Perceptual distance preference in learning sound correspondences during dialect acquisition

Sound correspondence (SC) relations occur when two sounds occupy corresponding positions in cognate words of historically related languages or varieties (Brown et al., 2013). For example, the English word *ten* and the German word *zehn* are cognates, and their onsets establish the SC /t : ts/. SCs have been examined in the domain of historical-comparative linguistics, as well as in the realm of second dialect acquisition (SDA) (Chambers, 1992; Rys, 2007). In the context of SDA, it has been found that learners establish connections between corresponding variants using shared lexico-semantic representations of cognates (Auer, 1993; Taeldeman, 1993). Factors influencing SC learning include mapping structure, word frequency, and phonotactics, etc. (Auer et al., 1998; Rys, 2007; Siegel, 2010).

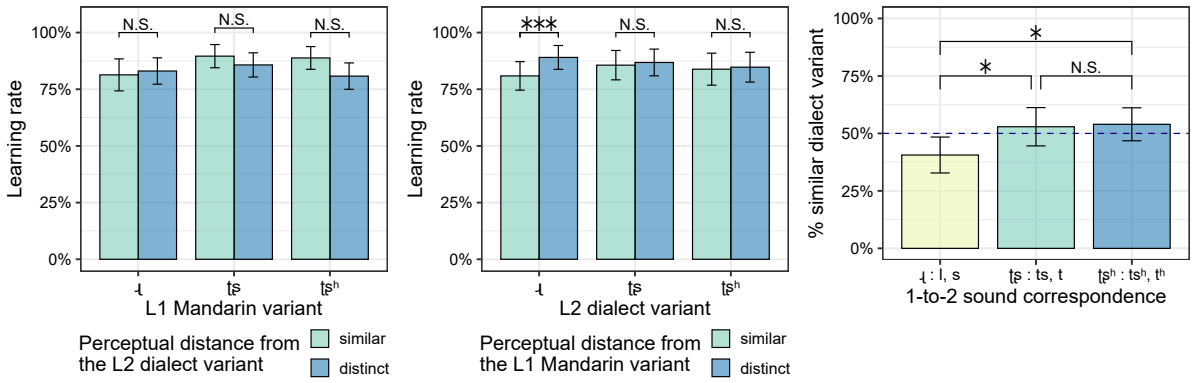
This study investigates a relatively unexplored aspect of SC learning in SDA, specifically focusing on the influence of perceptual distance. We aim to examine how perceptual distance affects the acquisition of new dialect phonological systems, similar to the process observed in second language acquisition (Best, 1994; Flege, 1995). Different proposals put forth distinct predictions regarding perceptual distance's effect on SC learning. Trudgill's (1986) salience hypothesis suggests that perceptually distinct dialect variants are more salient and learnable, while empirical studies have yielded mixed results (e.g., Gómez et al., 2020; Kerswill & Williams, 2002). Conversely, the substantive bias hypothesis in phonological learning posits that perceptually similar patterns are phonetically grounded and more learnable (Moreton & Pater, 2012; Wilson, 2006). The prevalence of similar SCs across languages also supports the notion that perceptually similar SCs are better learned (Brown et al., 2013).

We conducted an artificial language learning study to test how perceptual distance affects SC learning. 179 native Mandarin speakers participated, learning artificial dialects through SCs created with Mandarin phonemes. We collected perception data first to create a Mandarin phoneme confusion matrix, based on which the perceptual distance of SCs in the artificial dialects was manipulated. Similar SCs consisted of frequently confused Mandarin phoneme pairs differing by 1 contrastive feature, while distinct SCs involved less confusable pairs with 2 or 3 feature differences. The confusion matrix data confirmed that more feature differences correlated with larger perceptual distance. Dialect cognates were created by replacing the onsets of monosyllabic Mandarin words with the corresponding dialect variants (e.g., cognate *ròu* "meat", Mandarin /ɻəu/ → dialect /ləu/). *Praat* software (Boersma & Weenink, 2021) was used to modify each of the four Mandarin tones into non-existing contour tones, enhancing the distinctiveness of the artificial dialects from Mandarin.

Three conditions with different mapping structures were used in the study. In the one-to-one (1-to-1) condition, each L1 Mandarin variant corresponded to a single L2 dialect variant. Participants learned three SCs with either smaller (/ɻ : l/, /ʈʂ : ts/, /ʈʂʰ : tsʰ/) or larger (/ɻ : s/, /ʈʂ : t/, and /ʈʂʰ : tʰ/) perceptual distance. In the one-to-two (1-to-2) condition, participants learned SCs with two L2 variants corresponding to each L1 variant. Half of the dialect cognates had similar variants (/l, ts, tsʰ/), while the other half had dissimilar ones (/s, t, tʰ/). In the two-to-one (2-to-1) condition, the L1 and L2 variants of the 1-to-2 SCs were flipped. Each SC had similar and distinct L1 variants in equal numbers of cognates corresponding to the same L2 variant. The distinct variant pair /ɻ : s/ had a larger perceptual distance equivalent to 3 feature differences compared to the other two pairs (/ʈʂ : t/, /ʈʂʰ : tʰ/) with only 2 differences.

During training, participants saw Chinese characters (n = 60) and heard the pronunciation of dialect cognates. In the testing phase, they were shown characters of old (n = 30) and new words (n = 60) and had to select the correct dialect cognate from words with L1 Mandarin or L2 dialect variants. 1-to-1 and 2-to-1 tests had two alternatives, while 1-to-2 had three due to the possibility of two L2 dialect variants for each word.

Statistical analysis involved mixed-effects regression models and post-hoc pairwise comparisons. Learning rates, specifically the rates of selecting L2 dialect variants, were examined. Overall, there were no significant differences in learning rates across the three mapping structures, indicating comparable learning of SCs regardless of pattern complexity in the current experiment. In the 1-to-1 condition (Fig. 1), all six SCs showed high learning rates (> 75%) with no influence of perceptual distance observed (β = -0.03, z = 0.51, SE = -0.05, p = 0.96). In the 2-to-1 condition (Fig. 2), perceptual distance impacted only the /l, s : ɻ/ SC, where the L2 dialect variant corresponding to the distinct L1 Mandarin variant /s/ was learned better than the one corresponding to the similar L1 variant /l/ (β = 1.13, SE = 0.27, z = 4.12, p < 0.001). Similarly, in the 1-to-2 condition (Fig. 3), participants chose the similar variant /l/ significantly less than chance in the /ɻ : l, s/ SC (β = -0.93, SE = 0.28, z = -3.35, p < 0.001), while the other SCs had equal selection of both L2 variants. The findings suggest that the impact of perceptual distance on SC learning is contingent upon the presence of structurally complex mappings, i.e., 2-to-1 or 1-to-2. Furthermore, perceptual distance appears to be influential only among pairs characterized by a significant perceptual disparity, e.g., the case of /ɻ : s/ compared to /ʈʂ : t/ or /ʈʂʰ : tʰ/.



Figures 1 & 2. Learning rates in the 1-to-1 (left) and 2-to-1 (middle) conditions;
Figure 3. The rates of choosing the similar dialect variants in the 1-to-2 condition (right).

Our findings suggest that SC learning prefers larger perceptual distance, although this effect is not robust and depends on two factors. Firstly, perceptual distinction is preferred only in the presence of complex mapping structures (many-to-one or one-to-many). Secondly, an activation threshold (e.g., equivalent to 3 feature differences) must be met to trigger this preference. Contrary to the substantive bias towards similar patterns in phonological learning, our study uncovers an opposing bias in SC learning. We propose that dialects, in comparison to languages, already possess high phonological and lexical similarity (Siegel, 2010), potentially overriding learners' sensitivity to varying perceptual similarity. Likewise, studies on cognate lexical learning demonstrate learners' insensitivity to the degree of segment overlap between cognates (Andras et al., 2022; Comesaña et al., 2012; Dijkstra et al., 1999). Instead, in SCs involving similar variant pairs, only the dialect variant that is sufficiently perceptually distinct emerges as more salient and, therefore, more learnable, partially supporting the predictions of the salience hypothesis (Trudgill, 1986). To our knowledge, our study is the first to confirm the effect of perceptual distance on SC learning in SDA, while controlling for other intervening factors.

Selected references: Brown, et al.. (2013). Sound Correspondences in the World's Languages. *Language*, *89*(1); Kerswill, P., & Williams, A. (2002). "Salience" as an explanatory factor in language change: Evidence from dialect levelling in urban England. *Lang Change*; Moreton, E., & Pater, J. (2012). Structure and Substance in Artificial-Phonology Learning, Part II: Substance. *Lang & Ling Compass*, *6*(11); Rys, K. (2007). *Dialect as a Second Language:…* [PhD thesis, Ghent Uni.]; Trudgill, P. (1986). *Dialects in contact*.