

Deriving uniform information density behavior in pragmatic agents

Benjamin Peloquin

Stanford University, Stanford, California, United States

Noah Goodman

Stanford University, Stanford, California, United States

Michael Frank

Stanford University, Stanford, California, United States

Abstract

The combinatorial expressivity of natural language enables speakers to communicate a single idea in myriad ways. How do speakers decide which utterance to use? Under the Uniform Information Density (UID) hypothesis, speakers should plan their utterances to minimize listener comprehension difficulty by spreading out new information, for example, by using complementizers or avoiding contractions before high-surprisal content. We explore how UID behaviors may result from pragmatic considerations (e.g., social reasoning in context) using a computational pragmatics model. We show that artificial pragmatic agents communicating under noise conditions exhibit key UID effects: (A) speakers provide cues before high surprisal content, (B) given a UID-cue, listeners infer oncoming content is high-surprisal, (C) synthetic corpora generated from speakers reflects a signature UID effect: a positive relationship between likelihood of optional elements and surprisal of oncoming content. Thus, UID may follow from more general principles of pragmatic communication in the presence of noise.