Individual differences in categorical perception and novel phonological contrast learning

Jieun Lee & Hanyong Park
University of Wisconsin-Milwaukee

Listeners differ in their sensitivity to acoustic cues for categorization and their use of multiple acoustic cues. For the English stop contrast, for example, some listeners showed more use of the secondary fundamental frequency (f0) cue at vowel onset along with the primary voice onset time (VOT) cue than other listeners, and such individual differences were shown in the visual analogue scaling (VAS) task (Figure 1) [1, 2]. We examined whether these individual differences are related to non-native category learning, specifically, the Korean three-way stop in Experiment 1. Although VOT and f0 play a role in the perception of both the Korean and English stop contrasts, the relative use of f0 is different; it is the primary cue for the Korean lenis/aspirated distinction but secondary for the English stop contrast. We hypothesized that English listeners displaying gradient response patterns in the VAS task, that is, who show greater use of f0 in their native stop contrast, would be better at learning the Korean contrast than the listeners displaying categorical response patterns. We also investigated whether the additional training, expected to downweight the VOT cue and upweight the f0 cue in the perception of the English contrast [3], could aid in their learning of the Korean contrast (Experiment 2). We predicted that such a training would decrease the possible learning gap due to the listeners' differences in their sensitivities to VOT and f0.

English native speakers who participated in Experiment 1 exhibited differences in the gradiency of response on the VAS task (Figure 2). Some judged the English stimuli varying in VOT and f0 categorically (relying on the two endpoints of the VAS scale as in Figure 2a), while others judged them in a much more gradient manner (using the entire scale as in Figure 2b). Based on their response patterns, participants were divided into two groups: *Gradient* (G group) and *Categorical* (C group). We created training stimuli from three Korean native speakers' productions, systemically varying in 7 steps in VOT and 5 steps in f0, resulting in a continuum of /p'a/-/pa/-/p^{h}a/. Both groups of learners received three-day computer-based auditory training with feedback, and each training session lasted for about 30 minutes. Overall, the G group outperformed the C group. The G group showed improvement in identifying the target stimuli each day and generalized their learning to untrained Korean stop contrast and talker stimuli. Importantly, the G group demonstrated more native-like use of the f0 cue after receiving the training, while the C group still exhibited confusion between lenis and aspirated stops (Figure 3). The quantified gradiency (the quadratic regression curves overlaid on the histograms of VAS response) predicted the participants' learning outcomes.

In Experiment 2, a new group of English native speakers took part in the same procedure as in Experiment 1. The only difference between the two experiments is that the participants received the short and simple two-alternative forced choice 'inhibition training' [4] ("Is this English /ba/ or /pa/?") before a training session each day. English training stimuli varying along 7 VOT values but only with two extreme f0 values were presented and the answers for feedback were determined only by their f0 values. The purpose of this training was to downweight VOT cue and shift listeners' attention to f0 differences between stimuli, resulting in changes in their native cue-weighting strategies. The cue-weighting strategy transfer from native to nonnative language was assumed [5]. We observed the positive effects of this training to the C group; they performed similarly to the G group. Compared to the C group in Experiment 1, the C group with the inhibition training demonstrated more native-like use of f0 (Figure 3). Our study showed the relation between individual differences in native cue-weighting strategies and the non-native novel contrast learning, and a possible type of training to overcome disadvantages due to the individual differences.

Figure 1. Illustration of the visual analogue scaling (VAS) task.

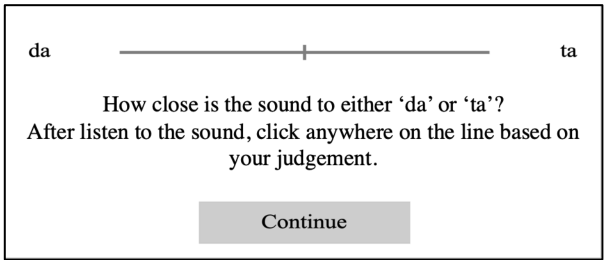


Figure 2. Examples of histograms of click locations of two response patterns for the VAS task: **categorical** listener (left) and **gradient** listener (right).

a. b.

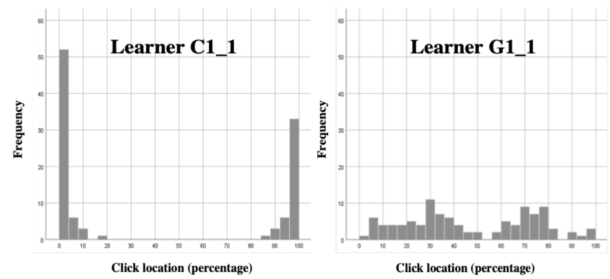
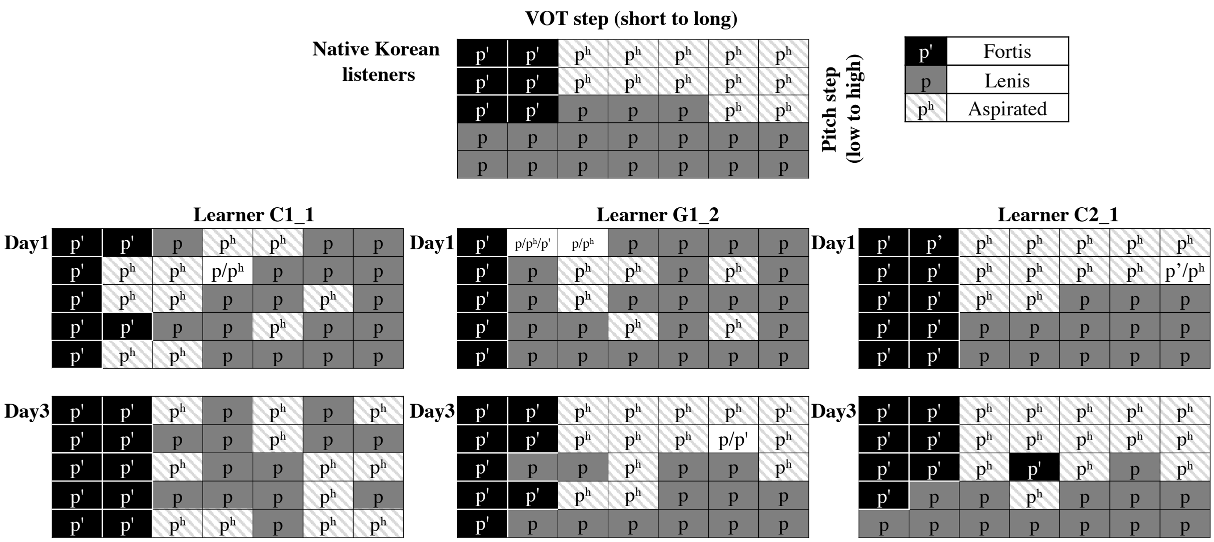


Figure 3. Native Korean listeners’ patterns in the target stimuli identification test (top) and mapping plots of responses of the tests after the first and third training sessions (bottom) by three participants: C1_1 from the Categorical group in Experiment 1 (left), G1_2 from the Gradient group in Experiment 1 (middle), and C2_1 from the Categorical group in Experiment 2 (right).



[1] Kapnoula, E. C., Winn, M. B., Kong, E. J., Edwards, J., & McMurray, B. (2017). Evaluating the sources and functions of gradiency in phoneme categorization: An individual differences approach. *Journal of Experimental Psychology: Human Perception and Performance*, *43*(9), 1594.

[2] Kong, E. J. & Edwards, J. (2016). Individual differences in categorical perception of speech: Cue weighting and executive function. *Journal of phonetics*, *59*, 40-57.

[3] Harmon, Z., Idemaru, K., & Kapatsinski, V. (2019). Learning mechanisms in cue reweighting. *Cognition*, *189*, 76-88.

[4] Kondaurova, M. V. & Francis, A. L. (2010). The role of selective attention in the acquisition of English tense and lax vowels by native Spanish listeners: Comparison of three training methods. *Journal of phonetics*, *38*(4), 569-587.

[5] Tremblay, A., Broersma, M., & Coughlin, C. E. (2018). The functional weight of a prosodic. cue in the native language predicts the learning of speech segmentation in a second language. *Bilingualism: Language and Cognition*, *21*(3), 640-652.