Language Science Colloquium 2016

Time: 11:00-12:30 pm
Location: 1511 Social and Behavioral Sciences Gateway Building
Lunch will be served.

T. Florian Jaeger

FRI, Dec 2nd

Title: *Consequences of uncertainty for speech production and perception*

Abstract: The talk is structured into two parts.

Perception. Talkers differ in how they map linguistic categories (e.g., phones or words) onto the acoustic signal. These differences do not seem to be reducible to physiological differences. Rather, differences are often conditioned on talker identity, including social factors (gender, age, dialect), and this conditioning can be arbitrary. This creates one of the fundamental challenges to speech perception. Effective understanding relies on good knowledge of the signal's statistics. Yet, these statistics differ by talker. Computationally, this means that listeners do not only need to infer linguistic categories (speech perception), but also the adequate model for the present talker (speech adaptation). The second part is crucial: an inability to learn and store talker-specific statistics would result in severe impairments to speech perception.

I discuss how both of these processes can be productively understood as a form of probabilistic inference under uncertainty. I illustrate how this provides a unifying framework for a large variety of properties of speech perception, including in particular incremental changes in listeners' interpretation of speech from a novel talker. I also present evidence that the systems underlying speech perception can be 'smart', automatically conducting inferences about the causes for unexpected pronunciations, using these inference to effectively generalize future input from the same or similar talkers, etc.

Production. I then illustrate the insights similar reasoning offers for speech production. I revisit the longstanding intuition that speech production trades off two often conflicting goals: the minimization of 'effort' and successful communication. At first blush, psycholinguistic research seems to call this intuition into question. This research has argued that articulatory planning is too complex, requiring automatic processes that are informationally encapsulated from communicative goals. I present evidence to the contrary: at least for healthy speakers, the perceived communicative success of previous articulations—as evidenced by the behavior of their listeners—, affects subsequent productions.

Speakers seem to be able to draw inferences about the likely cause of a miscommunication, adapting articulation in a way targeted at the (inferred) problem. Crucially, these adaptations seem to facilitate comprehension, though this facilitation only becomes apparent once we take into account the computational consequence of noisy motor and perceptual processes: changes to articulation need to be understood with reference to the distribution of articulations that speakers produce (and that listeners use to interpret the acoustic signal).

Research. Human Language Processing Lab, Brain and Cognitive Sciences, University of Rochester. https://www.hlp.rochester.edu/

Research in my lab seeks to understand the chain of processes from intention to articulation, and from acoustic signal to intention recognition. We approach these questions guided by computational frameworks. These frameworks characterize and quantify the problems that the neural systems underlying speech production and perception have to solve. They also allow us to understand the effect of tasks and goals on speech production/perception. My research has so far exclusively focused on healthy individuals, though I hope that this presentation will stimulate conversation about how this work and research on pathologies can inform each other.

I focus on the consequences of *uncertainty* for both speech production and perception. The speech signal is not only noisy, but also subjectively non-stationary (not just the signal, but also its underlying statistics vary over time). Understanding the consequences of this, I hold, is essential in understanding how the brain represents and processes speech.

The case studies I present use web-based crowdsourcing paradigms. These paradigms allow fast and cheap recruitment from a wide range of participant populations.

Selected relevant readings from my lab:

Buz, E., Tanenhaus, M. K., and Jaeger, T. F. 2016. Dynamically adapted context-specific hyper-articulation: Feedback from interlocutors affects speakers' subsequent pronunciations. Journal of Memory and Language 89, 68-86. [10.1016/j.jml.2015.12.009] [pdf]

Kleinschmidt, D. and Jaeger, T. F. 2015. Robust speech perception: Recognizing the familiar, generalizing to the similar, and adapting to the novel. Psychological Review 122(2), 148-203. [10.1037/a0038695] [pdf]

Qian, T., Jaeger, T. F., and Aslin, R. 2016. Incremental implicit learning of bundles of statistical patterns. Cognition 157, 156-173. [10.1016/j.cognition.2016.09.002] [pdf]