Reshaping Phonetic Categories: Adaptation to Phonetic Variation in Nonnative-accented Speech

Purpose

Unfamiliar speech variation, for instance, foreign accents, can be difficult to process initially, yet listeners can quickly overcome this difficulty and achieve enhanced comprehension. However, the mechanisms behind this adaptation remain unclear. Recent literature has suggested that perceptual adaptation to foreign-accented speech is in part due to reorganization of phonetic category structure. Alternatively, some studies suggest that at least part of accent accommodation involves a relaxation of phonemic categorization criteria in the word recognition process (Zheng & Samuel, 2020). The present study will examine rapid speech adaptation as a consequence of exposure to nonnative-accented input, focusing on native English speakers' perception of the /θ/-/s/ contrast in Mandarin-accented English. The goal is to determine what, if anything, is learned by listeners that could potentially facilitate adaptation to similar variation across different contexts. Specifically, this study aims to distinguish between two theoretically distinct mechanisms: learning of distributional properties in the speech stimuli versus adjustment in decisional bias.

Literature Review

Past work on the perception of foreign-accented speech has demonstrated a remarkably rapid adaptation to foreign-accented talkers. With less than 1 minute of exposure to non-native speech, listeners were able to overcome the initial decrease in processing speed for non-native compared to native-accented speech (Clarke and Garrett, 2004). It has also been shown that listeners use their lexical knowledge to aid in the interpretation of ambiguous speech sounds. When exposed to an ambiguous fricative midway between /f/ and /s/, listeners subsequently classify other ambiguous sounds along the /f/-/s/ continuum in accordance with the words they were exposed to in the exposure phase of the study (Norris et al., 2003). However, the mechanisms behind this perceptual adaptation remain unclear. Recent studies have critically suggested that listeners can modify existing phonetic representations and form a separate sound-to-category mapping for foreign-accented talkers (Dahan et al., 2008 & Xie et al., 2017). In the proposed study, we will attempt to replicate these results and extend them to study the perception of the /θ/-/s/ contrast in Mandarin-accented speech by native English speakers. This contrast is of particular interest: Mandarin speakers generally encounter difficulties with the production of the English dental fricative /θ/ (e.g., *th*ank) as dental fricatives do not exist in Mandarin Chinese. Native-Mandarin speakers will typically substitute /θ/ with two similar alveolar fricatives, /s/ and /z/ (Han, 2013). Production data show that approximately 15.17% of /θ/ is substituted by [s] (e.g., *th*ank → *s*ank) among Chinese students' production of English (Zhang & Xiao, 2014). Despite that well-recognized difficulty in production, past work did not specify the exact distributional properties

that led to ambiguity in the fricative /θ/-/s/ contrast. Jongman et al. (2000) suggests that spectral peak location, spectral moments, and both normalized and relative amplitude can serve to distinguish all four places of fricative articulation and this study will examine these properties in native English speakers' perceptions of Mandarin-accented speech.

Perceptual adaptation as a result of the variation found in the production of the /θ/-/s/ contrast has been studied by Zheng and Samuel (2020), where it was shown that after brief exposure to English produced by a Mandarin-accented speaker, participants displayed an increased acceptance of not only Mandarin-accented /θ/ words (i.e., increased accuracy for /θ/), but nonwords (e.g., *s*ief accepted as *th*ief) as words after the exposure phase. They attributed natural accent adaptation to a relaxation of phonemic boundaries. However, their results do not preclude the possibility that listeners may adapt to this contrast via distributional learning. One confound in this work is that the distributions between /θ/ and /s/ in their auditory stimuli appear to be almost indistinguishable. As such, we cannot be certain that there was indeed a salient signal from which listeners could learn from and cannot distinguish between two theoretically distinct underlying mechanisms responsible for the observed adaptation to the Mandarin-accented /θ/ sound. To address this issue, we will extend the original study in two ways. First, we will select two speakers, one whose production displays clear distributional differences in the /θ/-/s/ contrast (Speaker 1) to ensure that there is some signal that participants can attune to as well as one who does not (Speaker 2). Second, we will examine the word recognition accuracy for /s/ words (e.g., *s*eed) and /s/ nonwords (e.g., *th*udden) as well. We predict that exposure to nonnative-accented words will elicit detectable perceptual adaptation in native listeners of English, resulting in changes in recognition accuracy for Mandarin-accented English words. If distributional learning is indeed the mechanism behind perceptual adaptation, then we expect those listening to Speaker 1's productions to exhibit increased word recognition accuracy for both categories (/θ/ and /s/) following accent exposure. Those exposed to Speaker 2's productions should show increased accuracy for /θ/ and decreased accuracy for /s/ in performance given that there are no distributional properties to learn from.

Method

Experimental Design

This study will adopt a 2 Speaker X 2 Test Phase (pretest vs. posttest) mixed design which will include three total phases: pretest, exposure, and posttest. During pretest and posttest, participants will be exposed to a different list of naturally produced Mandarin-accented English words by one of two speakers and asked to complete an auditory lexical-decision task. The critical items are words and nonwords containing /θ/ or /s/. The order of the sets of stimuli in the pretest and posttest will be counterbalanced across participants. During exposure, which will take place in between the pretest and the posttest, participants will be exposed to Mandarin-accented sentences produced by the same speaker heard during the pretest. Critically, during exposure, listeners will hear a set of critical /θ/ words that are substituted with [s] in Mandarin-accented speech.

Participants.
96 monolingual English speakers with no hearing or visual problems will be recruited for this study from the University of California, Irvine. All participants will be screened to ensure that they are naïve to the Mandarin language and have little to no previous exposure to Mandarin-accented English. Half of the participants (n=48) will hear auditory stimuli produced by Speaker 1 and the other half from Speaker 2. Participants will receive monetary reward for their participation.

Materials.
We will have 5 female native-Mandarin speakers record all speech materials from which 2 will be selected based on their speech patterns and their productions used as the auditory stimuli. For the exposure phase, 96 sentences will be selected. All of the sentences will contain predictable lexical contexts and have an adequate number of instances of the critical /θ/ phoneme, providing ample opportunity for learning of the accent variation. To test natural accent adaptation, 40 critical /θ/ words and 40 /s/ words as well as 80 corresponding nonwords will be selected. Nonwords corresponding to the critical words will be made by replacing the /θ/ with an /s/ (e.g., le*th*al → lesal) and /s/ nonwords created by replacing /s/ with /θ/ (e.g., spa*ce* → spa*the*). An additional 80 filler words will be selected and 80 filler nonwords will be created by changing one phoneme in a filler word to another phoneme that is not difficult for Chinese speakers (e.g., pro*g*ram → pro*b*ram). A total of 320 words and nonwords will be used across both test phases. Two counterbalanced lists will be created for use in the pretest and posttest: half of the items (20 critical /θ/ words, 20 /s/ words, 40 filler words and their respective corresponding nonwords) will make up the lists. Both lists will be well matched in syllable length, frequency, and difficulty.

Procedure.
During the pretest and posttest, participants will be presented one item at a time and will be asked to indicate if the word they hear is a real English word by pressing a yes/no button as quickly and accurately as possible. During the exposure phase which will take place between the pretest and the posttest, participants will be presented with sentences spoken by the same non-native speakers who produce the materials in the pretest and posttest. The entire study will take approximately 30 minutes to complete.

Responsibilities

Regularly meet with faculty mentor and/or communicate between meetings through email with updates and any difficulties.

Experiment design

- Create auditory stimuli
- Create behavioral experiments in Finding-Five (a platform for conducting academic research online)
- Pilot study to refine details as needed

Data collection

- Recruit participants
- Screen possible participants for eligibility
- Grant compensation for participation

Data Analysis

- Analyze participant responses using R
- Interpret results

Dissemination

- Write up a paper on the research findings
- Prepare and present a poster/talk at local/national academic conferences if possible

Timeline

Time	Goal
Weeks 1-2	Finalize and record stimuli and design pilot study. Begin creating experimental materials using FindingFive platform and recruiting participants.
Week 3	Run pilot study and make adjustments as needed.
Weeks 4-5	Begin collecting data and continue recruiting participants.
Weeks 6-8	Continue collecting data and begin data cleaning and analysis.
Weeks 9-10	Write paper on research findings.

References

Clarke, C. M., & Garrett, M. F. (2004). Rapid adaptation to foreign-accented English. *The Journal of the Acoustical Society of America*, *116*(6), 3647-3658.

Dahan, D., Drucker, S. J., & Scarborough, R. A. (2008). Talker adaptation in speech perception: Adjusting the signal or the representations?. *Cognition*, *108*(3), 710-718.

Han, F. (2013). Pronunciation problems of Chinese learners of English. *ORTESOL Journal*, *30*, 26-30.

Jongman, A., Wayland, R., & Wong, S. (2000). Acoustic characteristics of English fricatives. *The Journal of the Acoustical Society of America*, *108*(3), 1252-1263.

Norris, D., McQueen, J. M., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive psychology*, *47*(2), 204-238.

Xie, X., Theodore, R. M., & Myers, E. B. (2017). More than a boundary shift: Perceptual adaptation to foreign-accented speech reshapes the internal structure of phonetic categories. *Journal of Experimental Psychology: Human Perception and Performance*, *43*(1), 206.

Zhang, Y., & Xiao, J. (2014). An Analysis of Chinese Students' Perception and Production of Paired English Fricatives: From an ELF Perspective. *Journal of Pan-Pacific Association of Applied Linguistics*, *18*(1), 171-192.

Zheng, Y., & Samuel, A. G. (2020). The relationship between phonemic category boundary changes and perceptual adjustments to natural accents. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *46*(7), 1270.