Similar Vowels in L1/L2 Production: Confused or Discerned in Early L2 English Learners with Different amount of Exposure

E-Chin Wu
Independent Researcher, Taiwan
r95142002@ntu.edu.tw

1. Introduction

In past studies of L2 sound acquisition, one issue that has received much attention is how L2 sounds are incorporated into learners’ phonological system in the presence of L1. The consequence was the development of various models to explain for incapability of native-like production in non-native languages. While these models may differ in their predictions, they all evolve around the concept of perceived similarity.

In the native language magnet theory (NLM) [1], L1 sounds are viewed as “magnets” that attract similar non-native sounds. As a result, incoming non-native sounds are less discernable or even replaced by L1 sounds when they are heard as similar to L1. On the contrary, non-native sounds are readily differentiated when they do not lie in the “magnetic field” of any L1 sound. In the perceptual assimilation model (PAM) [2], incoming foreign sounds undergo a process of mapping. That is, naïve listeners try to map non-native sounds into pre-existing L1 categories according to their ideal of similarity. Again, when naïve listeners believe there is a perfect match between an L1 and a non-native sound, the non-native sound is assimilated to the L1 category in question. In SLM [3], it is explicitly stated that if the perceived phonetic dissimilarity between L1 and L2 sounds are large enough, new phonetic categories would be established. Otherwise, L1 sounds are likely to be substitute for L2 ones.

At large, whether the end result is to establish new sound categories or to replace incoming sounds with L1 sounds is largely dependent on perceived similarity. Thus, knowing what perceived similarity is equivalent to when applying these models is an important agenda. Yet, none of these models seem to stretch the topic much further, and what it means to be “similar” is not straightforward. At best, “new” and “similar” sounds, terms coined by [4], roughly correspond to the non-overlapping and overlapping sets of phoneme inventory proposed for each language being compared. Therefore, if language A and language B are both said to have a category for /u/, the /u/ sound in these languages is classified as “similar”.

According to such standards, Mandarin and English have three monothongs that can be categorized as similar; namely, /i/, /a/, and /u/. While cross-language comparisons have been rare, scattered remarks from different scholars may shield some light on how the Mandarin and English versions of these sounds differ. [5] commented that Mandarin /u/ is more posterior compared with the English /u/. In [6], Mandarin /u/ was found to be both lower in vertical height and more anterior. Personal observations suggested that Mandarin /i/ is less tensed than the English one. Transforming these remarks into acoustic correlates of formant values, we see that when Mandarin is compared to English: (1) /u/ is lower in terms of F2, (2) /a/ is higher in F1 and higher in F2, and (3) /i/ may be lower in F1, lower in F2, or both, depending on which parameters represent tenseness.

Back to the issue of similarity, if learning L2 sounds similar to L1 is a plain matter of categorization or assimilation of L2 sounds to L1, no acoustic difference should be found between similar sounds produced for either L1 or L2 as they are classified as a single phonetic category. Applying this to the present study, we would expect Mandarin children learning English to have only one phonetic category for the similar vowels found between Mandarin and English despite their subtle differences. In [4] and [7], results obtained from adults do show such a tendency.

However, there is a catch to this prediction. In the present study, the subjects analyzed were children, meaning that age of learning (AOL) was early. Compared with [4] or [7], whose subjects have learned English at adulthood, children learners are said to retain higher degree of discriminability in sounds. It would therefore be even more interesting to see whether groups of Mandarin children with two different levels of exposure to English would turn out to be the same as the adults studied in previous research.

2. Method

2.1. Subjects

In this study, there were a total of three groups with each having six children subjects. The three groups of children were recruited from three different kindergartens in Taiwan, depending on the amount of English instructions used in class. The first group of children went to a kindergarten that used English instructions all day long everyday. This group, the whole-day English exposure group, will henceforth be called the whole-day group. The second group of children were those who went to a kindergarten that used English instructions only half of a day everyday. Thus, their exposure to English is only half compared with the whole-day group. Consequently, this group, the half-day English exposure group, will henceforth be referred to as the half-day group. Finally, the last group of
subjects were those who went to a kindergarten that used no English instructions. This group, the no English exposure group, will henceforth be called the non-exposure group.

The age of these children at the time they were being tested was around 6 to 7 years old. They all came from Mandarin speaking households. For the children who went to kindergartens that taught English, i.e., the whole-day and half-day group, they started to learn English around the age of three to four. They used English only during class. For the children who went to kindergartens that do not use any English instructions, i.e., the non-exposure group, it was made sure that they did not receive any English instructions somewhere out of class for more than two hours. It was this group that served as the monolingual reference for Mandarin. Ideally, a group of English monolinguals should have been recruited to serve as the reference for English, but as such a group was inaccessible at the time, no monolingual English group was included.

2.2. Stimuli
Stimulus designing was a vexatious job. As we would like to compare Mandarin vowels to English ones, controlling for the phonetic environments the vowels occurred in was necessary to avoid variation due to context differences. Yet, the differences in the phonotactics of Mandarin and English made it difficult to come up with words pairs that had strictly similar environments. Children’s limited foreign language vocabulary added on to the challenges of conceiving exactly equivalent phonetic environments.

Compromising, 18 word pairs containing the vowels /i/, /a/, and /u/ with similar onsets (i.e., being with either the same place of articulation or manner) were chosen for the children to produce. In other words, there were five word pairs for every vowel compared. However, not all of the children were able to correctly produce every word presented to them. Thus, there were children who could only produce 3 pairs of words for each vowel. In the end, for the whole-day group, 25 token pairs were collected for /i/, 30 token pairs were collected for /a/, and 31 token pairs were collected for /u/. For the half-day group, 24 token pairs were collected for /i/, 26 token pairs were collected for /a/, and 23 token pairs were collected for /u/. For the non-exposure group, 24 token pairs were collected for /i/, 26 Mandarin tokens were collected for /a/, and 23 token pairs were collected for /u/.

2.3. Procedure
The subjects were tested individually in a room separated from others in the kindergarten. Stimuli were presented to them via a picture naming task. Oral cues were given to them when they had trouble figuring out what the picture was about, but the oral cues given never contained the target words to avoid any effect of imitation. English was always presented first to avoid possible influences of the dominating language. Before recording, the experimenter asked the children some language background questions such as their age, when they started to learn English, and what language they used at home. Recording sessions lasted for about five to ten minutes.

2.4. Measurements
Recordings were labeled and formant (F1, F2, and F3) values were measured in Praat [8]. The formant values measured were F1, F2, and F3. To eliminate possible influences of the preceding and following consonants, the formant values analyzed were obtained from 20 msec of the middle part of each vowel.

3. Result

3.1. Are /i/, /a/, and /u/ distinguished for Mandarin and English?
Figures 1 and 2 visually show the mean F1 and F2 values for the whole-day group and half-day group, respectively. As there is no standard way of presenting F3 on a vowel space graph, the values obtained for F3 were plotted separately in Figures 3, 4, and 5.

![Whole-Day](image1)
![Half-Day](image2)

Figures 1 (left-hand column) & 2 (right-hand column): The vowel spaces of Mandarin and English for the vowels /i/, /a/, and /u/ for the whole-day group and the half-day group, respectively.

Paired t tests were run on the F1, F2, and F3 of each vowel of each group to see if any difference in formant value existed between the similar vowels of Mandarin and English. It turned out that there were discrepancies.

Looking at the whole-day group first, we see that the vowel spaces of Mandarin and English did not have much overlapping, which meant that similar vowels from different languages were distinguished. Most conspicuously, all vowels seemed to be differentiated in terms of F2. Statistical analyses double confirmed this observation.

With regards to /i/, a significant difference was found for the F2 values [t(24) = -3.56, p < .05], but not for the F1 or F3 values. Specifically, English had a F2 value larger than that of Mandarin, indicating that English /i/ was more anterior compared with that of Mandarin. Paired t tests implemented on the vowel /a/ showed that Mandarin and English were not only different in F2 [t(29) = 6.09, p < .001], but also in F1 [t(29) = 3.56, p < .05] and F3 [t(29) = -2.32, p < .05]. This indicated that the English /a/ produced by children with whole-day English exposure was not only more posterior in the oral cavity, but also higher in terms of vertical height. Lower values in F3 may indicate that more lip rounding was involved in the production of Mandarin /a/, for literature views F3 as a correlate to the degree of lip spreading [9]. For the vowel /u/, significant differences were observed both in F2 [t(30) = -5.91, p < .001] and F3 [t(30) = 4.58, p < .001]. This illustrated that the whole-day group children were discerning between their English and Mandarin /u/s by making their English one more anterior and rounded.

For the half-day group, we see that apart from /i/, the rest of the vowels also seemed to be distinguished. Statistical results revealed the following: With regards to the vowel /i/, no formant value significantly deviated from each other. This particular finding was not compatible to that of the whole-day
group as the whole-day group did tell their Mandarin and English /i/ sounds apart in terms of F2. For the vowel /a/, while the whole-day group could distinguish between their Mandarin and English /a/ sounds by placing the English /a/ in a more fronted position and adding more lip rounding, the half-day group did not. For the vowel /u/, similar to the whole-day group, the half-day group also produced a more fronted /u/ compared with the non-exposure group. However, the vowel /i/ was the only vowel where similar discrimination could be observed between the whole-day and half-day groups.

3.2. Do differences exist between the same language produced by different groups?

Shown in Figure 6 are the English vowel spaces produced by the whole-day group and half-day group. Displayed in Figure 7 are the Mandarin vowel spaces produced by the whole-day, half-day, and non-exposure groups. Independent samples t tests were run to see if any discrepancy existed between the English vowels produced by the whole-day group and the half-day group. Results showed that they did not differ in any formant value measured for English /a/, but they did differ in the F1 values for /i/ (t(47)=2.25, p < 0.05) and F2 values for /u/ (t(52)= -2.03, p < 0.05). For the half-day group, /i/ was produced with lower F1, whereas /u/ was produced with higher F2.

For Mandarin, one-way ANOVA tests were executed on the data to see if the three groups produced their Mandarin /i/, /a/, and /u/ in identical manners. No difference was found among the production of Mandarin /i/ across all three groups. However, one-way ANOVAs implemented on the vowel /a/ revealed significant differences in F2 and F3 values [F2: F(83)= 5.54, p < .05; F3: F(83)= 12.98, p < .001]. Further post hoc tests reported that the group with ad hoc F2 value was the half-day group. Specifically, this group seemed to have a more fronted /a/ compared with the rest. The group with eccentric F3 value was that of the whole-day group. In particular, their F3 was much lower than the rest, indicating that they used a more rounded lip shape to produce their Mandarin /a/. Dissimilitude was also revealed in the vowel /u/ [F2: F(74)= 29.42, p < .001]. This time the unusual group was the non-exposure group. Compared with the other groups, the non-exposure group produced a much more fronted /u/.

4. Discussion

This study showed that Mandarin children learning English in Taiwan were able to consistently produce different formant values for similar vowels found in their L1 and L2. Specifically, their productions were in accordance with the language-specific traits described in previous research. An effect of exposure was also presented. This was manifested by the whole-day group showing overall more discriminability than the half-day group, where overall discriminability was obtained by calculating the differences in F1, F2, and F3 values for their Mandarin and English /i/, /a/, and /u/.

If we take the differences found in acoustic measurements in children’s production of similar vowels as evidence that children had established new phonetic categories for their L2, we are readily to say that SLM is correct in assuming that a new phonetic category can be established for an L2 sound. We also observed the phenomenon of deflection, which was another characteristic of L2 learning predicted by SLM. In details, the whole-day group had lower F3 values for their /a/ sounds compared with the non-exposure group, whose values represented the monolingual norms. This deflection in L1 was probably made by the whole-day group to maintain phonetic contrast between the /a/ categories established for Mandarin and English, respectively. Analogously, the half-day group made their L1 /a/ higher in F2 value so as to maintain enough
phonetic distance from their L2 /a/, resulting in divergence from their monolingual counterpart. The same was observed in the production of F2 for Mandarin /a/. Conspicuously, both the whole-day and half-day groups moved their F2 value downwards so as to make their L1 /u/ much further back in the vowel space compared to the non-exposure group.

One question arises from this point. In the case of Mandarin /a/, it is unclear why the whole-day group and the half-day group used different acoustic features to maintain phonetic distance. Their penchant may have something to do with the amount of English exposure they received as this was the only factor that differentiated the two groups, but the relation between the two is by no means straightforward and thus would require further research to disambiguate.

While the results found in the present study supported some hypotheses of the SLM, they also evoked a number of challenges to the model. First, it is predicted in SLM that new established categories were the results of perceived perceptual distance. While new categories were found to be established in the present study, we are uncertain if such discriminability in production was due to perceived perceptual differences per se as we did not test whether similar vowels were also distinguished in terms of perception for these children. Hence, the present study could not support that the impetus to new category establishment is caused by perceived perceptual differences as stated in SLM. If discrimination of sounds in production is due to perceived similarity of sounds, we predict that children would also discern similar sounds in a perception experiment. If false, perceived similarity may not be the sole trigger to category establishment.

Second, SLM predicts that similar and new sounds would be treated differently. As a rule of thumb, it is more likely to establish a new phonetic category for a new L2 sound and to have a pre-existing category modified to represent both the incoming similar L2 sound and the already existing L1 sound. However, results from the present study suggested that new categories could be founded for L2 sounds that were presumably “similar” to the L1 ones in the sense that they are represented by the same IPA symbols in both languages.

The implications of this are three folds. First, with regards to similarity, giving it a multi-dimensional definition would seem more adequate than just defining it in terms of perception alone. As indicated by this study, similarity could also be conceptualized in terms of acoustics because the acoustic measurements displayed also seemed to provide insights to how sounds were organized by learners. A similar point was made by [10] who pointed out that dimensions such as gestures, acoustics, and abstract phonology should be considered in the determination of cross-language phonetic distance. Second, also pertaining to the concept of similarity is the unit defining it. In previous studies, it was segments that served as the units being compared. While scholars are fully aware of the fact that perceived similarity among sounds represented by same IPA symbols can in fact be different for different language pairs being compared, they have not come up with a systemic way of quantifying or normalizing these differences so as to make comparisons beyond two languages to be unbiased. In particular, it would seem unjust to say that speakers of one language did a better job in learning a target language than the other when the phonetic distances the languages had with the target language were in fact asymmetrical. A potential way to solve this problem is to decompose segments into features, as proposed by [11]. Such an approach, according to [3], is also subjected to a number of problems, but is not unfeasible. With regards to English learning, large scale comparisons of English with other languages (either by multiple dimension scaling or acoustic feature comparing) would be an inchoate step to take to induce a comprehensive similarity pattern in both perception and production.

Finally, the last issue discussed here is the role of L1. In models such as SLM and PAM, L1 is given a rather ubiquitous role. Specifically, it is believed that L2 always goes through the sieve of L1. It is derived from this concept that similar sounds from L2 or a non-native language are less likely to establish a new category. Yet, we found that different formant values were used for similar vowels. We interpret this as evidence that category assimilation is not the only mode L2 learners resort to in the disposing of similar sounds, which implies that though L1 is pivotal in the disposition of L2, its effect is not altogether invincible.

5. Conclusion

In this study, early L2 learners were found capable of discerning similar L1 and L2 vowels acoustically. It was also found that the amount of exposure to L2 imposed a positive impact on degree of L1/L2 discriminability. These results provided support and evoked challenges to SLM. Bolsters to SLM included the fact that L2 learners do establish new categories, for L2 and L1 categories do go through category deflections. Challenges to SLM involve how the definition of similarity could be refined. Specifically, acoustic properties may be added to the definition of similarity and featural level units may be exploited. Finally, the role of L1 may be considered to be abated somewhat in L2 learning models.

6. Reference


