

Article

# Acoustic Analyses of L1 and L2 Vowel Interactions in Mandarin–Cantonese Late Bilinguals

Yike Yang 

Department of Chinese Language and Literature, Hong Kong Shue Yan University,  
North Point, Hong Kong; yyang@hksyu.edu; Tel.: +852-2104-8264

**Abstract:** While the focus of bilingual research is frequently on simultaneous or early bilingualism, the interactions between late bilinguals' first language (L1) and second language (L2) have rarely been studied previously. To fill this research gap, the aim of the current study was to investigate the production of vowels in the L1 Mandarin and L2 Cantonese of Mandarin–Cantonese late bilinguals in Hong Kong. A production experiment was conducted with 22 Mandarin–Cantonese bilinguals, as well as with 20 native Mandarin speakers and 21 native Cantonese speakers. Acoustic analyses, including formants of and Euclidean distances between the vowels, were performed. Both vowel category assimilation and dissimilation were noted in the Mandarin–Cantonese bilinguals' L1 and L2 vowel systems, suggesting interactions between the bilinguals' L1 and L2 vowel categories. In general, the findings are in line with the hypotheses of the Speech Learning Model and its revised version, which state that L1–L2 phonetic interactions are inevitable, as there is a common phonetic space for storing the L1 and L2 phonetic categories, and that learners always have the ability to adapt their phonetic space. Future studies should refine the data elicitation method, increase the sample size and include more language pairs to better understand L1 and L2 phonetic interactions.

**Keywords:** acoustic analysis; speech production; vowel; bilingualism; Mandarin; Cantonese



**Citation:** Yang, Y. Acoustic Analyses of L1 and L2 Vowel Interactions in Mandarin–Cantonese Late Bilinguals. *Acoustics* **2024**, *6*, 568–578. <https://doi.org/10.3390/acoustics6020030>

Academic Editors: Georgios P. Georgiou and Jian Kang

Received: 24 April 2024

Revised: 12 June 2024

Accepted: 12 June 2024

Published: 17 June 2024



**Copyright:** © 2024 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

While most investigations on bilinguals' speech have focused on the characteristics of either their first language (L1) or their second language (L2), recent studies have suggested potential interactions between bilingual speakers' L1 and L2 phonetic systems [1,2]. Late bilinguals are the ideal population for studying L1 and L2 speech interactions because they start to learn their L2 after puberty and naturally avoid any maturational constraints related to language acquisition. This study explored the interactions of Mandarin and Cantonese monophthongs in Mandarin–Cantonese late bilinguals via acoustic analyses to provide a better understanding of the issue of bidirectional influence.

### 1.1. Interactions between L1 and L2 Speech in Late Bilinguals

The term 'late bilinguals' refers to people who commence learning their L2 after having fully acquired their L1. Investigations of late bilinguals' speech development have mainly focused on their acquisition of L2 speech, based on the assumption that the L1 of bilinguals does not alter. However, more recent studies have shown that the L2 of late bilinguals interferes with the full-fledged L1, which, in turn, causes the L1 of bilinguals to differ from the L1 of monolinguals [3,4]. This study focuses on the speech development of late bilinguals, who are usually immigrants who have relocated to a new environment in which their L2 is the dominant language, and in which their L1 is no longer used or is used less frequently [5,6]. Starting with the same level of L1 competence, late bilinguals do not have the potential maturational constraints in child and adolescent language development [7], and are therefore an ideal population for investigating L1 and L2 interactions.

Studies of the L1 and L2 in bilingual language development have developed in two separate directions. Until very recently, only a few studies have touched upon the issue of L1 and L2 interactions and confirmed the bidirectional influences of the L1 and L2 [8,9]. To the best of our knowledge, there are only three studies that have examined the L1 and L2 speech interactions [1,2,10]. Ref. [10] tested the production of voice-onset-time (VOT) by late Dutch–German bilinguals while [1] investigated the perception of VOT by English–Spanish bilinguals. Both studies suggest interactions between the bilinguals’ L1 and L2, but it is unclear whether similar findings can be observed when vowels are considered, which is the gap the current study aimed to fill. Refs. [1,10] interpreted their results as providing supporting evidence for Flege’s Speech Learning Model (SLM), as introduced below.

### 1.2. Theoretical Framework

The L1 and L2 speech interactions can be accounted for by the SLM [11] and its revised version, the Revised Speech Learning Model (SLM-r) [12], according to which there is a common phonetic space in a bilingual speaker’s mind. This space stores the phonetic categories of both the L1 and the L2; the L1 and L2 categories can thus mutually influence each other, in the process of category assimilation or category dissimilation.

The category assimilation hypothesis (CAH) in the SLM claims that in the common space, an L2 sound that is perceived as being similar to an L1 sound does not form a new category and is understood as a variant of the L1 sound at an allophonic level; that is, the cross-linguistic equivalence between the two sounds has been established. In this case, the phonemic variants for interdialectal contact are called diaphones [13], and the CAH advocates that only one single phonetic category is used to process the two linked diaphones. This mapping of diaphones will eventually give rise to a new merged category in mental representation of a bilingual, which will be realised differently from either the L1 sound or the L2 sound in production, and this phenomenon has been documented in several studies. For example, ref. [3] examined the post-vocalic /r/ of German–English bilinguals and discovered an influence from the L2, resulting in the assimilation of the consonant pair, which lends support to the CAH.

SLM also postulates the category dissimilation hypothesis (CDH). A new category will be established if an L2 sound is absent in the L1 system, which will make the combined phonetic space more crowded, and as a result, the phonemes will tend to disperse to compensate so that the phonetic contrast can be maintained. When category dissimilation operates, neither the newly established L2 category nor the closest L1 category will be identical to the categories for monolinguals, and consequently, both categories may shift away from their original phonetic space. One support for the CDH can be found in [14], which reported that Spanish–Catalan bilinguals have developed two categories to accommodate the mid-back vowels in the two languages, respectively.

Moreover, the SLM and SLM-r posit that the capacity for speech learning remains intact over the lifespan. This claim is particularly relevant to and can be tested by the current study. If it is the case, there must be L1 and L2 speech interactions even for late learners. The current study also extends the SLM and SLM-r to explore any L2-induced change(s) in the bilinguals’ L1.

### 1.3. The Current Study

Although Mandarin and Cantonese are two varieties (dialects) of Chinese, they have different phonological systems and are mutually unintelligible [15]. While there is no consensus regarding the numbers or distribution of monophthongs in these two varieties [16–20], it is acknowledged that Mandarin and Cantonese share three peripheral vowel pairs, namely, /a/, /i/ and /u/ [21], which were chosen as the target vowels to be examined in this study.

Thus far, the investigations on bilingual speakers’ vowel systems have largely concerned bilingual children (e.g., [22,23]). It remains to be explored as to whether there will

be interactions between the late bilinguals' L1 and L2 vowel systems, given that late bilinguals are mature learners with a full-fledged L1, which is very different from the case of bilingual children. The current study aimed to investigate the production of monophthongs in the L1 Mandarin and the L2 Cantonese of Mandarin–Cantonese bilinguals in Hong Kong (henceforth 'bilinguals') to answer the following research questions:

- (1) Are the L1 Mandarin vowels of the bilinguals influenced by Cantonese after years of immersion?
  - (a) Have the F1 and F2 in bilinguals' Mandarin undergone assimilation, namely, become more similar to their Cantonese counterparts and more deviant from those of the monolingual Mandarin speakers?
  - (b) Have the F1 and F2 in bilinguals' Mandarin undergone dissimilation, namely, become more deviant to their Cantonese counterparts and even shifted away from those of the monolingual Mandarin speakers?
- (2) Have the bilinguals reached native-like competence in L2 Cantonese in terms of vowel production?
  - (a) Have the F1 and F2 in bilinguals' Cantonese undergone assimilation, namely, become more similar to their Cantonese counterparts and more deviant from those of the monolingual Mandarin speakers?
  - (b) Have the F1 and F2 in bilinguals' Cantonese undergone dissimilation, namely, become more deviant to their Cantonese counterparts and even shifted away from those of the monolingual Mandarin speakers?
- (3) Are there any interactions between the bilinguals' L1 and L2 vowels?

## 2. Materials and Methods

### 2.1. Informants

Three groups of informants were recruited to participate in this study via online advertisements. In the pre-screening stage, all the potential informants were first invited to provide their background information pertaining to their language use, based on which we invited eligible informants. The bilinguals group consisted of 22 new immigrants (19 females, three males; aged  $30.14 \pm 4.30$ ) who spoke Mandarin as their L1 and had been exposed to Cantonese since their arrival in Hong Kong. The bilinguals had all arrived in Hong Kong after puberty (average age:  $22.73 \pm 4.21$ ) and their average length of residence was  $7.41 \pm 3.11$  years. To assess their language profile of Cantonese and Mandarin, the bilinguals completed a language background questionnaire prior to the recording session. The questionnaire was an adapted version of the Bilingual Language Profile [24], which was used to collect information about the participants' language history, language use, language proficiency and language attitudes, and the results were converted into scores for each subsection. The scores showed that, the participants were fluent Cantonese speakers, although they were more dominant in Mandarin at the time of the experiment. The Mandarin baseline group consisted of 20 native speakers of Mandarin (11 females, nine males; aged  $24.75 \pm 3.65$ ), who were born and raised in Mandarin-speaking regions and had little exposure to Cantonese. There were 21 native speakers of Cantonese in the Cantonese baseline group (ten females, 11 males; aged  $20.78 \pm 2.56$ ), and they were born and brought up in Hong Kong, where Cantonese is the native and dominant language. No participants had any history of speaking, hearing, or language difficulties.

### 2.2. Materials and Procedures

The vowels /a/, /i/ and /u/ were chosen as the target vowels for the following reasons. Firstly, as these three vowels are shared by and are commonly used in Mandarin and Cantonese, it was possible to use them to conduct cross-linguistic comparisons. Moreover, these vowels are peripheral vowels, which are ideal for measuring the vowel space of the informants. In addition, all three vowels are monophthongs, the subtle differences in which are easier to capture compared to the differences in more complicated diphthongs

or triphthongs. The vowels were embedded in either the first or second syllable of disyllabic words. Because tones have been shown to influence the production of vowels [25], we restricted our target stimuli to the first tone (T1) in Mandarin and Cantonese to minimise the coarticulation effect [26]. The words appeared in the subject or object position in different sentences. For the native speakers of Cantonese and Mandarin, each vowel appeared 15 times. For the bilinguals, the vowels appeared ten times in each language. In total, 3165 vowel tokens were collected: 3 vowels  $\times$  15 times  $\times$  20 Mandarin speakers + 3 vowels  $\times$  15 times  $\times$  21 Cantonese speakers + 3 vowels  $\times$  10 times  $\times$  22 bilinguals  $\times$  2 languages.

To make the data collection setting more naturalistic, the speech data were collected in dialogues between the experimenter and the participants, wherein the experimenter always asked the precursor questions, and the participants were instructed to answer the questions using the provided sentences naturally. An example of the Cantonese questions and answers is presented in (1) below, with the target syllable in the answer underlined ('gaa1'):

- (1) Question: 嗰個嘉賓拎乜嘢?  
 go2 go3 gaa1ban1 ling1 mat1je5  
 that CL guest carry what  
 'What did the guest carry?'  
 Answer: 嗰個嘉賓拎膠樽。  
 go2 go3 gaa1ban1 ling1 gaau1zeon1  
 that CL guest carry plastic\_bottle  
 'The guest carried a plastic bottle.'

The Mandarin and Cantonese speakers attended the recording session for their respective language, while the bilinguals were recorded in both Mandarin and Cantonese. Since exactly the same materials were used in each language, it was possible to compare the vowel production by bilingual and native speakers directly.

This project was approved by the Human Research Ethics Committee of Hong Kong Shue Yan University (Reference number: HREC 22-05 (M12)). All the participants gave their written informed consent prior to the recording sessions.

### 2.3. Data Analysis

To process the data, the vowel portions of the target syllables were manually segmented by trained phoneticians and the values of the first and second formants (F1 and F2) were extracted over the midpoint of each vowel with a script in Praat [27]. Individual differences among speakers are huge, and to eliminate the potential effect of inter-speaker variation on our data analysis, we adopted the Lobanov's approach [28] to normalise each speakers' F1 and F2 values individually, using Equation (1):

$$F_{n[V]}^N = (F_{n[V]} - \text{MEAN}_n) / S_n \quad (1)$$

where  $F_{n[V]}^N$  is the normalised formant value of the  $N^{\text{th}}$  formant for the vowel  $V$ ,  $F_{n[V]}$  stands for the original formant values that are measured in Hz, and  $\text{MEAN}_n$  and  $S_n$  are the mean and standard deviation (SD) of the  $N^{\text{th}}$  formant of the target speaker, respectively. To make the F1 and F2 values comparable to the findings in previous studies [23,29], the normalised formant values were then rescaled to Hz following [30] with Equations (2) and (3) below:

$$F'_1 = 250 + 500(F_1^N - F_{1\text{MIN}}^N) / (F_{1\text{MAX}}^N - F_{1\text{MIN}}^N) \quad (2)$$

$$F'_2 = 850 + 1400(F_2^N - F_{2\text{MIN}}^N) / (F_{2\text{MAX}}^N - F_{2\text{MIN}}^N) \quad (3)$$

where  $F'_i$  is the rescaled formant,  $F_i^N$  is the Lobanov normalised formant value, and  $F_{i\text{MIN}}^N$  and  $F_{i\text{MAX}}^N$  are the minimum and maximum values of  $F_i^N$ , respectively, across the dataset of the target speaker.

Next, the rescaled F1 and F2 values were then analysed with linear mixed-effects modelling using the 'lme4' package [31] in R [32,33], with the formant values (F1 or F2) as the dependent variables, vowel, speaker group (or language) as the fixed effects, and speaker and repetition as the random effects.

In addition, to measure the relative difference within each proposed pair accurately, we calculated the Euclidean distances of the Mandarin and Cantonese monophthongs based on the rescaled F1 and F2 values with Equation (4) [34]:

$$s(m, c) = \sqrt{(F1m - F1c)^2 + (F2m - F2c)^2} \quad (4)$$

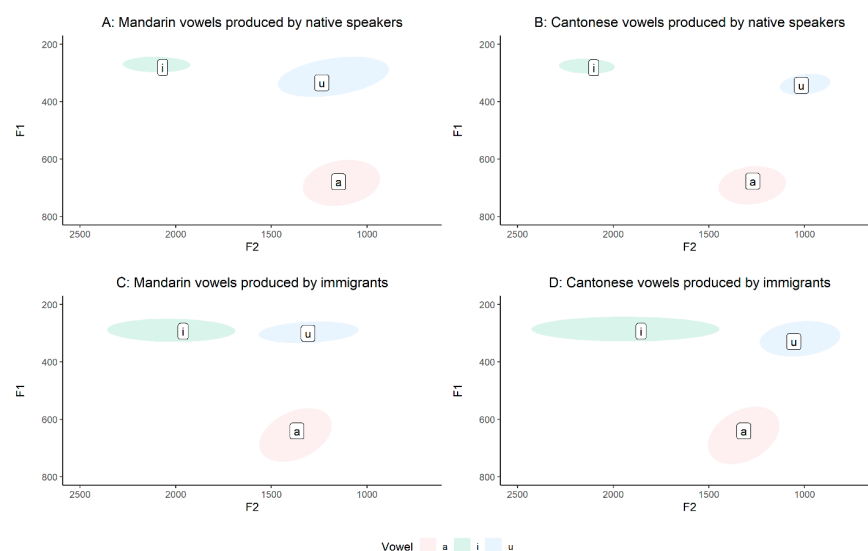
where  $s$  is the distance between two points in a two-dimensional Euclidian vowel space defined by F2 on the x axis and F1 on the y axis, and  $m$  and  $c$  each represent a specific monophthong in Mandarin and Cantonese. For the native speakers, the average F1 and F2 values for each vowel were used to calculate the Euclidean distances. For the bilinguals, the distances between the vowel pairs were calculated for each speaker.

### 3. Results

In this section, we will first present an overview of the vowel production data and then report on the statistical analyses of the F1 and F2 values of the vowels produced by native speakers in Section 3.1.1. The comparisons of native speakers and bilinguals for each language are presented separately in Sections 3.1.2 and 3.1.3 (Questions 1 and 2). Next, we compare the F1 and F2 of the vowels in the bilinguals' Mandarin and Cantonese in Section 3.1.4 (Question 3), which is followed by an interim summary of the F1 and F2 values. Finally, the Euclidean distances of the vowels are calculated and compared in Section 3.2.

#### 3.1. F1 and F2 of the Vowels

An overview of the vowel production in Mandarin and Cantonese by the bilingual and monolingual speakers is plotted in Figure 1, in which the vowel letters represent the average F1 and F2 values of each vowel and the circles indicate approximately 67% of the vowel ellipses for each vowel category.



**Figure 1.** F1 and F2 of the vowels produced by native speakers and bilinguals. (A,B) represent Mandarin and Cantonese vowel production of native speakers, respectively, while (C,D) show Mandarin and Cantonese vowels produced by immigrants. The circles indicate 67% of the vowel ellipses for each vowel category.

### 3.1.1. Vowel Production by Native Speakers

We first fitted models for the F1 and F2 of the vowels produced by native speakers of Mandarin and Cantonese. There was a main effect of vowel ( $\chi^2(2) = 2834, p < 0.001$ ) but no main effect of language ( $\chi^2(1) = 1.46, p = 0.226$ ) on F1, suggesting that the three vowels /a/, /i/ and /u/ were distinguishable in height, and that native speakers of Mandarin and Cantonese showed no height difference when producing these three pairs of vowels. For the F2 values, there were main effects of vowel ( $\chi^2(2) = 2265.8, p < 0.001$ ) and language ( $\chi^2(1) = 48.316, p < 0.001$ ), as well as an interaction of vowel and language ( $\chi^2(2) = 36.023, p < 0.001$ ). Specifically, the vowels within a pair differed in the degree of backness: /i/ of both languages overlapped in backness ( $p = 0.115$ ); /a/ was more back in Mandarin ( $p < 0.001$ ) and /u/ was more back in Cantonese ( $p < 0.001$ ).

### 3.1.2. Mandarin Vowel Production by Native Speakers and Bilinguals

Next, we compared the F1 and F2 of the Mandarin vowels produced by native speakers and bilinguals. With regard to the F1 values of the vowels in Mandarin, there were main effects of vowel ( $\chi^2(2) = 3348.2, p < 0.001$ ) and speaker group ( $\chi^2(1) = 4.426, p = 0.035$ ), as well as a two-way interaction of vowel and speaker group ( $\chi^2(2) = 54.933, p < 0.001$ ). A post hoc analysis of the group effect showed that the native speakers of Mandarin had larger F1 values compared to the bilinguals, suggesting that the vowels produced by the native speakers were generally lower than those produced by the bilinguals. For the specific vowels, the F1 values of /a/ and /u/ were significantly lower for the native speakers ( $ps < 0.001$ ) while the bilinguals and the native speakers had comparable F1 values for /i/ ( $p = 0.092$ ).

For the F2 values of the vowels in Mandarin, there were main effects of vowel ( $\chi^2(2) = 2010.3, p < 0.001$ ) and speaker group ( $\chi^2(1) = 6.853, p = 0.009$ ), as well as a two-way interaction of vowel and speaker group ( $\chi^2(2) = 141.01, p < 0.001$ ). Post hoc tests showed that the vowels /a/ and /u/ produced by native speakers were more back than those produced by the bilinguals ( $ps < 0.001$ ). Conversely, the vowel /i/ was more back in the bilinguals' Mandarin ( $p < 0.001$ ).

### 3.1.3. Cantonese Vowel Production by Native Speakers and Bilinguals

This subsection presents the analyses of the F1 and F2 of the Cantonese vowels produced by the native speakers and the bilinguals. For the F1 values of the vowels in Cantonese, there was a main effect of vowel ( $\chi^2(2) = 3677.5, p < 0.001$ ) and a two-way interaction of vowel and speaker group ( $\chi^2(2) = 48.571, p < 0.001$ ). A post hoc analysis showed that the native speakers of Cantonese had larger F1 values for /a/ and /u/ but smaller F1 values for /i/ compared to the bilinguals ( $ps < 0.001$ ).

For the F2 values, there were main effects of vowel ( $\chi^2(2) = 2525.6, p < 0.001$ ) and speaker group ( $\chi^2(1) = 5.260, p = 0.022$ ), as well as a two-way interaction of vowel and speaker group ( $\chi^2(2) = 178.93, p < 0.001$ ). The results of the post hoc tests suggested no difference between the two speaker groups in the F2 of the vowels /a/ ( $p = 0.120$ ) or /u/ ( $p = 0.265$ ). For the vowel /i/, the native speakers of Cantonese had smaller F2 values compared to the bilinguals ( $p < 0.001$ ).

### 3.1.4. Vowel Production by Bilinguals

Lastly, we fitted models for the F1 and F2 of the Mandarin and Cantonese vowels produced by the bilinguals. For the F1 values, there was a main effect of vowel ( $\chi^2(2) = 2955.2, p < 0.001$ ) but no effect of language ( $\chi^2(1) = 1.085, p = 0.298$ ). The two-way interactions between vowel and language reached significance ( $\chi^2(2) = 21.619, p < 0.001$ ). The vowels /a/ and /i/ shared similar F1 values in the bilinguals' Mandarin and Cantonese, but the vowel /u/ was produced higher in Mandarin than it was in Cantonese ( $p < 0.001$ ).

For the F2 values, there were main effects of vowel ( $\chi^2(2) = 1330.8, p < 0.001$ ) and language ( $\chi^2(1) = 38.829, p < 0.001$ ), as well as a two-way interaction of vowel and language ( $\chi^2(2) = 44.186, p < 0.001$ ). The post hoc analyses revealed no differences in the F2 values

of the vowel /a/ between the two languages. However, the vowels /i/ and /u/ had higher F2 values in Mandarin than they did in Cantonese ( $ps < 0.001$ ), suggesting that they were more back in the bilinguals' Cantonese.

### 3.1.5. Interim Summary

Table 1 provides a summary of the F1 and F2 statistics reported in this section. The native speakers did not show any difference in the formants of the vowel /i/, but the vowel /a/ was more back and the vowel /u/ was more front in Mandarin. The bilinguals differed from the native speakers in all three vowels, suggesting cross-linguistic influences from both Mandarin and Cantonese.

**Table 1.** Summary of the F1 and F2 statistics.

	/a/	/i/	/u/
Baseline groups	More back in Mandarin	No difference	More front in Mandarin
Mandarin	Lower and more back for native speakers	More front for native speakers	Lower and more back for native speakers
Cantonese	Lower in native speakers	More front and higher for native speakers	Lower for native speakers
Bilinguals	No difference	More front in Mandarin	Higher and more front in Mandarin

### 3.2. Euclidean Distances of the Vowels

As the native speakers of Mandarin and Cantonese only produced vowels in their respective language, it was impossible to compare the Euclidean distances of the vowels produced by each speaker directly. Instead, we calculated the Euclidean distances of the vowels produced by the native speakers based on the average F1 and F2 values of each speaker group. For the bilinguals, we calculated the distances of each vowel pair based on the average F1 and F2 values of the informant's Mandarin and Cantonese and listed the average distances and SDs.

The Euclidean distances of the vowels are presented in Table 2. For the native speakers, the distance between the Cantonese /i/ and Mandarin /i/ was the smallest, but the bilinguals showed a much larger distance between their Cantonese /i/ and their Mandarin /i/. Both native speakers and bilinguals exhibited the largest distance for the vowel /u/. With regard to the vowel /a/, both groups showed a moderate distance.

**Table 2.** Euclidean distances of the vowels produced by native speakers and bilinguals.

	/a/	/i/	/u/
Native speakers	121.52	33.29	219.74
Bilinguals (SDs)	103.45 (84.11)	172.59 (151.96)	284.98 (147.17)

## 4. Discussion

This study investigated the production of L1 and L2 vowels by Mandarin–Cantonese bilinguals in Hong Kong and addressed three research questions: (1) Are the L1 Mandarin vowels of the bilinguals influenced by Cantonese after years of immersion? (2) Have the bilinguals reached native-like competence in L2 Cantonese in terms of vowel production? (3) Are there any interactions between the bilinguals' L1 and L2 vowels?

With regard to the first research question concerning Mandarin vowel production, the data suggested differences in the Mandarin vowels produced by the bilinguals compared to those produced by the native speakers. Specifically, in the bilinguals' production, the front vowel /i/ became more back and the back vowel /u/ became more front, and the low vowel /a/ became higher, suggesting a more crowded Mandarin vowel space for the bilinguals. In addition, the vowel /a/ was also more front for bilinguals, making it further from

the same vowel produced by the native speakers. According to our analysis of native speakers' production, the vowel /a/ was more back in Mandarin than it was in Cantonese. It is possible that, due to the extensive exposure to Cantonese, the bilinguals shifted their way of producing the Mandarin vowel /a/ towards that of the Cantonese vowel /a/, although there was a difference in the backness of the vowel /a/ in Cantonese and in Mandarin. In this case, the Mandarin /a/ was assimilated to the Cantonese /a/ in the bilinguals' phonetic space, resulting in only one merged category to represent these two vowels.

Next, for the second research question on Cantonese vowel acquisition, we demonstrated that the bilinguals failed to successfully acquire the Cantonese vowels through immersion in the language. In the bilinguals' production, the vowel /i/ was more back and lower, and both vowel /a/ and vowel /u/ were higher compared to the vowels produced by the native Cantonese speakers. Note that the bilinguals were advanced Cantonese learners and their average length of residence in a Cantonese-speaking region was 7.41 years at the time of the recording. Despite their exposure to the target language, they had not yet succeeded in producing native-like vowel formants. This might be explained by the maturational constraints [35] or age effects [36] involved in language learning. It has been advocated that one should start to acquire an L2 as early as possible in order to attain native competence in the L2. The target group in this study consisted of late Mandarin–Cantonese bilinguals who had started to learn Cantonese after puberty (the average age of acquisition was 22.73), which may have prevented them from becoming successful learners of L2 Cantonese. However, as there is a lack of research on early Mandarin–Cantonese bilinguals' acquisition of Cantonese vowels, more data should be obtained before we can provide support for the claims regarding the maturational constraints or age effects. Another possible direction for future research could be to investigate whether the L2 Cantonese speech was accented, given the observed differences in the vowel formants, which would contribute to our understanding of the source of foreign accent in L2 speech and the relationship between accentedness and acoustic distances [37].

As shown above, the vowels in the bilinguals' L1 Mandarin and L2 Cantonese were generally deviant from the vowels in the corresponding native language. Our final research question concerned interactions between L1 and L2 vowel systems, the evidence for which was abundant because both category assimilation and dissimilation could be identified in the bilinguals' vowel production. The bilinguals' Cantonese /a/ showed no difference in terms of backness compared to the vowel /a/ produced by the native Cantonese speakers, suggesting that the bilinguals appeared to have successfully formed an /a/ category in their L2 Cantonese. As suggested in [12], such L2 category formation is not an easy task for L2 speakers because there is already a full-fledged L1 category in place, and the quality and quantity of L2 input usually vary. Moreover, in the bilinguals' production, the Mandarin /a/ was more front than the same vowel produced by the native Mandarin speakers; that is, the Mandarin /a/ produced by the bilinguals had shifted away from the native norms. Furthermore, in the bilinguals' own production of the vowel /a/, no distinction between Mandarin and Cantonese was made, indicating that the two vowel categories had merged into one. The question that then arises concerns which factors may have contributed to the bilinguals' L2 category formation and L1–L2 category merging. For the vowel /i/, the native speakers of Mandarin and Cantonese did not show any differences in the F1 or F2, and the Euclidean distance between the Mandarin /i/ and Cantonese /i/ was extremely small. With regard to the bilinguals, their production of /i/ differed from the /i/ production of the two native groups, and their own Mandarin /i/ and Cantonese /i/ productions were also different from each other. As shown in Figure 1, the /i/ tokens produced by the native speakers were consistent, but the /i/ tokens produced by the bilinguals were variable in both languages. A similar phenomenon was observed for the /u/ tokens, as the bilinguals also attempted to maintain a contrast between Mandarin and Cantonese. While the reason that the bilinguals tended to merge some categories but simultaneously preferred to split other categories remains to be explored, it is clear that there were interactions between the bilinguals' L1 and L2 vowel categories, thus replicating previous



findings regarding interactions between L1 and L2 consonants and extending the applicability of SLM and SLM-r to vowel categories.

The findings of this study must be seen in light of some limitations. Firstly, in terms of the elicitation method, the vowels were produced in sentences as responses to precursor questions asked by the experimenter. It has been demonstrated that different elicitation methods will influence how vowels are produced, leading to varying degrees of individual differences [38]. The issue of phonetic accommodation is also worth noting because there is recent evidence that both the L1 and L2 undergo phonetic accommodation [39], so that the L1 and L2 vowels produced in this study may have been affected by the experimenter's vowels. Future studies could consider using the more traditional read speech approach, which would reflect the participants' actual production. Secondly, the collected data were conversational in nature, and the phenomenon of vowel reduction is not uncommon in naturalistic speech [40,41], which may have had some negative impacts on the quality of the vowel production. It is, therefore, necessary for future studies to include words in isolation [23,30] or to prepare the target words in the focused position [42,43] to make it possible to elicit the vowels that are uttered more clearly. In addition, as a pilot study exploring L1 and L2 vowel interactions, this study had a relatively small sample size, with 3,165 tokens of three peripheral vowels in Cantonese and Mandarin. To have a better understanding of L1 and L2 vowel trajectories, it would be useful to consider more vowels and to include other language pairs that are more typologically different.

## 5. Conclusions

In summary, this study investigated the vowel production of Mandarin–Cantonese bilinguals and revealed vowel category assimilation and dissimilation in the participants' L1 and L2 vowels, thus indicating interactions between their L1 and L2 vowel systems. In general, the findings are in line with the hypotheses of SLM and SLM-r in that L1–L2 phonetic interactions are inevitable because there is a common phonetic space for storing the L1 and L2 phonetic categories, and that learners always have the ability to adapt their phonetic space. Future studies should refine the data elicitation method, increase the sample size and include more language pairs to better understand L1 and L2 phonetic interactions.

**Funding:** The work described in this paper was partially supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project Number: UGC/FDS15/H15/22) and an ASA International Student Grant from the Acoustical Society of America.

**Institutional Review Board Statement:** The study was conducted in accordance with the Declaration of Helsinki, and approved by the Human Research Ethics Committee of Hong Kong Shue Yan University (Reference number: HREC 22-05 (M12); date of approval: 1 June 2022).

**Informed Consent Statement:** Informed consent was obtained from all informants involved in the study.

**Data Availability Statement:** The original contributions presented in the study are included in the article; further inquiries can be directed to the corresponding author.

**Acknowledgments:** The author would like to thank all the informants for their participation in the recording experiment.

**Conflicts of Interest:** The author declares no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

## References

1. Gorba, C. Bidirectional influence on L1 Spanish and L2 English stop perception: The role of L2 experience. *J. Acoust. Soc. Am.* **2019**, *145*, EL587–EL592. [CrossRef] [PubMed]
2. Yang, Y. First Language Attrition and Second Language Attainment of Mandarin-Speaking Immigrants in Hong Kong: Evidence from Prosodic Focus. Ph.D. Thesis, The Hong Kong Polytechnic University, Hong Kong, 2022. Available online: <https://theses.lib.polyu.edu.hk/handle/200/11661> (accessed on 1 July 2022).

3. Ulbrich, C.; Ordin, M. Can L2-English influence L1-German? The case of post-vocalic /r/. *J. Phon.* **2014**, *45*, 26–42. [[CrossRef](#)]
4. Chang, C.B. A novelty effect in phonetic drift of the native language. *J. Phon.* **2013**, *41*, 520–533. [[CrossRef](#)]
5. Köpke, B.; Schmid, M.S. Language attrition: The next phase. In *First Language Attrition: Interdisciplinary Perspectives on Methodological Issues*; Schmid, M.S., Köpke, B., Keijzer, M., Weilemar, L., Eds.; John Benjamins: Amsterdam, The Netherlands, 2004; pp. 1–43.
6. de Leeuw, E.; Opitz, C.; Lubiska, D. Dynamics of first language attrition across the lifespan. *Int. J. Biling.* **2013**, *17*, 667–674. [[CrossRef](#)]
7. Flege, J.E.; Birdsong, D.; Bialystok, E.; Mack, M.; Sung, H.; Tsukada, K. Degree of foreign accent in English sentences produced by Korean children and adults. *J. Phon.* **2006**, *34*, 153–175. [[CrossRef](#)]
8. Meir, N.; Walters, J.; Armon-Lotem, S. Bi-directional cross-linguistic influence in bilingual Russian-Hebrew children. *Linguist. Approaches Biling.* **2017**, *7*, 514–553. [[CrossRef](#)]
9. Hopp, H.; Bail, J.; Jackson, C.N. Frequency at the syntax–discourse interface: A bidirectional study on fronting options in L1/L2 German and L1/L2 English. *Second Lang. Res.* **2020**, *36*, 65–96. [[CrossRef](#)]
10. Stoehr, A.; Benders, T.; van Hell, J.G.; Fikkert, P. Second language attainment and first language attrition: The case of VOT in immersed Dutch–German late bilinguals. *Second Lang. Res.* **2017**, *33*, 483–518. [[CrossRef](#)] [[PubMed](#)]
11. Flege, J.E. Second Language Speech Learning: Theory, Findings, and Problems. In *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*; Strange, W., Ed.; York Press: Timonium, MD, USA, 1995; pp. 233–277.
12. Flege, J.E.; Bohn, O.-S. The Revised Speech Learning Model (SLM-r). In *Second Language Speech Learning: Theoretical and Empirical Progress*; Wayland, R., Ed.; Cambridge University Press: Cambridge, UK, 2021; pp. 3–83. [[CrossRef](#)]
13. Weinreich, U. On the Description of Phonic Interference. *Word* **1957**, *13*, 1–11. [[CrossRef](#)]
14. Simonet, M. Production of a Catalan-specific vowel contrast by early Spanish-Catalan bilinguals. *Phonetica* **2011**, *68*, 88–110. [[CrossRef](#)]
15. Zhang, X. Dialect MT: A case study between Cantonese and Mandarin. In Proceedings of the Coling 1998, Montreal, QC, Canada, 10–14 August 1998; pp. 1460–1464. [[CrossRef](#)]
16. Lee, W.; Zee, E. Standard Chinese (Beijing). *J. Int. Phon. Assoc.* **2003**, *33*, 109–112. [[CrossRef](#)]
17. Zee, E. Chinese (Hong Kong Cantonese). *J. Int. Phon. Assoc.* **1991**, *21*, 46–48. [[CrossRef](#)]
18. Lin, Y.-H. *The Sounds of Chinese*; Cambridge University Press: Cambridge, UK, 2007.
19. Bauer, R.S.; Benedict, P.K. *Modern Cantonese Phonology*; Walter de Gruyter: Berlin, Germany, 1997. [[CrossRef](#)]
20. Duanmu, S. *The Phonology of Standard Chinese*; Oxford University Press: Oxford, UK, 2007.
21. Shi, F.; Peng, G.; Liu, Y. Vowel Distribution in Isolated and Continuous Speech: The Case of Cantonese and Mandarin. In *The Oxford Handbook of Chinese Linguistics*; Wang, W.S.-Y., Sun, C., Eds.; Oxford University Press: New York, NY, USA, 2015; pp. 459–473. [[CrossRef](#)]
22. Yang, J.; Fox, R.A. L1–L2 interactions of vowel systems in young bilingual Mandarin-English children. *J. Phon.* **2017**, *65*, 60–76. [[CrossRef](#)]
23. Yang, J. Vowel development in young Mandarin-English bilingual children. *Phonetica* **2021**, *78*, 241–272. [[CrossRef](#)] [[PubMed](#)]
24. Birdsong, D.; Gertken, L.M.; Amengual, M. Bilingual Language Profile: An Easy-to-Use Instrument to Assess Bilingualism. COERLL, University of Texas at Austin. 2012. Available online: <https://sites.la.utexas.edu/bilingual/> (accessed on 1 July 2022).
25. Hoole, P.; Hu, F. Tone-Vowel Interaction in Standard Chinese. In Proceedings of the First International Symposium on Tonal Aspects of Languages, Beijing, China, 28–30 March 2004; pp. 89–92.
26. Bergmann, C.; Nota, A.; Sprenger, S.A.; Schmid, M.S. L2 immersion causes non-native-like L1 pronunciation in German attriters. *J. Phon.* **2016**, *58*, 71–86. [[CrossRef](#)]
27. Boersma, P.; Weenink, D. Praat: Doing Phonetics by Computer. 2015. Available online: <http://www.praat.org/> (accessed on 1 October 2022).
28. Lobanov, B.M. Classification of Russian Vowels Spoken by Different Speakers. *J. Acoust. Soc. Am.* **1971**, *49*, 606–608. [[CrossRef](#)]
29. Torres, C.; Li, W.; Escudero, P. Acoustic, phonetic, and phonological features of Drehu vowels. *J. Acoust. Soc. Am.* **2024**, *155*, 2612–2626. [[CrossRef](#)]
30. Yang, J.; Fox, R.A. Acoustic development of vowel production in native Mandarin-speaking children. *J. Int. Phon. Assoc.* **2019**, *49*, 33–51. [[CrossRef](#)]
31. Bates, D.; Mächler, M.; Bolker, B.; Walker, S. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **2015**, *67*, 1–48. [[CrossRef](#)]
32. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2018. Available online: <https://www.r-project.org> (accessed on 1 December 2022).
33. RStudio Team. *RStudio: Integrated Development for R*; RStudio, Inc.: Boston, MA, USA, 2016. Available online: <http://www.rstudio.com/> (accessed on 1 December 2022).
34. Wright, R.; Souza, P. Comparing identification of standardized and regionally valid vowels. *J. Speech Lang. Hear. Res.* **2012**, *55*, 182–193. [[CrossRef](#)]
35. Newport, E.L. Maturational Constraints on Language Learning. *Cogn. Sci.* **1990**, *14*, 11–28. [[CrossRef](#)]
36. Flege, J.E.; Yeni-Komshian, G.H.; Liu, S. Age Constraints on Second-Language Acquisition. *J. Mem. Lang.* **1999**, *41*, 78–104. [[CrossRef](#)]

37. Saito, K.; Trofimovich, P.; Isaacs, T. Second language speech production: Investigating linguistic correlates of comprehensibility and accentedness for learners at different ability levels. *Appl. Psycholinguist.* **2016**, *37*, 217–240. [[CrossRef](#)]
38. Munro, M.J. Variability in L2 Vowel Production: Different Elicitation Methods Affect Individual Speakers Differently. *Front. Psychol.* **2022**, *13*, 916736. [[CrossRef](#)] [[PubMed](#)]
39. Ulbrich, C. Phonetic Accommodation on the Segmental and the Suprasegmental Level of Speech in Native–Non-Native Collaborative Tasks. *Lang. Speech* **2024**, *67*, 346–372. [[CrossRef](#)] [[PubMed](#)]
40. Oh, S. Phonetic and phonological vowel reduction in Brazilian Portuguese. *Phonetica* **2021**, *78*, 435–465. [[CrossRef](#)] [[PubMed](#)]
41. Sabev, M. Unstressed vowel reduction and contrast neutralisation in western and eastern Bulgarian: A current appraisal. *J. Phon.* **2023**, *99*, 101242. [[CrossRef](#)]
42. Georgiou, G.P.; Giannakou, A. Acoustic Characteristics of Greek Vowels Produced by Adult Heritage Speakers of Albanian. *Acoustics* **2024**, *6*, 257–271. [[CrossRef](#)]
43. Yang, Y.; Chen, S. Does prosody influence segments differently in Cantonese and Mandarin? A case study of the open vowel /a/. In Proceedings of the Speech Prosody 2022, Lisbon, Portugal, 23–26 May 2022; pp. 674–678. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.