Native phonological experience modulates cross-situational learning of non-native tones

Yin-To Chui[1] & Quentin Zhen Qin[2]

[1,2]*Speech, Learning, and the Brain (SLaB) Lab, Division of Humanities, The Hong Kong University of Science and Technology*

[1]*ytchuiac@connect.ust.hk,* [2]*hmzqin@ust.hk*

Abstract

Cross-situational word learning (CSWL) allows learners to associate words and meanings despite ambiguity in individual learning moments, through statistical tracking across multiple exposures (Yu & Smith, 2007; Smith & Yu, 2008). Recent literature indicates that CSWL performance is influenced by phonological factors, where the shorter acoustic distance between minimal pairs compared with non-minimal pairs hindered learning of the former (Escudero et al., 2016; Mulak et al., 2019). Cross-linguistic studies further showed additional L1 influence on CSWL (e.g., Ge et al., 2024), particularly when minimal pairs involve "difficult" phonological contrasts relative to learners' L1 system as posited by speech perception models (Tuninetti et al., 2020). This study extends this line of inquiry and examines how native (Cantonese) tonal language background influences non-native (Mandarin) tonal CSWL performance. Fifty-four participants (27 Cantonese and 27 Mandarin speakers) performed CSWL of 9 Mandarin word-referent pairs, consisting of three sets of tonal minimal triplets (level-rising-falling; T1-T2-T4). Participants underwent a training phase, where they saw objects paired with auditory pseudowords and selected the correct referent from two visual options (2AFC). In a subsequent testing phase, participants selected the correct referent from three options (3AFC), comprising two types of distractors: tonal minimal triplets (MT) and non-minimal triplets (non-MT). Results of the test phase indicated a significant group effect and trial type effect, where Mandarin participants showed better performance overall, and participants showed overall better performance on non-MT trials. Crucially, a significant interaction between group and trial type was observed, where Cantonese speakers exhibited a disproportionately larger drop in performance for MT trials than Mandarin speakers. Analysis of tone confusion patterns revealed Cantonese speakers specifically struggled with T1, frequently confusing it with T4, while Mandarin speakers showed no systematic confusions across tones. These results extend previous findings of phonological influence on CSWL performance to non-native tone learning.

Keywords

Cross-situational word learning (CSWL), Statistical tone learning, Speech perception