Multidimensional Scaling Representation of Speech Sound Dissimilarities

Leonardo Araujo, Hani Yehia and Thaïs Cristófaro

Contribution

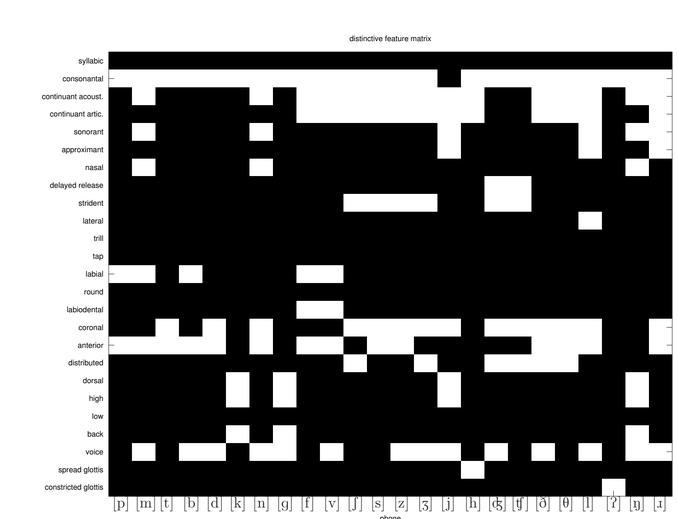
We introduce a method for measuring dissimilarity between phones using their distinctive features. The proposed measure is used to create a dissimilarity matrix and then a multidimensional scaling (MDS) procedure is used to achieve a suitable representation of the phones in a vector space.

Dissimilarity

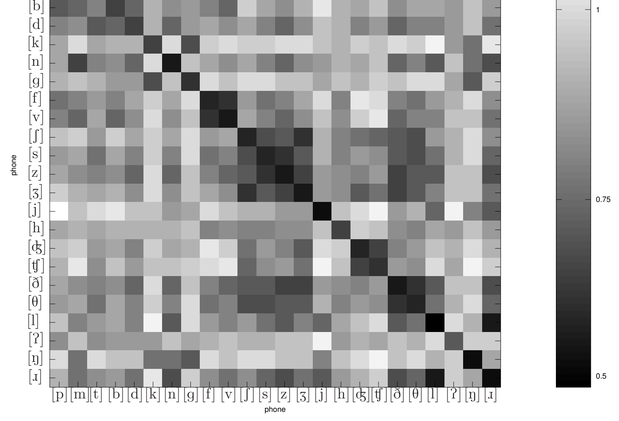
There are many different ways to express the distance between objects and stimuli. It is important then to choose a suitable metric which leads to a meaningful description of a data space, as wrong descriptions of facts may lead to false results and wrong interpretations.

The distinctive feature theory[1] provides a way of characterizing speech sounds based on articulatory, acoustical and perceptual attributes. In this theory, there is a unique representation, based on presence or absence of features, for each speech sound. We use here the theory of distinctive features to create a measure of dissimilarity: a distance measure between two segments defined as the number of features that they do not share.

The figure below shows the distinctive features for the English Language[5],



and the next figure presents the dissimilarity matrix created from the distinctive features shown above.



Multidimensional scaling

A multidimensional scaling (MDS) consists of finding a representation of objects in a vector space such that the distance between those representations are in accordance with the dissimilarities presented by the input matrix. "This configuration reflects the 'hidden structure' in the data, and often makes the data much easier to comprehend." [2]

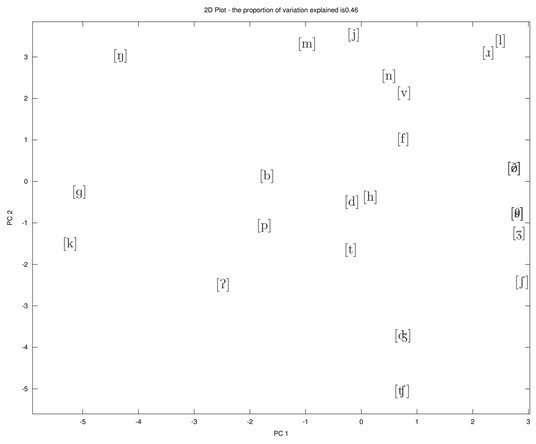
Comparison

Miller and Nicely[4] presented an experiment where a series of stimuli, sixteen English consonants, were presented to subjects who were forced to guess them at every sound. The stimuli were presented under different filtering (low and high pass) and various signal-to-noise ratios (SNR). The results were given as a confusion matrix.

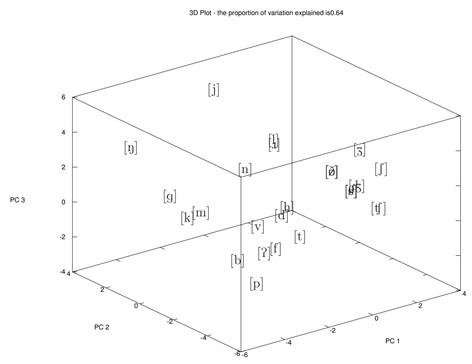
The results from Miller and Nicely were used to create a dissimilarity matrix. An MDS was also performed in order to compare these results with those achieved through the feature theory approach.

Results

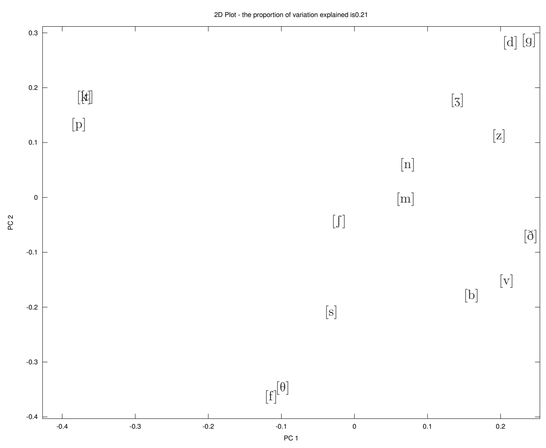
Figures (a) and (b) present MDS results from the distinctive feature data for the English consonants. Figures (c) and (d) show results obtained from the data of Miller and Nicely[4]. (e), (f), (g), (h), (i) and (j) are MDS graphics created from the distinctive feature data for vowels of different languages.



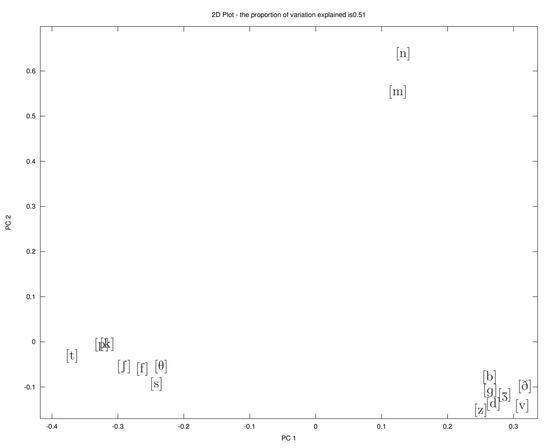
a) English Consonants (2D plot)



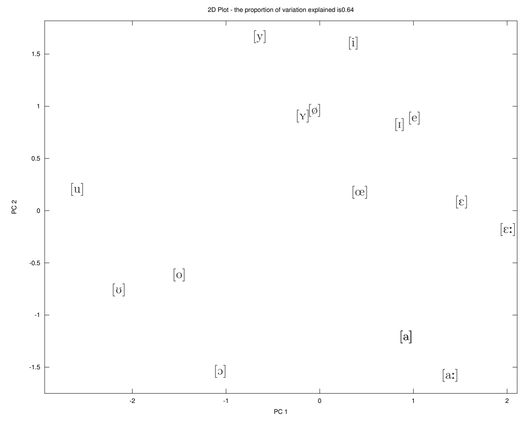
b) English Consonants (3D plot)



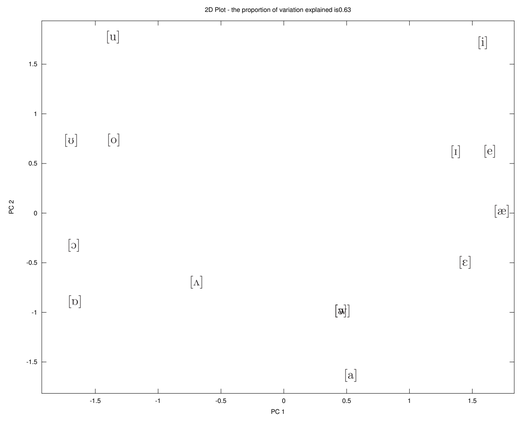
c) 16 English Consonants (Miller and Nicely)



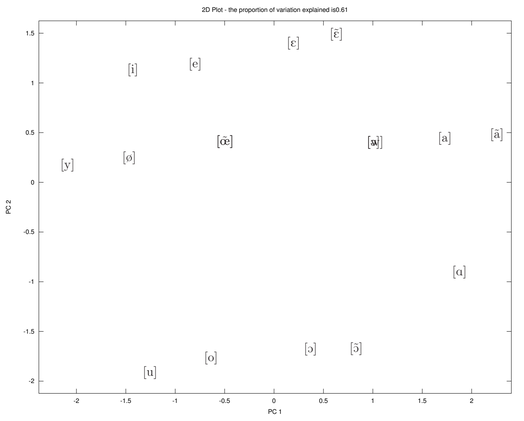
d) 16 English Consonants (Miller and Nicely)



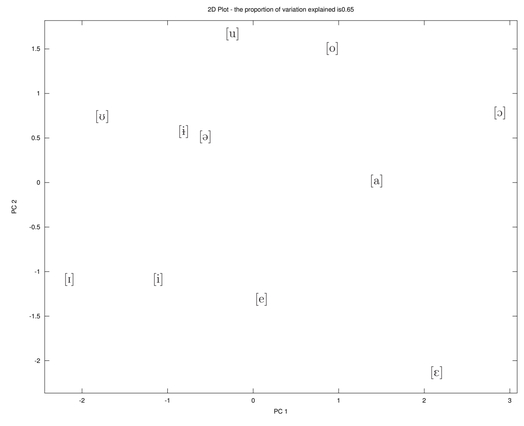
e) German Vowels



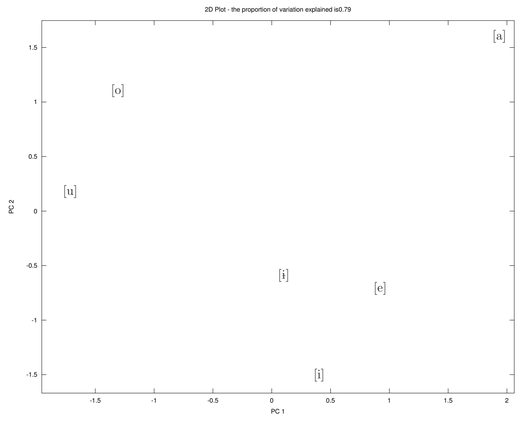
f) English Vowels



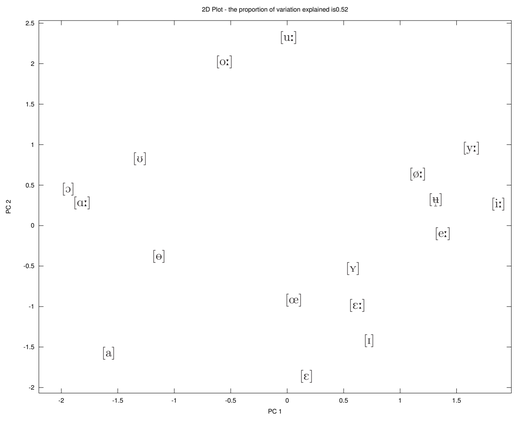
g) French Vowels



h) Portuguese Vowels



i) Russian Vowels



j) Swedish Vowels

References

[1] N. Chomsky, M. Halle. The Sound Pattern of English *Harper and Row, (1968)*

[2] J. B. Kruskal, M. Wish. Multidimensional Scaling *Sage (1978)*

[3] G. Young, A. Householder. Discussion of a set of points in terms of their mutual distances In *Psychometrika 3:19-22 (1938)*

[4] G.A. Miller, P.E. Nicely. An Analysis of Perceptual Confusions Among Some English Consonants In *The Journal of the Acoustical Society of America 27, 2 : 338-352 (1955)*

[5] UCSB and the Department of Linguistics. `http://www.linguistics.ucsb.edu/projects/featuresoftware/` Accessed in June 2010.

Funding

This work was supported by CNPq.