# STYLISTIC VARIATION IN VOWEL PRODUCTION: EVIDENCE FROM CHINA ENGLISH

Qian Wang

The Chinese University of Hong Kong
wangqian@hotmail.com.hk

## ABSTRACT

Previous research has shown that speech style triggers vowel variation in both length and acoustic quality. This study investigates vowel variability in three different speech styles (word reading, passage reading, and conversational interview) based on the acoustic data obtained from 20 speakers of China English (CE), a newly emerging variety of English in China. The influences of speech style on vowel duration and vowel space were examined using linear mixed effects models, which exclude the influences of individual differences and word frames. The interaction effects of vowel tenseness and speech style were also included in this model. The results suggest that the main effect of speech style and vowel tensity on vowel duration was significant. However, the effect of speech style on vowel space was not significant for the CE speakers. These findings shed some light on the vowel realization patterns of CE speakers in different styles and reveal some differences in stylistic vowel variation between L1 and L2 English vowels.

**Keywords**: vowels, speech style, duration, dispersion, phonetic variation, China English

## 1. INTRODUCTION

In traditional research of spoken language, speech style is usually controlled in order to exclude potential variations in sound quality which are caused by different speech styles. While controlling for speech style is fairly normal when investigating certain social and phonological variables, a better understanding of the effect speech style has on phonetic production would provide a more holistic picture of how sounds are realized across different styles, thus offering a direct comparison of stylistically divergent phonetic data. In light of this, this study examines the stylistic variation in vowels produced by CE speakers in three speech styles (i.e. word reading in citation form, passage reading, and conversational interview) by evaluating the influence of speech style on vowel length and vowel space. In addition, the interaction between vowel tensity and speech style in influencing the vowel length in different speech styles was also examined.

## 2. LITERATURE REVIEW

### 2.1. Stylistic variation

Previous studies have identified a variety of differences in acoustic realizations of speech sounds in different speech styles. For example, speakers may speak more clearly, more slowly [2, 11, 28], louder and with a higher voice pitch [2, 11] in clear speech than in conversational speech. Stylistic variation in speech is also prominently manifested in the characteristics of vowel modification; in particular, an expanded vowel space [2, 33], greater dynamic formant movement and longer vowel duration [11, 12, 13, 28] in clear speech than in conversational speech. Lindblom [24, 25] found that speakers may economize their vocal effort in connected speech, resulting in reduced speech (‘hypospeech’). Moreover, he argued that reduced speech and clear speech lie along a continuum of contextually determined variability because speech motor control is output-oriented and plastic.

### 2.2. Vowel variation

The acoustic properties of vowels vary depending on a wide range of intrinsic and extrinsic variables. Vowel-intrinsic variables include tensity, length, stress, and voice register [4, 18, 19, 22] and vowel-extrinsic variables include phonological context, which may cause coarticulation effects [17, 30], speech register and speech style [9, 20, 21, 28].

*2.2.1 Vowel tensity and vowel length*

Vowel length and vowel tensity are two interrelated constructs. Vowel length contrast is often accompanied by a difference in vowel tensity, and vowel tensity can also serve as a cue to indicate a vowel length contrast [26].

In English, spectral difference is the primary acoustic cue for the tense and lax vowel contrast ([16]). That is, tense vowels (e.g. /i:/ and /u:/) are produced with more extreme articulatory movements, resulting in larger vowel spaces than lax vowels (e.g. /ɪ/ and /ʊ/) ([16, 23]). In addition, vowel length also plays an important role in representing the English tense and lax vowel contrast (tensity difference). According to Leung [23], tense vowels are typically longer than their lax vowel counterparts because tense vowels require longer excursions for the articulators to reach the more extreme target positions ([16]). Thus, lax vowels (being intrinsically short) are more resistant to further compression in duration under the influence of the contextual factors (e.g. consonantal context and speaking rate), as compared to the tense vowels [14, 22, 30]. Some empirical studies also found that vowel shortening usually co-occurs with vowel centralization within the vowel space ([10, 26]). That is, when vowel duration is shortened, speakers may not be able to reach the intended vowel target, which arguably reflects a general process of 'vowel undershoot'.

*2.2.2 Speech style and vowel variation*

Apart from the phonetic variables, some paralinguistic factors also influence vowel articulation in a variety of ways. For example, many empirical studies have shown that acoustic qualities of vowels vary in different speech styles [9, 10, 11, 13, 21]. Although speech styles were operationalized in different ways in different studies, for example, 'elicited' vs. 'spontaneous' [9, 10] and 'clear' vs. 'conversational' speech [11, 13]), the studies yield very similar results: vowels have more extreme spectral properties and longer duration in clear speech than in conversational speech (e.g. [3, 9, 12, 18]). As longer duration and extreme spectral properties are also the acoustic characteristics for distinction between tense and lax vowels, the interaction effects of vowel tensity and speech style on vowel duration change is worth further study.

Numerous studies have shown that vowel reduction across different speech styles is a systematic phenomenon because the speech task is associated with different demands on the speech mechanism, which may lead to changes in vowel production [10, 21]. For example, DiCanio et al. [9] found that acoustic variations of vowels do exist across speech styles, and that they are not entirely reducible to durational differences. Spectral variation is also observable in accordance with the style change.

**2.3. Speech styles and L2 speakers**

As the previous review shows, vowel centralization and shortening in conversational speech as compared to clear speech have been widely observed in first language (L1) speech data. However, little research has been done to explore: 1. whether second language (L2) speakers also possess similar stylistic variation patterns in vowel length and vowel space to those observed in L1 speakers, and 2. whether L2 speakers have a similar interrelated relationship between vowel tensity and vowel length to that observed in L1 speakers.

Traditional L2 studies suggest that L2 speakers' interlanguage is more permeable to the target language in more formal speech tasks, or as more attention is paid to the form rather than the content (e.g. [27, 34]). Thus, it can be hypothesized that L2 Chinese speakers of English, who have been widely reported to have centralized vowel space and mergers of length and tensity vowel contrasts, should exhibit more disperse vowel space and maintain the length and tensity contrast better in elicited speech than in conversational speech [5,7].

The primary goal of this study is to explore the vowel variation of CE in three different speech styles with a special focus on vowel length and vowel space. In addition, the interaction effects of vowel tensity and speech style on vowel duration change were also examined.

## 3. RESEARCH METHODS

**3.1. Speakers**

Twenty CE speakers were recruited from the university students in a high-caliber university where English is used as the language of instruction for courses across different disciplines. By definition, CE refers to the highly intelligible English spoken by well-educated and highly-proficient English speakers in China. Thus, the participants in this study are highly-proficient English speakers who were born and raised in mainland China and who had lived and studied exclusively in mainland China before coming to Hong Kong for further study. These participants all self-claimed Mandarin as their L1. I intentionally excluded those who have high proficiency level in their home dialect to avoid the influence from different Chinese dialects. The participants had a mean age of 21.8 years at the time of recording (ranging from 18 to 26 years), and they all achieved an overall score of 7.0 or above in IELTS with at least 6.0 in the speaking component or the equivalent. The years of learning English of the participants varied from 11 to 21 years, with a mean of 14.6 years. Only female speakers were included in this study to avoid the variation caused by gender.

**3.2. Data collection**

Speech data were obtained from three different speech tasks (word reading, passage reading and conversational interview). The passage was adapted from a well-known fable by Aesop '*The Boy Who Cried Wolf*' in Deterding [7]. This passage has been widely used as a diagnostic tool to describe and measure the pronunciation features of new English varieties. Thirty-two words containing the 11 target monophthongs were selected from the passage for speakers to read in citation form. Clearly-spoken stressed instances of each monophthong were used. The words were all phrased in a carrier sentence 'I say_____' to minimize participants' over-emphasis of the isolated words. The conversational data were collected through an interview in which the speakers were asked questions about their English learning experience. On average, three tokens of each vowel were selected and analysed from interview speech of each speaker. As far as possible, the conversational vowels were extracted from the stressed syllables and have the phonological contexts resembling those in passage reading and word reading. The speech data were recorded on a Sony Digital Audio Tape recorder using a Shure microphone with a sampling frequency of 44100 Hz.

### 3.3. Data analysis

*3.3.1. Formant frequency values*

The frequency values of the first two formants (F1 and F2) were measured at the steady state around the vowel temporal midpoint in Praat. The raw frequency values were then converted into a Bark scale, proposed by Zwicker and Terhardt [35] and used by many phonetic studies [1, 6]. Using this conversion, the distance between the plotted vowel formant values is more similar to the distances between the auditorily perceived vowel qualities.

*3.3.2. Vowel space measures*

In order to examine how speech style influences vowel space and dispersion, two measures of vowel space size were calculated from the formant data: vowel space dispersion (mean Euclidean distance of the individual vowels from the center of the vowel space (cf. [3, 33]) and vowel space area (area of the vowel polygon based on the first two formant values of four corner vowels /i:, u:, ɑ:, æ/ (cf. [3, 29, 33]). Instead of using raw formant frequency values, the Bark value was used when measuring vowel space and vowel dispersion.

### 3.3.3. Vowel duration

Vowel duration measures were made from the spectrogram and waveform in Praat. The vowel onset and offset were manually identified by the author.

## 4. RESULTS

### 4.1 Vowel duration

To explore the effects of vowel tensity and speech style on vowel duration, a linear mixed effects model was performed in R ([32]) with two fixed effects (speech style and vowel tensity) and two random effects (speaker and word frame). Two-way interactions were also included in this full model. The results suggest that there was a significant effect of vowel tensity on vowel duration (AIC = 5888.51 vs. 5878.72; $\chi^2 = 11.79$, $p < .001$). The average ratio between lax vowels (mean duration = 113.3 ms) and tense vowels (mean duration = 142.1 ms) was 1:1.26.

The main effect of speech style on duration was also significant, as shown in Figure 1. The vowels in the word list reading have a significantly longer duration than vowels in conversational speech (AIC = -5888.51 vs. -5849.11; $\chi^2 = 41.40$, $p < .001$) as well as vowels in passage reading (AIC = -5888.51 vs. -5833.31; $\chi^2 = 104.37$, $p < .001$). There was no significant durational difference observed for vowels in passage reading and conversational speech (AIC = -5888.51 vs. -5890.16; $\chi^2 = 0.35$, $p = .55$). The durational ratio between vowels in conversational speech (mean duration = 111.9 ms) and word reading in citation form (mean duration = 156.8 ms) was 1:1.40. However, this effect was asymmetrical, as there was a significant interaction effect of vowel tensity and speech style on vowel duration for word reading and conversational speech (AIC = -5888.51 vs. -5885.09; $\chi^2 = 5.43$, $p = .02$). That is, tense vowels in conversational speech (mean duration = 123.6 ms) were 58.9 ms shorter than their counterparts in word reading (mean duration =182.5 ms), and the average ratio of vowel duration in two speech tasks was 1:1.48. However, lax vowels in conversational speech (mean duration = 102.4 ms) were only 34.1 ms shorter than their counterparts in word reading (mean duration = 136.5 ms). Thus, the durational ratio of lax vowels in two speech styles was 1:1.33. This result suggests that, compared to tense vowels, lax vowels may be more resistant to the compression in length when the speech style changes.

### 4.2 Vowel space

As shown in Figure 2, vowel tokens from the three speaking styles cluster together to varying degrees, regardless of the tensity of vowels. As shown in Figure 3, the vowel polygons in three speech styles that were drawn based on the mean frequency values of the first two formants of four corner vowels /i:, u:, ɑ:, æ/ did overlap to a great extent.

**Figure 1:** Vowel duration in three speech styles.

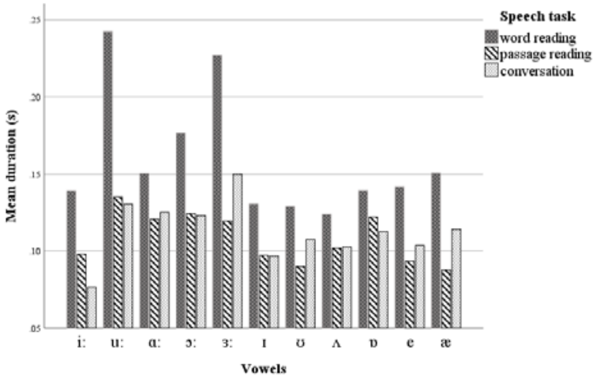
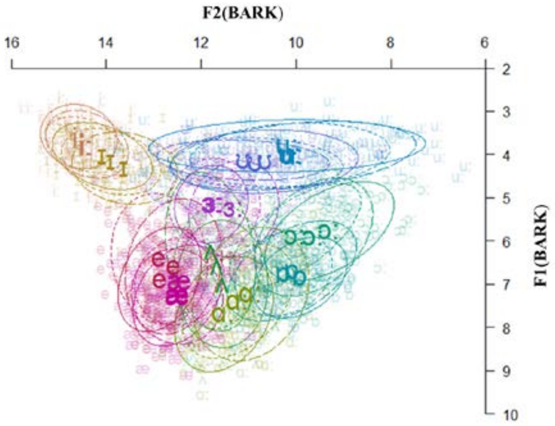


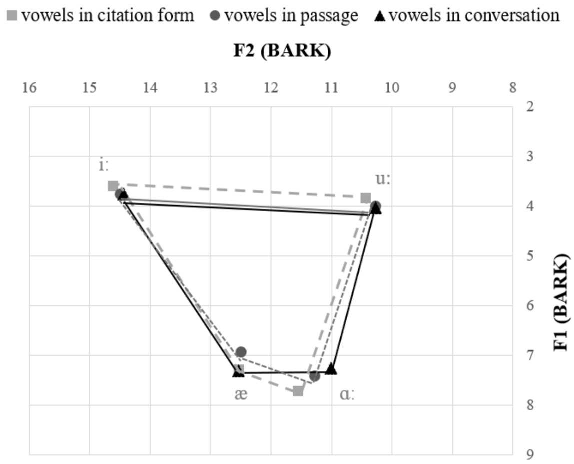

**Figure 2:** Vowel scatter plots in three styles.




As for the measurement of vowel space expansion, the results of each individual speaker's vowel space dispersion for four corner vowels in three speaking styles were calculated. A linear mixed effects model was performed to explore the effect of speech style on vowel dispersion. 'Speech style' and 'vowel' were entered as fixed effects (without interaction term), and 'speaker' was entered as the random effect. The results suggest that there was no significant effect of speech style on the degree of vowel dispersion (AIC = 507.30 vs. 507.49; $\chi^2$ = 2.19, $p$ = .14), but that a significant effect of individual vowels on the degree of vowel dispersion was found (AIC = 507.30 vs. 832.82; $\chi^2$ = 331.52, $p$ < .001). In other words, the degree of vowel dispersion in CE is not influenced significantly by speech style, but different vowels may have varying degrees of dispersion in different speech styles.

A linear mixed effects model was constructed with 'speech style' as the fixed effect and 'speaker' as the random effect to explore the effect of speech style on the vowel space area in three speech styles. The result suggests that the main effect of speech style on vowel space area was not statistically significant. The difference between word list reading and passage reading (AIC = 383.98 vs. 382.82; $\chi^2$ = .84, $p$ = .36), as well as the difference between word list reading and conversational speech (AIC = 383.98 vs. 382.94; $\chi^2$ = .96, $p$ = .33), are both statistically insignificant.

**Figure 3:** Vowel space area in three speech styles.




## 5. DISCUSSION & CONCLUSIONS

The results from the present study suggest that the main effect of vowel tensity on vowel duration was significant in CE vowels, which means that tense vowels were significanty longer than their lax counterparts in CE. This finding contradicts the previous findings on Chinese L2 English learners which stated that they tend to merge the length contrasts between tense and lax vowels ([1, 5]). This interesting finding may provide some support to the claim that CE, which is different from L2 learner English, has its own stabilized sound system and internalized phonological rules, as do other inner-circle English varieties.

In addition, the main effect of speech style on vowel duration was significant; i.e., vowels in the word list reading were significantly longer than those in passage reading and those in conversational speech. However, vowel space and vowel dispersion did not show any significant difference between the three speech styles. This result did not conform to previous studies on L1 English speakers [21, 23], which found salient vowel space expansion in clear speech, as in conversational speech. Based on this result, we can speculate that speech style may exert different influences on L1 and L2 speech production, but further empirical research needs to be conducted to validate this speculation since the current data only focused on the effects of style on vowel acoustics. Thus, this study not only contributes to our knowledge on the acoustic quality of CE vowels, but also provides a better understanding of vowel variability caused by speech styles in L2 spoken language.

## 6. REFERENCES


[1] Ao, R., Low, E. L. 2015. A description of the Yunnan English accent. *World Englishes* 34(3), 336-354.

[2] Bradlow, A. R., Kraus, N., Hayes, E. 2003. Speaking clearly for children with learning disabilities: Sentence perception in noise, *J. Speech Lang. Hear.* 46, 80–97.

[3] Bradlow, A. R., Torretta, G. M., Pisoni, D. B. 1996. Intelligibility of normal speech. I. Global and fine-grained acoustic-phonetic talker characteristics. *Speech Communication* 20, 255–272.

[4] de Jong, K., Zawaydeh, B. 2002. Comparing stress, lexical focus, and segmental focus: patterns of variation in Arabic vowel duration. *J. Phonetics* 30, 53–75.

[5] Chang, J. 2001. Chinese speaker. In: M. Swan & B. Smith (eds), *Learner English: A teacher' guide to interference and other problems (2nd ed.*). Cambridge, England: Cambridge University Press, 310-324.

[6] Deterding, D. 2003. An instrumental study of the monophthong vowels of Singapore English. *English World-Wide* 24(1), 1–16.

[7] Deterding, D. 2006a. 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.

[8] Deterding, D. 2006b. The pronunciation of English by speakers from China. *English World-Wide* 27(2), 175-198.

[9] DiCanio, C., Nam, H., Amith, J. D., Castillo García, R., Whalen, D. H. 2015. Vowel variability in elicited versus spontaneous speech: evidence from Mixtec. *Journal of Phonetics* 48, 45–59.

[10] DiCanio, C., Whalen, D.H. 2015. The interaction of vowel length and speech style in an Arapaho speech corpus. *Proceedings of the 18th International Congress of the Phonetic Sciences*, Glasgow, Scotland.

[11] Ferguson, S. H., Kewley-Port, D. 2002.Vowel intelligibility in clear and conversational speech for normal-hearing and hearing-impaired listeners. *J. Acoust. Soc. Am*. 112, 259–271.

[12] Ferguson, S. H., Kewley-Port, D. 2007. Talker differences in clear and conversational speech: Acoustic characteristics of vowels. *J. Speech Lang. Hear. Res*. 50, 1241–1255.

[13] Ferguson, S. H., Quené, H. 2014. Acoustic correlates of vowel intelligibility in clear and conversational speech for young normal-hearing and elderly hearing-impaired listeners. *J. Acoust. Soc. Am*. 135(6), 3570-3584.

[14] Gopal, H. S. 1990. Effects of speaking rate on the behavior of tense and lax vowel durations. *J. Phonetics* 18, 497–518.

[15] Hazan, V., Baker, R. 2011. Acoustic-phonetic characteristics of speech produced with communicative intent to counter adverse listening conditions. *J. Acoust. Soc. Am*. 130, 2139–2152.

[16] Hillenbrand, J. M., Getty, L. A., Clark, M. J., Wheeler, K. 1995. Acoustic characteristics of American English vowels. *J. Acoust. Soc. Am.* 97, 3099–3111.

[17] Hillenbrand, J. M., Clark, M. J., Nearey, T. M. 2001. Effects of consonant environment on vowel formant patterns. *J. Acoust. Soc. Am*. 109(2), 748–763.

[18] Hirata, Y., Tsukada, K. 2009. Effects of speaking rate and vowel length on formant frequency displacement in Japanese. *Phonetica* 66, 129–149.

[19] Johnson, K., Martin, J. 2001. Acoustic vowel reduction in Creek: Effects of distinctive length and position in the word. *Phonetica* 58, 81–102.

[20] Keating, P. A., Huffman, M. K. 1984. Vowel variation in Japanese. *Phonetica* 41, 191–207.

[21] Kuo, *C,* Weismer G. 2016. *Vowel reduction across tasks for male speakers of American English*. *J. Acoust. Soc. Am*. 140, 369-383.

[22] Lehnert-Lehouillier, H. 2010. A cross-linguistic investigation of cues to vowel length perception. *J. Phonetics* 38, 472–482.

[23] Leung, K. K. W., Jongman, A., Wang, Y., Sereno, J. A. 2016. Acoustic characteristics of clearly spoken English tense and lax vowels. *J. Acoust. Soc. Am.* 140(1), 45-58.

[24] Lindblom, B. 1983. *Economy* of *speech gestures.* In: P. F. MacNeilage *(*ed.), *The Production* of *Speech.* New York*:* Springer-Verlag, 217–246.

[25] Lindblom, B. 1990. Explaining phonetic variation: A sketch of the H&H theory. In: W. J. Hardcastle and A. Marchal (eds.), *Speech Production and Speech Modeling*. Netherlands: Kluwer Academic Press, 403–439.

[26] Maddieson, I. 1984. *Patterns of Sounds*. Cambridge University Press.

[27] Mckay, S. L., Hornberger, N. H. (eds.) 1995. *Sociolinguistics and language teaching*. New York: Cambridge University Press.

[28] Moon, S.-J., Lindblom, B. 1994. Interaction between duration, context, and speaking style in English stressed vowels. *J. Acoust. Soc. Am*. 96(1), 40–55.

[29] Neel, *A. T.* 2008. *Vowel space characteristics and vowel identification accuracy, J. Speech Lang. Hear. Res.* 51(3), 574–585.

[30] Ohde, R. N., Sharf, D. J. 1975. Coarticulatory effects of voiced stops on the reduction of acoustic vowel targets. *Acoust. Soc. Am*. 58, 923-927.

[31] Port, R. F. 1981. Linguistic timing factors in combination. *J. Acoust. Soc. Am.* 69, 262–724.

[32] R Core Team. 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

[33] Smiljanic, R., Bradlow, A. R. 2005. Production and perception of clear speech in Croatian and English. *J. Acoust. Soc. Am.* 118, 1677–1688.

[34] Tarone, E. 1979. Interlanguage as chameleon. *Language Learning* 29, 181-191.

[35] Zwicker, E., Terhardt, E. 1980. Analytical expressions for critical-band rate and critical bandwidth as a function of frequency. *J. Acoust. Soc. Am.* 29(5), 1523–1525.