

Speech Processing does not Involve Acoustic Maintenance

Spencer Caplan

University of Pennsylvania, Philadelphia, Pennsylvania, United States

Alon Hafri

University of Pennsylvania, Philadelphia, Pennsylvania, United States

John Trueswell

University of Pennsylvania, Philadelphia, Pennsylvania, United States

Abstract

What happens to the acoustic signal after it enters the mind of a listener during real-time speech processing? Since processing involves extracting linguistic evidence from multiple, temporally distinct sources of information, successful communication relies on a listener's ability to combine these potentially disparate signals. Previous work has shown that listeners are able to maintain, and rationally update, some type of intermediate representations over time. However, exactly what type of information is being maintained—be it acoustic-phonetic or rather a probability distribution over phonemes—has been underspecified. In this paper we present a perception experiment aimed at identifying the internal contents of intermediate representations in speech processing. Using an accent-adaptation paradigm, we find that listeners adapt to modulated acoustic signal when the corresponding orthography is provided before the audio, but not when audio follows the orthography. This supports the position that intermediate representations are uncertainty-distributions over discrete units (e.g. phonemes) and that, by default, speech processing involves no maintenance of the acoustic-phonetic signal.