Imitation and perception of individual accented features

Jessamyn Schertz, Department of Language Studies, University of Toronto Mississauga

Accent imitation is a complex behavior, requiring accurate perception of the accented features, followed by accurate (re-)production of the same features. Previous work has shown that the linguistic status of a feature (Olmstead et al., 2013; Wagner et al., 2019) and perceptual sensitivity to the feature (Llompart and Reinisch, 2018) influence imitation. This is the first in a set of studies designed to compare imitation of individual accented features independently, by comparing imitation of artificial accents that differ minimally in a single feature, and examine the extent to which individual variation in imitation can be explained by variability in perception of these features. We explore the role of two factors in predicting success in imitation: 1) the linguistic status of the feature (stop aspiration/voicing, /u/-fronting, and /i/-/ɪ/ merger) and 2) ability to perceive the difference in accents (via accurate classification of a new talker).

Stimuli: We created four pairs of accents, each differing in a single feature (Table 1). We expected that accented features that were potentially contrast-threatening, such as /ɪ/-raising, where /ɪ/ is pronounced as [i], would be more salient and thus more imitated than those that were not, such as /u/-fronting, which does not result in a change in phonemic category. To create each set of stimuli, we recorded a native speaker of English reading sentences containing multiple instances of each target sound, then manipulated the relevant feature (VOT or vowel quality) to create a pair of baseline "accents" differing only in the relevant feature. To avoid having the talkers sound identical, we scaled the formants and f0 of each "talker" (i.e. Accent A sounded like a prototypical male, and B sounded female). There were three sentences per accent pair.

Participants and procedure: 21 native speakers of English living in Toronto participated in the study. Trials were blocked by feature type, and each trial consisted of three stages. In the **Exposure** phase, participants heard two talkers, represented by aliens, saying the same sentence in alternation, three times. In the **Imitation** phase, they heard the same sentences in alternation again and were asked to imitate each sentence after they heard it. In the **Test** phase, they then heard a new talker and were asked to decide whether they sounded like Talker A or B.

Analysis: Target sounds in the imitations were measured for aspiration (voiceless stops), presence of prevoicing (voiced stops) and F1/F2 (vowels). Mixed-effects models tested whether values on the relevant features differed in imitation of the two accents, and whether accuracy in classification of new talkers differed across accented conditions.

Results: As shown in Figures 1 and 2, participants imitated the relevant differences in the two contrast-threatening features (/ɪ/-raising and deaspiration), and were also above chance in classifying novel talkers in these conditions (all statistically significant). However, participants showed less imitation, and less accurate performance in the test phase, in the deaspiration condition than in the /ɪ/-raising condition. Only about half of the participants showed imitation of deaspiration. Furthermore, imitation appeared to be categorical rather than gradient, e.g. with participants pronouncing a deaspirated /p/ identically to their English /b/ category, even with prevoicing in some circumstances. The non-contrast-threatening feature conditions showed no imitation, and participants were at chance in the test phase. Together, these results support the idea that the type of imitation elicited in this task is heavily influenced by phonological categories, with participants showing little sensitivity to fine phonetic detail, in contrast to results on shadowing tasks (e.g. Nielsen, 2011). The group-level production-perception correspondence suggests that imitation is linked to accurate perception of differences in a given feature; however, on an individual level, imitation was not correlated with accuracy in the test phase. We are exploring other predictors of this individual variability, including language background and base perceptual sensitivity to the acoustic dimension, in follow-up work.

Feature	Accent A (canonical)	Accent B
/ɪ/-raising (i.e. /i/-/ɪ/ merger)	/ɪ/ → [ɪ]	/ɪ/ → [i]
/u/-fronting	/u/ → [u]	/u/ → [ʉ]
/p t k/ deaspiration	/p t k/ → [pʰ tʰ kʰ]	/p t k/ → [p t k]
Prevoiced /b d g/	/b d g/ → [b̥ d̥ g̊]	/b d g/ → [b d g]

Table 1: Contrast-threatening (grey) and non-contrast-threatening (white) features. For each feature, two sets of accents were created, differing only in the feature shown here.
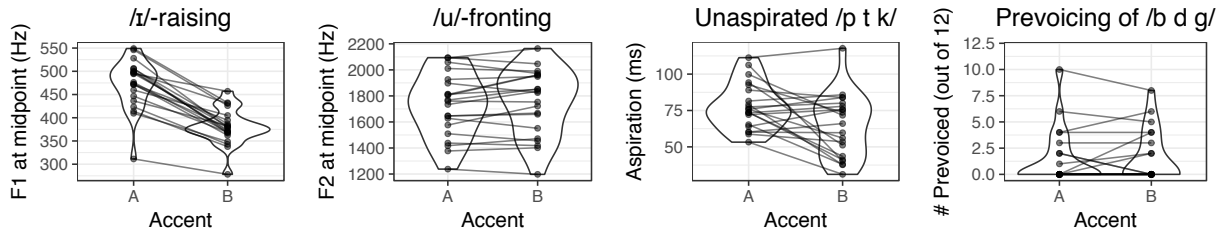

Figure 1: Imitation results. Each panel shows a different feature condition. Lines show each participant's mean value when imitating Accent A vs. Accent B. For example, for the /ɪ/-raising condition, all participants produced lower F1 (i.e. a higher vowel) when imitating Accent B (which had a raised /ɪ/ relative to Accent A).
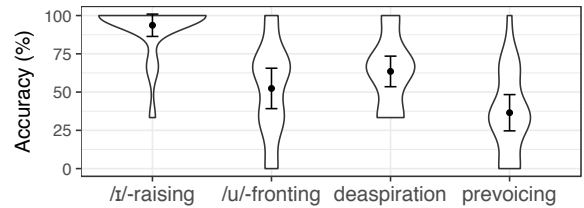

Figure 2: Test results. Density plots showing participants' accuracy in classifying new talkers in each condition (error bars show +/- $2SE$ of by-participant means).

References

Llompart, M. and Reinisch, E. (2018). Imitation in a second language relies on phonological categories but does not reflect the productive usage of difficult sound contrasts. *Language and Speech*, 62(3):594–622.

Nielsen, K. (2011). Specificity and abstractness of VOT imitation. *Journal of Phonetics*, 39(2):132–142.

Olmstead, A. J., Viswanathan, N., Aivar, M. P., and Manuel, S. (2013). Comparison of native and non-native phone imitation by English and Spanish speakers. *Frontiers in Psychology*, 4:475.

Wagner, M. A., Broersma, M., McQueen, J. M., and Lemhöfer, K. (2019). Imitating speech in an unfamiliar language and an unfamiliar non-native accent in the native language. *Proceedings of ICPhS*, pages 1362–1366.