Exploring the front fricative contrast in Greek: A study of acoustic variability based on cepstral coefficients

Jason Lilley / Laura Spinu / Angeliki Athanasopoulou

Nemours Biomedical Research / University of Toronto / UC San Diego

laura.spinu@utoronto.ca

Abstract

We explore the factors underlying the difficulty in acoustic classification of front fricatives by taking a closer look at acoustic variability in the production of Greek speakers. We apply a novel classification tool based on cepstral coefficients in order to classify front fricatives from an experimental corpus with 29 subjects, and employ statistical methods to classify the place of articulation. Our method yields the best correct classification rates reported to date with front fricatives.

BACKGROUND: FRONT FRICATIVES

▷ Classification methods based mainly on traditional acoustic measures for fricatives (e.g. spectral peak location, amplitude, and duration) generally yield lower success rates with front fricatives [f, v, θ, ð], compared to sibilant fricatives [s, z, ʃ, ʒ] in English, e.g. 66% vs. 88% (Jongman et al. 2000).

▷ Acoustic investigation employing both traditional and more innovative measures did not find any cues "even modestly invariant for place of articulation in non-sibilants" (Forrest et al. 1988, Tomiak 1990, McMurray & Jongman 2011, Jongman et al. 2000, Nissen & Fox 2005, Kong et al. 2014).

▷ Articulatory study using magnetic resonance imaging (Narayanan et al. 1995): labiodental fricatives exhibited the most variability across speakers, "the vocal tract and tongue shapes for the labiodentals exhibited wide variabilities. Hence, it is not possible to posit generalized aerodynamic characteristics for the labiodentals with the currently available data."

OUR STUDY: EXPAND CLASSIFICATION METHOD TO GREEK

Goal: investigate if a cepstral coefficient-based classification method, previously successful with obstruents (Bunnell et al. 2004), vowels (Ferrage & Pellegrino 2010) and Romanian fricatives (Spinu & Lilley, 2016) is also successful with fricatives from Greek.

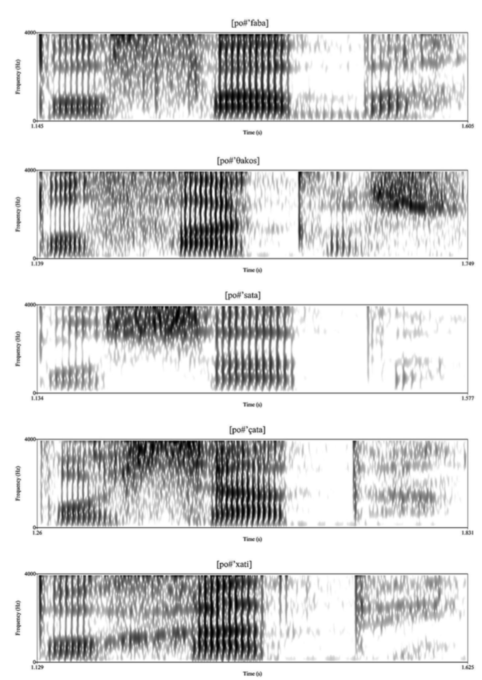
In this study we analyze data from Greek fricatives from five places of articulation and two voicing values.

We tackle the challenging front fricative classification problem.

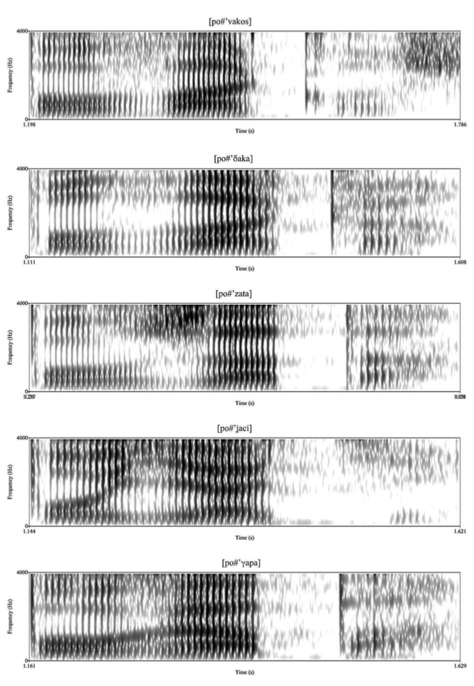
METHOD: PRODUCTION EXPERIMENT

Stimuli: 5 places of articulation, voiced and voiceless

	labial	interdental	dental	palatal	velar
voiceless	f	θ	s	ç	x
voiced	v	ð	z	ʝ	ɣ



(a) Voiceless fricatives.



(b) Voiced fricatives

Vowels:	each fricative matched with each of the vowels [a], [o], [u]
Word type:	60 pseudo-words, two for each pair of fricative+V
Form:	CVCV or CVCVC depending on suffix used
Position:	end of sentence

(c) Word properties

Experimental Procedure

1. Program - familiarization (Invtool/ModelTalker - Bunnell et al. 2007)
2. Words - familiarization
3. Test phase

Subjects

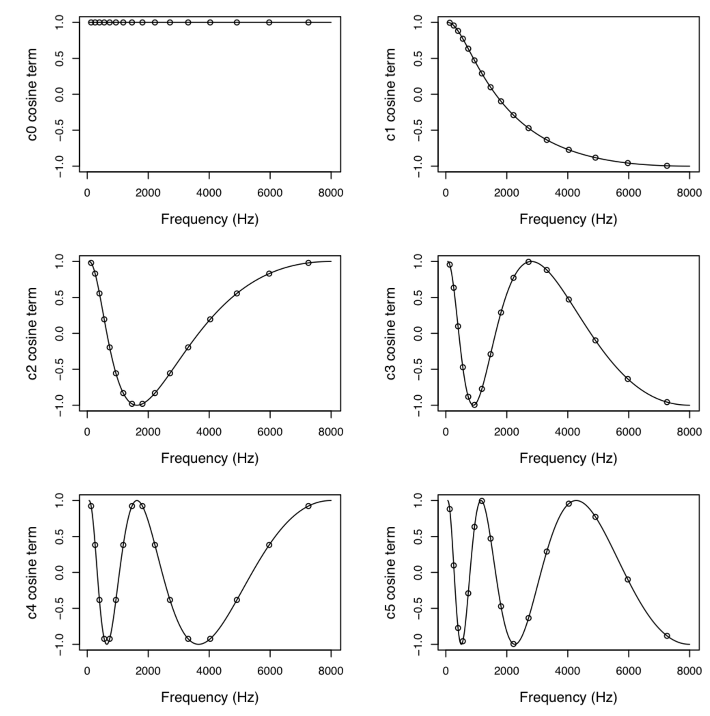
- 29 monolingual native speakers of Greek, male and female
- Speakers of the standard dialect
- Age range 20.8 - 26.6 years with a mean of 23.8 years
- Recruited at the University of Ioannina

Total number of fricatives for analysis: 1,651

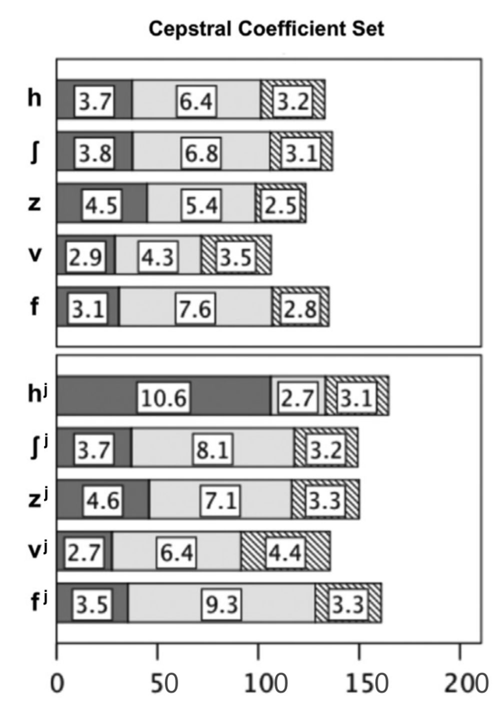
ANALYSIS: CEPSTRAL COEFFICIENTS (CC) AND HMM REGIONS

Fricative boundaries automatically aligned w/HMMs; 30% hand-adjusted.

Six Bark-frequency cepstral coefficients (CC 0-5) were extracted from 20-ms-wide Hamming windows spaced 10 ms apart (Forrest et al., 1988; Halberstadt & Glass, 1997). Hidden Markov Models (HMMs) were used to divide the fricatives & adjacent vowels into 3 regions of internally minimized variance. Mean CCs from each region were then used for classification (logistic regression).



(a) Cosine waves used to calculate cc 0-5.



(b) HMM-based regions (Spinu & Lilley 2016).

RESULTS

Table 1: *Correct classification of place (%). Overall correct: 89.0%*

	Dental	Interdental	Labiodental	Palatal	Velar
Dental	**93.6**	1.8	0.6	2.7	1.2
Interdental	2.5	**81.3**	13.5	0.0	2.8
Labiodental	1.2	12.2	**82.6**	0.3	3.7
Palatal	2.7	1.2	0.6	**93.4**	2.1
Velar	0.6	1.5	3.0	0.9	**93.9**

Table 2: *VOICED ONLY. Correct classification of place (%). Overall: 96.4%*

	Dental	Interdental	Labiodental	Palatal	Velar
Dental	**100.0**	0.0	0.0	0.0	0.0
Interdental	0.6	**90.0**	8.1	0.0	1.3
Labiodental	0.0	5.4	**92.8**	0.0	1.8
Palatal	0.0	0.0	0.0	**99.4**	0.6
Velar	0.0	0.6	0.0	0.0	**99.4**

Table 3: *VOICELESS ONLY. Correct classification of place (%). Overall: 96.4%*

	Dental	Interdental	Labiodental	Palatal	Velar
Dental	**98.8**	0.0	0.0	0.0	1.2
Interdental	0.6	**88.6**	9.6	0.0	1.2
Labiodental	0.0	12.3	**87.7**	0.0	0.0
Palatal	0.0	0.0	0.6	**98.8**	0.6
Velar	0.0	0.0	0.6	0.0	**99.4**

Table 4: *Front fricatives only.*

ALL	Interdental	Labiodental	**VOICED**	Interdental	Labiodental	**VOICELESS**	Interdental	Labiodental
Interdental	**86.5**	13.5	Interdental	**95.0**	5.0	Interdental	**88.0**	12.0
Labiodental	13.1	**86.9**	Labiodental	2.4	**97.6**	Labiodental	11.7	**88.3**

Dimension reduction. The top 5 predictors alone (based on ANOVA) yielded approximately 70% overall correct classification of front fricatives.

CONCLUSION

Best ever reported correct classification rate for front fricatives.

Narayanan et al. (1995): "The variations in the labiodentals are not surprising: the tongue, which is the principal articulator for the other fricatives, is relatively unrestricted for the labiodentals".

Next: cross-validation and unsupervised dimension reduction.