INTERSPEECH 2023
20-24 August 2023, Dublin, Ireland

First Language Effects on Second Language Perception: Evidence from English Low-vowel Nasal Sequences Perceived by L1 Mandarin Chinese Listeners

Sijia Zhang[1]

[1]University of British Columbia, Canada
`sijia01@student.ubc.ca`

Abstract

First language (L1) sound systems can shape second language (L2) perception of non-native phonological contrasts. This study examines how L1 Mandarin listeners perceive English low vowel + nasal (VN) sequences that are not contrastive in Mandarin. A speeded AX discrimination task tested listeners' low-level processing of English VN sequences, and a perceptual similarity rating task assessed listeners' higher-level phonological knowledge. Despite high discrimination accuracy, reaction times showed that L1 Mandarin listeners had a harder time processing nasal-different sequences compared to L1 English listeners. Moreover, L1 Mandarin listeners perceived vowel-different sequences as more distinct than English listeners. Overall, listeners' L1 phonological system influences L2 perception, suggested by group-level differences in both perceived phonological distinctiveness and phonetic discrimination.

Index Terms: second language perception, Mandarin, vowel nasal sequences, phonological contrast, phonetic processing

1. Introduction

First language (L1) sound systems can influence listeners' perception of non-native speech sounds [1]. In particular, the presence or absence of a phonological contrast in the L1 has an impact on one's perceptual ability. At an auditory level, listeners might have difficulty distinguishing between sounds that are not contrastive in their L1, such as the perception of English /l/ and /r/ by Japanese listeners [2], and Hindi dental and retroflex stops perceived by American English listeners [3]. At a phonological level of processing, [4] [5] and [6] assessed the perceptual consequences of L1 phonological relations between sounds. Listeners tend to perceive sounds in allophonic alternations as more similar than those that are phonemically contrastive. For instance, L1 Spanish listeners perceived [ð] and [d] as more similar than L1 English listeners, as [ð] and [d] are allophones in Spanish and are distinct phonemes in English [4].

In second language (L2) acquisition, the revised Speech Learning Model (SLM-r) posits that L2 learners start by automatically linking non-native sounds that they hear to an existing L1 category [7]. According to the SLM-r, L2 learners gradually discern the phonetic differences between L1 and L2 sounds as they receive more L2 input. The likelihood that a new phonetic category is formed for an L2 sound depends on the perceived dissimilarity of the L2 sound from its closest L1 sound, as well as the quality and quantity of the input received by the L2 learners [7] (p.33).

L2 perception of non-native nasal contrasts is often challenging [8, 9], especially when different L1 phonological restrictions apply concerning the surrounding segments. In the case of Mandarin and English, Mandarin has multiple phonological restrictions on low-vowel + nasal sequences ($V_{[+low]}N$, e.g., ***ham***[æm] and ***dawn***[ɑn]) that do not apply in English. North American English has two low vowel phonemes, the low front vowel /æ/ and the low back vowel /ɑ/. The three nasal contrasts /m, n, ŋ/ are all allowed in the coda position. In contrast, Mandarin allows the alveolar nasal /n/ and the velar nasal /ŋ/ in coda position, but not the bilabial nasal /m/. The low vowels [a] and [ɑ] are allophones in Mandarin, and their backness is conditioned by the place of the following nasal coda [10, 11]. Namely, the low front vowel must precede the alveolar nasal (i.e., [an] but *[aŋ]), and the low back vowel must precede the velar nasal (i.e., [ɑŋ] but *[ɑn]) (Rhyme Harmony as in [10]). While the only contrastive $V_{[+low]}N$ permitted in Mandarin are [an] and [ɑŋ], English permit all the combinations of $V_{[+low]}N$ sequences as contrastive (i.e., [æm, æn, æŋ, ɑm, ɑn, ɑŋ]).

Production errors have been reported with respect to English vowel + nasal (VN) sequences in general learned by Mandarin speakers (e.g., the coda [n, ŋ] and the place of vowels are confusable) [12, 13]. How all the possible English $V_{[+low]}N$ sequences in particular are perceived by L1 Mandarin listeners has not been systematically investigated in the current literature.

This study aims to examine how Mandarin-speaking adults' L1 background influences their perception of $V_{[+low]}N$ contrasts in English. A speeded AX discrimination task is adopted to tap a phonetic level of processing [14], while a perceptual similarity task is used to tap phonological level of processing [4, 6]. Two demographic groups of L1 Mandarin listeners are tested, in order to further examine whether different amounts of L2 experience influence the two levels of processing.

Compared to English listeners, we expect that Mandarin listeners will overall find it harder to distinguish between $V_{[+low]}N$ sequences that are not contrastive in their L1 (e.g., [ɑn] vs. [ɑŋ]). Sequences that differ in nasals will be perceived as more similar than those that differ in vowels. Mandarin listeners, especially those with limited L2 English input, may associate certain perceptually similar nonnative $V_{[+low]}N$ contrasts with a single L1 category [7, 1]. Thus, L1 Mandarin listeners may perceive higher similarity between English $V_{[+low]}N$ sequences that are non-contrastive in Mandarin than L1 English listeners.

2. General Procedure

Two groups of Mandarin-speaking participants originally from Northern regions of China[1] were recruited through Chinese social media. 24 of the L1 Mandarin-speaking participants were living in China and learned English as L2 (Chi-M group) (8

---

[1] Speakers of Northern Mandarin varieties (e.g., Beijing Mandarin) are expected to maintain the /n/ and /ŋ/ coda nasal contrast followed by the low vowel in Mandarin [15, 16]

10.21437/Interspeech.2023-1123

male, 16 female, mean age = 25.2 years, range = 20-55). 26 of the L1 Mandarin speakers were living in Canada and learned English as L2 (Can-M group) (10 male, 16 female, mean age = 24.9 years, range = 22-30). All did not speak a Chinese dialect other than Mandarin, and reported that they learned North American variety of English (mean age of acquisition of Chi-M = 7.7 years, Can-M = 4.7 years). Participants in the Can-M group lived in Canada for 3 to 8 years (mean age of arrival = 19.23, range = 15-26). 26 of L1 English-speaking participants of a North American English variety were recruited from a Canadian University's participant pool as the L1 control group (4 male, 22 female, M = 20.8 years, range = 18-25). The English control (Eng) group reported no proficiency in any Chinese languages.

The Can-M group was assumed to have more exposure to English than the Chi-M group. Individual L2 experience and proficiency level were quantitively measured through the Lexical Test for Advanced Learners of English (LexTALE) [17] and the questionnaire Bilingual Language Profile (BLP) [18]. The means of the LexTALE scores differed across the three groups (Chi-M = 55.81, Can-M = 72.83, Eng = 90.96), suggesting that Can-M had more advanced English lexical knowledge and higher L2 proficiency compared to Chi-M. The negative means of the BLP scores showed that both Chi-M and Can-M groups were Mandarin-dominant rather than English-dominant (Chi-M = -141.76, Can-M = -60.48), yet Chi-M had a greater degree of dominance in Mandarin than Can-M.

Participants completed two online perception tasks: a speeded AX discrimination task, and a perceptual similarity rating task. The order of the two tasks was counterbalanced across participants. All the instructions presented to the L2 participant groups were in both English and simplified Chinese characters. The whole study lasted about 40 minutes.

3. Speeded AX discrimination task

In this task, participants heard two auditory items on each trial, with an inter-stimulus interval (ISI) of 250 msec. Participants responded whether the two items were the same or different as fast as they could. The response time window had a limit of 3000 msec. To encourage faster responses, a feedback window with the overall accuracy rate and response time was given after each response. The relatively short ISI and speeded response time window had a low memory load and aimed to tap low-level phonetic processing [14, 6, 4].

3.1. Materials

All of the auditory stimuli were recorded by an L1 female speaker of Canadian English. 9 of the "different" item pairs used in the auditory stimuli contained all the combinations of $V_{[+low]}N$ sequences (i.e., [æm, æn, æŋ, ɑm, ɑn, ɑŋ]) which differed only in either the vowel (e.g., [æm] vs. [ɑm]) or the nasal (e.g., [ɑn] vs. [ɑŋ]). The stimuli used for the 6 "same" $V_{[+low]}N$ item pairs were physically identical. In addition, the stimuli consisted of pairs of low vowels (e.g., [ɑ] vs. [æ]), and pairs that contained the mid front vowel [e] and one of the nasals (e.g., [em] vs. [en]) in separate blocks.[2] The items in each "different" pair were presented in both orders and each stimulus was presented three times. Each of the "same" pair stimuli was presented 3-9 times to match the total number of its "different" pair counterpart in each block. This led to 156 trials in total. The stimuli were presented in three blocks (i.e., vowel only, mid-vowel+nasal, and $V_{[+low]}N$) and were randomized within each block. The RMS intensity of all stimuli were normalized to 65 dB.

	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 $V_{[+low]}N$ sequences.*

3.2. Analysis and results

Reaction times (RTs) were calculated from the onset of the second stimulus for each trial. Responses received before the onset of the second stimulus were removed from the data (1% of the responses).

3.2.1. Accuracy

All the language groups showed high accuracy on average in discriminating the $V_{[+low]}N$ sequences. As presented in Table 1, nasal-different pairs (i.e., nasal condition) had slightly lower average accuracy than vowel-different pairs (i.e., vowel condition). A Bayesian mixed-effects model was fitted in Stan using the brms package [19] in R [20] to analyze the accuracy of responses. The fixed effects included dummy-coded Condition (nasal, vowel, same, ref level: nasal), Helmert-coded Group (ENG vs. mean of Can-M and Chi-M, Can-M vs. Chi-M), the interaction effect of Condition and Group, and the Order of items in each pair (Order1 vs. Order2). The outcome variable was dummy-coded Correct Responses (True = 1, False = 0). By-item and by-subject random intercepts were included, as well as by-item random slopes for Group and by-subject random slopes for Condition. A Bernoulli distribution was assumed for the correct responses. Weakly informative priors were used for all population-level parameters (centered at 0 and had a 0.25 standard deviation). Hamiltonian Monte-Carlo sampling was used to draw samples from the posterior distribution (four chains, each with 2000 iterations and 1000 warm-up). We report the mean of the posterior distribution, the 95% credible interval (CrI), and the probability of direction (PD). A 95% CrI that does not encompass 0 is considered as evidence for a meaningful effect, and a PD that is greater than 95% provides strong evidence of a non-null effect with a particular direction (positive or negative) [21].

The model revealed weak evidence for a difference between nasal and vowel conditions in their response accuracy (β = 0.71, CrI = [-0.55, 1.96], PD = 87.78%). The model also provided weak evidence for a difference between the English group and the mean of the two Mandarin groups (β = 0.61, CrI = [-0.61, 1.88], PD = 84.12%), and weak evidence for a difference between the two Mandarin groups (β = 0.78, CrI = [-0.48, 2.08], PD = 89.28%). These result indicate that the accuracy was high in all groups and conditions.

3.2.2. Reaction times

Reaction times to the "different" pairs reflect how difficult to hear the difference between the stimuli [22]. Figure 1 shows RTs for correctly responded "different" pairs comparing nasal-

---

[2] The reason to include these two types of pairs was to examine how well low vowels and coda nasals were distinguished, independent from a $V_{[+low]}N$ environment ([e] was chosen as a non-low vowel). We do not report the results here due to space limit.

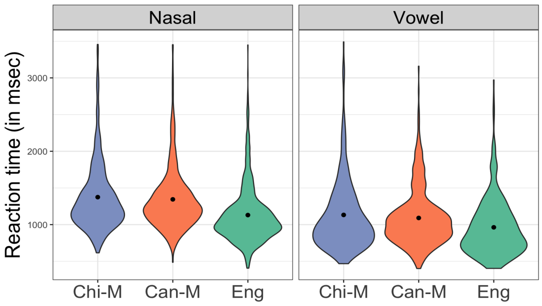

Figure 1: *Reaction times to correct "different" pairs in AX task for different pair types (nasal vs. vowel) and listener groups.*

different vs. vowel-different pairs across the three language groups. A Bayesian mixed-effects model was fitted, with weakly informative priors used for all parameters. Similar to the model for accuracy, fixed effects included dummy-coded Condition (nasal, vowel, ref level: nasal), Helmert-coded Group, the interaction effect of Group and Condition, and Order. By-item and by-subject random intercepts were included, as well as by-item random slopes for Group and by-subject random slopes for Condition.

The model provided strong evidence that the RTs were longer for nasal pairs than vowel pairs (β = -221.54, CrI = [-281.88, -151.42], PD = 100%). We also observed strong evidence that the English group had shorter RTs than the mean of the two Mandarin groups (β = -232.18, CrI = [-352.23, -110.26], PD = 100%), yet little-to-no evidence for a difference between the two Mandarin groups (β = -34.27, CrI = [-174.53, 115.84], PD = 72.92%). Nevertheless, we found strong evidence for an interaction effect between Group (ENG vs. mean of Chi-M and Can-M) and Condition (β = 79.62, CrI = [10.28, 147.97], PD = 98.85%), suggesting that the difference in RTs between English and the mean of the two Mandarin groups was greater for nasal pairs than vowel pairs (as seen in Figure 1). There was weak evidence for an order effect (β = -8.41, CrI = [-62.50, 45.07], PD = 62.38%).

Together, these results show that all listeners were able to accurately discriminate $V_{[+low]}$N sequences at a phonetic level. However, the longer reaction times suggested a harder processing task for L2 listeners.

4. Perceptual similarity rating task

In this task, participants heard two auditory items on each trial, with an ISI of 500 msec. Participants rated how similar the two items were on a scale of 1 to 5 (1 = "very different", 5 = "very similar"). There was no limit on the response time window, and no feedback was given when the response was received. The relatively longer ISI was designed to tap a higher level of processing that accessed phonological representations.

4.1. Materials

The same set of $V_{[+low]}$N sequences recorded for the AX task were used in this task. The stimuli contained three types of item pairs, namely, 9 critical pairs that differed only in either the nasal (e.g., [ɑn] vs. [ɑŋ]) or the vowel (e.g., [æm] vs. [ɑm]), 6 non-critical pairs that differed in both the nasal and the vowel (e.g., [ɑn] vs.[æm]), and 6 physically identical pairs. Items in the critical and non-critical pairs were presented in both orders. The stimuli of all the three pair types were presented three times, resulting in 108 trials in total. The order of all the stimuli was randomized, presented in one single block.

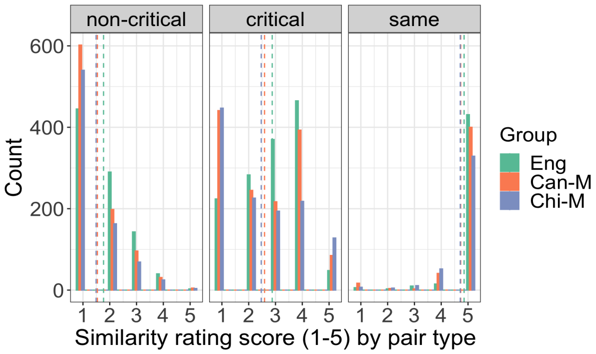

Figure 2: *Overall distribution of similarity ratings across the three trial types by listener group. The dotted lines indicate the average rating scores for each type per group.*

4.2. Analysis and results

Figure 2 shows the distribution of similarity ratings across the three pair types, where all listener groups perceived the same pairs as highly similar (average = 4.77), and the non-critical pairs as relatively distinct (average = 1.60). To further analyze rating data of the critical pairs (average = 2.65), a Bayesian mixed-effects ordinal regression (i.e., cumulative link) model was adopted. The model captures the cumulative likelihood of responses that fall in a certain range of ordered response categories. Compared to a linear regression model which assumes the dependent variables are continuous, the ordinal regression takes into account that the distributions across ordinal categories are uneven, and it can generate predictions that fall within the response scale [23, 24].

The model was fitted in Stan using the brms package [19] in R [20], with weakly informative priors used for all parameters. Similar to the models described in section 3.2, fixed effects included dummy-coded Condition (nasal, vowel, ref level: nasal), Helmert-coded Group, the interaction effect of Group and Condition, and Order. By-item and by-subject random intercepts were included, as well as by-item random slopes for Group and by-subject random slopes for Condition. The link function parameter was set to "logit" to fit an ordered logistic regression model, and the threshold was "flexible" by default without adding any further constraints to the intercepts [23].

The model provided strong evidence that the listeners on average perceived the vowel pairs with lower similarity compared to the nasal pairs (β = -2.97, CrI = [-3.76, -2.11], PD = 100%). However, we only observed weak evidence of Group effects for nasal pair perception. Specifically, there was no big difference in the ratings between the English group and the mean of the two Mandarin groups (β = 0.32, CrI = [-0.51, 1.33], PD = 85.52%), nor between the two Mandarin groups (β = 0.28, CrI = [-0.50, 0.98], PD = 76.65%).

For the interaction effects between Group and Condition, we found strong evidence that the differences in similarity ratings between the English group and the mean of the two Mandarin groups were greater for vowel pairs compared to nasal pairs (β = 1.50, CrI = [0.50, 2.54], PD = 99.83%). The interaction effect can be clearly seen from Figure 3 and Figure 4 (explained below). There was weak evidence that the differences in rating between Chi-M and Can-M vary for vowel pairs compared to nasal pairs (β = -0.02, CrI= [-1.17,1.18], PD =

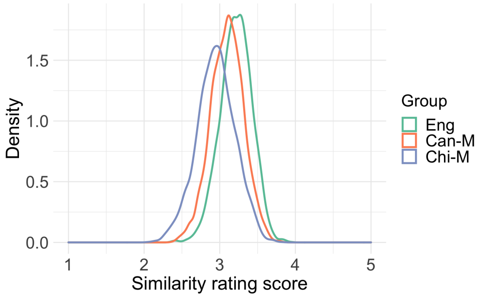

Figure 3: *Average rating distributions for nasal-different pairs estimated by the model.*

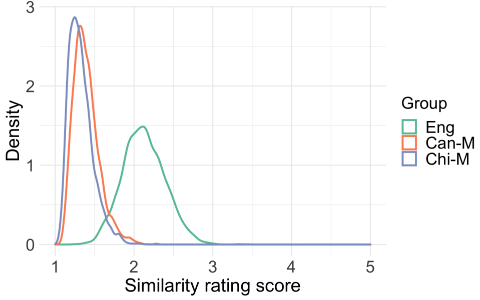

Figure 4: *Average rating distributions for vowel-different pairs estimated by the model.*

51.85%).

Based on the posterior distribution generated from the model, we plotted the estimated distributions of average rating scores by the three groups for nasal pairs (Figure 3) and vowel pairs (Figure 4). In Figure 3, despite some overlap, the average ratings by English listeners were a bit higher than the Can-M group, and both are higher than the Chi-M group. The difference among the three groups was not very large (all most likely in the range of 3 - 3.5), as suggested by the weak evidence from the model. In Figure 4, the average ratings by English listeners were higher than the two Mandarin groups. The variability of the mean distribution was smaller for the two Mandarin groups than the English group, as the distribution of the English group was more spread out. Compared to nasal pairs (Figure 3), the ratings for vowel pairs were lower for all groups, especially the two Mandarin groups (Figure 4). The rating differences between English and Mandarin listeners were greater in Figure 4 than Figure 3, consistent with the strong interaction effect interpreted from the model. Together, these results show that although there was no strong group effect on perceiving the nasal-different $V_{[+low]}N$ sequences, both groups of Mandarin listeners perceived vowel-different $V_{[+low]}N$ sequences as more distinct than English listeners. For all groups, nasal-different $V_{[+low]}N$ sequences were perceived as more similar than vowel-different ones.

5. Discussion and conclusions

Through the examination of how English $V_{[+low]}N$ sequences are perceived by L1 Mandarin listeners, the current study provides insights into how perception is shaped by listeners' L1 sound system given the evidence from different processing levels. At a phonetic level, despite high discrimination accuracy across listener groups, Mandarin listeners exhibited longer time processing non-contrastive sequences in their L1 compared to English listeners, especially in discriminating nasal-different $V_{[+low]}N$ sequences. Speeded discrimination task with a shorter ISI is often assumed to be less sensitive to language-specific factors, as reported in studies such as [14] and [6]. However, the results of the current study provide evidence that low-level processing can be warped by the L1 sound system, aligning with the findings of [4] and [5], among others.

With respect to the phonological level of processing, interestingly, we found that Mandarin listeners overall perceived the $V_{[+low]}N$ sequences that were non-contrastive in their L1 as less similar compared to English listeners. Based on the assumptions in the SLM-r and the findings in [5] and [6], we would expect that L2 listeners, especially those less-experienced ones, would perceive higher similarity between these English $V_{[+low]}N$ sequences. This is because they might have not yet formed L2 categories for non-contrastive sound sequences like [ɑn] and would likely associate them with an existing L1 category (e.g., [ɑŋ]). Given the current results, one explanation might be that some L2 sound sequences were perceived as comparatively novel to the Mandarin listeners and were not able to be sufficiently associated or assimilated to any of the existing L1 category, which is described as the "uncategorized" case by the L2 Perceptual Assimilation Model (PAM-L2) in [1]. A preliminary analysis of the similarity ratings by individual item pairs suggested that nasal-different pairs containing the labial nasal [m] (e.g., [æm] vs. [æn] and [ɑm] vs. [ɑn]) were perceived as more distinct by L2 listeners, possibly due to the fact that coda [m] is completely banned in Mandarin and thus [Vm] is hard to assimilate to any L1 category. To further test the L2 categories of $V_{[+low]}N$ with different nasal codas, a repetition priming paradigm can be adopted to examine their phonological representations at the lexical level.

Another possibility to explain the similarity rating results here has to deal with the design of the task and the types of knowledge that were successfully tapped. [4] pointed out that the similarity task would invite listeners' metalinguistic knowledge for off-line judgements. In other studies like [25], a vowel identification task was explicitly used to tap listeners' metalinguistic knowledge. It is possible that L2 listeners in the current design might perform the task as in a "test" mode and hence their metalinguistic awareness was implicitly tapped [26]. L2 listeners might be more motivated to judge the different pairs as more distinct than L1 listeners as a demonstration of their L2 knowledge. Notably, vowel-different pairs were particularly rated as more distinct by L2 listeners than L1 listeners, which might be due to the fact that vowel cues are arguably more perceivable than nasal cues, especially those in coda positions [27, 28]. When English loanwords are adapted into Mandarin, the place of coda nasal often changes to agree with the place of the preceding back vowel, arguably because nasals are weaker in phonetic salience than vowels [29].

The current study compares data from two different L2 demographic populations, and examines the potential variations in their perception caused by the differences in L2 experiences. According to the SLM-r, L2 listeners are more likely to discern the phonetic differences between L1 and L2 sound categories as more L2 input is received. The LexTALE and BLP were jointly used to assess L2 learners' relevant experience in English, and both data suggested clear group-level differences between L2 learners in China and L2 learners in Canada. While the results from the two perceptual studies did not show robust differences between the two demographic groups, individual correlations between perceptual data and language experience measures will be further explored in a follow-up study to see the potential impact of L2 experience on perception.

6. References

[1] C. T. Best and M. Tyler, "Nonnative and second-language speech perception," in *Language experience in second language speech learning*, O. Bohn and M. Munro, Eds. John Benjamins Amsterdam, 2007, pp. 13–34.

[2] H. Goto, "Auditory perception by normal Japanese adults of the sounds 'l' and 'r'," *Neuropsychologia*, 1971.

[3] J. F. Werker, J. H. Gilbert, K. Humphrey, and R. C. Tees, "Developmental aspects of cross-language speech perception," *Child development*, pp. 349–355, 1981.

[4] A. Boomershine, K. C. Hall, E. Hume, and K. Johnson, "The impact of allophony versus contrast on speech perception," *Contrast in phonology*, pp. 143–172, 2008.

[5] T. Huang and K. Johnson, "Language specificity in speech perception: Perception of Mandarin tones by native and nonnative listeners," *Phonetica*, vol. 67, no. 4, pp. 243–267, 2010.

[6] K. Johnson and M. Babel, "On the perceptual basis of distinctive features: Evidence from the perception of fricatives by Dutch and English speakers," *Journal of phonetics*, vol. 38, no. 1, pp. 127–136, 2010.

[7] J. E. Flege and O.-S. Bohn, "The revised speech learning model (slm-r)," in *Second Language Speech Learning: Theoretical and Empirical Progress*, R. Wayland, Ed. Cambridge University Press, 2021, p. 3–83.

[8] J. D. Harnsberger, "On the relationship between identification and discrimination of non-native nasal consonants," *The Journal of the Acoustical Society of America*, vol. 110, no. 1, pp. 489–503, 2001.

[9] C. R. Narayan, "The acoustic–perceptual salience of nasal place contrasts," *Journal of Phonetics*, vol. 36, no. 1, pp. 191–217, 2008.

[10] S. Duanmu, *The Phonology of Standard Chinese*. OUP Oxford, 2007.

[11] M. Luo, "Processes and consequences of co-articulation in Mandarin v1n.(c2) v2 context: Phonology and phonetics." in *INTERSPEECH*, 2020, pp. 641–645.

[12] Y. Liu, "Oral gestural reduction of English nasal coda produced by Mandarin English-as-a-foreign-language learners," *The Journal of the Acoustical Society of America*, vol. 140, no. 4, pp. 3339–3339, 2016.

[13] J. G. Hansen, "Linguistic constraints on the acquisition of English syllable codas by native speakers of Mandarin Chinese," *Applied Linguistics*, vol. 22, no. 3, pp. 338–365, 2001.

[14] J. F. Werker and J. S. Logan, "Cross-language evidence for three factors in speech perception," *Perception & Psychophysics*, vol. 37, no. 1, pp. 35–44, 1985.

[15] Y. Chen and S. Guion-Anderson, "Perceptual confusabiltiy of word-final nasals in Southern Min and Mandarin: Implications for coda nasal mergers in Chinese." in *ICPhS*, 2011, pp. 464–467.

[16] M. Wu, M. Sloos, and J. van de Weijer, "The perception of the English alveolar-velar nasal coda contrast by monolingual versus bilingual Chinese speakers," in *2016 10th International Symposium on Chinese Spoken Language Processing (ISCSLP)*. IEEE, 2016, pp. 1–5.

[17] K. Lemhöfer and M. Broersma, "Introducing LexTALE: A quick and valid lexical test for advanced learners of english," *Behavior research methods*, vol. 44, pp. 325–343, 2012.

[18] D. Birdsong, L. M. Gertken, and M. Amengual, *Bilingual language profile: An easy-to-use instrument to assess bilingualism*. COERLL, University of Texas at Austin, 2012, https://sites.la.utexas.edu/bilingual/.

[19] P.-C. Bürkner, "brms: An r package for Bayesian multilevel models using Stan," *Journal of statistical software*, vol. 80, pp. 1–28, 2017.

[20] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2022. [Online]. Available: https://www.R-project.org/

[21] B. Nicenboim and S. Vasishth, "Statistical methods for linguistic research: Foundational ideas—part ii," *Language and Linguistics Compass*, vol. 10, no. 11, pp. 591–613, 2016.

[22] R. N. Shepard, D. W. Kilpatric, and J. P. Cunningham, "The internal representation of numbers," *Cognitive psychology*, vol. 7, no. 1, pp. 82–138, 1975.

[23] P.-C. Bürkner and M. Vuorre, "Ordinal regression models in psychology: A tutorial," *Advances in Methods and Practices in Psychological Science*, vol. 2, no. 1, pp. 77–101, 2019.

[24] J. K. Kruschke and T. M. Liddell, "Bayesian data analysis for newcomers," *Psychonomic bulletin & review*, vol. 25, no. 1, pp. 155–177, 2018.

[25] J. Cabrelli, A. Luque, and I. Finestrat-Martínez, "Influence of L2 English phonotactics in L1 Brazilian Portuguese illusory vowel perception," *Journal of Phonetics*, vol. 73, pp. 55–69, 2019.

[26] D. Birdsong, *Metalinguistic performance and interlinguistic competence*. Springer Science & Business Media, 2012, vol. 25.

[27] Z. Bond, "Listening to elliptic speech: Pay attention to stressed vowels," *Journal of Phonetics*, vol. 9, no. 1, pp. 89–96, 1981.

[28] Y. Chen, V. Kapatsinski, and S. Guion-Anderson, "Acoustic cues of vowel quality to coda nasal perception in Southern Min," in *Thirteenth Annual Conference of the International Speech Communication Association*, 2012.

[29] F.-F. Hsieh, M. Kenstowicz, and X. Mou, "Mandarin adaptations of coda nasals in English loanwords," *Loan phonology*, vol. 307, p. 131, 2009.