Sibilants in Gujarati Phonology

Purnima Thakur
The Graduate Center, CUNY
pthakur@gc.cuny.edu

BACKGROUND

Gujarati, a regional language spoken in India (classified as a New Indo-Aryan language) is not rich in native fricatives. Most analyses of Gujarati assume alveolar /s/, alveo-patal /ʃ/, and glottal /h/.

**Three differing views on the status of sibilants in Gujarati.**

1. Single sibilant /s/ and allophonic [ʃ]: Sibilant contrast historically neutralized in all environments. Alveo-palatal [ʃ] appears as an allophone of /s/ before front vowels /i,ɛ,e/ and glide /j/ (Turner, 1921; Grierson, 1931).

2. Phoneme /s/ and quasi-phoneme /ʃ/: Sibilant contrast neutralized in all environments except before front vowels and glide. /s/ and /ʃ/ contrast before front vowels /i,ɛ,e/ and glide /j/, and elsewhere only /s/ appears (Pandit, 1954).

3. Phonemes /s/ and /ʃ/: The sibilants are considered truly contrastive, due to the existence of some minimal pairs like [sal] 'hindrance'~ [ʃal] 'shawl', [pas] 'near' ~ [paʃ] 'bondage', and [oresjo] 'pebble' ~ [kələʃjo] 'big glass' (Adenwala, 1965; Dave, 1977; Masica, 1991)

ISSUE

Does Gujarati have three distinct fricative systems synchronically with individual speakers acquiring different ones?

Three distinct fricative systems in Gujarati

(i)	s	h	[ʃ] is an allophone of /s/
(ii)	s (ʃ)	h	/ʃ/ is a quasi-phoneme
(iii)	s ʃ	h	/s/ and /ʃ/ are contrastive

- Traditional models of phonological relationships, based on a binary distinction between *predictably distributed (allophonic)* and *not predictably distributed (contrastive)*, can successfully predict two of the three possible fricative systems.
- They, however, don't easily capture a possible third system – one with *quasi-phonemic distinctions* between the two sibilants.
- A probabilistic measure, rather than a binary measure is needed to define predictability of distribution.
- The Probabilistic Phonological Relationship Model (PPRM) proposed by Hall (2009), based on a continuous scale of predictability, calculates the precise degree to which two sounds in a language are predictably distributed.

Non-overlapping distributions ————— Overlapping distributions
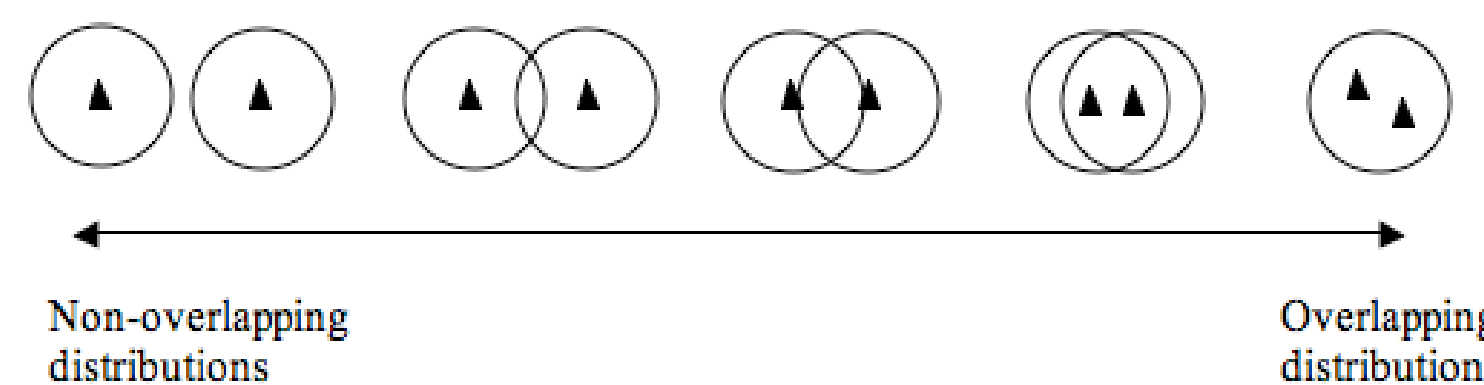
- It can help capture the Gujarati situation more appropriately.

THE PPRM

The Probabilistic Phonological Relationship Model

- It accounts for *intermediate phonological relationships* between pairs of segments that fall somewhere between "fully contrastive" and "fully allophonic."
- It is based on the twin concepts of *probability* and *entropy* (uncertainty), a concept borrowed from Information Theory.
- Three basic components of the PPRM:
  (i) PROBABILITY of occurrence of one of the two sounds in a particular environment; a *bias* towards X vs. Y in a given environment.
  (ii) ENVIRONMENTAL ENTROPY: measures degree of *uncertainty* of choice between a pair of sounds, X and Y, in a certain environment.
  (iii) SYSTEMIC ENTROPY: measures *uncertainty* of choice between two segments, X and Y, across all environments or the entire system.

PROBABILITY of occurrence of *X* in environment *e*, limiting possibilities to *X* and *Y*:

$$p(X; X, Y \mid e) = N_{X/e} / (N_{X/e} + N_{Y/e})$$

Probability Range: 0 to 1

ENVIRONMENTAL ENTROPY	SYSTEMIC ENTROPY
$H = -\sum p_i \, log_2 \, p_i$	$\Sigma \, (H(e) * p(e))$
p_i is the probability of each segment (X and Y) in the given environment.	a. $H(e) = -\sum p_i \, log_2 \, p_i$ b. $p(e) = N_e / \Sigma \, N_{e \, \epsilon \, E}$

Entropy Range: 0 to 1 (0 no uncertainty = allophony, 1complete uncertainty = contrast)

- Entropy and Probability both are defined over pairs of segments in a language system.
- While two probabilities are required to understand the relationship between a pair of segments, entropy offers a single measure to capture the relationship.
- They are used in conjunction to examine the notion of phonological contrast.

PRODUCTION STUDY

Hypotheses

(i) Allophonic distribution (Single sibilant /s/ and allophonic [ʃ])
  [ʃ]/ ___ {i, ɛ, e, j}
  [s] elsewhere
(ii) Quasi-phonemic contrast (Phoneme /s/ and quasi-phoneme /ʃ/)
  [s, ʃ]/ ___ {i, ɛ, e, j}
  [s] elsewhere
(iii) Perfect contrast (Phonemes /s/ and /ʃ/)
  [s] and [ʃ] contrast in all given environments

PRODUCTION STUDY

Participants: 20 multilingual Gujarati speakers (18 - 29 o.a.)

Word reading task: Participants read out words embedded in a carrier phrase
Həve _____ bolo ("Now ____say")

Materials: 90 familiar Gujarati words targeting sibilants /s/ and /ʃ/ in word-initial, word-medial, and word-final positions.

A. Simple onset contrasts (40 words)

front vowels	ʃijalo	'cold'	sitaro	'planet'
	ʃekvũ	'to roast'	sevvũ	'to serve'
back vowels	ʃubʰ	'auspicious'	sukʰ	'happiness'
	ʃodʰək	'examiner'	sogən	'promise'
low vowel	ʃãnt	'quiet'	sãŋ	'evening'
central vowel	ʃəŋka	'doubt'	səŋkʰja	'number'

B. Medial onset contrasts with glide (10 words)

udeʃʃə 'advice'   aləsjə 'laziness'

C. Simple coda contrasts (20 words)

syllable-final	məʃkəri	'fun'	təskəri	'theft'
word-final	vivəʃ	'dependant'	divəs	'day'

D. Coda contrasts with plosives (20 words)

syllable-final	raʃṭrə	'nation'	vəstrə	'clothing'
	prəṭiʃṭʰa	'prestige'	vjəvəstʰa	'arrangement'
word-final	bʰraʃṭ	'corrupted'	trəst	'frightened'
	vəriʃṭʰ	'senior'	gruhəstʰ	'householder'

Data Analysis

- Based on perceptual judgments of four raters (author and three others), utterances were coded on a scale of 1-2-3 (1 = [s], 3 = [ʃ], 2 = ambiguous token). An equal number of items with [s] and [ʃ] were included in the analysis.
- "Majority wins" strategy applied to ambiguous tokens showing 75-25% split: if 3 out of 4 raters agreed token was [s], it was coded as 1
- Ambiguous tokens showing 50-50% split were discarded along with the corresponding minimal pair word.
- The PPRM was used to calculate *for each Gujarati speaker*,
(a) ENVIRONMENTAL ENTROPY (H(e)): *uncertainty* of choice between alveolar [s] and alveo-palatal [ʃ] in each of the given environments.
(b) PROBABILITY (P): *bias* towards alveolar [s] versus alveo-palatal [ʃ] in each environment based on occurrence of each of the sibilants.
(c) SYSTEMIC ENTROPY (H): *total uncertainty* of choice between sibilants across all environments.

RESULTS

- Systemic entropies (H) ranged between 0.482 and 1.000
- 12 speakers showed *near perfect contrast* (H 1.000 - 0.943)
- 8 speakers displayed a *quasi-phonemic contrast* (H 0.857 - 0.482)
- No evidence of *allophonic distribution* as none had a systemic entropy score of 0.000
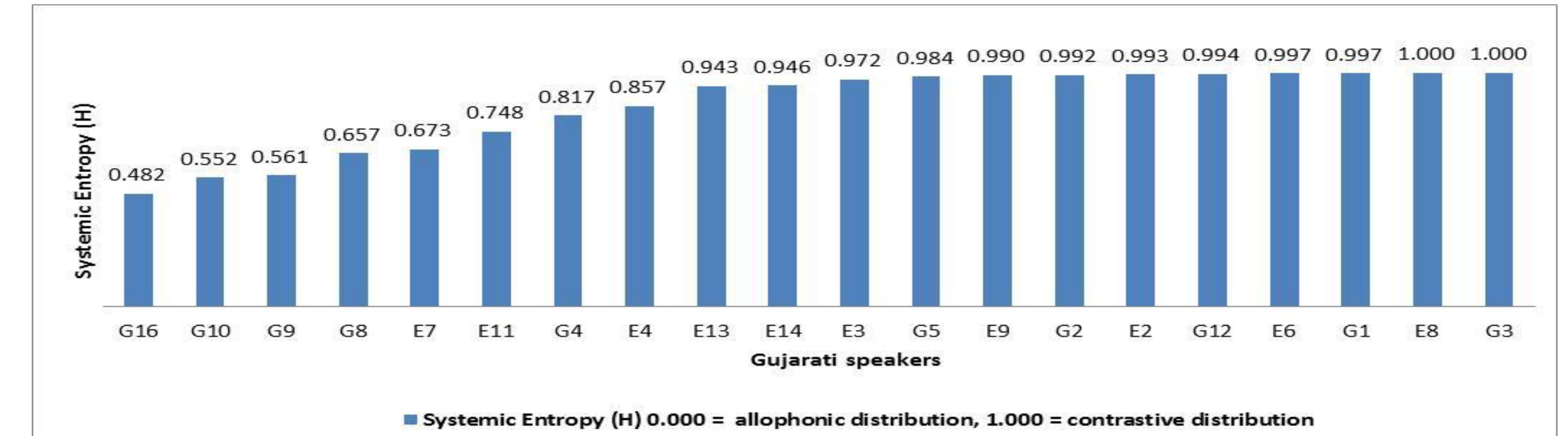
Near Perfect Contrast Group
• 8 speakers contrasted sibilants in all given environments (they showed a slight [s]- or [ʃ]-bias in certain environments (H(e) > 0.971 < 1.000), which is treated as noise in the data)
• Remaining 4 speakers contrasted sibilants in almost all environments, except one – a strong bias towards either [s] or [ʃ] in these environments was noted.

!! Interesting Observations
Phonetic conditioning
- Speaker G5: strong [ʃ]-bias when followed by velar /k/ (H(e) 0.881) → Sanskrit Ruki Rule.
- E14: strong [ʃ]-bias when followed by glide /j/ (H(e) 0.722) → /j/ trigger for palatalization.
Not readily explained
- E3: strong [s]-bias when followed by retroflex /ṭ, ṭʰ/ (H(e) 0.8111), a context for palatalization.
- E13: strong [ʃ]-bias in the environment of low vowels H(e) 0.5444)

Quasi-phonemic Contrast Group
Front vowels /i, ɛ, e/
- Only 2 speakers showed the predicted contrast
- Of the remaining, 3 were [s]-biased and 3 [ʃ]-biased.
Glide /j/
- No contrast observed, all 8 speakers produced [ʃ], indicating this is an environment for palatalization.
Other environments
- 4 speakers were [s]-biased and 4 were [ʃ]-biased

!! Interesting Observations
3 speakers (G10, G16 and G9) showed a distribution similar to that predicted for the allophonic group:
• always produced [ʃ] in context of glide /j/
• almost always produced [s] in other contexts
    HOWEVER, G10 showed contrast in the context of back vowels; and G16 and G9 in the environment of codas with plosives.

DISCUSSION AND CONCLUSION

- Evidence was found for at least two possible fricative systems in Gujarati: *near perfect,* and *quasi-phonemic,* and a tendency towards a possible third: *allophonic.*
- However, the quasi-phonemic speakers' patterns do not match the predicted quasi-phonemic system.
- Lack of 'clear' evidence of allophonic system does not imply a lack of this system – monolingual Gujarati speakers (not tested in this study) are more likely to have single sibilant /s/, due to minimal exposure to languages like Hindi and English, which contrast the two sibilants
- Gujarati does not have a single fricative system. The sibilants are not clearly contrastive or allophonic, they have an intermediate phonological relationship. The PPRM, which uses probabilistic rather than binary measures, successfully calculates the exact degree to which the sibilants are predictably distributed..
- The production study taps into the *use* of grammatical knowledge. To confirm the *precise detail* of the sibilant grammar, a future study should also include a perception experiment, with a discrimination task.

REFERENCES

- Adenwala, M. (1965). The structural analysis of the phonology and morphemics of Gujarati. (Master's thesis). Available from State University of New York at Buffalo.
- Dave, R. V. (1977). Studies in Gujarati Phonology and Phonetics . (Doctoral dissertation). Available from Cornell University.
- Grierson, G. A. (1931). On the modern Indo-Aryan Vernaculars . Bombay, 1931-33.
- Hall, K. C. (2009). A Probabilistic Model of Phonological Relationships from Contrast to Allophony . (Doctoral dissertation). Available from The Ohio State University.
- Masica, C. P. (1991). The Indo-Aryan languages. Cambridge; New York: Cambridge University Press.
- Pandit, P.B. (1954). *Indo-Aryan sibilants in Gujarati*. Indian Linguistics, 14, 503-511.
- Turner, R. L. (1921). Gujarati Phonology . The Journal of the Royal Asiatic Society of Great Britain and Ireland, 3, 329-365.