INDIVIDUAL VARIABILITY IN L1 CATEGORY COMPACTNESS ON L2 PRODUCTION COMPACTNESS AND ACCURACY

Weiyi Zhai[a,c], Meghan Clayards[a,b,c], Heather Goad[a,c]

[a] Department of Linguistics, McGill University; [b] School of Communication Sciences and Disorders, McGill University; [c] Centre for Research on Brain, Language and Music, Montreal, Canada

weiyi.zhai@mail.mcgill.ca; meghan.clayards@.mcgill.ca; heather.goad@mcgill.ca

ABSTRACT

Individuals with the same native language (L1) differ in how successfully they acquire the sounds of a second language (L2). One source of individual differences explored in recent studies is L1 production variability. In this study, we examined whether individual variation in the compactness of Japanese L1 categories can be directly correlated with the compactness and accuracy of speakers' L2 English categories. The F2 and F3 values of 30 Japanese speakers' productions of L1 /ɾ/ were used to calculate individual L1 compactness. The F2 and F3 values of the same speakers' productions of L2 English /l/ and /ɹ/ were used to calculate speakers' L2 category compactness and production accuracy (as compared to L1 English). We found no correlation between speakers' L1 category compactness and L2 production compactness or accuracy.

Keywords: category compactness, production variability, second language acquisition

1. INTRODUCTION

1.1. Background

The speech of second language (L2) learners is often marked by non-native-like perception and pronunciation. Theories of L2 acquisition attribute learners' difficulties acquiring the sound system of the target language to cross-language interference where speakers' native language (L1) sound systems act as a filter that influences the production and perception of novel L2 sounds [5, 3]. However, it has been shown that speakers with the same L1 background and with similar levels of L2 proficiency can vary greatly in how native-like their L2 productions are (e.g. [9]). Research has also shown that L1 speakers differ from each other in L1 production [12, 13, 11].

Kartushina and Frauenfelder [8] investigated the role of L1 category variability in the production and perception of L2 vowels. They quantified variability in production as a single "Compactness Score" (CS). CS is calculated as the area of an ellipse using the standard deviation of the F1 and F2 values of a vowel as the two axes. Speakers who produced more variable, less consistent L1 vowels had higher compactness scores. The authors focused on native Spanish speakers' production of the French /e/ - /ɛ/ contrast which is assimilated to a single /e/ category in Spanish. Spanish speakers who produced more compact L1 /e/ also produced the French /e/ - /ɛ/ contrast more accurately as defined by a smaller acoustic distance between the L2 and native French speakers' productions of the same vowel in F1-F2 space. This study demonstrated that for native-like production of L2 phones, it is advantageous for L2 speakers to have more compact L1 categories.

Huffman and Schuhmann [7] showed that the link between L1 category variability and L2 production accuracy extends beyond the F1-F2 space to voice onset time (VOT). They measured English speakers' production of VOT in L1 word-initial stops and L2 Spanish stops. They measured the standard deviation of learners' L2 VOT over the course of a semester in an introductory Spanish class. They found that the learners' L1 English production compactness in stop VOT at the beginning of the semester correlated with the speakers' production accuracy of L2 Spanish VOT at the end of the semester as in [8]. Additionally, they found that speakers with more accurate L2 VOTs are also more compact in their L2 VOT productions.

Kartushina and Frauenfelder [8] suggested two possible explanations for the positive correlation between compact L1 categories and accurate L2 productions. First, having more compact categories could mean a greater distance in acoustic or articulatory space between L1 categories and new categories, making the new categories easier to learn [5]. Second, individuals' L1 category compactness could reflect cross-language articulatory skill. More compact categories could demonstrate higher precision and higher articulatory skill, and that this skill can vary across individuals. Articulatory precision should give speakers an advantage in learning any novel L2 phones.

If articulatory precision were the explanation, we would expect to find a correlation between speakers' L1 and L2 category compactness in addition to L1 compactness and L2 accuracy. The study of VOT [7] found that speakers with more compact L1 categories are more accurate in L2 productions and that speakers with more accurate L2 VOTs are also more compact

in their L2 VOTs, providing some support for this hypothesis. However, [8] tested the correlation between L1 and L2 category compactness directly and did not find a link. Therefore, it is still unclear whether there is such a link, and thus what the nature of the relationship is between L1 compactness and L2 production accuracy.

1.2. Current study

The current study aims to further explore the relationship between individual-specific L1 category compactness and L2 production. We do so by replicating the production experiment in [8] to extend it to the F2 - F3 space for consonants. We will test whether the compactness of native Japanese speakers' productions of the L1 flap /ɾ/ category is linked to production accuracy and compactness in the L2 English /l/ - /ɹ/ contrast by measuring F2 and F3. This contrast is notoriously difficult for Japanese L1 learners [1, 14, 6] and while large individual differences have been observed [6], the source(s) of these differences remain unclear. In addition to shedding light on sources of individual variability in Japanese learners of English, this will also allow us to replicate the link between L1 compactness and L2 accuracy found for vowels [8] and VOT [7] with a new set of categories and contrast. Secondly, we will test whether there is a direct link between L1 compactness and L2 compactness which would support a role for generalized articulatory skill in L2 learning. Additionally, as a preliminary exploratory analysis, we will look for any possible effects of coarticulation from the phonetic context on the compactness and accuracy of L2 productions. Since previous research did not control for the phonetic environment of the productions [3, 7], participants could have more or less compact categories due to more or less coarticulation. By exploring coarticulatory influences, we can determine whether speakers who are less compact in a category are so due to them having a general lack of articulatory precision, or having distinct productions of the same category depending on the phonetic context.

2. METHODS

2.1. Participants

After applying exclusion criteria, productions from 31 native Japanese speakers (19M, 12F) currently residing in Japan and 5 native English speaking controls from the United States were included in the analysis (3M, 2F). As a measure of proficiency, Japanese participants were asked to read a short English passage. Two sentences from the middle of the recording were used to assess global accent. Sentences were rated by 10 native English speakers on a scale from least (1) to most (9) fluent, with participants receiving an average of 5.7 (SD = 1.5) for English fluency.

None of the native English speakers self-reported any prior experience with Japanese or fluency in another language.

2.2. Stimuli and procedure

To elicit L1 production of Japanese liquids, a word list was created with 10 /ɾ/-initial two-mora target words. The full word list is available at the OSF: https://osf.io/wu7rq/?view_only=3735fd1c2fed46e4a525824007715207. The word list included the target L1 phone in a variety of following vowel environments. All words had an initial high pitch accent. Similarly, to elicit productions of the English liquids for the native Japanese and native English speakers, a word list was created with 20 /ɹ/-initial and 20 /l/-initial one-syllable words. The target liquids were initial singletons to match the profile of the Japanese targets and to control for any effects due to syllable position. Additionally, 30 Japanese and 94 English words were included as fillers for the Japanese and English production tasks respectively. Filler words matched the profile of the target words.

The experiment was created and hosted on Gorilla Experiment Builder [2]. All participants completed the experiment on their own laptops or desktop computers. Prior to completing experimental tasks, all participants filled out a language questionnaire. Native English participants completed the English production task. Native Japanese participants first completed the Japanese production task, followed by the English passage reading task, and finally the English production task.

The production task was disguised as a picture-naming memory task to avoid participants putting any deliberate focus on the target sounds. Each trial began with a familiarization phase in which participants were shown three picture-word pairings and were asked to memorize the pairings. In each trial, there was one target word and two filler words, and the position of the target and filler words was randomized between trials. The words were presented only orthographically in the familiarization phase. Participants were then presented with each of the three pictures individually in random order and asked to name the picture out loud.

For L1 productions, participants had 5s to familiarize themselves with the three picture-word pairings and 5s to record each word. For the L2 English productions, Japanese participants had 15s to familiarize themselves with the picture-word pairings and 5s to record each word.

2.3. Measurement and normalization

Formant values and duration data were extracted from the recordings with Praat [4]. For each speaker's productions, F1, F2, and F3 values of the liquid productions were measured at the earliest measurable point in the steady state of the approximant. To compare speakers across genders and different vocal tract sizes while preserving formant variation, Nearey Normalization [10] was used on all formant values.

2.3.1. Computation of category compactness

Compactness scores (CS) were calculated for the target liquids for each participant in their L1 and L2. Because the distribution of target phones was assumed to be elliptical and normally distributed, the formula for compactness scores (CS) was derived from the equation for the area of an ellipse in (1) following [8].

(1) $$\text{area} = ab\pi$$

a = ½ the length of the ellipse's major axis, b = ½ the length of the ellipse's minor axis

This is translated into F2-F3 space in (2).

(2) $$CS = \sigma F2 \sigma F3 \pi$$

σF2 = 1 standard deviation of the mean of normalized F2,
σF3 = 1 standard deviation of the mean of normalized F3

The area calculated in CS corresponds to a 66% confidence ellipse: the smallest ellipse that captures 66% of the data. Two sets of compactness scores were calculated for each of the Japanese participants. This includes two L1 compactness scores (L1CS) for the L1 /ɾ/ and two L2 compactness scores (L2CS) for the L2 /ɹ/ and /l/ categories. Higher CS values indicate less compact categories and lower CS values indicate more compact categories.

2.3.2. Computation of L2 production accuracy

Following [8], Japanese participants' L2 production accuracy was calculated as the Euclidean distance (EuD) in the F2-F3 space between each Japanese speakers' production and the average value of the native English-speaking controls' approximant productions in the same target word. This allowed us to control for any co-articulatory effects from the surrounding context. The EuD is the distance from one point to another in a two-dimensional space, calculated as in (3).

(3) $$x = \sqrt{(F2_i - F2_j)^2 + (F3_i - F3_j)^2}$$

Here, x is the EuD, F2i and F3i represent the Nearey normalized second and third formant values of a given Japanese participant's production, F2j and F3j represent the normalized formant values of the monolingual English participants. Greater EuD values represent lower production accuracy and smaller EuD values represent higher production accuracy.

3. RESULTS

For Japanese L1 productions, there were a total of 277 tokens of word-initial /ɾ/. The L1 compactness score (L1CS) for the Japanese liquids ranged from 0.0013 to 0.4361 (SD = 0.075).

For English L1 productions, there were a total of 106 tokens of word-initial /l/ and 110 tokens of word-initial /ɹ/. Figure 1 shows the F2 and F3 values for the L1 Japanese and L1 English productions.

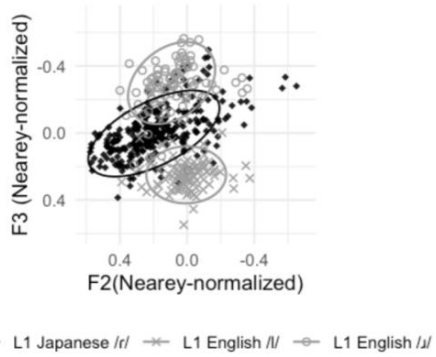

Figure 1: Scatter plot of L1 liquid productions.

For English L2 productions, 566 /ɹ/-initial and 533 /l/-initial L2 English productions were measured. Figure 2 shows the F2 and F3 values of all L2 English productions.

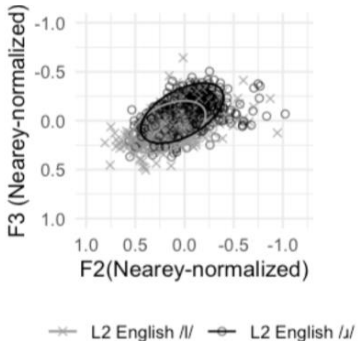

Figure 2: Scatter plot of L2 productions of English liquids.

Paired Pearson correlation analysis between speakers' L1 compactness scores and L2 EuD scores did not find evidence of a correlation ($r(28) = 0.46$, p

= 0.65). No correlation was found after excluding 5 data points that are visually identified as outliers (r(23) = 0.17 , p = 0.86; Fig. 3).

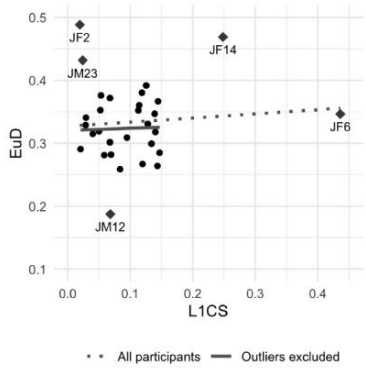

Figure 3: Correlation between L1CS and EuD with and without outliers. Each point in the graph represents a participant. Outliers are labeled.

We also did not find a correlation between speakers' L1 compactness scores and their L2 compactness scores (r(28) = 0.52, p = 0.61). No correlation was found after excluding 4 data points identified as outliers by visual inspection (r(24) = 0.40, p = 0.693).

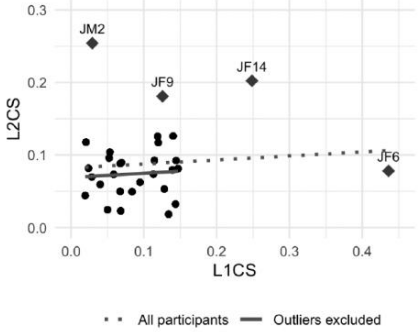

Figure 4: Correlation between L1CS and L2CS with and without outliers. Each point in the graph represents a participant. Outliers are labeled.

In addition, we did not find correlations between Japanese speakers' English fluency in the global accent test and L2 production accuracy (r(28) = -0.04, p = 0.83) or L2 compactness (r(28) = 0.04, p = 0.39).

A generalized linear mixed-effects model was built with L2 speakers' EuD values as the response variable. The model found that Japanese speakers produce more accurate English liquids when the liquid is followed by a back vowel compared to a non-back vowel (β = -0.04, σ = 0.02, $p < 0.05$). In addition, we also found a significant interaction between L1CS and following vowel backness (β = -0.19, SE = 0.06, $p < 0.005$). This shows that the effect of L1CS on EuD is modulated by the backness of the following vowel.

4. DISCUSSION AND CONCLUSIONS

Our first research question aimed to assess whether individual differences in native Japanese speakers' L1 category variability impact the production accuracy of L2 liquid consonants. We measured native Japanese speakers' L1 category compactness for /ɾ/ in the F2-F3 space. The results showed no evidence of a direct relationship between L1 category compactness and production accuracy in the L2. This result is inconsistent with what was found in [7] and [8]. This may be because having a more compact L1 /ɾ/ category in Japanese does not provide an advantage for native-like production of L2 English liquids, or we may have failed to find a link for other reasons. The second research question was whether there is a direct correlation between L1 and L2 category compactness. In [7], an indirect link was found between L1 compactness and L2 accuracy and between L2 accuracy and L2 compactness. However, like in [8], we did not find a correlation between L1 and L2 category compactness in the current study.

While we must be cautious in interpreting these null results, they could indicate that articulatory skill may not be the best explanation for the individual variation observed in L1 and L2 category compactness. It is possible that compactness in F2 - F3 space should be measured differently from compactness in the F1 - F2 space in [8]. Future research can explore different ways of measuring compactness for different categories.

In terms of coarticulatory influence of the following vowel backness, we found that L2 productions are more accurate when the liquid is followed by a back vowel. This is because English liquids have lower F2 compared to the Japanese flap [Fig. 1]. When the F2 of L2 productions are lowered due to coarticulation, the productions are also more accurate. The interaction between L1CS and vowel backness shows that for speakers with more compact L1 categories, the effect of back vowel coarticulation is small and speakers with less compact L1 categories are more prone to the effect of coarticulation. This result suggests that one reason behind speakers' category compactness in both L1 and L2 is speakers' resistance to coarticulation. Although this result does not explain the possible relationship between L1 compactness and L2 accuracy, it is an interesting area for future follow-up studies.

In conclusion, the results of this study indicate no direct relationship between L1 category compactness and L2 production accuracy. However, category compactness might be influenced by the degree of coarticulation that varies on an individual level.

5. REFERENCES

[1] Aoyama, K., Flege, J. E., Guion, S. G., Akahane-Yamada, R., & Yamada, T. 2004. Perceived phonetic dissimilarity and L2 speech learning: The case of Japanese /r/ and English /l/ and /r/. *Journal of Phonetics, 32*(2), 233-250.

[2] Anwyl-Irvine, A.L., Massonié, J., Flitton, A., Kirkham, N.Z., Evershed, J.K. 2019. Gorilla in our midst: an online behavioural experiment builder. *Behavior Research Methods.*

[3] Best, C. T., & Tyler, M. D. 2007. Nonnative and second-language speech. *Language experience in second language speech learning: In honor of James Emil Flege, 13*-34.

[4] Boersma, P., Weenink, D. 2022. Praat: doing phonetics by computer [Computer program]. Version 6.2.12, retrieved 17 April 2022 from http://www.praat.org/

[5] Flege, J. E. 1995. Second language speech learning: Theory, findings, and problems. *Speech perception and linguistic experience: Issues in cross-language research, 92*, 233-277.

[6] Hattori, K., & Iverson, P. (2009). English /r/-/l/ category assimilation by Japanese adults: Individual differences and the link to identification accuracy. *J. Acoust. Soc. Am., 125*(1), 469-479.

[7] Huffman, M. K., Schuhmann, K. S. 2020. The relation between L1 and L2 category compactness and L2 VOT learning. *In Proceedings of Meetings on Acoustics 179ASA* (Vol. 42, No. 1, p. 060011). Acoustical Society of America.

[8] Kartushina, N., Frauenfelder, U. H. 2014. On the effects of L2 perception and of individual differences in L1 production on L2 pronunciation. *Frontiers in Psychology, 5*, 1246.

[9] Moyer, A. 2004. Age, accent and experience in second language acquisition. Multilingual Matters.

[10] Nearey, T. M. 1978. Vowel space normalization in synthetic stimuli. *J. Acoust. Soc. Am. 63*(S1), S5-S5.

[11] Newman, R. S., Clouse, S. A., & Burnham, J. L. (2001). The perceptual consequences of within-talker variability in fricative production. *J. Acoust. Soc. Am., 109*(3), 1181-1196.

[12] Perkell, J. S., Guenther, F. H., Lane, H., Matthies, M. L., Stockmann, E., Tiede, M., & Zandipour, M. (2004). The distinctness of speakers' productions of vowel contrasts is related to their discrimination of the contrasts. J. Acoust. Soc. Am., 116(4), 2338-2344.

[13] Perkell, J. S., Matthies, M. L., Tiede, M., Lane, H., Zandipour, M., Marrone, N., ... & Guenther, F. H. (2004). The distinctness of speakers'/s/—/ʃ/contrast is related to their auditory discrimination and use of an articulatory saturation effect.

[14] Sheldon, A., & Strange, W. 1982. The acquisition of/r/and/l/by Japanese learners of English: Evidence that speech production can precede speech perception. *Applied psycholinguistics, 3*(3), 243-261.