

Letter shapes phonology: Feature economy and informativeness in 43 writing systems

Yoolim Kim

Wellesley College, Wellesley, Massachusetts, United States

Marc Allasonnière-Tang

CNRS, Paris, France

Helena Miton

Santa Fe Institute, Santa Fe, New Mexico, United States

Olivier Morin

PSL University, Paris, France

Abstract

Differentiating letter shapes accurately is an increasingly crucial competence. Are letters as distinctive as they could be? We used a unique dataset of crowdsourced letter descriptions across 43 writing systems to produce a comprehensive typology of letter shapes for these diverse scripts. We extracted from 19,591 letter classifications, contributed by 1,683 participants, enough features to provide a unique description of all letters in each system. We show that scripts, compared to phoneme inventories, are feature-extensive: they use additional features to do what could be done with a lower number of features, used more efficiently. Compared to 516 phoneme inventories from the P-base dataset, our 43 scripts have lower feature economy (fewer symbols for a given number of features) and lower feature informativeness (a less balanced distribution of feature values). Letter shapes, we argue, having more degrees of freedom than speech sounds, use features in a more wasteful way.