Three Men Walk into A Bar: Quantifying Phonological Distance Between Languages

What is similarity? It's not that we do not know what similarity is. We do know, but we do not know what we know. For example, is Korean more similar to Japanese or Mandarin? Can you explain your choice?

Many researchers have studied the similarity between languages, yet none of them has quantified this similarity. The main goal of this research is to build a single, universal similarity metric for all the phonological properties that speakers may use to determine similarity between languages. In order to reach this goal, two experiments were conducted with Hebrew speakers.

In the preliminary experiment, 132 Hebrew speakers rated their familiarity with a diverse group of selected languages, and the results were used in the analysis of the main experiment.

Before the main experiment, the degree of similarity between the selected languages was quantified using the proposed metric, by dividing the number of shared phonological properties by the total number of phonological properties examined in the experiment. The properties were both segmental (e.g., the appearance of retroflex consonants, the appearance of nasalized vowels) and prosodic (e.g., the consonant-vowel ratio, rates of speech in syllables-per-second).

362 subjects participated in the main experiment, in which participants determined which of the selected languages were more similar to one another. The results show that the proposed metric based on the shared number of properties largely predicts the participants’ determination of similarity between languages (i.e., they choose the more similar answer) – a correct outcome in 74.4% of the cases.

The talk will also discuss the remaining 25.6%, as well as some other possible non-phonological factors.

Click here to see the colloquium program for the fall-winter semester

YOU ARE ALL INVITED