

Spectral Moments as a source of speaker discriminant information

Nikita Suthar¹, Peter French²

University of York, UK

nikita.suthar|peter.french @york.ac.uk

Abstract (max 300 words)

Forensic speaker discriminant studies aim to determine features and combinations of features that can differentiate speakers from one another. Formant analysis has previously figured in such studies (Cao & Dellwo, 2019; Jessen, 2008; Nolan F., 1983; Rose, 2002). The current work takes formant analysis further by adding spectral moments to formant centre frequency values, - which have been the mainstay of much previous research to increase the discriminant power of formants (Nittrouer, 1995, Forrest et al.,1988). The study has centred round value of four primary spectral moments, each concerning the distribution of within-formant energy: centre of gravity, standard deviation of the energy variance across the spectrum, skewness, and kurtosis.

Forty-five female Marwari monolinguals from the Bikaner district (India) were recruited. The recordings were collected from spontaneous and non-spontaneous speech and focused on 8 different vowels. Three types of data were collected; (i) a list of 80 words (10 tokens per vowel) that the participants were asked to read aloud; (ii) a picture description task; (iii) free conversation where two participants were paired up and asked to have a natural conversation on a topic of their choice or from a provided list. Marwari language was used as a testbed and in theory the analysis could be carried out with any other language.

Spectral moments were extracted with the help of a Praat script. An ANOVA conducted in R showed significant vowel differences.. Table 1 summarises the discriminant analysis results and the classification rates of the individual features for the wordlist and story data. All investigated spectral measures increased the classification rates by a minimum factor of 2.5, suggesting that spectral moments carry valuable speaker-specific information. While the results show a clear pattern, the exact relationship between spectral moments and human vocal tract needs further exploration.

Acoustic Measures	Wordlist		Story	
	<i>Classification Rate</i>	<i>Times above chance</i>	<i>Classification Rate</i>	<i>Times above chance</i>
F1+F2+F3+F4	15%	6.5 times	11%	4.5 times
COG: F1-F4	15%	6.5 times	12%	5 times
SD F1-F4	9%	3.5 times	9%	3.5 times
Kurtosis F1-F4	8%	3 times	9%	3.5 times
Skewness F1-F4	7%	3 times	7%	2.5 times

Table 1 Discriminant Analysis of Spectral Moments and Centre Formant Frequencies for Wordlist and Story Data

[Abstract Word count: 293 (excluding table)]

Reference:

- Cao, H., & Dellwo, V. (2019). The role of the first five formants in three vowels of mandarin for forensic voice analysis. Melbourne: *International Congress of Phonetic Sciences*, 617-621. Retrieved June 2021, from University of Zurich: <https://www.zora.uzh.ch/id/eprint/177494/>
- Forrest, K., Weismer, G., Milenkovic, P., & Dougall, R. N. (1988). Statistical analysis of word-initial voiceless obstruents: Preliminary data. *The Journal of the Acoustical Society of America*, 84(1), 115–123.
- Jessen, M. (2008). Forensic phonetics. *Language and Linguistics Compass*, 2(4), 671–711.
- Nittrouer, S. (1995). Children learn separate aspects of speech production at different rates: Evidence from spectral moments. *Journal of the Acoustical Society of America*, 97(1), 520–530. <https://doi.org/10.1121/1.412278>
- Nolan, F. (1983). *The phonetic bases of speaker recognition*. Cambridge: Cambridge University Press.
- Rose, P. (2002). *Forensic Speaker Identification*. London and New York: CRC Press.