INDONESIAN JOURNAL OF APPLIED LINGUISTICS
Vol. 16, No. 1, May 2026, pp. 143-158

Available online at:
https://ijal.upi.edu/index.php/ijal/article/view/393
https://doi.org/10.17509/hhwnng14

Acoustic analysis of English vowel contrast by age and gender among Mainland Chinese speakers

Huihui Wang[1], Stefanie Pillai[2*] and Yunisrina Qismullah Yusuf[3]
[1, 2]Faculty of Languages & Linguistics, Universiti Malaya, Kuala Lumpur, Malaysia
[1]International College of Engineering, Changsha University of Science and Technology, Changsha, P.R. China
[3]Department of English Education, Faculty of Teacher Training and Education, Universitas Syiah Kuala, Banda Aceh, Indonesia

ABSTRACT

The rapid growth of English language learners in Mainland China and of Chinese students in English-medium institutions of higher education has sparked scholarly interest in Chinese English (CE). Yet, from a phonetic perspective, few studies have examined how age and gender affect the production of CE vowels, despite research showing that these factors are important for understanding phonetic variation. This includes differences in English vowel quality and length. To address this gap, this quantitative study acoustically examined vowel contrast in English monophthongs among 23 male and female participants from two age groups comprising English majors and teachers from a public university in Mainland China. Participants read two passages from which 3236 monophthongs were extracted. The first (F1) and second (F2) formants and durations from the recordings were acoustically measured using Praat and statistically analysed using R. The findings revealed a consistent lack of vowel quality and length contrast across age and gender groups for typical vowel pairs. These findings echo previous studies suggesting that this pattern is becoming a common feature of CE. In terms of the vowel space, the vowels produced by older and female speakers were more spread out in the vowel space compared to those of younger speakers. This could indicate more carefully produced pronunciation among these groups of speakers. The findings of this study enhance our understanding of the evolutionary process of CE, an emerging variety of English, and can inform English pronunciation teaching and learning in Mainland China.

Keywords: Age; Chinese English; gender; monophthongs; vowel contrast

Received:	**Revised:**
4 December 2025	14 March 2026
Accepted:	**Published:**
4 May 2026	25 May 2026

How to cite (in APA style):
Wang, H., Pillai, S., & Yusuf, Y. Q. (2026). Acoustic analysis of English vowel contrast by age and gender among Mainland Chinese speakers. *Indonesian Journal of Applied Linguistics*, *16*(1), 143-158. https://doi.org/10.17509/hhwnng14

INTRODUCTION

There has been a dramatic growth in the number of English language learners in Mainland China (Schneider, 2014; Tao et al., 2025; Zhang & Lütge, 2024) as well as in the number of Chinese students studying abroad in English-medium institutions (Education Malaysia Global Services, 2025; Market Report China, 2026; Miao & Wang, 2024; Williams, 2025). Previous studies on CE suggest that it is an emergent variety of English (Deterding, 2006a; Wang, 2023; Wang & Pillai, 2025; Zeng et al., 2025) within the Kachruvian Expanding Circle of World Englishes (Kachru, 1990). Despite being classified as an Expanding Circle variety, it is posited that the variety of English produced by Mainland Chinese speakers will continue to develop into a variety with its own identity over the years (Deterding, 2006a).

Thus, it is no surprise that scholarly interest has increased in how Mainland Chinese speakers use English. This includes the features of their English pronunciation, which is assumed to be influenced by Mandarin (Albrecht et al., 2026; Wang & Pillai, 2025; Zeng et al., 2025) and a result of the limited use of English in everyday contexts in China (He, 2020; Wang & Pillai, 2025). The salient

*Corresponding author
Email: stefanie@um.edu.my

features of CE pronunciation reported in previous studies include a smaller vowel space, the absence of contrast between typical vowel pairs (Wang, 2023), and the addition or reduction of vowels (Albrecht et al., 2026; Deterding, 2006a; Li & Sewell, 2012; Wang & Pillai, 2025; Zeng et al., 2025). In addition, Deterding (2006a), Wang (2023) and Zeng et al. (2025) found that Mainland Chinese speakers lack durational contrast between vowel pairs.

Recent studies have begun examining age- and gender-related differences in vowel contrast among CE speakers with different language backgrounds. Zeng et al. (2025), for example, found that although both male and female Yi speakers tended to conflate /ɪ/-/iː/, male Yi speakers distinguished between the /ʌ/-/ɑː/ and /ɒ/-/ɔː/ pairs. However, Wang and Pillai (2025) found that age, rather than gender, had an effect on vowel conflation in CE. Similarly, Albrecht et al. (2026) reported that younger Chinese speakers conflated the pairs /ɪ/–/iː/ and /ʊ/–/uː/. The lack of vowel contrast has also been reported in Outer Circle varieties of English. Among these is Malaysian English, where there is conflation of traditional vowel pairs, /ɪ/ – / iː/, /e/ – /æ/, and /ʌ/ -/ɑː/, with some attempt to maintain durational contrast by more fluent speakers of English (Pillai et al., 2010; Tan & Low, 2010). Similar features have also been documented in Singapore English (Tan & Low, 2010). In contrast, in Hong Kong English, no durational contrast was found between long-short vowels in the KIT-FLEECE (/ɪ/ – / iː/) pair by Chan and Fuchs (2023).

While existing literature has revealed the conflation of vowel pairs in CE, there are gaps that require attention. First, the use of a wordlist to elicit data in phonetic studies, including for CE vowels (e.g., Feng & Wang, 2024; Gong et al., 2020; Wang & Pillai, 2025; Zeng et al., 2025). It is likely that hypercorrection will occur most often in wordlist and minimal-pair tasks. Therefore, the inconsistent findings regarding vowel conflation in CE may have arisen from the elicitation methods used.

Second, there is still a need for more empirical evidence on gender differences among Mainland Chinese CE speakers, given that work on language change has shown gender differences in speech patterns among females and males. For example, females are more likely to pioneer linguistic change compared to males (Maclagan, 2000). Similarly, features of CE pronunciation across age groups have not received much attention despite shifting foreign policies in China over the years (Wang & Pillai, 2025).

Third, in the context of World Englishes, the emergence of CE as a distinct variety calls for a shift in perspective, i.e., away from the traditional deficit-oriented approach that seeks to discover and understand patterns of use. In the case of pronunciation, this shift questions the need for prescriptive norms, for example, of pronunciation models, typically based on Inner Circle norms (e.g., Received Pronunciation and General American English pronunciation). Vowel contrast in English is of particular interest, as the lack of contrast has been reported across many varieties, suggesting it is a shared feature (Mesthrie & Bhatt, 2008). What distinguishes the varieties is the particular pairs that are conflated and whether they are distinguishable in terms of vowel quality and/or length (Loakes et al., 2024; Mesthrie & Bhatt, 2008). Subsequently, the extent to which English vowels are contrasted (e.g., in terms of age and gender) and across different varieties of English, including CE, is likely to affect their vowel inventories in terms of the quality of the monophthongs produced and the number of monophthongs. Thus, findings on vowel contrast in CE can not only contribute to our understanding of how vowels are produced in CE but also allow for comparisons with other varieties of English. These contribute to establishing whether vowel contrast is indeed a key pronunciation feature of the English language (Chan, 2025). Such knowledge could inform pedagogy by identifying which areas of English pronunciation to focus on. To address the gaps noted, the present study aimed to investigate vowel contrast in the production of English monophthongs by Mainland Chinese English speakers, with respect to age and gender.

METHOD

Research Design

This study was quantitative in nature, aiming to acoustically determine vowel contrast in terms of vowel quality and duration for monophthongs produced by CE speakers across different age and gender groups. Since we were measuring the durations and the first (F1) and second (F2) formants of the vowels produced by the speakers, a quantitative approach allowed us to compare measurements within and across groups statistically and to present the results in relevant figures.

Participants

To address the research aim, 23 CE speakers (12 males and 11 females) aged 20-78 years (average age: 41 years) were recruited from a public university in Changsha (the capital city of Hunan Province) through convenience and snowball sampling. We initially aimed to recruit more speakers, but due to COVID-19 restrictions at the time of data collection, it was not possible to obtain a larger sample from this university or other institutions. We acknowledge this as a limitation of our study. All 23 participants who consented to participate in the study reported being free of hearing and speech disorders. These speakers were assumed to be relatively proficient in English in the Chinese context, as all were English majors,

professors, or associate professors in English-related fields.

These speakers were further divided into two age groups: those born before 1980 and those born in or after 1980: 11 speakers aged 44 years and above and 12 speakers aged 23 to 43 years at the time of data collection. This division reflects the differing educational policies and English learning circumstances to which the two groups were exposed. The older age group represented individuals who were born and grew up before the open-door policy. This policy refers to the economic policy implemented by Deng Xiaoping, which opened China to the world. The younger age group was those born and exposed to the fast-paced globalisation in China from the 1980s onward, which aligns with the changes in China. Before the 1980s, British English was adopted as the only pedagogic model in Mainland China. However, American English was introduced after China and the United States established their diplomatic relationship in the 1970s (He, 2020).

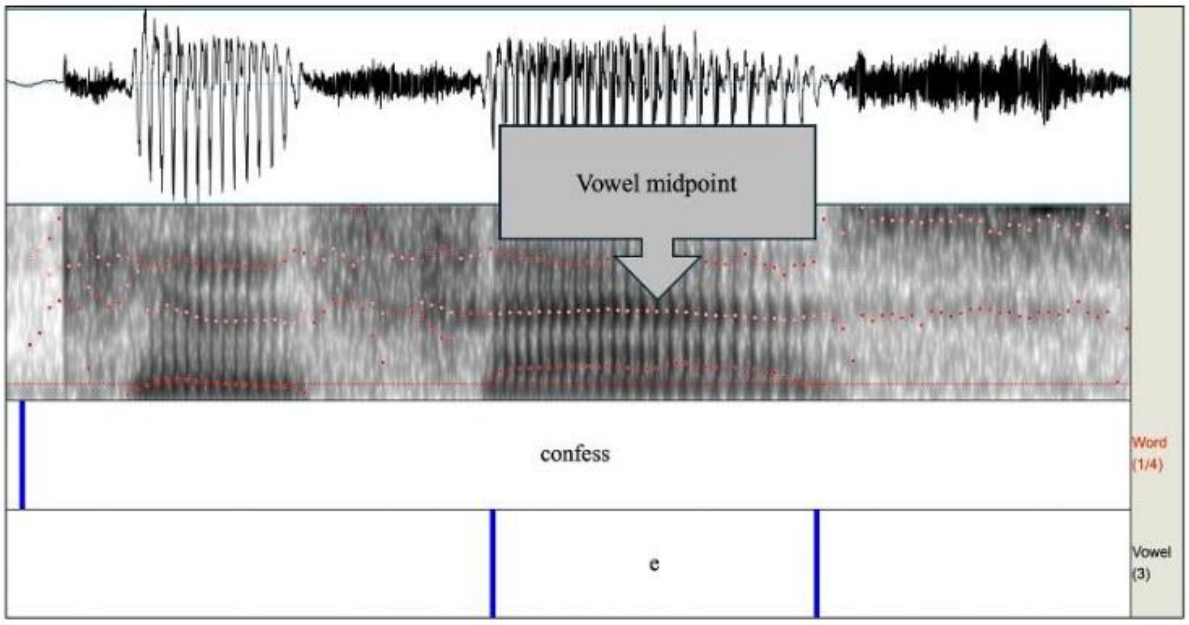
Data Collection and Analysis

This study is part of a larger study on CE vowels, which commenced upon obtaining permission from the university to carry out the study. As part of this study, a questionnaire was used to collect demographic data and participants' language backgrounds (see Appendix). Participants provided informed consent for the questionnaire and recordings. Participants were recorded reading two reading passages, *The North Wind and the Sun* (*NWS*) and *The Boy Who Cried Wolf* (*Wolf*). The *NWS* passage is commonly used in phonetic studies, especially for illustrating the International Phonetic Alphabet (e.g., Ido, 2025; Rao & Shaw, 2024), including for the study of English vowels (e.g., Sukaton et al., 2023; Yazawa, 2022; Yazawa et al., 2023). However, as noted by Deterding (2006b), this text has limitations for phonetic research; hence, he proposed the *Wolf* passage for analysing English sounds, which is why we used both passages.

The selection of the monophthongs from the two passages was based on typical vowel pairs that are contrasted in English: /ɪ/-/iː/, /e^{i}/-/æ/, /ʌ/-/ɑː/, /ʊ/-/uː/ and /ɒ/-/ɔː/. Each passage was read twice by the participants, and the speakers were recorded with a TASCAM DR-600 (16-bit, 44.1kHz) and RODE NTDG4 microphone. Praat Version 6.0 (Boersma & Weenink, 2023) was then used to annotate the sound files and measure the vowels. The target vowels were identified, segmented, and labelled based on auditory perception and inspection of their spectrograms and waveforms, as shown in Figure 1.

Figure 1

Segmentation of the Vowel [e] from the Word 'confess' in Praat



Once the vowel was segmented, the duration of the vowel from onset to offset was measured in milliseconds using a Praat Script[ii], while the F1 and F2 of the vowels were measured in Hertz (Hz) using another Praat Script[iii]. The latter calculated the formant values at the midpoint of the segmented and labelled vowel (see Figure 1). The midpoint was used to minimise coarticulatory effects (Whalen et al., 2018; Yusuf et al., 2021; Zeng et al., 2025). The F1 and F2 values in Hz were then converted to a Bark scale (Zwicker & Terhardt, 1980) to visually represent the vowels in a perceptual space (Albrecht et al., 2026) using R (RStudio Team, 2022). The F1 and F2 of vowels are generally used to classify vowel quality, as F1 inversely reflects the articulatory and perceptual dimension of vowel

height, while F2 directly corresponds to the advancement and retraction of the vowel (Georgiou, 2022).

To test the vowel contrast between the F1 and F2 of each pair, the vowel category was included as a fixed variable in each model. Accordingly, the F1 and F2 frequencies and vowel duration were adopted as the dependent variables, while age, gender, vowel, and their three-way interactions were included as fixed effects in the models. The goodness-of-fit of the models (Lüdecke et al., 2021) was assessed using R's performance package to detect collinearity, singularity, and heteroscedasticity. Assumptions were verified by visual inspection by plotting residual plots, histograms, and Q-Q plots to check the distribution of the data. Additionally, following Barr et al. (2013), by-participant and by-token random slopes and their intercepts for all the relevant fixed effects were included. Outliers exceeding 2.5 standard deviations were addressed (Wu et al., 2023). Significance for main effects and interactions was determined via likelihood-ratio tests using the mixed () function (Winter, 2019). Degrees of freedom were approximated using the Kenward-Roger function in the *lmerTest* package (Kuznetsova et al., 2017), and *p*-values were adjusted via Tukey's method. An alpha level of 0.05 was set for all tests.

FINDINGS AND DISCUSSION

Frequency of Monophthongs

Table 1 displays the frequency counts of the monophthongs produced by the speakers for the target words that were extracted from the recordings. The numbers indicate the total number of vowels for each age and gender group.

Table 1
Frequency of monophthongs from the recordings

Target words	Vowels	Old	Young	Female	Male
considered, did, fist, this, chicken, convinced	[ɪ]	165	129	139	155
succeeded, immediately, even, sheep, feast	[iː]	166	111	132	145
then, attempt, confess, shepherd, get, pleasure, successful	[e]	102	121	85	138
traveller, wrapped, plan, exactly, actually, began	[æ]	228	176	194	210
first, heard, concern, third	[ɜː]	98	75	82	91
Sun, other, up, company, fun, much, duck, come	[ʌ]	255	201	211	245
could, foot, good, looking, took	[ʊ]	107	69	72	104
	[ʊˈ]*	22	23	32	13
blew, two, afternoon, soon	[uː]	109	74	73	110
	[uːˈ]*	42	34	50	26
shone, off, flocks, hot, not, bother	[ɒ]	122	79	63	138
	[ɑːˈ]**	44	24	49	19
hard, last, dark, after, afternoon	[ɑː]	100	80	74	106
	[æˈ]**	26	17	31	12
North, more, thought, course, before	[ɔː]	241	196	211	226
	Total:	1827	1409	1498	1738

Note:
[uːˈ] and [ʊˈ] refer to the fronted variants of [uː]/[ʊ].
[aːˈ] refers to the General American English realisation [æ] in the words after, afternoon, and last.
[ɒˈ] refers to the General American English realisation [ɑː] in the words bother, flocks, hot, and not.

A total of 3236 tokens were analysed for this study. Among all the target words used in this study, the words, *after*, *afternoon*, *last* and *bother, flocks, hot, not* differ in pronunciation, but were closer to Standard Spoken British English, i.e., approximating Received Pronunciation (RP) e.g., [ˈɑftə], [ˈɑːftəˈnuːn], [lɑːst], [ˈbɒðə], [flɒk], [hɒt] and [nɒt] compared to General American English Pronunciation (GA), [ˈæftɚ], [æftɚˈnuːn], [læst] [ˈbɑːðɚ], [lɑːk], [hɑːt] and [nɑːt]. The difference was ascertained via auditory examination of words and with reference to their related spectrograms. As an example, Figure 2 illustrates the difference in the initial vowel in the word after, where the higher F2 for the more American English realisation of the word suggests a more fronted vowel compared to the initial vowel in British English. As shown in Table 1, most speakers realised the first vowel in each of these words closer to the British English pronunciation than to the American English pronunciation.

Within-Group Comparisons

Older Age Group

Figure 3 displays ellipses that represent the dispersion of different vowel categories, alongside the vowel variants across the F1–F2 space, with each 95% ellipse corresponding to one vowel or one vowel variant in the older age group. On the whole, there is a lack of categorical separation between the vowel pairs [ɪ] – [iː], [ʌ] – [ɑː], and [ʊ] – [uː] in the older age group.

Figure 2
Comparison of British English and American English Realisation for the Word 'after'

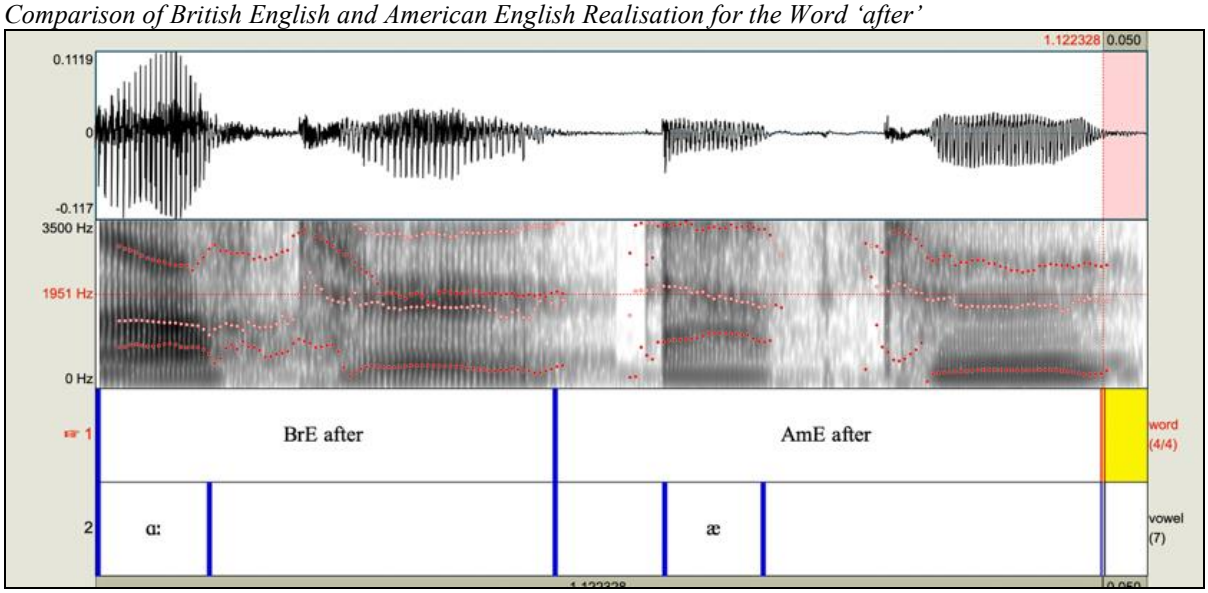
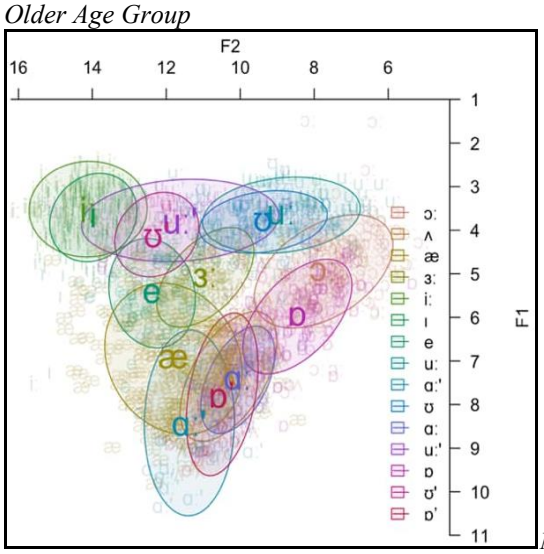


Figure 3
Vowel Distribution of Monophthong Pairs in the Older Age Group



Note:
[ɑ:'] refers to the General American English realisation [æ] in the words after, afternoon, and last.
[ɒ'] refers to the General American English realisation [ɑ:] in the words bother, flocks, hot, and not.

Despite the overlap between the vowel pairs [ɒ] – [ɔ:] and [e] – [æ], differences in vowel height can be observed in the older age groups. In addition, [ʊ'] and [u:'] fronting is visible compared to [ʊ] and [u:], and this is consistent with the significant differences in the F2 of [ʊ] and [ʊ'] as well as [u:] and [u:'] as shown in Table 2. A main effect of vowel was found for both F1 and F2 in the older age group (see Table 2).

Based on Figure 3, [ɑ:'] had a similar vowel placement as [æ], but [ɑ:'] was clearly lower than [æ], and hence there was a statistical significance for F1 between [ɑ:'] and [æ] (see Table 2).

Table 2
Pairwise Comparison of F1 and F2 in the Older Age Group

Vowel Pair	β	SE	*t*	*p*-value	Sig.
F1					
[i:] – [ɪ]	-0.24	0.21	-1.13	1.00	No
[æ] – [e]	1.22	0.20	6.11	< 0.0001	Yes
[æ] – [ɑ:']	-1.59	0.23	-6.84	< 0.0001	Yes
[ɑ:] – [ʌ]	0.08	0.19	0.41	1.00	No
[ɒ'] – [ʌ]	0.20	0.21	0.98	1.00	No
[u:] – [ʊ]	-0.13	0.22	-0.60	1.00	No
[u:] – [u:']	-0.12	0.14	0.84	1.00	No
[ʊ] – [ʊ']	-0.07	0.16	-0.44	1.00	No
[ɒ] – [ɔ:]	1.04	0.21	4.91	0.0005	Yes
Overall effect of vowel (F1): $\chi^2(14) = 796.78, p < .001$					
F2					
[i:] – [ɪ]	-0.23	0.26	0.88	1.00	No
[æ] – [e]	-0.21	0.25	-0.86	1.00	No
[æ] – [ɑ:']	-1.29	0.28	4.66	0.0007	Yes
[ɑ:] – [ʌ]	-0.14	0.23	-0.62	1.00	No
[ɒ'] – [ʌ]	-0.14	0.25	-0.54	1.00	No
[u:] – [ʊ]	-0.32	0.27	-1.20	1.00	No
[u:] – [u:']	-2.21	0.16	-13.74	< 0.0001	Yes
[ʊ] – [ʊ']	-2.20	0.18	-12.58	< 0.0001	Yes
[ɒ] – [ɔ:]	0.91	0.26	3.45	0.06	No
Overall effect of vowel (F2): $\chi^2(14) = 624.88, p < .001$					

In addition, [ɒ'] in *bother* merged with the [ʌ] – [ɑ:] with no statistical significance for F1 between [ɒ'] and [ʌ] and between [ɑ:] – [ʌ]. Similarly, there was no statistical significance for the back vowels [ʊ] and [ʊ'] and [u:] and [u:'] as well as [u:] – [ʊ]. As can be seen in Table 2, the same findings applied to the F1 of [i:] – [ɪ], indicating that all these pairs were produced similarly in terms of vowel height.

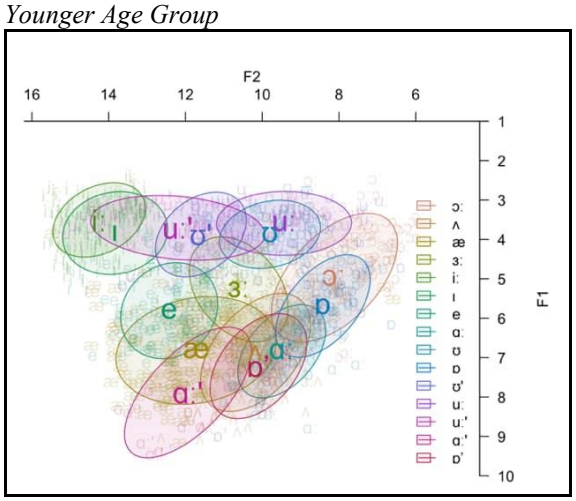
There was an evident contrast in terms of vowel fronting with statistical significance for F2 between the vowel pairs [ʊ] and [ʊ'] and [uː] and [uː']. Thus, it can be surmised that there was a tendency for older speakers to display vowel fronting for the vowels [ʊ] and [uː], as well as maintain observable and statistical height difference only for [e] – [æ] and [ɑː′] – [æ], while conflating the other pairs.

Younger Age Group

Figure 4 shows ellipses that symbolise the dispersion of different vowel categories, together with the vowel variants across the F1–F2 space, with each 95% ellipse equal to one vowel or one vowel variant in the younger age group. As displayed in Table 3, the younger speakers tended to merge vowel pairs, with no significant difference found for vowel height for the vowel pairs, with only the F1 for [e] – [æ] being significantly different. Similarly, as shown in Table 3, there were no statistical differences for F2 for [uː] – [ʊ], [iː] – [ɪ], [ɒ] – [ɔː] and [ɑː] – [ʌ] as well as for [e] – [æ] and [ʊ] – [ʊ'']. Additionally, fronting of [ʊ] and [uː] was observed in the younger age speakers, with significant differences in the F2 of [ʊ] and [ʊ'] and [uː] and [uː'] (see Table 3).

Figure 4

Vowel Distribution of Monophthong Pairs in the Younger Age Group



Female Group of Speakers

Figure 5 shows the dispersion of different vowel categories, together with the vowel variants across the F1–F2 space, with each 95% ellipse equal to one vowel or one vowel variant in the female group of speakers. As shown in Figure 5, female speakers appeared to merge the [ɪ] – [iː], [ʌ] – [ɑː], and [ʊ] – [uː] vowel pairs. In contrast, although the vowel pairs including [ɒ] – [ɔː] and [e] – [æ] overlapped, the F1 differences can be observed in Figure 5. The statistical results suggested that the vowel had a main effect on F1 (see Table 4). The post-hoc analyses shown in Table 4 show that no statistical significance was observed in F1 between [iː] - [ɪ], [ɑː] – [ʌ] and [uː] - [ʊ]. However, statistical significance was found for vowel height between the pairs [e] – [æ], [ɑː′] – [æ] and [ɒ] – [ɔː].

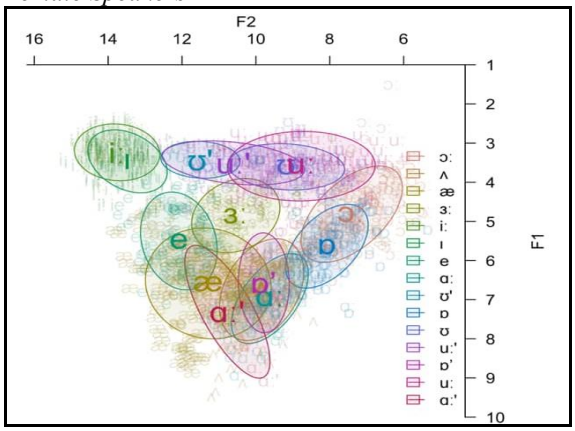
Table 3

Pairwise Comparison of F1 and F2 in the Younger Age Group

Vowel Pair	β	SE	*t*	*p*-value	Sig.
F1					
[iː] – [ɪ]	-0.32	0.22	-1.43	0.98	No
[æ] – [e]	0.84	0.21	4.03	0.01	**Yes**
[æ] – [ɑː′]	-0.86	0.27	-3.24	0.09	No
[ɑː] – [ʌ]	-0.01	0.21	0.03	1.00	No
[ɒ′] – [ʌ]	0.29	0.25	1.16	1.00	No
[uː] – [ʊ]	-0.20	0.24	-0.84	1.00	No
[uː] – [uː′]	-0.22	0.17	-1.25	1.00	No
[ʊ] – [ʊ′]	-0.08	0.19	-0.44	< 0.0001	**Yes**
[ɒ] – [ɔː]	0.78	0.23	3.45	0.07	No
Overall effect of vowel (F1): $\chi^2(14) = 506.78$, $p < .001$					
F2					
[iː] – [ɪ]	0.43	0.25	1.73	0.91	No
[æ] – [e]	-0.33	0.23	-1.45	0.97	No
[æ] – [ɑː′]	0.93	0.28	3.29	0.08	No
[ɑː] – [ʌ]	-0.59	0.22	-2.62	0.37	No
[ɒ′] – [ʌ]	0.01	0.27	0.06	1.00	No
[uː] – [ʊ]	-0.33	0.26	-1.27	0.99	No
[uː] – [uː′]	-2.37	0.19	-12.70	< 0.0001	**Yes**
[ʊ] – [ʊ′]	-1.37	0.20	-7.02	<0.0001	Yes
[ɒ] – [ɔː]	0.59	0.25	2.37	0.54	No
Overall effect of vowel (F2): $\chi^2(14) = 469.20$, $p < .001$					

Figure 5

Vowel Contrast of Monophthong Pairs among Female Speakers



In terms of vowel fronting/ retraction, a main effect of vowel was identified in F2. A post hoc analysis yielded no statistical difference in F2 between all the typical vowel pairs (see Table 4). In addition, [ɒ'] in bother also merged with the [ɑː] – [ʌ] vowel pair. However, no main effect was found

in F1 between [ɒ'] – [ʌ] (see Table 4). Additionally, [u:]-[ʊ] fronting was observed in the female group. Although a post-hoc analysis showed no main effect in F1 between [ʊ] – [ʊ'] and [u:] - [u:'], the F2 difference between the vowels in each of these pairs was statistically significant. This indicates that fronting was more salient than height for these pairs. Additionally, statistical differences in F2 were found for [æ] – [ɑ:'].

Table 4
Pairwise Comparison of F1 and F2 for Female Speakers

Vowel Pair	β	SE	*t*	*p*-value	Sig.
F1					
[i:] – [ɪ]	-0.31	0.23	-1.38	0.99	No
[æ] – [e]	1.19	0.22	5.52	<0.0001	Yes
[æ] – [ɑ:']	-1.57	0.25	-6.41	<0.0001	Yes
[ɑ:] – [ʌ]	-0.28	0.21	-1.30	0.99	No
[ɒ'] – [ʌ]	0.24	0.22	1.09	1.00	No
[u:] – [ʊ]	-0.46	0.25	-1.87	0.86	No
[u:] – [u:']	-0.46	0.17	-2.80	0.25	No
[ʊ] – [ʊ']	-0.04	0.17	-0.22	1.00	No
[ɒ] – [ɔ:]	1.12	0.23	4.80	0.0006	Yes
Overall effect of vowel (F1): $\chi^2(14) = 478.03$ $p < .001$					
F2					
[i:] – [ɪ]	0.25	0.27	0.95	1.00	No
[æ] – [e]	-0.13	0.26	-0.50	1.00	No
[æ] – [ɑ:']	1.18	0.28	4.20	0.004	Yes
[ɑ:] – [ʌ]	-0.64	0.25	-2.60	0.38	No
[ɒ'] – [ʌ]	-0.08	0.26	-0.31	1.00	No
[u:] – [ʊ]	-0.34	0.29	-1.17	1.00	No
[u:] – [u:']	-2.80	0.18	-15.39	<0.0001	Yes
[ʊ] – [ʊ']	-1.79	0.18	-9.80	<0.0001	Yes
[ɒ] – [ɔ:]	0.77	0.28	2.80	0.27	No
Overall effect of vowel (F2): $\chi^2(14) = 578.43$, $p < .001$					

Male Speakers

Figure 6 illustrates the vowel contrast in the male group of speakers. As shown in Table 5, the results suggest the vowel effect in both F1. A convergence was observed between [ɪ] – [i:], [ʌ] – [ɑ:], and [ʊ] – [u:], with post hoc analyses suggesting that no statistical differences for F1 were found in all the vowel pairs except for [e] – [æ] and [ɒ] – [ɔ:]. However, although [e] – [æ] and [ɒ] – [ɔ:] differed statistically in terms of F1, the vowels in each pair did not differ significantly for F2 (see Table 5). Additionally, for F2, no statistical significance emerged between all but three of the vowel pairs, supporting their convergence. Similar to the two age groups and the female speakers, [ʊ] and [u:] fronting was found among the male speakers, with significant difference for F2 between[æ] – [ɑ:'], [u:] – [u:'] and [ʊ] – [ʊ'] (see Table 5).

Figure 6
Vowel Contrast of Monophthong Pairs among Males

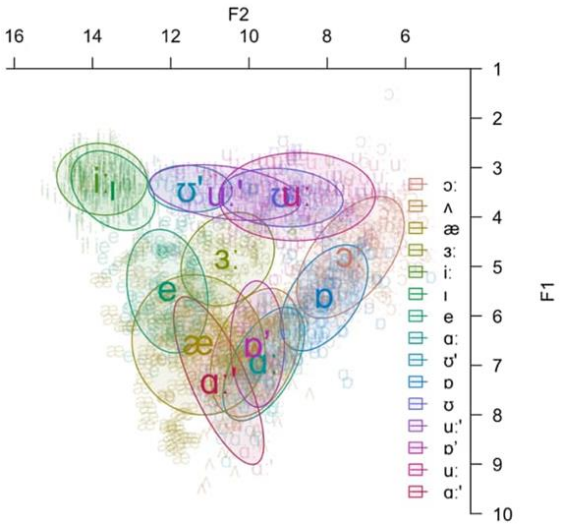


Table 5
Pairwise Comparison of F1 and F2 for Male Speakers

Vowel Pair	β	SE	*t*	*p*-value	Sig.
F1					
[i:] – [ɪ]	-0.25	0.21	-1.21	1.00	No
[æ] – [e]	0.85	0.20	4.37	0.004	Yes
[æ] – [ɑ:']	-0.58	0.26	-2.02	0.79	No
[ɑ:] – [ʌ]	0.32	0.19	1.71	0.93	No
[ɒ'] – [ʌ]	-0.11	0.23	-0.49	1.00	No
[u:] – [ʊ]	-0.06	0.22	-0.29	1.00	No
[u:] – [u:']	-0.12	0.15	-0.82	1.00	No
[ʊ] – [ʊ']	-0.03	0.18	-0.19	1.00	No
[ɒ] – [ɔ:]	0.79	0.21	3.79	0.002	Yes
Overall effect of vowel (F1): $\chi^2(14) = 518.67$, $p < .001$					
F2					
[i:] – [ɪ]	0.37	0.26	1.45	0.98	No
[æ] – [e]	-0.35	0.24	-1.46	0.98	No
[æ] – [ɑ:']	1.14	0.30	3.77	0.02	Yes
[ɑ:] – [ʌ]	-0.10	0.23	-0.46	1.00	No
[ɒ'] – [ʌ]	-0.17	0.27	-0.62	1.00	No
[u:] – [ʊ]	-0.43	0.26	-1.61	0.95	No
[u:] – [u:']	-1.59	0.18	-8.67	<0.0001	Yes
[ʊ] – [ʊ']	-1.91	0.21	-9.20	<0.0001	Yes
[ɒ] – [ɔ:]	0.77	0.26	2.98	0.19	No
Overall effect of vowel (F2): $\chi^2(14) = 430.44$, $p < .001$					

Within and Between-Group Comparisons

Figures 7 and 8 present the vowel charts for each group based on age and gender across the F1–F2 space. To display the vowel space of the speakers more explicitly, the vowel variants including [ʊ'], [u:'], [ɑ:'], and [ɒ'] were excluded from the vowel chart. Figure 7 shows substantial overlap between the two age groups. However, the vowels produced

by older speakers were spread more widely across the vowel space. In contrast, the production of the younger speakers tended to be more clustered, indicating a possible reduction in vowel dispersion in the younger speakers. This corresponds to the findings that younger speakers tended to merge more vowel pairs, which may result in a more clustered vowel space.

Figure 7
Vowel Distribution by Age Groups

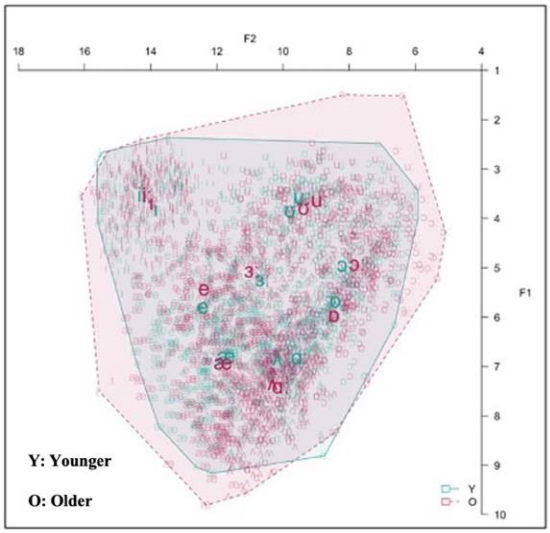


Figure 8
Vowel Distribution by Gender

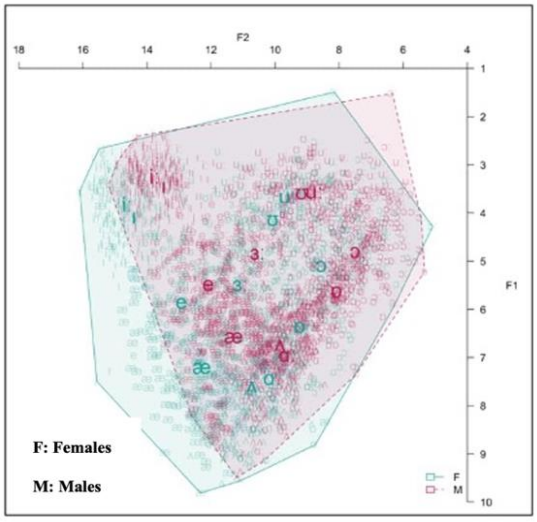


Figure 8 shows the differences in vowel distribution space between males and females. The production of female speakers generally occupied a larger vowel space and exhibited a more peripheral vowel placement. This suggests that the vowel space of the female speakers was more expanded while the vowel space of the male speakers was more compressed, particularly in the mid and low vowel regions. In addition, as illustrated in Figure 8, the vowels produced by the female speakers were evidently more fronted than those produced by the male speakers.

In terms of F1, a significant main effect was observed in vowel and gender, with a significant interaction between vowel and gender as well as between vowel and age (see Table 6). In addition, a three-way interaction between vowel versus age versus gender was also observed in F1 (see Table 6). These findings suggest that the effects of age, gender, and their interaction were vowel-based. Similarly, the statistical results for F2 also support the age-based effect on vowels, the gender effect, and their interactions with the front/back dimension. As presented in Table 6, a significant main effect was observed for vowel and gender, with significant interactions between vowel and gender, and between vowel and age. In addition, a three-way interaction between vowel, age, and gender was also observed for F2.

The post-hoc analyses presented in Table 6 show that no significant differences were observed for the F1s of any of the vowels between older and younger female speakers. However, for the F1s between older and younger male speakers, except for [ʌ] and [ɑː] (see Table 6). Nevertheless, in general, it can be surmised that the age effect in vowel height remained subtle across the gender groups. For F2, as can be seen in Table 6, no significant difference emerged between younger and older females for all vowels except for [ɜː] and [uː]. Likewise, older males had lower F2 than younger males for [uː] and [ʊ]. All the other vowels showed no significant age differences among male speakers (see Table 6). These results indicate that age-related F2 changes are limited across the gender groups, except for the back vowels [ʊ] and [uː].

There was, however, a significant interaction between gender and age for most vowels, as shown in Table 7. Older females exhibited significantly higher F1 values than older males for most vowels and vowel variants except for [uː], [ɑː] and [ɔː]. In the younger age group, younger females had higher F1 than younger males for almost all vowels, and the difference between the F1s of younger females and males was significant, except for [e] and [uː] (see Table 7). These findings, therefore, support the gender effect in vowel height in both age groups for most vowels.

In relation to F2, older females generally exhibited higher F2 values than older males for most vowels except for [e], where no significant difference was observed between the older male and female speakers. Younger females had higher F2 than younger males except for [uː] and [ʊ], where the difference between younger male and female speakers was not significant (see Table 7), which, on the whole, supports the effect of gender in the front and back dimensions.

Table 6
Post-Hoc Analysis of the Interaction between Age and within each Gender Group for F1 and F2

Vowel	Older Females vs Younger Females β	SE	*t*	*p*-value	Sig.*	Older Males vs Younger Males β	SE	*t*	*p*-value	Sig.*
F1										
[iː]	-0.04	0.18	-0.22	1.00	No	-0.03	0.18	-0.17	1.00	No
[ɪ]	-1.09	0.18	-0.61	0.93	No	-0.07	0.17	-0.38	0.98	No
[e]	-0.12	0.20	-0.61	0.93	No	-0.43	0.18	-2.42	0.08	No
[æ]	-0.09	0.17	-0.54	0.95	No	0.40	0.17	2.43	0.08	No
[ɜː]	-0.12	0.20	-0.58	0.94	No	-0.24	0.19	-1.22	0.62	No
[ʌ]	0.29	0.17	1.71	0.33	No	0.71	0.16	4.38	0.0004	Yes
[uː]	0.29	0.21	1.42	0.49	No	-0.07	0.19	-0.37	0.98	No
[ʊ]	-0.07	0.21	-0.31	0.99	No	0.08	0.19	0.41	0.98	No
[ɒ]	0.22	0.22	1.00	0.75	No	0.31	0.18	1.74	0.31	No
[ɔː]	-0.22	0.17	-1.28	0.58	No	0.21	0.16	1.27	0.59	No
[ɑː]	0.39	0.21	1.90	0.23	No	0.77	0.19	4.15	0.0004	Yes
F2										
[iː]	0.24	0.19	-1.25	0.59	No	0.20	0.19	1.08	0.70	No
[ɪ]	-0.04	0.19	-0.23	1.00	No	-0.40	0.17	-2.34	0.10	No
[e]	-0.36	0.21	-1.72	0.32	No	0.26	0.19	1.43	0.49	No
[æ]	-0.14	0.18	-0.78	0.87	No	0.41	0.17	2.36	0.10	No
[ɜː]	0.79	0.21	3.75	0.002	Yes	-0.16	0.21	-0.77	0.87	No
[ʌ]	0.23	0.18	-1.30	0.57	No	0.15	0.17	0.88	0.81	No
[uː]	0.78	0.22	3.52	0.003	Yes	-1.31	0.20	-6.57	<0.0001	Yes
[ʊ]	0.44	0.22	1.98	0.20	No	-1.03	0.20	-5.16	<0.0001	Yes
[ɒ]	0.20	0.24	0.84	0.84	No	-0.06	0.19	-0.31	0.99	No
[ɔː]	-0.24	0.18	-1.39	0.51	No	-0.40	0.18	-1.39	0.51	No
[ɑː]	0.52	0.22	2.37	0.87	No	0.41	0.17	2.36	0.10	No

Main effect and interactions:

F1
Vowel: $\chi 2\ (10) = 642.31$, $p < 0.001$
Gender: $\chi 2\ (1) = 21.39$, $p < 0.001$
Vowel x Gender: $\chi 2\ (10) = 108.72$, $p < 0.001$
Vowel x Age: $\chi 2\ (10) = 104.62$, $p < 0.001$
Vowel x Age x Gender: $\chi 2\ (10) = 34.57$, $p < 0.001$

F2
Vowel: $\chi^2\ (10) = 211.59$, $p < 0$
Gender: $\chi 2\ (1) = 33.85$, $p < 0.001$
Vowel x Gender: $\chi 2\ (10) = 44.45$, $p < 0.001$
Vowel x Age: $\chi 2\ (10) = 90.53$, $p < 0.001$
Vowel x Age x Gender: $\chi 2\ (10) = 195.13$, $p < 0.001$

**Sig. refers to Significant*

Duration

Figure 9 presents the durational distribution of monophthongs across age groups using violin plots, with the X-axis indicating the vowels and the Y-axis showing the vowel duration. Both the older and younger speakers exhibited slightly longer duration in the long vowels and the vowel variants such as [iː], [uː], [ɔː], [ɑː], and [uː'] compared to their short counterparts. Further, the durations of the vowel variants [ɑː'], [ɒ'], [ʊ'] and [uː'] are similar to those of their shorter counterparts [æ], [ʌ], [uː] and [ʊ], respectively, in both age groups. However, these differences were not significant across age groups (older age group: $\chi^2\ (14) = 28.12$, $p = 0.014$; younger age group: $\chi^2\ (14) = 34.67$, $p = 0.002$). Post hoc analyses found no significant differences between vowel pairs [ɪ] – [iː], [ʌ] – [ɑː], [ʊ] – [uː], [ɒ] – [ɔː], and [ʊ'] – [uː'] across age groups ($p > 0.05$). Similar findings were obtained for [ʌ] – [ɒ'], [uː'] – [uː], [ʊ] –[ʊ'], and [æ] – [ɑː'] ($p > 0.05$). Overall, the older speakers exhibited a similar durational pattern to the younger speakers.

Figure 10 shows the durational distribution of monophthongs by gender, with the X-axis indicating the vowels and the Y-axis showing the vowel duration. The female speakers exhibited slightly longer durations in the long vowels, [iː], [uː], [ɔː], [ɑː], and [uː'] compared to their shorter counterparts. In addition, the durations of the vowel variants [ɑː'] and [ɒ'] were similar to those of their vowel counterparts [æ] and [ʌ], respectively. A main effect of vowel duration in the female group was found ($\chi^2\ (14) = 25.99$, $p = 0.026$).

Table 7
Post-Hoc Analysis of the Interaction across Age and Gender

Vowel	Older Females vs Older Males β	SE	*t*	*p*-value	Sig.*	Younger Females vs Older Males β	SE	*t*	*p*-value	Sig.*
F1										
[iː]	0.52	0.17	3.16	0.01	Yes	0.53	0.19	2.78	0.03	Yes
[ɪ]	0.59	0.17	3.6	0.003	Yes	0.64	0.19	3.41	0.006	Yes
[e]	0.52	0.19	2.77	0.03	Yes	0.21	0.19	1.12	0.68	No
[æ]	0.45	0.16	2.86	0.03	Yes	0.94	0.18	5.29	<0.0001	Yes
[ɜː]	0.69	0.19	3.74	0.002	Yes	0.57	0.21	2.73	0.04	Yes
[ʌ]	0.81	0.16	5.23	<0.0001	Yes	1.23	0.18	7.02	<0.0001	Yes
[uː]	0.27	0.18	1.45	0.47	No	-0.1	0.21	-0.47	0.97	No
[ʊ]	0.49	0.18	2.7	0.04	Yes	0.64	0.22	2.95	0.02	Yes
[ɒ]	0.71	0.18	3.87	0.001	Yes	0.8	0.22	3.68	0.002	Yes
[ɔː]	0.13	0.16	0.84	0.84	No	0.55	0.18	3.15	0.01	Yes
[ɑː]	0.37	0.19	2	0.19	No	0.75	0.21	3.62	0.003	Yes
F2										
[iː]	0.71	0.17	4.11	0.0006	Yes	0.91	0.19	4.73	0.0001	Yes
[ɪ]	0.74	0.17	4.3	0.0003	Yes	1.14	0.19	6.07	<0.0001	Yes
[e]	0.47	0.2	2.39	0.08	No	0.73	0.2	3.62	0.003	Yes
[æ]	0.86	0.16	5.26	<0.0001	Yes	1.27	0.18	7.09	<0.0001	Yes
[ɜː]	0.98	0.19	5.07	<0.0001	Yes	0.83	0.22	3.85	0.001	Yes
[ʌ]	0.99	0.16	6.16	<0.0001	Yes	1.14	0.17	6.52	<0.0001	Yes
[uː]	1.78	0.20	9.13	<0.0001	Yes	-0.31	0.23	-1.39	0.51	No
[ʊ]	1.51	0.19	7.82	<0.0001	Yes	-0.04	0.23	0.17	1	No
[ɒ]	1.19	0.19	6.2	<0.0001	Yes	1.13	0.21	5.45	<0.0001	Yes
[ɔː]	1.06	0.16	6.51	<0.0001	Yes	0.91	0.18	4.96	<0.0001	Yes
[ɑː]	0.53	0.2	2.69	0.04	Yes	1.16	0.21	5.51	<0.0001	Yes

**Sig. refers to Significant*

Figure 9
Duration of Monophthongs between Age Groups

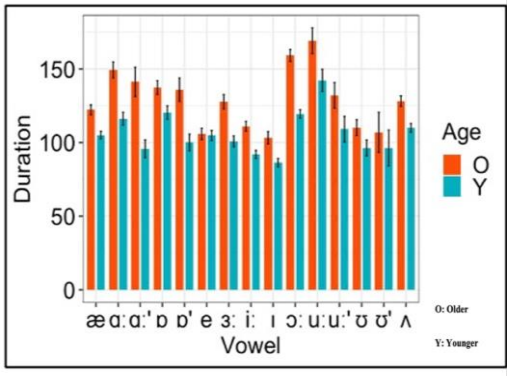
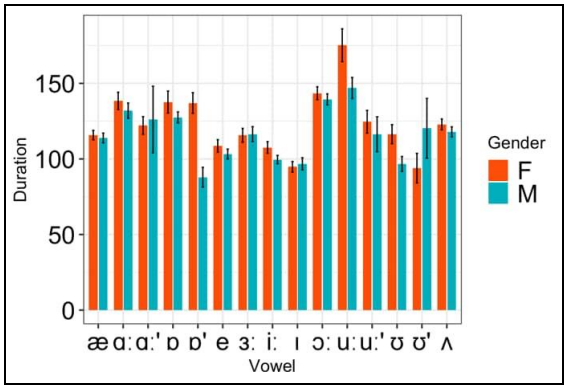


Figure 10
Duration of Monophthong and Monophthong Variants between Males and Females



However, post hoc analyses revealed no significant differences in duration for the vowel pairs in the female group. In addition, no significant differences emerged for the vowels and their variant counterparts ($p > 0.05$). Similarly, the statistical results yielded a main effect for vowels in the male group ($\chi 2$ (14) = 45.84, $p < 0.001$), but no significant differences were observed for the vowel pairs or for the vowels and their variant counterparts.

Linear mixed-effect models were used to examine the gender and age effects as well as their interactions. Significant main effects in duration were found in vowel (χ^2 (14) = 51.74, $p < 0.001$) and age (χ^2 (1) = 8.04, $p = 0.005$). Significant vowel × age (χ^2 (14) = 43.91, $p < 0.001$) and vowel × age × gender (χ^2 (14) = 25.80, $p = 0.027$) interactions were found. The post hoc analyses revealed that most contrasts were not statistically significant ($p > 0.05$), suggesting that durations were generally comparable across groups.

Discussion of Vowel Conflation in CE

The key finding related to vowel contrast is that some vowel pairs that are traditionally distinctive in Inner Circle varieties, such as [ɪ] – [iː], [ʊ] – [uː] and [ʌ] – [ɑː], were conflated across all CE speakers in this study. This echoes previous studies suggesting that vowel conflation is emerging as a distinctive feature of CE (Albrecht et al., 2026;

Deterding, 2006a; Wang & Pillai, 2025; Zeng et al., 2025). However, the current study differs from previous studies in the specific vowel pair that emerges. As previously mentioned, Wang (2023), for example, reported the merger of the vowel pairs [iː] – [ɪ], [uː] – [ʊ], and [æ] – [e]. On the contrary, the current findings suggested that [e] – [æ] were distinguished, which aligns with Wang and Pillai (2025) and Zeng et al. (2025). Wang (2023) suggested that the vowel merger between these vowels in CE can be attributed to their absence in Mandarin (Luo et al., 2020), and thus, speakers might assimilate both monophthongs to the Mandarin mid vowel /ə/, which is between [æ] and [e] in English. However, all speakers in the present study used Mandarin as their dominant language and still contrasted this vowel pair. The inconsistency in the results may be due to differences in the speakers recruited. In Wang's (2023) study, although the speakers received an overall IELTS score of 7, most were non-English majors, suggesting potentially less English exposure and use.

Another key finding of the current study is that older speakers tended to merge [iː] – [ɪ], [uː] – [ʊ], and [ʌ] – [ɑː], while the younger speakers appeared to merge all vowel pairs except for [æ] – [e]. This result aligns with Wang and Pillai (2025), who also found more evidence of vowel pair conflation in the younger generation. The generational difference in this study could be attributed to changes in diplomatic relations with the West, which affected pedagogical models in Mainland China (He, 2020). At the same time, this difference also suggests that an overall lack of vowel contrast between vowel pairs may be developing into a common pattern in CE. In CE, the more obvious vowel conflation in the younger group leads to a smaller vowel inventory and, therefore, a more concentrated vowel space. The findings for the younger speakers are consistent with previous studies reporting that CE has a smaller vowel space, presumably in contrast to a first-language (L1) variety of English (e.g., Zeng et al., 2025).

In terms of gender, the vowels produced by female speakers occupied a slightly larger space. We hesitate to speculate on the reason for this finding at this stage, but it may reflect more careful production by females reported in the literature (Maclagan, 2000). Interestingly, females tended to produce monophthongs more anteriorly than males in the current study, which aligned with the more forward monophthong articulation in Wang and Pillai (2025) and is consistent with Zeng et al. (2025), who reported that women and men differed in distinguishing vowel pairs, aligning with the cross-linguistic tendency for women to produce more peripheral vowels.

Vowel conflation in terms of vowel quality is further reinforced by the overall lack of durational contrast across all typical vowel pairs found in the present study. The loss of durational contrast is expected, as Mandarin lacks vowel-duration contrast (Luo et al., 2020). These findings, however, contradict previous studies that suggested that durational contrast was maintained in CE to varying degrees (e.g., Feng & Wang, 2024; Gong et al., 2020; Wang, 2023; Wang & Pillai, 2025; Zeng et al., 2025). These contradictory findings might be attributed to the contexts in which data were elicited across the studies. Previous studies tended to use word lists to elicit data, whereas the current study adopted the reading passage context to elicit the speech data. Coarticulation and the speech rate in the reading passages may also have contributed to varying realisations of vowel length contrast between typical vowel pairs.

On the whole, the conflation of vowel contrast appears to be developing into a feature of CE, similar to reports of this in other Asian varieties of English (e.g., Hong Kong, Malaysian and Singapore Englishes). Other findings from this study include a tendency to realise the first vowel in words like afternoon and last as [ɑː] in British English rather than [æ] in American English, although there were a few instances of the latter (Table 1). This could be attributed to the coexistence of two pedagogical models in English language education in Mainland China and to the norms favoured by teachers. Another finding was the fronting of the back vowels /uː/ and /ʊ/ in words like *two*, *good*, *soon,* and *took*. This feature has been widely documented in other varieties of English, indicating a diachronic process in progress; such fronting has been documented in a wide range of other varieties of English, including American English (Havenhill, 2024).

From a practical point of view, this study encourages a multilingual lens for investigating features of CE, not from a deficit (i.e., 'errors', 'wrong English' or 'bad English') perspective, but as developing patterns of use. Rather, in addition to first-language influence, these patterns can also be explained by the concept of "Transnational Attraction" (Schneider, 2014, p. 28), which refers to the use of English features for practical communication without strict adherence to native-speaker norms, typically based on standard varieties of British and American English (Wang & Pillai, 2025). From this view, reduced vowel distinctions and the loss of durational contrasts may reflect functional adaptations rather than phonological limitations. In multilingual contexts where intelligibility and efficiency are prioritised, maintaining vowel contrasts may have limited communicative benefit. Speakers may therefore adopt a simpler vowel system that supports ease of production and mutual understanding. The stability

of these features across generations suggests that they may be becoming established characteristics of CE, shaped not only by L1 influence but possibly by a more pragmatic motivation for communicative language use.

Thus, it is posited that the variety of English produced by Mainland Chinese speakers will continue to emerge and become an established variety with its own identity over the years and across generations (Deterding, 2006a).

CONCLUSION

In sum, this study reveals a more nuanced picture of vowel contrast in CE by considering age and gender differences among 23 speakers from a public university in Changsha. The findings on the CE vowels enhance our understanding of the patterns that persist in CE and highlight the possibility that the lack of quality and durational vowel contrast may be developing into a common feature of CE. Based on the acoustic and statistical findings, vowel conflation was common among age and gender groups, with younger speakers appearing to have a stronger tendency for vowel conflation. Although this study does not directly contribute to English teaching and learning in Mainland China, the current findings can raise awareness among English language educators, not just in China (as many Chinese students now study abroad), about how CE is produced and challenge the deficient view of their spoken English.

Despite these insights, we acknowledge that this study has limitations. First, the sample size was admittedly small due to data-collection constraints during the COVID-19 pandemic and restrictions on person-to-person contact. Second, this restriction also affected the sample, which could have been more balanced across age and gender groups. Therefore, the findings cannot be taken as representative of all Mainland Chinese English speakers. Future studies should include a broad cross-section of mainland Chinese English users further to examine the developing patterns of pronunciation in CE. Nevertheless, this study serves as a springboard for future studies to investigate social variables and the production of CE vowels and other aspects of CE pronunciation.

Beyond internal linguistic factors, the stability of these merged features, particularly among younger generations, must be considered within the broader sociolinguistic context of English use in Mainland China. While English education is widespread, the practical need for phonetic accuracy aligned with a specific native-speaker norm may be low in contexts where English is primarily used as a lingua franca among speakers of various L1s. This environment may inadvertently favour a simpler, more efficient phonological system, contributing to the establishment of CE features such as a reduced vowel inventory and the loss of durational contrast. This dynamic development is central to understanding the trajectory of CE as an emerging World English.

ACKNOWLEDGEMENTS

This study was supported by a Universiti Malaya Faculty of Languages & Linguistics Research Grant [UMG005N-2024]. We would like to thank Dr. Ren Xin for her assistance in providing the R scripts for statistical analyses in R programming. We are also grateful for the insightful comments and suggestions from the reviewers and editors, which helped us to enhance this paper.

REFERENCES

Albrecht, S., Beccard, M., & Schmied, J. (2026). The Chinese English vowel space across three provinces. *Asian Englishes*, *28*(1), 20–36. https://doi.org/10.1080/13488678.2025.2551884

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Boersma, P., & Weenink, D. (2023). *Praat*: *Doing phonetics by computer* [Computer program]. http://www.praat.org/

Chan, J. Y. H. (2025). Towards a feature-based L2 pronunciation teaching approach: variations in Hong Kong English and ELF intelligibility. *Journal of English as a Lingua Franc*a, *14*(1–2), 1–32. https://doi.org/10.1515/jelf-2024-0003

Chan, H. L., & Fuchs, R. (2023). Revisiting the vowels of Hong Kong English—The post-handover generation. *Proceedings of the 20th International Congress of Phonetic Sciences (ICPhS 2023). Prague, Czech Republic*. https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2023/full_papers/997.pdf

Deterding, D. (2006a). The pronunciation of English by speakers from China. *English World-Wide*, *27*(2), 175–198. https://doi.org/10.1075/eww.27.2.04det

Deterding, D. (2006b). The North Wind versus a Wolf: Short texts for the description and measurement of English pronunciation. *Journal of the International Phonetic Association*, *36*(2), 187–196. https://doi.org/10.1017/S0025100306002544

Education Malaysia Global Services (EMGS). (2025). *International student data*.

https://educationmalaysia.gov.my/more/student-data/international-student-data

Feng, H., & Wang, L. J. (2024). Acoustic analysis of English tense and lax vowels: Comparing the production between Mandarin Chinese learners and native English speakers. The *Journal of the Acoustical Society of America*, *155*(5), 3071–3089. https://doi.org/10.1121/10.0025931

Georgiou, G. P. (2022). Acoustic markers of vowels produced with different types of face masks. *Applied Acoustics*, *191*, 108691. https://doi.org/10.1016/j.apacoust.2022.108691

Gong, J., Xue, D., Bellamy, W., Wang, F., & Ji. X. L. (2020). Temporal and formant trajectory analysis of English tense-lax vowels produced by native Chinese speakers. *Proceedings of the Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), Auckland, New Zealand* (pp. 589–594). http://www.apsipa.org/proceedings/2020/pdfs/0000589.pdf

Havenhill, J. (2024). Articulatory and acoustic dynamics of fronted back vowels in American English. *The Journal of the Acoustical Society of America*, *155*(4), 2285–2301. https://doi.org/10.1121/10.0025461

He, D. Y. (2020). *China Englishes* in *World Englishes: Education and use in the professional world.* Springer Nature Singapore.

Ido, S. (2025). Uzbek. *Journal of the International Phonetic Association*, *55*(1–2), 152–68. https://doi.org/10.1017/S0025100324000148

Kachru, B. B. (1990), World Englishes and applied linguistics. *World Englishes*, *9*(1), 3–20. https://doi.org/10.1111/j.1467-971X.1990.tb00683.x

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Li, S. Q., & Sewell, A. (2012). Phonological features of China English. *Asian Englishes*, *15*(2), 80–101. https://doi.org/10.1080/13488678.2012.10801331

Loakes, D., Clothier, J., Hajek, J., & Fletcher, J. (2024). Sociophonetic variation in vowel categorisation of Australian English. *Language and Speech*, *67*(3), 870-890. https://doi.org/10.1177/00238309231198520

Lüdecke, D., Ben-Shachar, M. S., Patil, I., Waggoner, P., & Makowski, D. (2021). An R package for assessment, comparison and testing of statistical models. *Journal of Open-Source Software*, *6*(61), 3319. https://doi.org/10.21105/joss.03139

Luo, J., Li, V. G., & Mok, P. P. K. (2020). The perception of Cantonese vowel length contrast by Mandarin speakers. *Language and Speech*, *63*(3), 635–659. https://doi.org/10.1177/0023830919879471

Maclagan, M. A. (2000). How long have women been leading language change? In J. Holmes (Ed.), *Gendered speech in social context: Perspectives from gown and town* (pp. 87–98). Victoria University Press.

Market Report China (2026). https://www.educationfair.net/market-reports/asia/china/

Mesthrie, R., & Bhatt, R. M. (2008). *World Englishes: The study of new linguistic varieties*. Cambridge University Press.

Miao, X., & Wang, Y. (2024, March 16). China still leading source of foreign students. *China Daily*. https://www.chinadaily.com.cn/a/202403/16/WS65f4e78ba31082fc043bcf71.html

Pillai, S., Mohd Don, Z., Knowles, G., & Tang, J. (2010). Malaysian English: An instrumental analysis of vowel contrasts. *World Englishes*, *29*(2), 159–172. https://doi.org/10.1111/j.1467-971X.2010.01636.x

Rao, D., & Shaw, J. A. (2024). Zhongjiang Chinese. *Journal of the International Phonetic Association*, *54*(3), 1071–99. https://doi.org/10.1017/S0025100324000203

RStudio Team. (2022*). RStudio: Integrated development environment for R* (version 1.1.383) [Computer software]. http://www.rstudio.com/

Schneider, E. W. (2014). New reflections on the evolutionary dynamics of world Englishes. *World Englishes*, *33*(1), 9–32. https://doi.org/10.1111/weng.12069

Sukaton, O. Z., Nugroho, D. P., & Ferdinandus, A. Y. (2023). How do I sound? Vowel production of multilingual students at Ma Chung University. *Rainbow: Journal of Literature, Linguistics and Culture Studies*, *12*(2), 124–133.

Tao, Y., Jalaluddin, I., & Kasim, Z. M. (2025). Striving for perfection: The influence of perfectionism on English language proficiency among Chinese non-English majors. *Indonesian Journal of Applied Linguistics*, *15*(1), 116–130. https://doi.org/10.17509/ijal.v15i1.79491

Tan, R. S. K., & Low, E. L. (2010). How different are the monophthongs of Malay speakers of Malaysian and Singapore English? *English World-Wide*, *31*(2), 162–189. https://doi.org/10.1075/eww.31.2.03tan

Wang, H., & Pillai, S. (2025). A preliminary study of the production of English monophthongs by mainland Chinese speakers based on age and gender. *Asian Englishes*, *27*(3), 682–699. https://doi.org/10.1080/13488678.2025.2517497

Wang, Q. (2023). An instrumental investigation of the vowel inventory of China English. *Asian Englishes*, *25*(1), 111–132. https://doi.org/10.1080/13488678.2022.2158265

Whalen, D. H., Chen, W. R., Tiede, M. K., & Nam, H. (2018). Variability of articulator positions and formants across nine English vowels. *Journal of Phonetics*, *68*, 1–14. https://doi.org/10.1016/j.wocn.2018.01.003

Williams, G. (2025, January 14). Chinese lifeline for struggling Malaysian varsities. *FMT*. https://www.freemalaysiatoday.com/category/opinion/2026/01/14/chinese-lifeline-for-struggling-malaysian-varsities

Winter, B. (2019). *Statistics for linguists*: *An introduction using R*. Routledge.

Wu, S., Liu, D., & Huang, S. (2023). The effects of over- and under-specified linguistic input on L2 online processing of referring expressions. *Journal of Psycholinguistic Research*, *52*(1), 283–305. https://doi.org/10.1007/s10936-022-09879-3

Yazawa, K. (2022). Investigating vowel productions with Japanese-English speech corpus. In *Proceedings of Data Science in Collaboration* Vol. 5 (pp. 9–18). https://tsukuba.repo.nii.ac.jp/records/2008343

Yazawa, K., Konishi, T., Whang, J., Escudero, P., & Kondo, M., (2023). Spectral and temporal implementation of Japanese speakers' English vowel categories: A corpus-based study. *Laboratory Phonology*, *14*(1), 1–33. https://doi.org/10.16995/labphon.6427

Yusuf, Y. Q., Fata, I. A., & Karwinda, S. (2021). Oral monophthong vowel qualities of the Jamee language in Aceh. *Indonesian Journal of Applied Linguistics*, *10*(3), 794–803. https://doi.org/10.17509/ijal.v10i3.31767

Zeng, J., Pillai, S., & Chew, S. Y. (2025). The production of English monophthongs by Chinese Yi and Han speakers: A comparative acoustic analysis. *English Today*, *41*(4), 225–232. https://doi.org/10.1017/S0266078425000094

Zhang, X., & Lütge, C. (2024). Understanding Chinese international students' language ideologies and multilingual practices in English-medium instruction programs in Germany. *Language and Education*, *38*(5), 881–899. https://doi.org/10.1080/09500782.2024.2348594

Zwicker, E., & Terhardt, E. (1980). Analytical expressions for critical bandwidths as a function of frequency. *The Journal of the Acoustical Society of America*, *68*, 1523–1525. https://doi.org/10.1121/1.385079

APPENDIX: QUESTIONNAIRE

I, __INSERT FULL NAME____, have read and understood all the terms and conditions of my participation in this study, and voluntarily agree to take part in the research.

SIGNATURE:_____ DATE: _____

Please answer each question by ticking (P) the appropriate number or filling in the blank.

1. Full Name: _____
2. Mobile Number: _____
3. Identity: (1) an undergraduate student (2) a master's student
   (3) a PhD candidate (4) an English teacher
4. Highest academic level achieved:
   (1) High School diploma (2) Bachelor's Degree
   (3) Master's Degree (4) Doctoral Degree
5. Were/Are you an English major during your undergraduate period?
   (1) Yes, please specify your name of degree and major (e.g. B*achelor of Arts, English)*:_____
   (2) No, please specify the name of your degree and major (e.g. *Bachelor of Science, Physics*):_____
6. Was/Is your master's programme related to English Language Studies?
   (1) Yes, please specify your name of degree and specific area of study (e.g. *Master of Arts, English Literature*): _____
   (2) No, please specify your degree and specific area of study (e.g. *Master of Science, Operational Research and Cybernetics*): _____
7. Was/Is your PhD programme related to English Language Studies?
   (1) Yes, please specify your area of study (e.g., *Semiotic Study of Political Memes*): _____
   (2) No, please specify your area of study (e.g., *Geographic Information System*): _____
8. Gender: (1) Female (2) Male
9. Age: (1) 18-30 (2) 31-45 (3) 46-60 (4) 61-72
10. Ethnicity: _____
11. Birth place: _____(city)_____(province)
12. Where did you mostly live between the ages of 0-12: _____(city)_____(province)
13. Where did you mostly live between the ages of 12-18: _____(city)_____(province)
14. Where have you spent most of your time after the age of 18:
    _____(city)_____(province)
15. Other places in mainland China that you have lived for more than six months (where and at what ages):
    _____
16. Other regions (including Taiwan, Hong Kong, and Macau) that you have lived for more than six months outside Mainland China (where and at what ages):
    _____
17. The most frequently used language that with your parents at home (if you use Chinese dialects, please specify what they are. For example, the Changsha dialect: _____
18. Other language that you use with your parents at home: _____
19. The most frequently used language currently spoken at home
    Spoke (For example, Changsha dialect and *Putonghua* with my parents; speak *Putonghua* with current family members): _____
20. Parents' home-domain language with their parents:
    Mother: _____
    Father: _____
21. Age of your first exposure to English: _____
22. Number of years learning English:
    (1) 5 years and below
    (2) 6-10 years
    (3) 11-14 years
    (4) 15 years and above
23. The level of your English proficiency:
    (1) CET 4, score:_____

(2) CET 6, score:_____
(3) TEM 4, score:_____
(4) TEM 8, score:_____
(5) IELTS, score:_____
(6) TOEFL, score:_____

[i] In this paper we use /e/ instead of /ɛ/ to represent the DRESS vowel. A discussion on the use of /e/ instead of /ɛ/ by the phonetician John Wells can be found at https://phonetic-blog.blogspot.com and

[ii] https://github.com/FieldDB/Praat-Scripts/blob/main/calculate_segment_durations.praat

[iii] https://github.com/FieldDB/Praat-Scripts/blob/main/collect_formant_data_from_files.praat