Multinomial logistic regression on the three sibilants in the aCa context
- 20 predictors (RMS, PeakF, LFT, HFT, M1–M4, VC/CV F2 and F3 transition coefficients), all z-score normalized by speaker

Model patterns in the aCa environment:

- Palatal–retroflex model confusions predominate
- Model confusions between alveolar and retroflex categories are rare

	alv.	ret.	pal.
alv.	96.1	1.2	2.7
ret.	0.4	69.2	30.3
pal.	3.8	30.1	66.2

Effects of lexical characteristics:

- Model accuracy was significantly higher on items with sibilant-contrast neighbors ($e^{\beta} = 1.386$, $z = 10.74$, $p < 0.001$), controlling for lexical frequency and neighborhood density
- Lexical frequency had a significant negative effect ($e^{\beta} = 0.89$, $z = -13.09$, $p < 0.001$), meaning lower frequency words were associated with higher model accuracy in distinguishing sibilant place of articulation

DISCUSSION

- The present data, combined with the general sparsity of minimal pairs in the Telugu lexicon,[8] point toward a sibilant system which is more reliably comprised of two categories than three
  - Notably, following the recording many speakers indicated that while they were taught three distinct pronunciations in school, they are only able to perceive or produce two
  - Speakers also have an awareness of which dialects are more or less likely to show the palatal → alveolar alternation
- Further examination of item-specific patterns is needed to account for the lexical variability in palatal similarity to alveolars and retroflexes

REFERENCES

[1]Krishnamurti, B. (2003); [2]Masica, C. P. (1993); [3]Sjoberg, A. (1962); [4]Bhaskararao, P., & Ray, A. (2017); [5]Maddieson, I., & Precoda, K. (1990); [6]Żygis, M., & Padgett, J. (2010); [7]Li, M., & Zhang, J. (2017); [8]Baker *et al.* (2002)

ACKNOWLEDGEMENTS

Thanks to Indranil Dutta and the Phonetics Laboratory at EFLU, Hyderabad for facilitating the recordings, and to the members of the KU Experimental Research Seminar for their helpful feedback.