Optimization-based modeling of suprasegmental speech timing

Andreas Windmann

I present a computational model of suprasegmental speech timing based on the assumption that speech patterns are shaped by global and local adjustments of trade-offs between conflicting demands of minimizing production effort and maximizing perceptual clarity. The model uses an optimization algorithm to determine durations of suprasegmental constituents of simulated utterances such that a composite cost function is minimized. The cost is a function of the constituent durations and encompasses different components that represent independently derived measures of speaker-based production efficiency, listener-oriented perceptual clarity as well as time conceptualized as a resource shared between both parties.

The trade-offs between these influences can be globally and locally adjusted by weights assigned to individual components within the composite cost function. I show that this approach facilitates modeling a hierarchy of interacting prosodic features of utterances, such as different degrees of prominence or effects of speaking rate and overall requirements of clarity. I outline the theoretical foundations and the architecture of the model and present results of simulation experiments, demonstrating that the model correctly predicts a range of suprasegmental timing phenomena that so far have not been addressed by a unified model.