Towards a measure of optimization in natural vowel systems

INTRODUCTION

Studies looking at the emergence and evolution of phonological systems have shown that, given sufficient evolutionary time, organizations of the articulatory space emerge in which phonemes are maximally distinctive (e.g. Steels, 1997; de Boer, 2000; Oudeyer, 2005; de Boer & Zuiddam, 2010). However, there has been little investigation into the typological description of articulatory 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. The aims of this research were to demonstrate (a) that such a measure is possible, and (b) that this measure could have practical uses within Evolutionary Linguistics.

METHODS

To develop a method for measuring vowel-space optimization, it is necessary to have a representative of how the vowels of a language relate to each other within the space formed by the oral cavity. Peterson and Barney (1952) showed that when the first and second formant frequencies of a set of vowel sounds are plotted on reversed logarithmic axes, a visualization of the vowel-space emerges. Plots of this kind thus provide a convenient way to observe the spatial relationships between a language’s vowel sounds.

Data collection

Since formant frequency data for the world’s languages are not readily available, the data had to be collected manually. Audio recordings were downloaded from the UCLA Phonetics Lab Archive (Lade-Foged & Blankenship, 2007) for acoustic analysis. A sample of 70 languages was selected at random, and the Praat software application (Boersma & Weenink, 2011) was used to extract the formant frequency data for each vowel in each of the 70 languages for a total of 415 vowels.

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. The mel scale (Stevens, Volkman, & Newman, 1937) was chosen simply because its calculation is the least computationally expensive; other scales (Bark, etc.) give the same final results. (Stevens, Volkman, & Newman, 1937) was chosen simply because its calculation is the least computationally expensive; other scales (Bark, etc.) give the same final results. Consequently, the vowels appear more or less distinctively than those of, for example, Swahili. Here, I introduce a methodology for measuring exactly this. The aims of this research were to demonstrate (a) that such a measure is possible, and (b) that this measure could have practical uses within Evolutionary Linguistics.

RESULTS

There is a high degree of variation in vowel-space optimization, which ranges from 0.0696 in the Azerbaijani language to 1.644 for the Nagasamartan language (mean = 2.087, SD = 1.556). See the plots of variation in Table 1. Optimization inversely correlates with vowel inventory size (r = -0.514, p < 0.000006), such that languages with large inventories of vowels tend to be less optimized. This should be obvious, since as we add more vowels into a system, the energy forcing them apart will increase.

DISCUSSION

This research has demonstrated that it is possible to construct a measure of vowel-space optimization — which to my knowledge has not been attempted previously. Such a measure may have practical uses in several areas of Linguistics. However, more work will be required to make more robust conclusions. In particular, these 70 languages represent just 1 or 2 per cent of global linguistic diversity.

Strengths

The optimization score does seem to intuitively fit with what unoptimized and optimized vowel-spaces ought to look like, and I therefore very tentatively suggest that this method could be of use to future linguistic research. Although further work will be required to improve on this method, it could have practical implications for understanding how vowel systems evolve in order to adapt to changing environmental, social, and cognitive demands.

Challenges

Firstly, the raw data used in this study are inherently fuzzy – speakers do not consistently produce vowels with precisely the same formant frequencies, which could amount to significant inaccuracies that skew the results. Secondly, the simulations assume that the vowels of the natural language delineate the maximum space available for a given speaker. This could introduce ceiling effects such that the randomized languages may only make use of a subset of the full articulatory space that is realistically possible. Thirdly, it is currently difficult to make comparisons between languages with different inventory sizes because the score is not inventory-size neutral.

Main conclusions

This paper has demonstrated a method for measuring the level of optimization in the perceptual vowel-spaces of natural languages, which could be of use to future linguistic research. Although further work will be required to improve on this method, it could have practical implications for understanding how vowel systems evolve in order to adapt to changing environmental, social, and cognitive demands.

REFERENCES