The online processing of non-native phonological contrasts in L2: from acoustics to lexicon

Sijia Zhang, Molly Babel, & Christopher Hammerly
Department of Linguistics, University of British Columbia

When processing spoken words, listeners continuously abstract the phono-lexical information from the acoustic details in the signal. For a second language (L2) listener, processing words with non-native contrasts can be particularly challenging as their L1 sound system can influence low-level perception [1; 2], as well as a higher level of phonological encoding in the lexicon [4; 5]. It is thus unclear whether L2 lexical confusion arises from inaccurate phonetic processing or higher-level encoding, or both. Although some studies showed a connection between L2 phonetic discrimination and phonological categorization [3], others found a discrepancy between the processing of phonetic and lexical information, such that L2 listeners showed accurate low-level processing yet poor lexical decision [6]. Using an eye-tracking study, [5] showed that L2 listeners leveraged the acoustic details early on in online processing, whereas experienced more lexical competition at a later time. Building upon [4; 5; 6], this study aims to (1) test how L2 listeners process a non-native contrast *online* when experiencing lexical competition, compared to L1 listeners; (2) untangle whether the difficulties in L2 lexical processing is attributed to imprecision at the acoustic-level perception or phonological-level encoding.

Methods. Beijing Mandarin (BM) permits a more restricted set of low-vowel + nasal (loVN) sequences compared to North American English (NAE). Among all the /æm, æn, æŋ, ɑm, ɑn, ɑŋ/, only /æn/ and /ɑŋ/ follow the phonological restrictions in BM [7]. While [8] and [9] reported errors in the production of English loVN contrasts by L1 BM speakers, the perception and processing of loVN remain understudied. In **Experiment 1**, 24 L1 BM-L2 NAE speakers and 23 L1 NAE speakers participated in an eye-tracking study with a visual world paradigm. On each trial, they saw four printed words on the screen (target, competitor, and two distractors), paired with an audio stimulus that consisted of the target word. Participants chose the word they heard and their eye movements were recorded. In the *critical* trials, the target and the competitor contained loVN that differed in the nasal of the first syllable (e.g., *con**[an]**tact-con**[aŋ]**quer*). Either the target contained a non-native loVN for L1 BM and the competitor contained a native loVN, or the other way around. The *baseline* trials included a native loVN in both the target and competitor. In **Experiment 2**, 50 L1 BM-L2 NAE speakers and 26 L1 NAE speakers completed a speeded AX discrimination task. They heard pairs of English loVN sequences in isolation and decided whether they were the "same" or "different" (vowel-differed e.g., [æm]-[ɑm], or nasal-differed e.g., [ɑn]-[ɑŋ]). Both the accuracy and reaction times were measured.

Results. For **Exp. 1** (Fig.1), an LME-based cluster permutation analysis [10] revealed that the fixation proportion to the target over the competitor (i.e., lexical competition effect) was significantly different between L1 NAE and L1 BM only in a later time window (420-980ms; cluster mass=1226.24, $p<0.001$), and the group difference was greater in the critical trials than the baseline (780-940ms; cluster mass=89.35, $p<0.001$). No difference was found between the two critical conditions. For **Exp. 2**, Bayesian mixed-effects models revealed weak evidence for a difference in the discrimination accuracy between L1 NAE and L1 BM ($\beta=0.61$, CrI=[-0.61,1.88], PD=84.12%), despite strong evidence that L1 NAE had shorter RTs than L1 BM ($\beta=-232.18$, CrI=[-352.23,-110.26], PD=100%). L2 listeners were able to accurately discriminate loVN at the phonetic level (Table 1), despite longer reaction times (Fig.2). Taken together, in L2 online processing, there is little difference in perceiving acoustic details in early time window compared to L1, yet more difficulties lie within later time window which involves higher-level phonological encoding in the lexicon. No distinction is encoded for such contrast by L2 listeners, as both non-native and native loVN activate each other, as opposed to asymmetric competition reported in [4] and [6]. This suggests an overall fuzzy lexical representation [11] established by advanced L2 learners who have little problem processing non-native phonetics.

[1] Werker, J. F., Gilbert, J. H., Humphrey, K., and Tees, R. C. (1981). Developmental aspects of cross-language speech perception. *Child development*, 52(1), 349–355. [2] Johnson, K. & Babel, M. (2010). On the perceptual basis of distinctive features: Evidence from the perception of fricatives by Dutch and English speakers. *Journal of phonetics*, 8(1), 127–136. [3] Boomershine, A., Hall, K. C., Hume, E., & Johnson, K. (2008). The impact of allophony versus contrast on speech perception. *Contrast in phonology*, 143–172. [4] Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of memory and language*, 50(1), 1-25. [5] Turner, J. (2022). Analysing the relationship between L2 production and different stages of L2 processing: Eye-tracking and acoustic evidence for a novel contrast. *Journal of Phonetics*, 91, 101134. [6] Darcy, I., Daidone, D., & Kojima, C. (2013). Asymmetric lexical access and fuzzy lexical representations in second language learners. *The mental lexicon*, *8*(3), 372-420. [7] Duanmu, S. (2007). *The Phonology of Standard Chinese*. OUP Oxford. [8] Hansen, J. G. (2001). Linguistic constraints on the acquisition of English syllable codas by native speakers of Mandarin Chinese. *Applied Linguistics*, 22(3), 338–365. [9] Liu, Y. (2016). Oral gestural reduction of English nasal coda produced by Mandarin English-as-a foreign-language learners. *The Journal of the Acoustical Society of America*, 140(4), 3339–3339. [10] Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG-and MEG-data. *Journal of neuroscience methods,* 164(1), 177-190. [11] Gor, K., Cook, S., Bordag, D., Chrabaszcz, A., & Opitz, A. (2021). Fuzzy lexical representations in adult second language speakers. *Frontiers in Psychology*, 12, 732030.

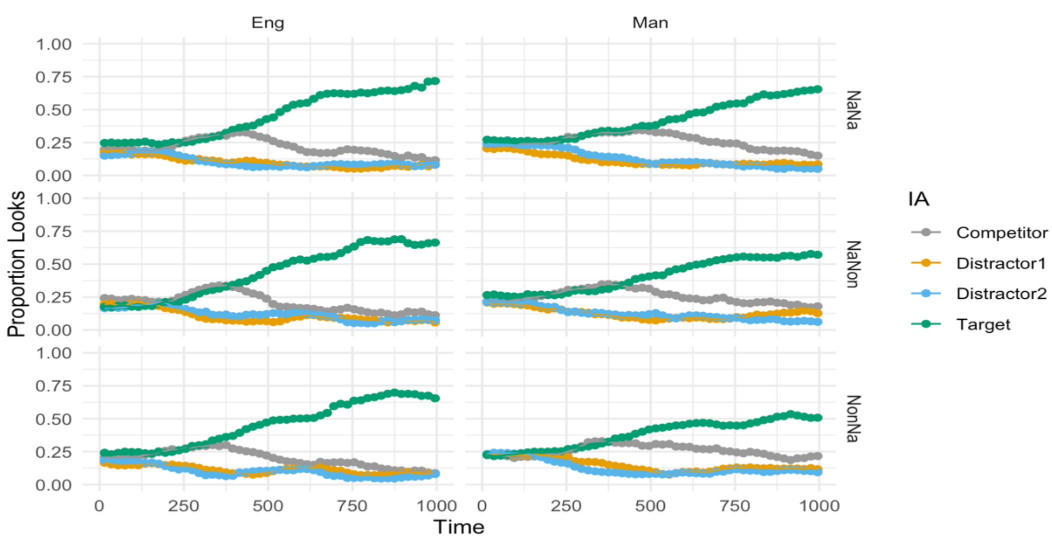

Fig. 1. The mean proportions of fixations to the four words in 0-1000 ms starting from the onset of the target word in the audio stimuli (aggregated into 20ms time bins). (NaNa: baseline; NaNon: target = native, competitor = non-native; NonNa: target = non-native, competitor = native).

	ENG	Can-M	Chi-M
Nasal	96.24%	93.46%	90.45%
Vowel	96.97%	99.14%	97.22%
Same	95.47%	96.07%	96.12%

Table 1. Average accuracy rate of discriminating loVN. (Can-M: L1 BM living Canada; Chi-M: L1 BM in China).

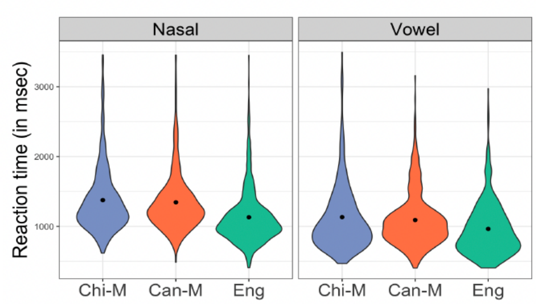

Fig. 2. Reaction times to correct different pairs (nasal- vs. vowel-differed).