THE ROLE OF L1 PRODUCTION COMPACTNESS ON THE L2 PRODUCTION ACCURACY

Na Zhi[1,2], Aijun Li[1], Yuan Jia[1]

[1.] **Institute of Linguistics, Chinese Academy of Social Sciences, Beijing, China**
[2.] **Capital Normal University, Beijing, China**
zhinacn@yeah.net; liaj@cass.org.cn; summeryuan_2003@126.com

ABSTRACT

This study explores the role of different speakers' L1 vowel compactness on their L2 vowel production efficiency. It is found in study [4] that Spanish speakers who have more compact L1 vowel space can perform better than those "dispersed" speakers in the distinction and production of French vowels, including both the L2 "assimilated" vowel-contrasts, ones similar to the L1 vowel categories, and the "uncategorized" L2 vowels, ones as unfamiliar and new sounds to L1 speakers. In light of [4], this paper examines the correlation between the L1 compactness variety and the L2 production accuracy with two groups of Chinese speakers, Beijing speakers and Jining-mandarin diglossia ones. The individual speakers' native production differences, their English vowel performance, and the relation between the two are presented in the study.

Key words: compactness, accuracy, Beijing speakers, Jining-mandarin diglossia speakers

1. INTRODUCTION

The study [4] proposed that native speakers can be highly variable in their L1 production, thus the same speaker may have inconsistent production in the sound of the same category. The intra-speaker variability in pronunciation can be indicated by the compactness degree of the within-category sound, and the global compactness of between-category sounds. The individual differences of L1 speakers based on their production compactness can lead us to classify some speakers as "stable and precise" ones, while other speakers as "overlapped and sloppy" ones. The previous speakers have compact F1/F2 acoustic space in the same-category vowel production and clear boundaries for vowels of different categories, while the latter, in contrast, have varied and dispersed vowel space within category, thus often present overlapped boundaries across categories.

In the study of [4], the individual differences in L1 production are found to be transferred to the speakers' L2 performance. The speakers who have careful and consistent production in L1 vowels often perform better in their L2 vowel production than dispersed and varied L1 speakers. The evidence is in favour of the surface transfer of L1 phonetics in L2 perception and production [2], and the belief that compact speakers have more blank and available acoustic space to accommodate the new target sounds. The difference in articulatory skill across speakers is also considered as a factor in determining the varied L2 production accuracy. Therefore, it gives sense that accurate L1 speakers are more capable in distinguishing the L1-L2 sound difference and in establishing L2 categories that are close to the targets than the less precise L1 speakers.

Due to the surface transfer of native sounds to L2 perception and production, the sounds which are perceived similar to native vowels are assimilated to the existing L1 category, e.g., to Spanish speakers, the French vowel contrast, /e-ɛ/ is similar to the Spanish vowel /e/, thus considered as the "assimilated" category; while the new L2 sounds, such as the French /ø-œ/ contrast, which cannot map to any native Spanish vowel category is classified as the "uncategorized" type. In [4], it is found that the L2 vowel contrasts can be distinguished better by speakers if their similar L1 vowels are produced in a compact and precise way, while the new L2 sounds are more accurately performed when the global compactness of L1 is small and stable.

1.1. The Compactness Score

The vowel compactness can be assessed based on two measures which calculate the F1/F2 acoustic space [4]. One measure is the compactness score for specific vowels (CS_V) with the following mathematical formula (1), where σF1 and σF2 is the standard deviation of the average F1 value, and that of the average F2 value, respectively.

$$CS_V = \pi\sigma F1\sigma F2 \quad (1)$$

The CSv computes the compactness of each vowel category produced by an individual speaker.

The other measure computes the compactness of the global vowel categories (CS_G). The global compactness score is the sum of CSv. For example, a Chinese mandarin speaker's CS_G can be derived through the sum of CSv of each vowel category, as presented in the following formula (2).

(2) CS_G = CSv(/a/) + CSv(/ɤ/) + CSv(/i/) + CSv(/ɿ/) + CSv (/ʅ/) + CSv (/u/) + CSv (/y/)

1.2. The Distance score

Mahalanobis distance score (DS) is employed for computing the distance between the L2 speakers' vowel performance and the production of the same vowel by native speakers. In study [4], the accuracy of French vowels produced by Spanish speakers was assessed by calculating the DS between the L2 vowel space (the dimensions defined by F1 and F2) and the target space of the same vowel by French native speakers. The vowel space of the target language was employed as the referential stimuli for judging the production accuracy of L2 vowels. The smaller the mean distance between the vowels of L2 speakers and the vowels of native French speakers, the more accurate the L2 production is considered.

This current study in light of [4] examines the role of individual difference in L1 production and its impact on L2 production with two groups of Chinese speakers, Beijing speakers and Jining-mandarin diglossia ones. Three English vowel contrasts: /ɪ/-/i/, /ʌ/-/a/ and /æ/-/e/ produced by two groups of Chinese speakers are studied, and their production accuracy is respectively computed to find the correlation between the L2 production efficiency and the speakers' L1 vowel compactness of /i/, /a/ and the global L1 vowel performance.

2. DATA SELECTION AND METHODOLOGY

2.1. Speakers

The data is selected from the corpus [3], in which speakers from different regions of China were instructed to read at a natural and moderate speed a list of words, sentences and one passage in their local dialect, mandarin and English.

In this study, we select 6 Beijing Chinese speakers and 6 Jining-mandarin diglossia speakers, with 3 males and 3 females in each group. Jining is a region in Shandong Province, located in the central part of China. The diglossia speakers, esp., the young are able to speak both mandarin and the local dialect fluently. Mandarin is used widely in their life at schools and offices, and presented in most TV programmes, while the Jining dialect is employed in their daily communication with families and local friends. For the diglossia speakers, both the mandarin and the Jining dialect are regarded as their first language.

The above 12 Chinese speakers are all college students with the age of around 20-years old, and they have taken at least 6 years of English-language classes since their initial acquaintance with English in Junior middle school. The 12 speakers have no experience in living in an English-speaking country, and they acquire the language mostly by following the classes. The Chinese participants are judged as similar in their levels of English pronunciation, according to a native English speaker's judgement on their oral performance.

For the referential stimuli of the English data produced by native English speakers, the recordings of three male UK speakers are employed. The 3 speakers are from the same region of UK in order to minimize the regional dialect influence in the native vowel production.

2.2. Monosyllabic words

Of the above 15 speakers' recordings, all the monosyllabic words are selected from the data of different languages. The aim of employing only the monosyllabic data is for a consistent criterion on word stress across languages. It is also believed that in the monosyllabic reading task, the speakers can present clear and stable vowel production, which can decrease the odds of formant changes as seen in connected and emotional speech [5].

Each monosyllabic word selected in the study contains one of the vowels of interest. For the mandarin data, the 7 cardinal vowels, /i/, /ɿ/, /ʅ/, /a/, /u/, /y/ and /ɤ/ are included; the six cardinal vowels of the Jining data, /i/, /ʅ/, /a/, /u/, /y/ and /ɤ/ are employed; for the English data, three vowel contrasts /ɪ/-/i/, /ʌ/-/a/, and /æ/-/e/ are particularly considered.

The phonetic context of each vowel is varied in different words, e.g. the English vowel /ʌ/ is situated in contexts, such as [b_d], [p_f], [d_k], [ð_s] and etc. However, the monosyllabic data which has a nasal coda, [m], [n], [ŋ], or a lateral coda [l], or which has an approximate onset, such as [w], [r], [j] are discarded to eliminate the significant contextual effect on vowel formants. The stimuli of different languages employed in the study are summarized:

Table 1: The sum of monosyllabic data in different languages.

Speaker / Language	Beijing (6 speakers)	Jining-mandarin (6 speakers)	UK (3 speakers)
Mandarin	606	606	
JN dialect		672	
English	1506	1506	393

Among the stimuli of each language, the different vowels are found in unequivalent number of words, due to the fact that some vowels

conditioned by the language phonotactic complexity, occur more frequently than other vowels in the language. The following Table 2 and Table 3 present the vowel occurrence in the adopted stimuli.

Table 2: The sum of respective vowels in mandarin and Jining dialect.

Chinese	/i/	/ʅ/	/ʮ/	/a/	/u/	/y/	/ɤ/
Mandarin (12 speakers)	168	60	84	264	360	228	48
JN dialect (6 speakers)	96		78	126	180	114	78

Table 3: The sum of respective vowels in English produced by UK, Beijing and Jining speakers.

English	/ɪ/	/i/	/ʌ/	/a/	/æ/	/e/	/ɒ/	/ɔ/
UK (3)	45	46	38	26	113	67	38	20
BJ (6)	330	162	180	90	264	276	84	60
JN (6)	330	162	180	90	264	276	84	60

2.3. Data analysis

All the vowels have been labelled and analyzed with Praat [1]. For each vowel token, the F1 and F2 values (in Hz) are the average of the respective values derived at the 10 equal points in the vowel steady state.

With Formula (1) and (2), the CS_V of each vowel category produced by the individual speaker, and the CS_G of the speaker in respective languages can be derived. It deserves notice that for Jining speakers both mandarin and Jining dialect are their first language, thus each Jining speaker's CS_G of L1 is the average value of their CS_G in mandarin and in the dialect. For the 12 Chinese speakers, the compactness value of the global L1 vowel production and of the specific vowel categories of /i/ and /a/ are presented in the following table.

Table 4: Each Beijing (BJ) and Jining (JN) speaker's L1 CS_G and the CS_V of /i/ and /a/.

compactness / speaker	L1 global	/i/	/a/
BJ-1	130147	16301	22755
BJ-2	64402	15813	7276
BJ-3	264178	15444	20076
BJ-4	377049	6265	9025
BJ-5	486333	34338	5102
BJ-6	449339	8251	8250
JN-1	295490	74366	26671
JN-2	175204	10058	10552
JN-3	119965	17233	30661
JN-4	334555	21307	29555
JN-5	293610	11132	15910
JN-6	288491	3474	13226

For assessing the accuracy of the 12 Chinese speakers' production of English vowel contrasts, /ɪ-/i/, /ʌ-/a/, and /æ-/e/, the Mahalanobis Distance Score between each vowel's F1/F2 formants produced by L2 speakers and those by the referential UK speakers are calculated.

To make the comparison across speakers of different genders, we firstly normalize the raw F1/F2 values (in Hz) through the Lobanov z-score procedure [6]. The normalization can eliminate the physiological variation in formant values across genders while preserving the linguistic difference. In the following step, with the normalized data of each vowel token, we compute the distance score in Matlab to compare each of the 12 Chinese speakers' L2 vowel space and the corresponding one of 3 native English speakers. The smaller the mean distance, the more accurate the L2 vowel production is considered.

The distinction of the three English vowel contrasts are estimated by calculating the sum of the respective distances, e.g., the /ɪ/-/i/ DS value is derived by summing up the value between /ɪ/ (L2 speaker)-/ɪ/ (native speaker) and the value between the /i/ (L2 speaker)-/i/ (native speaker). The same computation is done to the /ʌ/-/a/, and /æ/-/e/ contrasts. In this way we present the performance of the three vowel contrasts by each individual L2 speaker, which are indicated by the DS values.

Table 5: The DS of English vowel contrasts by Beijing (BJ) and Jining (JN) speakers.

DS / Speaker	/ɪ/-/i/	/ʌ/-/a/	/æ/-/e/
BJ-1	8.19	7.79	6.70
BJ-2	38.37	10.57	11.29
BJ-3	8.81	10.50	8.11
BJ-4	23.81	10.06	13.06
BJ-5	13.10	8.19	9.77
BJ-6	11.20	4.50	13.45
JN-1	8.06	9.02	10.08
JN-2	8.16	18.07	8.63
JN-3	9.28	18.45	7.83
JN-4	7.53	13.45	9.94
JN-5	11.97	7.29	7.07
JN-6	5.96	9.03	9.49

In the above table, the distinction of the "assimilated" English vowel contrasts, /ɪ-i/ and /ʌ-a/, as well as the "uncategorized" vowel contrast /æ-e/ by different Chinese speakers are presented. It can be seen that the DS is varied in different speaker's production. The smaller the DS value, the better the distinction of the contrast is achieved.

3. RESULTS AND DISCUSSION

The study of [4] proposed that the L2 production accuracy is related to the L1 vowel compactness. Speakers who have more compact L1 vowel space

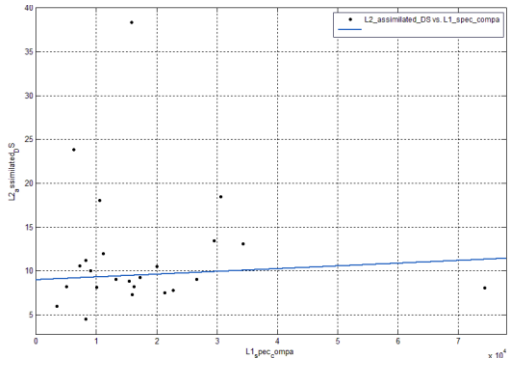
will perform better in the distinction of similar vowel categories. While the speakers who have compact global L1 vowel space can achieve more accuracy in distinguishing the uncategorized vowels. For both Beijing and Jining speakers, the English vowel contrasts of /ɪ-i/, and /ʌ-a/ are the similar categories comparing to their L1 /i/ and /a/ vowels, respectively.

We firstly evaluate such proposal by computing the correlation between each speaker's /a/ production compactness and the assimilated English contrasts of /ʌ-a/, and the same method is applied to the relation between Chinese /i/ compactness and the same speaker's English /ɪ-i/ contrast distinction. As in [4], to analyze the assimilated and uncategorized English vowels produced by Chinese speakers, the DSs are fitted to a general linear mixed-effects model, represented with the formula (3), where k is the slope, and b is a constant.

(3) $$DS = k*CS + b$$

Through computing the correlation between each speaker's distinction score of assimilated vowel contrasts and the L1 vowel compactness, we derive the result of Figure 1, which revealed that the coefficients (with 95% confidence bounds) k is 3.135e-05, and b is 8.998.

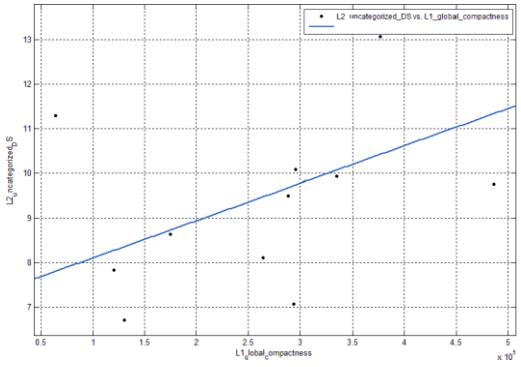
Figure 1: Regression line for the correlation between 12 speakers' L1 compactness of /i/ and /a/, and the DS of L2 vowel contrasts, /ɪ-i/ and /ʌ-a/.



The positive value of k indicates that the smaller the L1 compactness value, the smaller the DS of the L2 vowel contrasts. The result reveals that the participants who have more compact L1 vowel production of /i/ and /a/ have more accurate L2 performance in vowel contrasts of /ɪ-i/ and /ʌ-a/.

In addition, the relation between the global L1 vowel compactness of each speaker and their English pronunciation accuracy of the uncategorized /æ-e/ is computed in the following Figure 2.

Figure 2: Regression line of 12 speakers' global compactness of vowel categories, and the DS of uncategorized vowel contrast /æ-e/.



The correlation result reveals that the coefficients (with 95% confidence bounds) k is 8.369e-06 and b is 7.268, which presents a close correlation between the L1 CS_G and the production efficiency of the L2 vowel contrast /æ-e/.

According to the above two correlation analyses, the study is in line with the proposal [4] that the individual speaker's native space in production can affect how the L2 sounds are accommodated and performed. Speakers who are more compact in L1 production have more blank and available region left between the categories for the new sounds to fall in, which helps in distinguishing L1-L2 difference.

4. CONCLUSION

This study examines the role of intra-speaker variety in L1 production on affecting the L2 vowel performance. With the data of two groups of Chinese speakers, we compute each individual's compactness degree of native vowel performance. It is found through the correlation analyses that the speaker's L1 compactness can predict their L2 vowel performance of similar categories, and the native speakers who are stable and precise in global vowel performance are good in the distinction of uncategorized or new vowel contrasts. The paper provides some evidence of surface transfer and the belief that consistent speakers have more available acoustic space to accommodate the new target sounds. The difference in individual's articulatory skill should also be considered in determining the varied L2 production accuracy.

5. ACKNOWLEDGEMENT

This work was supported by the National Basic Research Program (973 Program) of China (No. 2013CB329301), CASS innovation project 'Key Laboratory of Phonetics and Speech Science'.

6. REFERENCES

[1] Boersma, P., Weenink, D. 1992-2015. *Praat: a System for Doing Phonetics by Computer*.

[2] Flege, J. E. 1995. Second language speech learning theory, findings, and problems. In: Strange, W. (ed.) *Speech Perception and Linguistic Experience: Issues in Cross-language Research*, 233-277.

[3] Jia, Y., Li, A. J., Tseng, C. Y. 2013. Construction of English learners from Chinese dialectal regions. *Chinese Journal of Phonetics 4*, 38-45.

[4] 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*, 1-17.

[5] Li, A. J., Cao, M. X., Fang, Q., Hu. F., Dang, J. W. 2013. Acoustic and Articulatory analysis on Chinese and Japanese Vowels in Emotional Speech. *Chinese Journal of Phonetics 4*, 125-142.

[6] Lobanov, B. M. 1971. Classification of Russian vowels spoken by different speakers. *Journal of the Acoustic Society of America 49*, 606-608.