THE EFFECT OF LEARNING CONTEXT ON MANDARIN LISTENERS' PERCEPTION OF ENGLISH VOWELS

Jiayi Lu[1] & Mitsuhiko Ota[2]

[1]Hong Kong Polytechnic University, [2]University of Edinburgh
jiayilu1998@163.com, mota@ed.ac.uk

ABSTRACT

Do developmental patterns in second language speech perception differ depending on whether the learning takes place in a predominantly naturalistic setting (second language acquisition [SLA]) or an instructed setting (foreign language acquisition [FLA])? We addressed this question by testing the predictions of the extended Perceptual Assimilation Model (PAM-L2) in both settings. Native adult speakers of Mandarin learning English in the UK (SLA learners; N=15) and in China (FLA learners; N=15) performed an ABX discrimination task on 15 pairs of British English vowels. The vowel pairs were classified into six PAM assimilation categories based on goodness-of-fit ratings between English and Mandarin vowels obtained from Mandarin speakers with little exposure to English (N=15). The order of discrimination accuracy for PAM categories did not differ between the two groups (i.e., SLA and FLA), demonstrating that learning context does not affect the relative difficulty with which adult learners perceive non-native vowels.

Keywords: vowel perception, second/foreign language learning, Perceptual Assimilation Model.

1. INTRODUCTION

1.1. Background

Attainment levels of second language (L2) speech are known to differ depending on whether the learning takes place in a predominantly naturalistic context (second language acquisition [SLA]), where the target language is the language of the community, or an instructed context (foreign language acquisition [FLA]), where the target language is not the community language [1, 2]. Such differences in attainment levels are usually attributed to the limited quantity and quality of the input FLA learners receive, their lack of opportunities to practice the L2, and the formal nature of the instruction they receive.

A less well understood question is whether different learning contexts, such as SLA and FLA, also lead to different developmental patterns in speech perception. More specifically, does the relative difficulty with which learners perceive L2 sounds depend on their contexts of learning? This issue has been raised by Best and Taylor [1] in their adaptation of the Perceptual Assimilation Model to the L2 context (PAM-L2). According to PAM-L2, a pair of L2 sounds that are assimilated to separate native categories (two-category [TC] assimilation) is easier to discriminate than a pair of L2 sounds that are assimilated to a single native category with different degrees of goodness-of-fit (category-goodness [CG]). A CG pair, in turn, is easier to discriminate than a pair of L2 sounds that are perceived to be equally good exemplars of a single native category (single-category [SC]).

These predictions have been confirmed separately in SLA [3, 4] and FLA [2, 5, 6, 7]. However, few studies have systematically tested the predictions of PAM's discrimination hierarchy in both SLA and FLA within a single study, nor have studies tested the perceptual accuracies of assimilation pairs other than TC, CG, and SC in these contexts.

Moreover, two studies investigating vowel perception in FLA have yielded results inconsistent with PAM. In Sun and van Heuven [8], Mandarin speakers learning English in an FLA context showed very good discrimination of an SC contrast (/æ-ɑː/), against PAM's prediction. Lai [9], who also studied Mandarin FLA learners of English, found that a vowel pair (/æ-ɛ/) classified as two uncategorized L2 sounds (uncategorized-uncategorized [UU]) was better discriminated than six other pairs classified as one uncategorizable and one native exemplar (uncategorized-categorized [UC]), despite PAM's prediction that UC should show better discrimination than UU. These results can be interpreted in different ways. It is possible that Mandarin learners of English in an FLA context do not follow PAM, thus presenting counter-evidence to the applicability of the model to FLA. Lai's [9] results can be partly interpreted as evidence that PAM makes incorrect predictions for the relative difficulties of UC and UU. Finally, we need to consider the possibility that these discrepancies are due to methodological idiosyncrasies in the two studies [8, 9].

The purpose of the current study is to directly compare the relative difficulties with which Mandarin learners in FLA and SLA contexts perceive English vowels belonging to different PAM assimilation types. In order to test a wide range of assimilation

types, we include vowel pairs that fall into the categories of UC and UU in addition to those belonging to TC, CG, and SC. These classifications were carried out using a goodness-of-fit task with naïve Mandarin-speaking listeners of English. Furthermore, we tested both SLA and FLA learners using the same methods. If L2 perception develops in a similar way regardless of the learning contexts, the same order of perceptual difficulty should be observed in both SLA and FLA groups.

2. EXPERIMENT I. GOODNESS-OF-FIT RATING

2.1. Participants and stimuli

The participants in Experiment 1 were 15 naïve Mandarin speakers living in China (age range = 20–25 years, mean age = 23.4 years, 7 females). They were all Mandarin monolinguals with limited exposure to English.

The perceptual stimuli were the following British English vowels: /ɪ, e, æ, ɒ, ʌ, ʊ, ɑː, ɜː, iː, ɔː, uː, ə, aɪ, eɪ, əʊ, aʊ, ɔɪ, ɪə, eə, ʊə/ [10]. The vowels were recorded by three female native speakers of Southern Standard British English (mean age = 20.7 years) in a /h_t/ context. The stimuli were denoised and normalized to an amplitude of 70 dB, after which the vowel portions were extracted as the final stimuli.

2.2. Procedure

The experiment was conducted online using Qualtrics. All participants were asked to complete the test alone in a quiet room using a laptop and wearing headphones with sound levels adjusted to a comfortable level. They also completed a language background questionnaire after the rating task.

In the rating task, the participants were asked to listen to each English vowel and choose the Mandarin vowel that they thought was closest to the English vowel. The Mandarin vowels provided were the following: /i, ə, ɚ, ɤ, a, ɔ, u, y, ai, au, ei, ie, ia, ou, ua, uo, ye, uei/ [11, 12, 13]. The option "none of the above" was also provided. After they selected a Mandarin vowel, they were asked to rate its goodness-of-fit with the English vowel on a scale of 1 (not similar) to 7 (very similar).

2.3. Results

For each English vowel, the proportion of the Mandarin vowel that was most commonly selected (modal assimilation proportion) and its goodness-of-fit rating were multiplied to calculate the fit index (FI). Table 1 presents the modal assimilation proportion, goodness-of-fit score, and FI of each English vowel, shown in descending order of the FI within each assimilation set. An assimilation set is defined as a collection of L2–L1 pairs in which two or more L2 sounds are assimilated into a single L1 sound category [8]. For example, the first four L2–L1 pairs in Table 1 involve different English vowels that are all assimilated into the same Mandarin vowel (/ei/). Vowels that do not belong to any assimilation set are listed at the end of the table in descending order according to the FI. The English vowels are divided into different categories based on the FI using the standard deviation (SD) as the criterion. English vowels with an FI more than 1 SD (1.0) above the mean FI (2.5) were classified as "good" instances of a Mandarin category. Those with an FI between the mean and 1 SD above the mean were classified as "fair" instances of a Mandarin category. Those with an FI between the mean and 1 SD below the mean had a "poor" fit. Finally, those with an FI lower than 1 SD below the mean were considered "uncategorized."

English-Mandarin vowel assimilation	Modal proportion	Rating score	Fit index	
eɪ-ei	0.93	4.5	4.2	Good
e-ei	0.73	4.4	3.2	Fair
eə-ei	0.67	4.4	3.0	Fair
ɪ-ei	0.47	4.3	2.0	Poor
æ-a	0.93	4.5	4.2	Good
ɑː-a	0.93	3.1	2.9	Fair
ʌ-a	0.47	3.3	1.6	Poor
ə-ɤ	0.60	3.0	1.8	Poor
ɜː-ɤ	0.47	3.3	1.5	Poor
ʊə-ɤ	0.47	3.3	1.5	Poor
ɒ-ɔ	1.00	3.1	3.1	Fair
ʊ-ɔ	0.27	2.8	0.8	Uncategorized
iː-i	1.00	3.1	3.1	Fair
aɪ-ai	1.00	3.1	3.1	Fair
uː-y	0.93	3.2	3.0	Fair
aʊ-au	0.93	3.2	3.0	Fair
ɪə-ie	0.87	3.2	2.8	Fair
əʊ-ou	0.67	3.1	2.1	Poor
ɔː-uo	0.40	3.5	1.4	Uncategorized
ɔɪ-uei	0.40	2.8	1.1	Uncategorized

Table 1: Assimilation sets and fit indexes (FIs) for modal assimilated English–Mandarin vowel pairs.

Based on these, we selected the English contrast /eɪ-æ/ as a case of TC contrast, /æ-ɑː/, /æ-ʌ/, and /ɑː-ʌ/ as CG contrasts, and /e-eə/ and /ɜː-ʊə/ as SC contrasts. The pair /ɔː-ɒ/ involves one uncategorized sound and one categorized sound but, as the latter elicited an L1 response that overlaps with another uncategorized sound (/ɔ/), it was classified as UC+OL (UC with an overlap). Other UC pairs, i.e., /ɔː-eɪ/, /ɔː-iː/, /ɔː-ə/, /ɔɪ-eɪ/, and /ʊ-ʌ/, were classified as UC-OL (UC without an overlap). Contrasts involving two uncategorized sounds, i.e., /ʊ-ɔɪ/, /ɔː-ʊ/, and /ɔː-ɔɪ/, were classified as UU-OL (UU without an overlap).

3. EXPERIMENT II. AXB DISCRIMINATION TASK

3.1. Participants and stimuli

The participants in Experiment 2 were 15 FLA Mandarin speakers (age range = 19–23 years, mean age = 20.8, 7 females) and 15 SLA Mandarin speakers (age range = 21–23 years, mean age = 22.5, 8 females). All FLA participants were non-English majors, and none had any experience of going abroad or receiving English lessons from a native English speaker. Both FLA and SLA participants in this study had studied English in the classroom for several years (FLA: mean = 12.2 years, SD = 2.8; SLA: mean = 12.6 years, SD = 1.6); in addition, the SLA group had attended a university in the UK for an average of 2.8 years (SD = 0.8). The FLA learners' mean exposure to English in the classroom was 3.1 hours per week, with little to no exposure to English outside the classroom. The SLA learners' mean English exposure in the classroom was 22.5 hours per week, with their exposure outside the classroom approximately half of what they received in class, although with large individual differences (mean = 11.8 hours/week, SD = 11.4).

The stimuli were the vowels recorded in Experiment 1 in the context of [h_t] denoised and normalized. Each AXB trial contained three vowels produced by three different speakers. The A, X, and B components were concatenated with a 1.5-second inter-stimulus interval.

3.2. Procedure

The participants were instructed to decide whether the first or third sound was more similar to the middle sound in each trial. They were told that different speakers produced the three sounds in each trial and that they needed to ignore the differences in the speaker's voice. There were 120 trials in total (15 pairs * 8 trials). The fifteen vowel pairs were presented in a random order across subjects.

3.3. Results

Figure 1 shows the discrimination results of the 15 English vowel pairs in the two learning context groups.

A mixed two-way ANOVA was conducted with group (two levels) as a between-subjects factor and vowel pair (15 levels) as a within-subjects factor. The results show a main effect of vowel pair, $F(14, 392) = 44.116$ ($p < 0.001$), but no main effect of group or an interaction.

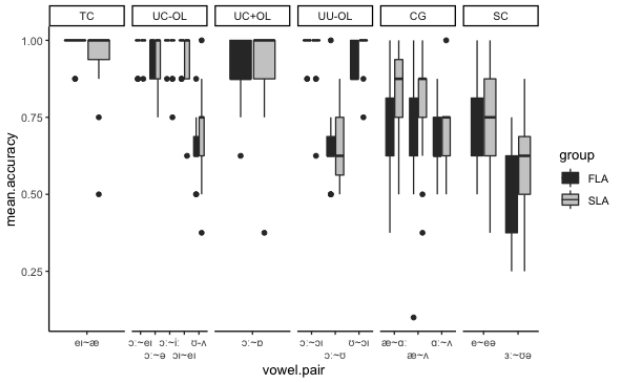

Figure 1: Discrimination accuracy of the 15 British English vowel pairs by SLA learners (N = 15) and FLA learners (N = 15).

A Tukey multiple pairwise-comparisons test was conducted to identify the sources of the vowel pair effect. The results are summarized in Figure 2.

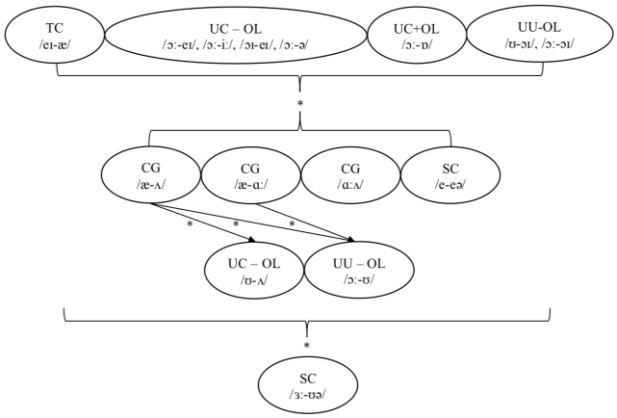

Figure 2: Summary of Tukey test results (* $p < 0.05$).

Discrimination levels were significantly higher for the TC pair /eɪ-æ/, most UC–OL pairs (/ɔː-eɪ/, /ɔː-iː/, /ɔɪ-eɪ/, /ɔː-ə/), the UC+OL pair /ɔː-ɒ/, and most UU–OL pairs (/ʊ-ɔɪ/, /ɔː-ɔɪ/) than all CG pairs (/æ-ɑː/, /æ-ʌ/, /ɑː-ʌ/), all SC pairs (/e-eə/, /ɜː-ʊə/), the UU–OL pair /ɔː-ʊ/, and the UC–OL pair /ʊ-ʌ/. Among the three CG pairs, only /æ-ʌ/ was significantly higher than the UC–OL pair /ʊ-ʌ/, and only /æ-ʌ/ and /æ-ɑː/ were significantly higher than the UU–OL pair /ɔː-ʊ/. The three CG pairs were all significantly higher than the SC pair /ɜː-ʊə/. Finally, the UU–OL pair /ɔː-ʊ/ and the UC–OL pair /ʊ-ʌ/ were significantly higher than the SC pair /e-eə/.

4. DISCUSSION

This study systematically examined the relative difficulties with which Mandarin listeners perceive English vowel contrasts depending on whether they are exposed to English in a predominately naturalistic context (SLA) or an instructed context (FLA). Based

on their discrimination of vowel pairs belonging to a range of PAM assimilation categories, no difference between the SLA and FLA groups emerged in our results, indicating that L2 learners follow similar developmental patterns in speech perception regardless of their learning contexts.

Our results also show that the majority of PAM-L2 predictions were borne out in both learning contexts. Pairs classified as TC, CG, or SC contrasts largely followed the predicted discrimination accuracy order of TC > CG > SC. Thus, the core of PAM-L2 applies not only to SLA but also to FLA. Note here that the contrast /æ-ɑː/, which was treated as a case of SC in Sun and van Heuven [8], was classified as CG in our study based on the outcome of Experiment 1. In [8], the accuracy level of that contrast was seen to be too high for an SC contrast, but this apparent counter-evidence to PAM-L2 is likely to be due to misclassification of /æ-ɑː/ for Mandarin listeners.

Not all PAM-L2 predictions were confirmed, however. Against the model, and as in Lai [9], two UU contrasts (/ʊ-ɔɪ/, /ɔː-ɔɪ/) were discriminated better than a UC contrast (/ʊ-ʌ/). Another point of discrepancy can be found in the UC contrasts with overlap (UC+OL) and without overlap (UC-OL). Recent work [14] proposes that UC+OL should have poorer discrimination than UC-OL. However, in the current study, the discrimination score for the UC+OL contrast /ɔː-ɒ/ was better than the UC-OL contrast /ʊ-ʌ/. If these findings are found to be generalizable to other L1–L2 pairs, they force a re-examination of the UC and UU categories in PAM-L2.

Finally, we note that we found no between-group difference in the overall discrimination accuracy of the target English contrasts. This is somewhat surprising considering that the SLA participants had lived in a target language environment for nearly three years on average, and we would, therefore, expect them to exhibit higher levels of attainment in their speech learning than the FLA participants. This suggests that the SLA learners in this study had not yet seen the benefits of an immersive linguistic experience (perhaps due to lower-than-expected use of English with native speakers); this also raises the possibility that SLA learners with longer or more intensive naturalistic target language exposure will exhibit a speech perception profile that is qualitatively different from that of FLA learners. Further work is required to determine long-term effects of learning contexts on L2 speech perception patterns.

5. REFERENCES

[1] Best, C.T., Tyler, M.D. 2007. Nonnative and second-language speech perception: Commonalities and complementarities, In: O.-S. Bohn & M. J. Munro (eds.), *Language Experience in Second Language Speech Learning: In Honor of James Emil Flege.* Amsterdam: John Benjamin, 13-34.

[2] Grimaldi, M., Sisinni, B., Gili Fivela, B., Invitto, S., Resta, D., Alku, P., Brattico, E. 2014. Assimilation of L2 vowels to L1 phonemes governs L2 learning in adulthood: a behavioral and ERP study. *Frontiers in Human Neuroscience*, 8.

[3] Bundgaard-Nielsen, R. L., Best, C. T., Tyler, M. D. (2011). Vocabulary size is associated with second-language vowel perception performance in adult learners. *Studies in Second Language Acquisition*, 33(3), 433-461.

[4] Guion, S. G., Flege, J. E., Akahane-Yamada, R., Pruitt, J. C. 2000. An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *The Journal of the Acoustical Society of America*, 107(5), 2711–2724.

[5] Ghaffarvand-Mokari, P., Werner, S. 2017. Perceptual assimilation predicts acquisition of foreign language sounds: The case of Azerbaijani learners' production and perception of Standard Southern British English vowels. *Lingua*, 185, 81–95.

[6] Lengeris, A., Hazan, V. (2007). Cross-language perceptual assimilation and discrimination of southern British English vowels by Greek and Japanese learners of English. *Proc. 16th ICPhS Saarbrücken*, 1641-1644.

[7] Rallo Fabra, L., Romero, J. 2012. Native Catalan learners' perception and production of English vowels. *Journal of Phonetics*, 40(3), 491–508.

[8] Sun, L., van Heuven, V. J. 2007. Perceptual assimilation of English vowels by Chinese listeners. *Linguistics in the Netherlands*, 24, 150–161.

[9] Lai, Y. H. 2010. English vowel discrimination and assimilation by Chinese-speaking learners of English. *Concentric: Studies in Linguistics*, 36(2), 157-182.

[10] Ogden, R. 2017. *Introduction to English Phonetics.* Edinburgh university press.

[11] Jia, G., Strange, W., Wu, Y., Collado, J., Guan, Q. 2006. Perception and production of English vowels by Mandarin speakers: Age-related differences vary with amount of L2 exposure. *The Journal of the Acoustical Society of America*, 119(2), 1118.

[12] Lee, W. S., Zee, E. 2003. Standard Chinese (Beijing). *Journal of the International Phonetic Association*, 33(1), 109-112.

[13] Wan, I. P., Jaeger, J. J. 2003. The phonological representation of Taiwan Mandarin vowels: A psycholinguistic study. *Journal of East Asian Linguistics*, *12*(3), 205-257.

[14] Faris, M. M., Best, C. T., Tyler, M. D. 2016. An examination of the different ways that non- native phones may be perceptually assimilated as uncategorized. *The Journal of the Acoustical Society of America*, 139(1), EL1–EL5.