Complexity/informativeness trade-off in the domain of indefinite pronouns

Milica Denic
University of Amsterdam, Amsterdam, Netherlands

Shane Steinert-Threlkeld
University of Washington, Seattle, Washington, United States

Jakub Szymanik
University of Amsterdam, Amsterdam, Netherlands

Abstract

The vocabulary of human languages has been argued to support efficient communication by optimizing the trade-off between complexity and informativeness (Kemp and Regier, 2012). The argument has been based on cross-linguistic analyses of vocabulary in semantic domains of content words such as kinship, color, and number terms. The present work extends this analysis to a category of function words: indefinite pronouns (e.g. someone, anyone, no-one, cf. Haspelmath, 2001). We establish the meaning space and feature-based representations for indefinite pronouns, and show that indefinite pronoun systems across languages optimize the complexity/informativeness trade-off. This demonstrates that pressures for efficient communication shape both content and function word categories, thus tying in with the conclusions of recent work on quantifiers (Steinert-Threlkeld, 2019). Furthermore, we argue that the trade-off may explain some of the universal properties of indefinite pronouns, thus reducing the explanatory load for linguistic theories.