

Perception and Production across L2 Phonological Contrasts in Bilinguals

Caihong Weng^{1,2}, Ioana Chitoran^{1,3}, Martijn Wieling^{2,4}, and Alexander Martin²

¹Department of Linguistics, Université Paris Cité, Paris, France ²Center for Language and Cognition, University of Groningen, Groningen, The Netherlands ³Institut Universitaire de France, Paris, France ⁴Vrije Universiteit Brussel, Brussel, Belgium

Phonological mergers offer a revealing window into perception-production relationships. While individual-level alignment has received growing attention in contexts of sound change and variation [2,5], less is known about how this relationship behaves when contrasts are socially stratified within a bilingual community. This study examines two mergers in a Quanzhou Southern Min (QSM)-Mandarin bilingual community: the /f~/x/ merger, associated with local Fujian identity but negatively evaluated locally as a marker of limited education, and /s~/ʃ/ merger, perceived as a broader southern accent feature. Using a perception-production design, we investigate whether perceptual discrimination predicts production behavior for each contrast, and whether this relationship is modulated by contrasts and by speaker characteristics such as Mandarin usage and generational exposure.

METHOD: 58 QSM-Mandarin bilinguals (younger generation 18–30: 27; older generation 31–55: 34) from Quanzhou, China completed a sentence-reading task and an ABX discrimination task. Production accuracy was assessed through an identification task in which 10 L1 Mandarin listeners categorized target fricative productions from isolated words. For the ABX task, we compared perception of a control contrast /m~/p/ (present in both languages: *Two-Category Assimilation*) to two target contrasts: /f~/x/ (/f/ is assimilated to QSM /h^w/ and /x/ to /h/: *Category-Goodness Assimilation*) and /s~/ʃ/ (both assimilated to QSM /s/: *Single-Category Assimilation*).

RESULTS: Production accuracy. For production, Mandarin fricatives present in the QSM inventory (/s/, /x/) were identified more accurately than absent ones (/f/, /ʃ/). Overall production accuracy was higher for /f~/x/ than /s~/ʃ/ (see Fig. 1). **Perception accuracy.** Both target contrasts showed significantly lower discrimination accuracy than the control contrast /m~/p/ (/f~/x/: $\beta = -2.38, p < .001$; /s~/ʃ/: $\beta = -2.88, p < .001$), with /s~/ʃ/ showing the lowest discriminability, in line with PAM predictions (see Fig. 2). **Perception-production link.** In an initial model, production accuracy increased with perceptual accuracy across contrasts ($\beta = 0.62, p < .001$; Fig. 3). Once *contrast* was included, the effect was no longer significant ($\beta = 0.15, p = .07$), and no perception $\times$ contrast interaction emerged ($\beta = 0.02, p = .89$), suggesting that the overall association reflects contrast-level differences. The three-way interaction between *perceptual accuracy*, *contrast*, and *Mandarin usage* was not significant ($\beta = 0.08, p = .15$), providing no evidence that any contrast-specific perception–production relationship is further modulated by Mandarin usage level. However, a significant three-way interaction between *perceptual accuracy*, *contrast*, and *generational exposure* (younger vs. older) emerged ($\beta = 0.67, p < .05$; Fig. 5), indicating that the perception–production relationship varies as a function of both phonological contrast and language experience.

CONCLUSION AND DISCUSSION: The two Mandarin fricative contrasts differed in both perception and production in line with PAM predictions, with /s~/ʃ/ showing lower discriminability and weaker production accuracy than /f~/x/. Although perceptual accuracy initially predicted production accuracy, this effect was substantially reduced when contrast was included, suggesting the apparent link is largely driven by contrast-level differences. Crucially, the perception-production relationship was stronger for younger than older speakers, and this generational exposure asymmetry was contrast-specific, pointing to the role of ongoing language shift in shaping perception-production relationship. Future work should combine acoustic analyses with speaker-level modeling to disentangle individuals from community-level patterns.

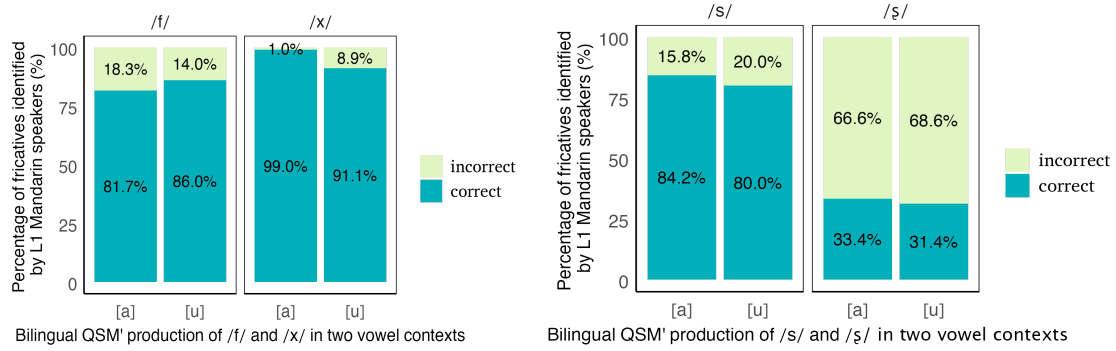


Fig 1: Percentages of two target contrasts (left: /f/~/x/; right: /s/~/ʃ/) productions across two vowel contexts by QSM-Mandarin bilinguals that were identified by L1 Mandarin listeners as either *correct* (matching the target) or *incorrect* (mismatching the target).

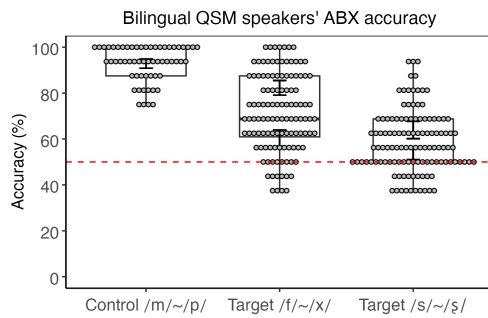


Fig 2: ABX discrimination accuracy (%) for bilingual QSM-Mandarin speakers on the control contrast (/m/~/p/) and the two target contrasts (/f/~/x/ and /s/~/ʃ/). Each point represents one participant. The dashed red line indicates chance-level performance (50%).

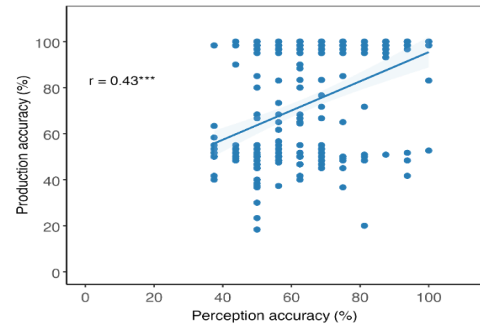


Figure 3: Relationship between perception accuracy and production accuracy. Each point represents one bilingual QSM speaker. Pearson correlation coefficients are indicated within the figure (* $p < .05$, ** $p < .01$).

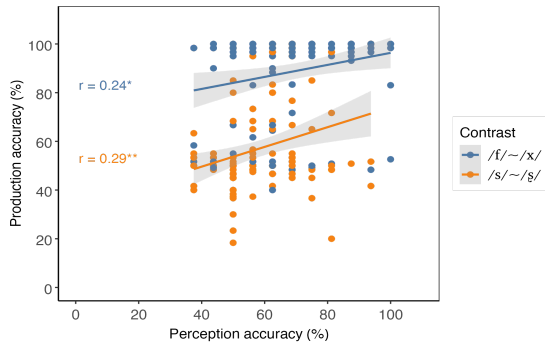


Figure 4: Perception-production relationship by generation group and contrast. Each point represents a speaker's perception (ABX discrimination accuracy) and production accuracy. Pearson correlation coefficients are indicated within the figure (* $p < .05$, ** $p < .01$).

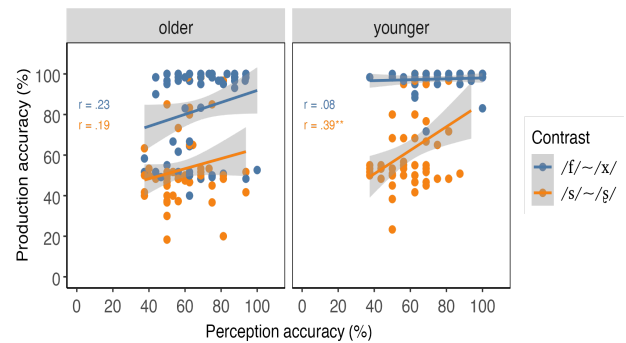


Figure 5: Perception-production relationship by generation group and contrast. Each point represents a speaker's perception (ABX discrimination accuracy) and production accuracy. Pearson correlation coefficients are indicated within the figure (* $p < .05$, ** $p < .01$).

References:

- [1] Best, C. T., McRoberts, G. W., & Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system *The Journal of the Acoustical Society of America*, 109(2), 775–794. [2] de Leeuw, E., Stockall, L., Lazaridou-Chatzigoga, D., & Gorba Masip, C. (2019). Illusory vowels in Spanish-English sequential bilinguals: Evidence that accurate L2 perception is neither necessary nor sufficient for accurate L2 production. *Second Language Research*, 37(4), 587-618. [3] Flege, J. E., Aoyama, K., & Bohn, O.-S. (2021). The revised Speech Learning Model (SLM-r) applied. In *Second Language Speech Learning: Theoretical and Empirical Progress* (pp. 84-118). Cambridge University Press. [4] Lindblom, B. (1990). "Explaining phonetic variation: A sketch of the H&H theory," in *Speech production and speech modelling*. Dordrecht: Springer Netherlands. 403–439. [5] Melnik-Leroy, G. A., Turnbull, R., & Peperkamp, S. (2021). On the relationship between perception and production of L2 sounds: Evidence from Anglophones' processing of the French /u/-/y/ contrast. *Second Language Research*, 38(3), 581-605.