Measuring the optimization of vowel spaces

A method for cross-linguistic analysis

INTRODUCTION

Simulations of the emergence and evolution of phonological systems have shown that, given sufficient time, organizations of the articulatory space emerge in which the phonemes are maximally distinctive (e.g. de Boer, 2000; Oudeyer, 2005; de Boer & Zuidema, 2010). However, there has been little investigation into the typological description of phonological optimization across the world's languages. It is not known, for example, how optimized natural vowel spaces actually are, or whether the vowels of, for example, English are more or less distinctive than those of, for example, Swahili. Here, I introduce a methodology for measuring exactly this.

METHODS

A convenient way to observe the spatial relationships between a set of vowel sounds is to plot the vowel's first and second formant frequencies on reversed logarithmic axes (Peterson & Barney, 1952). Here we use this basic idea to observe the spatial distribution of vowel sounds in a language.

Data collection

Audio recordings were downloaded from the UCLA Phonetics Lab Archive for acoustic analysis. A sample of 70 languages was selected at random, and the Praat software application was used to extract the formant frequency data for each vowel in each of the 70 languages.

Geographical distribution of languages in the sample.

Transformation to a psychoacoustic scale

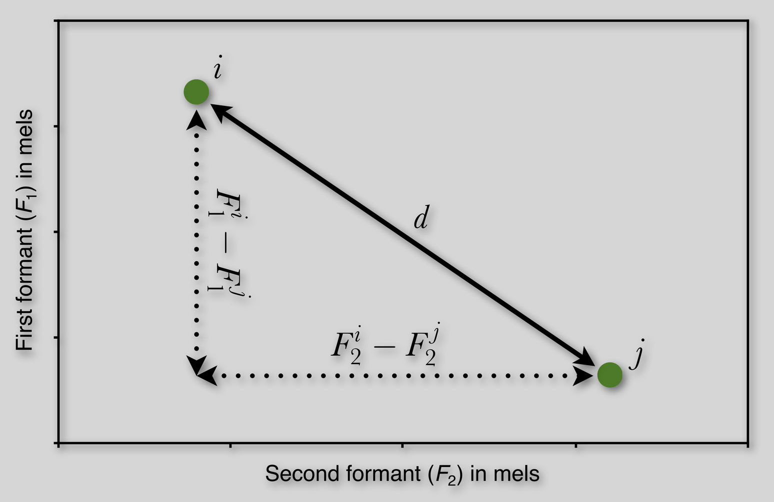
The vowels' formant frequencies were transformed to a psychoacoustic scale. This is necessary because the human auditory system works logarithmically, such that high frequency sounds appear closer together than low frequency sounds. Here we used the mel scale (Stevens, Volkmann, & Newman, 1937) but other scales (Bark, etc.) give the same final results.

Measuring the distance between vowels

The Euclidean distance between a pair of vowels i and j can be calculated by using the formant values as Cartesian co-ordinates in two-dimensional space. Thus,

$$d = \sqrt{(F_1^i - F_1^j)^2 + (F_2^i - F_2^j)^2}$$

gives the distance d between vowel i (F_1^i, F_2^i) and vowel j (F_1^j, F_2^j). Since the third formant (F_3) is also important in the perception of vowels, the vowel space was extended into three dimensions, with F_3 plotted on the z-axis, in order to capture the three most salient vowel properties.



The Euclidean distance d between vowels i and j.

Measuring optimization

Following Liljencrants and Lindblom (1972), the inverse-square law from theoretical physics is used to get a sense of how optimally distributed the vowels are. The optimal state is the one in which the potential energy in the system is minimized. Assuming that this law can accurately model vowel systems (in the sense that vowel systems seek to maximize perceptual contrast), the potential energy in a vowel system can be calculated thus:

$$E = \sum_{i=1}^{n-1} \sum_{j=0}^{i-1} \frac{1}{\sum_{f=1}^{3} (F_f^i - F_f^j)^2}$$

where energy E is the sum of the inverse of the square of the Euclidean distance between vowels i and j for every possible pairing of n vowels. When the total energy in a vowel system is minimized, the optimization of the vowel system is maximized, and vice versa. Therefore, by taking the inverse of E, we derive a number that directly corresponds to vowel space optimization.

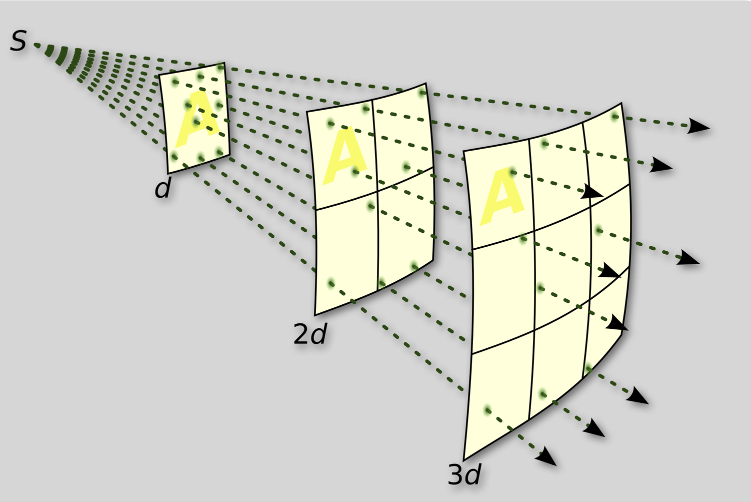
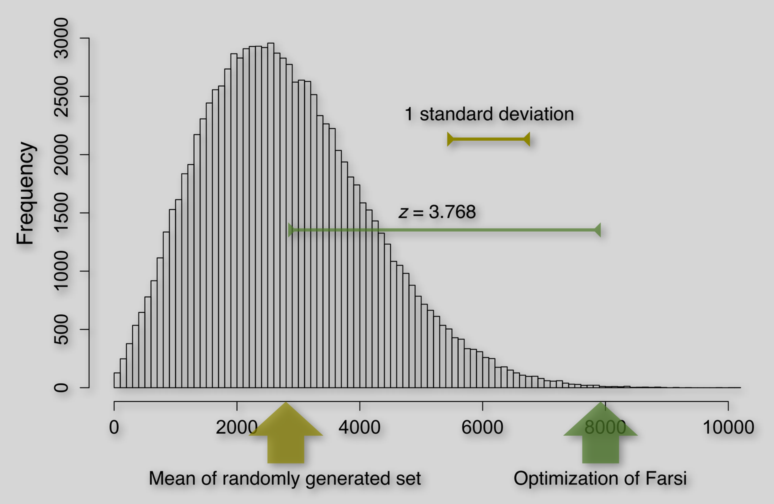


Illustration of the inverse-square law. The force of repulsion decreases with the square of the distance from the source.

Standardization of the measure

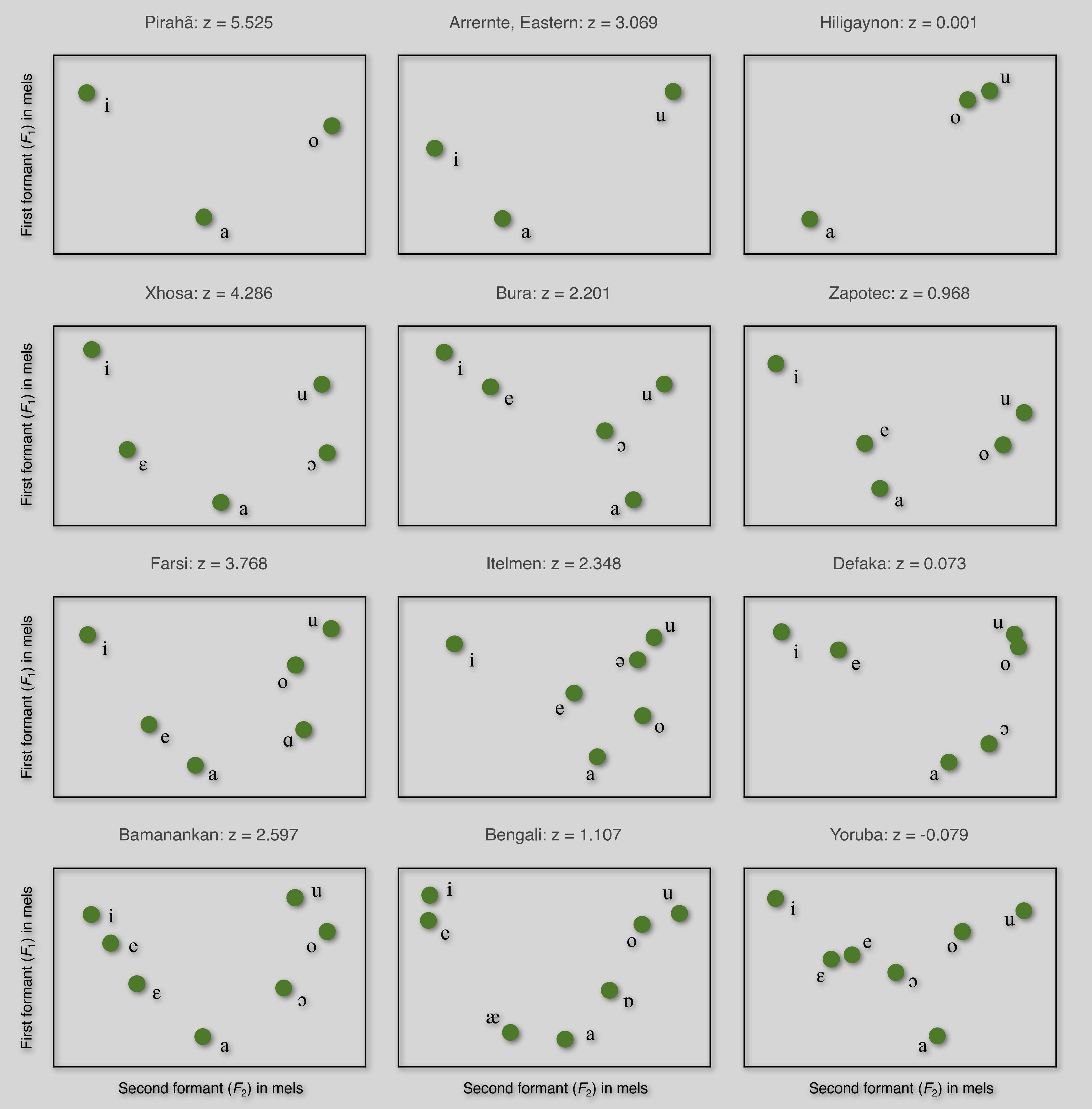
To standardize this measure, and test for statistical significance, a Monte Carlo technique was used. For a given language, we generate 100,000 randomized vowel sets, and calculate the optimization of each set as described above. We then calculate a standard score (z-score) by comparing the natural vowel set against the mean and standard deviation of the randomized ones. A z-score greater than 0 suggests that the vowels are further apart than one would expect by chance.



Distribution of optimization scores for 100,000 randomly generated vowel systems. The natural vowel system (Farsi) is significantly more optimized than the generated ones.

RESULTS

Overall, there was a high degree of variation in levels of vowel space optimization; the z-scores ranged from -0.969 for the Azerbaijani language to 6.144 for the Nyangumarta language (mean = 2.087, SD = 1.556). The plots in the panel below show a sample of 12 of the 70 languages analysed. The optimization of each vowel system is reflected in its associated z-score (N.B. the 3D vowel spaces are flattened into 2D for illustration).



DISCUSSION

The method described here captures two key properties of an optimized system: effectiveness and order. The inverse-square law tells us how effective (i.e. how perceptually distinctive) the distribution of vowels is given the finite space in which they exist. The application of Monte Carlo techniques tells us how ordered (i.e. how non-random) the vowel system is by comparing the natural system against ones which are known to be stochastic in nature. Furthermore, the optimization score does seem to intuitively fit with what unoptimized and optimized vowel spaces ought to look like.

This research has demonstrated that it is possible to measure the optimization of natural vowel systems – which, to my knowledge, has not been attempted previously. But what kind of uses could such a measure have? Coming at language from an evolutionary perspective, I am interested in how culture and language impact upon each other. For example, Lupyan and Dale (2010) have recently shown that demographic properties, such as population size, correlate with the level of grammatical structure in a population's language. I am interested in looking for similar correlations between phonological structure and culture. Nevertheless, I imagine there might be other potential uses in linguistics for a measure of vowel system optimization.

Without being able to visualize linguistic data in spectrograms, vowel plots, histograms, and schematic diagrams, determining a method such as this would be infinitely more difficult, and so this poster demonstrates the importance of visualization in linguistics.

REFERENCES

de Boer, B. (2000). Self-organization in vowel systems. *Journal of Phonetics*, *28*, 441–465.

de Boer, B., & Zuidema, W. (2010). Multi-agent simulations of the evolution of combinatorial phonology. *Adaptive Behavior*, *18*, 141–154.

Liljencrants, J., & Lindblom, B. (1972). Numerical simulation of vowel quality systems: The role of perceptual contrast. *Language*, *48*, 839–862.

Lupyan, G., & Dale, R. (2010). Language structure is partly determined by social structure. *PLoS ONE*, *5*, e8559.

Oudeyer, P.-Y. (2005). The self-organization of speech sounds. *Journal of Theoretical Biology*, *233*, 435–449.

Peterson, G. E., & Barney, H. L. (1952). Control methods used in a study of the vowels. *Journal of the Acoustical Society of America*, *24*, 175–184.

Stevens, S. S., Volkmann J., & Newman, E. (1937). A scale for the measurement of the psychological magnitude pitch. *Journal of the Acoustical Society of America*, *8*, 185–190.

This research was supported by a Carnegie-Cameron Taught Postgraduate Bursary.

Many thanks are due to Prof. Jim Hurford and Dr Hannah Cornish for their useful suggestions.

Jon William Carr
Language Evolution and Computation Research Unit
School of Philosophy, Psychology and Language Sciences
University of Edinburgh
j.w.carr@sms.ed.ac.uk