

German-speaking learners' perception of Korean stops at novice and advanced beginner levels: The role of L1 regional variation

Yeongeun Choi¹, Stephan Schmid², Eva Maria Luef³, Paola Escudero¹

¹Western Sydney University, ²University of Zurich, ³Charles University

y.choi@westernsydney.edu.au; stephan.schmid@uzh.ch; EvaMaria.Luef@ff.cuni.cz; paola.escudero@westernsydney.edu.au

A sound category is the result of the way in which speakers of a given language map continuous acoustic signals onto one or more discrete sound units [1, 2]. This language-specific nature of sound categories may lead learners from different language backgrounds to perceive the same sounds differently in nonnative speech perception. The Second Language Linguistic Perception (L2LP) model [3, 4] posits that listeners—particularly at the initial stage—‘fully copy’ the sound categorization acquired in their first language (L1) into second language (L2) speech perception. Crucially, this L1 transfer is not limited to a single language variety, as regional varieties within an L1 often differ in cue weighting for the same L1 sound categories (e.g., [5]); and in the L2LP model, this may yield differential starting points for listeners from different regions in perceptual tasks for an L2 sound (e.g., [6]). In this talk, we argue based on data from German, Austrian, and Swiss learners of Korean that variety-specific cue weighting in the L1 can lead to distinct categorization of the same L2 sound, and this regional variation may affect acquisitional development across different stages of learning.

In word-initial position, Standard German plosives are distinguished by voice onset time (VOT): lenis /b d g/ with short-lag VOT and fortis /p^h t^h k^h/ with long-lag VOT [7]. Austrian German plosives also distinguishes two VOT-based categories, but the VOT boundary is lowered from 40ms to 30ms [8]. In contrast, Swiss German plosive categories are based on closure duration rather than on VOT [9], though the graphemes <p t k> are sometimes aspirated in proper nouns and in loanwords [10]. When German-speaking learners listen to word-initial Korean plosives, they are given two perceptual tasks to fully categorize them: (1) a target-like VOT boundary shift between two VOT-based categories (short vs. long VOT: fortis /p* t* k*/ vs. lenis /p t k/ and aspirated /p^h t^h k^h/) and (2) a new creation of two categories based on fundamental frequency (F0) of the following vowel (low vs. high F0: lenis vs. fortis/aspirated) (cf. [11]). Given the acoustic properties, L2-to-L1 perceptual mapping is established differently across L1 varieties: Swiss learners effectively start with a quasi-single plosive category in terms of VOT and F0, while German and Austrian learners start with two VOT-based categories.

81 German-speaking learners of Korean from Germany, Austria, and Switzerland at novice and advanced levels and 15 native Korean controls participated. An AX discrimination task tested listeners' auditory-driven discrimination of (1) VOT, using stimuli in which F0 was held constant (110-170Hz, 20Hz steps), and (2) F0, using stimuli in which VOT was held constant (0-90ms, 30ms steps), as well as (3) integrated VOT and F0 cues, using natural stimuli. Responses were analyzed using generalized linear mixed-effects models (GLMM).

Overall, all learner groups were successful in categorizing Korean plosives based on VOT, but not on F0. For VOT-based discrimination, Swiss learners at both levels achieved higher perception accuracy than the other groups (Fig. 1), suggesting that setting a new boundary based on familiar VOT may be easier than shifting an existing VOT boundary, as in the Austrian case. For F0-based perception, Swiss and Austrian learners initially employed the target cue more effectively than German learners; however, at the advanced stage, the decline of correct F0 perception was greater in these two groups (Fig. 2). Likely, groups for whom VOT shifts are required tend to allocate greater perceptual attention to VOT, resulting in less reliance on F0: for Swiss learners, successful VOT categorization made VOT a reliable cue; for Austrian learners, relatively frequent miscategorization necessitated continued corrective attention to VOT. This biased cue weighting was also mirrored in the perception of integrated

cues in natural stimuli (Fig. 3). To sum up, although learners failed to establish a three-way categorization of Korean plosives up to the advanced beginner stage, we expect that more L2 input will drive re-weighting of VOT and F0 cues, but in an L1 variety-specific fashion.

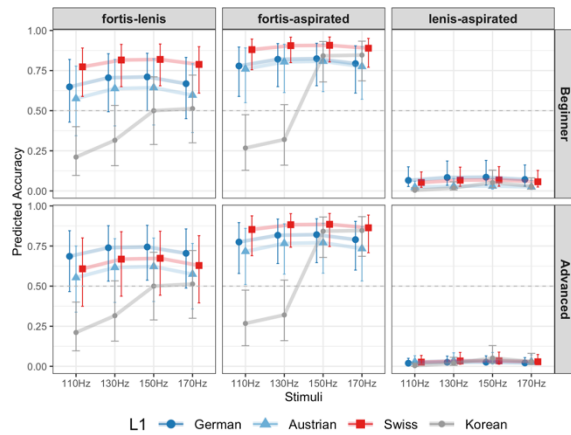


Fig. 1 GLMM prediction of accuracy by stimulus for each pair type in the VOT-based discrimination session (with manipulated F0).

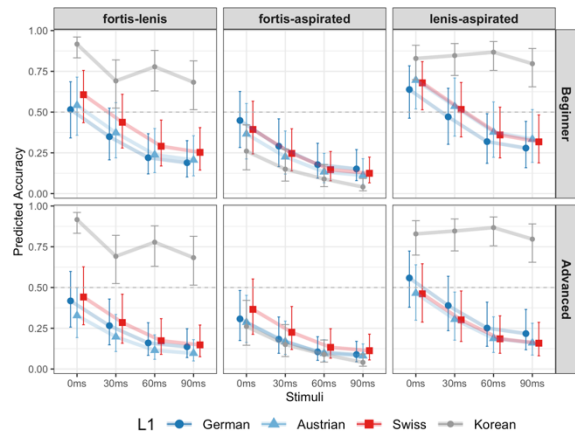


Fig. 2 GLMM predictions of accuracy by stimulus for each pair type in the F0-based discrimination session (with manipulated VOT).

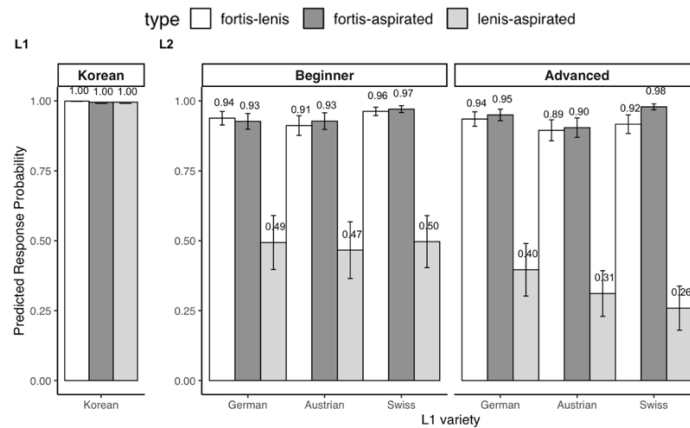


Fig. 3 GLMM predictions of mean accuracy for the original tokens

References

- [1] Boersma, P. (1998). *Functional Phonology: Formalizing the interactions between articulatory and perceptual drives* (Doctoral dissertation). University of Amsterdam.
- [2] Escudero, P. & Boersma, P. (2004). Bridging the gap between L2 speech perception research and phonological theory. *Studies in second language acquisition*, 26(4), 551-585.
- [3] Escudero, P. (2005). *Linguistic Perception and Second Language Acquisition: Explaining the Attainment of Optimal Phonological Categorization* (Doctoral dissertation). Utrecht University.
- [4] Escudero, P. & Yazawa, K. (2024). The Second language linguistic perception model (L2LP). In M. Amengual (Ed.), *The Cambridge handbook of bilingual phonetics and phonology* (pp. 173-195). Cambridge University Press.
- [5] Kong, E. J. & L, H. (2021). Relationship between executive function and cue weighting in Korean stop perception across different dialects and ages. *Phonetics and Speech Sciences*, 13(3), 21-29.
- [6] Choi, T-H., Kim, G-H., & Han, J-I. (2013). Perception and production of English stops by tonal and non-tonal Korean dialect speakers. *The Journal of the Acoustical Society of America*, 134(6), EL541-EL546.
- [7] Kleber, F. (2018). VOT or quantity: What matters more for the voicing contrast in German regional varieties? Results from apparent-time analyses. *Journal of Phonetics*, 71, 468-486.
- [8] Luef, E. M. (2020). Development of voice onset time in an ongoing phonetic differentiation in Austrian German plosives: Reversing a near-merger. *Zeitschrift Für Sprachwissenschaft*, 39(1), 79-101.
- [9] Enstrom, D. H. & Spörri-Bütler, S. (1981). A voice onset time analysis of initial Swiss-German stops. *Folia Phoniatrica*, 33(3), 137-150.
- [10] Zebe-Sheng, F., Watter, C., Schmid, S. & Ladd, D. R. (2025). Transitory and sustained Cf0 effects: Evidence from Swiss German. *Journal of Phonetics*, 113, 1-17.
- [11] Kang, K.H. & Guion, S. G. (2008). Clear speech production of Korean stops: Changing phonetic targets and enhancement strategies. *The Journal of the Acoustical Society of America*, 124(6), 3909-3917.