From abstract to episodic: Representations in speech production

Jonathan C. Paramore

A growing body of work shows that speakers retain rich phonetic detail and extralinguistic social information in mental representations of speech. At the same time, phenomena such as the sensitivity of phonological productivity to lexical type frequency and the permeation of regular sound changes across the lexicon indicate that speech is stored in relatively abstract categories as well. In this talk, I ask how abstract such categories can be. Drawing on air pressure data from two production experiments on Pakistani Punjabi, I argue that speech representations can be completely *covert*, containing phonological features that never surface faithfully in any allomorph.

In Pakistani Punjabi, predictably nasal vowels occurring before nasal consonants, [ṼːN], are produced with the same degree of nasality as contrastive nasal vowels, /Ṽː/. However, predictably nasal [Ṽː**N**] and contrastive nasal /Ṽː/ behave differently in their interaction with a nasal harmony process, in that only contrastive nasal /Ṽː/ triggers harmony. I argue that this asymmetry requires an analysis in which (i) pre-nasal vowels are underlyingly oral /Vː**N**/, and (ii) only underlyingly nasal /Ṽː/ triggers harmony.

Nevertheless, recent work on phonological learnability argues that URs at this level of abstraction may not be learnable, as they entail a prohibitively large hypothesis space of potential URs. To address this challenge, I propose a novel learning algorithm that structures the search space based on UR→SR disparities, allowing it to be explored efficiently without exhaustive enumeration. This approach mitigates the computational challenges posed by highly abstract URs and enables efficient learning of patterns like those found in Punjabi.