

Zipfian distributions facilitate word segmentation in infants

Lucie Wolters

The Hebrew University of Jerusalem, Jerusalem, Israel

Mitsuhiko Ota

University of Edinburgh, Edinburgh, United Kingdom

Inbal Arnon

Hebrew University, Jerusalem, Israel

Abstract

The word-frequency distributions infants hear during language learning are highly skewed (Zipfian). Previous work suggests that such skewed distributions facilitate speech segmentation, a crucial milestone in language acquisition. However, the experimental studies supporting this have only examined individuals aged 10 years or older, and it is not clear whether this effect arises from accumulated linguistic experience or is already present in the early stages of learning. To address this, we ran a word-segmentation study with 8-month-old infants (N=60) by exposing them to a continuous speech stream with a skewed or uniform frequency distribution of artificial words. At test, infants in the skewed condition exhibited a larger looking time difference between heard and unheard words than infants in the uniform condition. These findings suggest that Zipfian distributions can facilitate word segmentation during early linguistic development, and highlight the importance of the distributional characteristics of linguistic input in natural language learning