A computational survey of gaps in obstruent inventories

Sheng-Fu Wang, New York University

Introduction Sound inventories of human languages are not made up of random bags of possible speech sounds. Proposals on inventory shapes such as feature economy and symmetry (e.g., Clements, 2003) predict that a stop inventory like /p, t, b, d/ is more likely than /p, t, b/ or /p, t, b, g/. Proposals on markedness and implicational universals (e.g., Greenberg 1970; Gamkrelidze, 1975) also make predictions on inventory shapes. For example, when a language lacks a voiced stop, it is more likely to be /g/ than /d/.

This study compares different models on organizing principles of inventory structure by examining the distribution of gaps in the inventories of obstruents across languages. A gap is referred to as the absence of an [α voice] stop/fricative in a certain place of articulation when a [−α voice] counterpart exists in the inventory. The computational task is a binary choice between the gap and an attested sound, which is referred to as a 'foil'. A foil is an attested stop/fricative that share the [α voice] feature with the gap, with a different place of articulation. A model is more successful if it more frequently identifies the gap correctly. Example pairs of gaps and foils are shown in (1)

(1) Wogeo obstruent inventory: /b, d, g, t, k, v, f, s/ gap-foil pairs: /p/-/t/, /p/-/k/, /z/-/v/

Two major types of models are tested: The MARKEDNESS and the FEATURE-SYSTEMIC models. The MARKEDNESS models predict that the gap is always more marked than the foils. Two types of MARKEDNESS models are tested. The *grounded markedness* model ranks the markedness of obstruents based on their constriction site (Gamkrelidze, 1975): A voiced obstruent is more marked when the constriction site is further back in the vocal tract. Conversely, a voiceless obstruent is more marked if it is fronter. The *typological frequency* markedness model uses the frequency of segments across inventories for ranking markedness: the less frequent sounds are more marked.

The FEATURE-SYSTEMIC models select the gap based on the following criterion: between two sounds, the gap is the sound that decreases the overall goodness of the inventory based on some feature-based metrics. Three metrics are used in this study: *feature entropy* (Mukherjee et al., 2007), *local feature symmetry*, and *global feature symmetry* (Dunbar and Dupoux, 2016). *Feature entropy* is taken as a measure of feature economy: the feature representation of an inventory is more economical if it can be expressed by fewer bits. *Local feature symmetry* measures the number of pairs of sounds in an inventory that differ only in one feature: an inventory is more locally symmetrical if it has more such pairs. *Global feature symmetry* measures the difference in the size of the plus and the minus feature values in a feature system: an inventory is more globally symmetrical if the difference is smaller.

To further investigate to what extent the place of articulation of gaps can be learned from the distribution of segments across inventories, artificial neural networks are trained to do the gap-prediction task. The input contains inventories represented as bags of segments, which in turn are represented as bags of feature values. The training objective to either choose a gap from two sounds given the knowledge of the inventory at issue (*Inventory* model), the gap and the foil (*Segment* model), or both (*Inventory*+*Segment* model).

Method This study uses the PHOIBLE database of phoneme inventories (Moran et al., 2014), which contains 2155 inventories, where segments are described by a phonetically detailed feature set. Only [-soronant] sounds are used in this study. After filtering out repetitive inventories, 1874 obstruent inventories remain, where gaps and foils are identified. Data points are generated by pairing a gap with a foil. Overall, there are 4705 data points, which are split into the training/development set (3714) and the test set (991). Test set results are used to compare models.

Results The MARKEDNESS models are more successful in accounting for the place of articulation of gaps than FEATURE-SYSTEMIC models. In the test set, the segment-based typological frequency model has a 68.9% accuracy, the place-based frequency model has a 61.5% accuracy, and the grounded markedness model has a 60% accuracy. All FEATURE-SYSTEMIC models perform around chance level. The best models are *Inventory+Segment* and *Segment* network models, where the accuracies are 73% and 75.6%.

Figure 1 shows model performance with different sound types. The *grounded markedness* has the most success in voiced stops, followed by voiceless stops. The better performing models, the *Inventory+Segment* and the *Segment* network models and the *segmental frequency* model have similar patterning, having more success with voiceless sounds than with voiced sounds.

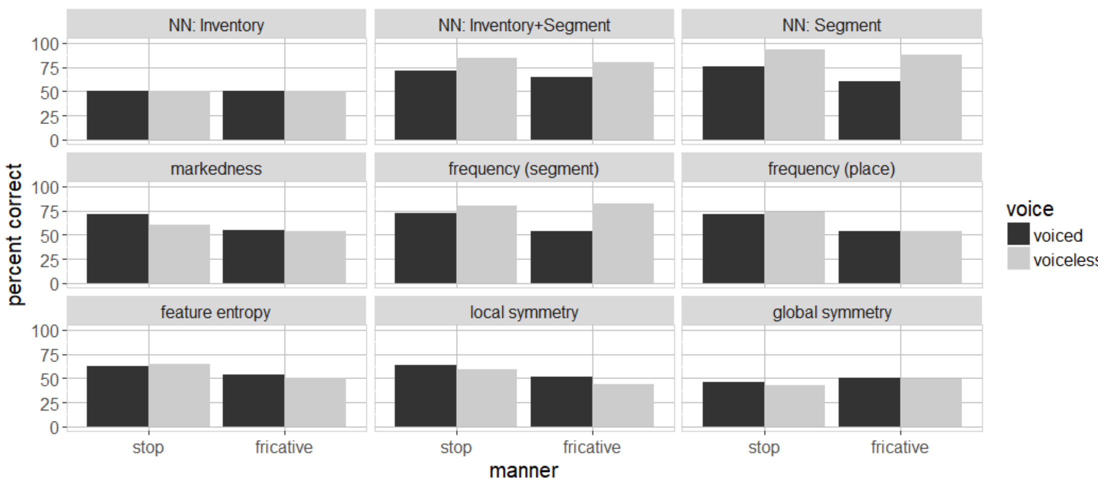

Figure 1: Results: model performance as a function of sound types

Discussion When a language lacks an [α voice] obstruent and the [−α voice] counterpart exists, the absent sound is likely to be more marked than an attested [α voice] obstruent with a different place of articulation. The markedness can either be grounded phonetic markedness crudely defined by the interaction of voicing and constriction site, or typological segment/place frequencies. On the other hand, the at-chance performance by the FEATURE-SYSTEMIC models show that whether the gap occurs at one place of articulation rather than another is not driven by a force to make the whole inventory more economical or symmetrical in its feature system. The success of the *Inventory+Segment* and *Segment* networks over the *Inventory* network also shows that segmental properties are more relevant in predicting gaps than inventory-level information. In other words, whether a segment is likely to be gapped or not depends more on the segment itself rather than the characteristics of the inventory.

The relatively worse performance of the FEATURE-SYSTEMIC models suggests that these metrics may only encode a preference for gapless inventories, but are agnostic to where the gaps are. A supplement analysis that compares gapped and gapless inventories, as defined in this study, shows that gapless inventories have more economical and globally symmetrical feature representations.

References • Clements, G. N. (2003). Feature economy in sound systems. *Phonology*. • Dunbar, E. and Dupoux, E. (2016). Geometric constraints on human speech sound inventories. *Frontiers in psychology*. • Greenberg, J. H. (1970). Some generalizations concerning glottalic consonants, especially implosives. *International Journal of American Linguistics*. • Gamkrelidze, T. V. (1975). On the correlation of stops and fricatives in a phonological system. *Lingua*. • Moran, S., McCloy, D., and Wright, R. (2014). PHOIBLE online • Mukherjee, A., Choudhury, M., Basu, A., and Ganguly, N. (2007). Modeling the co-occurrence principles of the consonant inventories: A complex network approach. *Int. Journal of Modern Physics C.*