

Facial Expressions and Phonetic Recalibration in Speech Perception

Listeners can adjust and recalibrate their phonetic boundaries based on exposure to new speech input (Norris et al., 2003). In this study, we investigate whether social factors external to the speech signal during exposure can affect this phonetic recalibration. Specifically, we test whether phonetic recalibration is modulated by the facial expression of the speaker. Existing studies show that speech production and perception are dynamically sensitive to social characteristics of the speaker (Niedzielski, 1997; Johnson et al., 1999; Babel 2012, i.a.), but it has not been studied whether perceptual learning (i.e., phonetic recalibration) is similarly sensitive to social factors.

During a training phase, participants were presented auditorily with (i) 60 words with a word-medial /d/ (e.g., *academia*), (ii) 60 with a word-medial /t/ (e.g., *politician*), and (iii) 60 filler words containing neither /d/ nor /t/. An additional set of 180 non-word fillers contained neither /d/ nor /t/. The auditory material was produced by a female native speaker of American English. The task of the participants was to make a lexical decision for the 360 spoken words and non-words. Crucially, the /t/ sounds in the t-words were carefully manipulated – in particular, by shortening VOT and closure length – to be ambiguous between /t/ and /d/, and this manipulation was verified in a separate norming study. The /d/ sounds were not manipulated. During this training phase, a picture of a woman was presented on the screen. In one between-subjects condition (Smile), the woman was smiling; in the other condition (No Smile), the same woman was not smiling. After the training phase, the participants performed a categorization task for tokens on an 11-step /ata/-/ada/ continuum to assess whether their category boundary between /t/ and /d/ had shifted. Since the /t/ sounds in the training are closer to /d/ than usual, if perceptual learning occurs, the category boundary should shift towards the /d/-end of the continuum.

Results from 18 female participants are shown in Figure 1. (Data collection is ongoing and the study will include a total of 32 female and 32 male participants.) Listeners in the No Smile condition showed a positive effect of perceptual learning, in that they tended to choose /t/ more often for higher continuum steps than a control group did ($z = 1.9$, $p = 0.06$), shifting the category boundary to the /d/-end. (The baseline was obtained from a separate group of female participants who did not undergo training.) Listeners in the Smile condition, on the other hand, showed no evidence for perceptual learning ($z = -0.9$, $p = 0.4$).

This finding is somewhat counter to studies on learning that report better learning outcomes with more attractive or likable instructors (Westfall et al., 2016), though Babel (2012) shows that greater likeability and attractiveness can sometimes result in reduced phonetic imitation. The current study provides a novel finding that phonetic recalibration is affected by speech-external social factors, though more research is needed to understand the role of specific facial expressions.

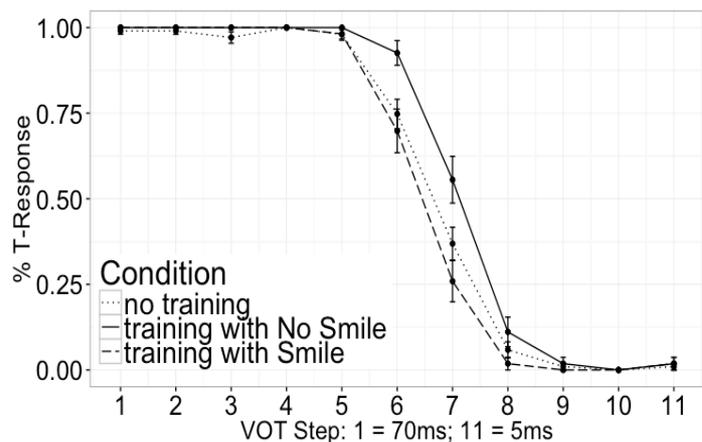


Figure 1. Preliminary results of the post-training categorization on an /ata/-/ada/ continuum