Interspeech 2018
2-6 September 2018, Hyderabad

Vowels and Diphthongs in Hangzhou Wu Chinese Dialect

Yang Yue[1], *Fang Hu*[2]

[1]Department of Linguistics, Graduate School, Chinese Academy of Social Sciences
[2] Institute of Linguistics, Chinese Academy of Social Sciences
kk24325535@qq.com, hufang@cass.org.cn

Abstract

This paper gives an acoustic phonetic description of the vowels and diphthongs in Hangzhou Wu Chinese dialect. Data from 12 speakers, 6 male and 6 female, were measured and analyzed. Monophthongs were investigated in CV, CVN, and CVC syllables; diphthongs were examined in terms of temporal organization, spectral properties, and dynamic aspects. Results suggest that falling diphthongs tend to have a single dynamic target, while rising diphthongs have two static spectral targets.

Index Terms: vowel production, monophthongs, falling diphthongs, rising diphthongs, Hangzhou Wu Chinese dialect

1. Introduction

Hangzhou was the capital of China during the historical period of Southern Song dynasty (A.D. 1127-1279) and is now the capital city of Zhejiang province. The Hangzhou dialect belongs to the Wu dialect family but has certain characteristics of Mandarin ([1], [2], [3], [4]). Previous studies on Hangzhou phonology were mostly impressionistic dialectological works, and little was based on phonetic data ([5], [6]). This paper attempts to give an outline of vowel phonology in Hangzhou on the basis of an acoustic phonetic description of vowel production.

Syllables in Chinese dialects are straightforward, as each syllable is a separate written unit. Chinese dialects usually have a simple syllable structure of CVC, where both initial and final consonants are optional, and the coda consonant is often limited to a few nasal consonants and in some dialects also a few unreleased stops. As compared to consonants, there is more complexity in vowels in Chinese dialects. There are monophthongs, diphthongs, triphthongs, and in some rare cases tetraphthongs ([7]). There is a long debate on the structure of rimes and vowels, and a key issue concerns the nature of diphthongs. For instance, there are various proposals for the phonology of vowels and diphthongs in Standard Chinese ([8], [9], [10], [11], [12], [13], [14], [15], [16]). However, previous studies are based on phonemic analyses in general, and thus have non-unique solutions ([17]). An insightful observation is that falling and rising diphthongs are different. Chao pointed out that falling diphthongs are true diphthongs in Wu dialects, but rising diphthongs are not ([18]). Recent phonetic researches from Chinese dialects proposed that falling diphthongs and monophthongs are single-event articulations, while rising diphthongs are sequences of two articulation events ([19], [20], [21], [22], [23], [24]). For instance, falling diphthong [ai] is not a sequence of [a] and [i], but a single dynamic articulatory event, and should thus be treated as being distinctive to the monophthongal vowel [a]; in contrast, rising diphthong [ia] is a sequence of [i] and [a].

Hangzhou has a rich vowel inventory, which is typical in northern Wu Chinese dialects. It is of theoretical interest to examine how phonological vowel oppositions are implemented in Hangzhou on one hand and to explore how physical aspects of vowel production play a role in the phonology in a particular language with a rich inventory of vowels and diphthongs on the other hand. Rising diphthongs are labeled as GV in this paper, and thus the three syllable types of Hangzhou were summarized as C(G)V, C(G)VN, and C(G)VC. As shown in Table 1, C(G)Vs are open syllables, and there are 8 monophthongs [ɿ ʮ a i u y ɛ ɔ], 2 falling diphthongs [ei ou], 10 rising diphthongs [ia iɔ iɛ ua uɛ uo yo ʮa ʮɛ ʮo], and 1 triphthong [uei]. Note that this paper follows Karlgren's notation ([25]) in the transcription of apical vowels: the alveolar unrounded [ɿ] and the alveolar rounded [ʮ]. C(G)VNs are syllables with nasal coda and the velar ŋ is the only legitimate nasal coda in Hangzhou. Vowel contrasts are substantially reduced in syllables with a nasal coda, and there are 5 monophthongs [i y a o ə] and 6 rising diphthongs [ia ua ʮa uə ʮə io]. C(G)VCs are checked syllables, and the glottal stop ʔ is the only legitimate coda C. And even less vowels are found in checked syllables: 2 monophthongs [o a] and 5 rising diphthongs [io iɛ ua yɛ ʮa]. Note that falling diphthongs only occur in open syllables, and are not allowed in syllables with a consonantal coda N or C. And note that the monophthong [o] in C(G)VN and C(G)VC is a monophthongal version of the falling diphthong [ou].

Table 1. *Hangzhou vowel inventory.*

Vowels	C(G)V	C(G)VN	C(G)VC
Monophthongs	a i u y ɛ ɔ ɿ ʮ	i(ŋ) y(ŋ) a(ŋ) ə(ŋ) o(ŋ)	a(ʔ) o(ʔ)
Falling diphthongs	ei ou		
Rising diphthongs	ia iɔ iɛ ua uɛ uo yo ʮa ʮɛ ʮo	ia(ŋ) ua(ŋ) ʮa(ŋ) uə(ŋ) ʮə(ŋ) io(ŋ)	io(ʔ) iɛ(ʔ) ua(ʔ) yɛ(ʔ) ʮa(ʔ)
Triphthongs	uei		

2. Methodology

12 native speakers, 6 male and 6 female, provided speech data. All of them were born and raised up in Hangzhou, and had no reported history of speech disorders. Meaningful monosyllabic

10.21437/Interspeech.2018-1225

words were used as test words and two words, preferably one with a labial initial and the other without an initial consonant, were used for each target vowel. And all test words preferably have a mid tone. The sounds were recorded into a laptop PC with DMX 6 Fire USB sound card through a SHURE SM86 microphone. Speakers were instructed to speak the test words in a natural way with a normal tempo. The sample rate is 22,050 Hz, and five repetitions were recorded.

The data were annotated and analyzed by PRAAT 6.0.19 ([26]). Monophthongs usually have a spectral steady state and were annotated as one segment; diphthongs are typically composed of two steady states and a transition, and were annotated as three segments, i.e. an onset, an offset, and the transition connecting them. The lowest three formants were extracted from the spectrogram in the mid-point of each target segment. Temporal structures were measured for each diphthong, and the second formant (F2) range and rate of change were then calculated for the characterization of spectral dynamics of diphthongs.

3. Results

3.1. Monophthongs

3.1.1. Monophthongs in CV syllables

Figure 1 shows the distribution of the 8 monophthongs [a i u y ɛ ɔ ɿ ʮ] in CV syllables in Hangzhou dialect in the acoustic vowel plane by using first formant (F1) for Y axis and second formant (F2) for X axis with the origin to the top right. The axes are Bark-scaled ([27]) but still labeled in Hz. Each 2-sigma confidence ellipse is based on 60 data points (2 targets × 5 repetitions × 6 speakers).

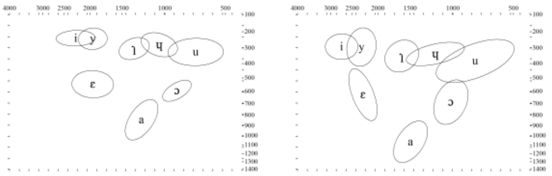

Figure 1: *2-sigma ellipses for the monophthongs in CV syllables in Hangzhou: male (left) and female speakers (right).*

It can be seen from the figure that Hangzhou monophthongs in CV syllables have a triangular distribution in the acoustic F1/F2 vowel plane. The monophthongs distinguish three levels of height and three levels of backness: [i u y ɿ ʮ] are high vowels, [ɛ ɔ] are mid vowels, and [a] is the low vowel; [i y ɛ] are front vowels, [ɿ ʮ a] are central vowels, and [u ɔ] are back vowels. Except for the high vowels, Hangzhou monophthongs have a predictable relationship between lip rounding and vowel backness, namely front vowels are unrounded and back vowels rounded. The high vowels distinguish in lip rounding: [i ɿ] are unrounded, and [y ʮ u] are rounded. It is common that apical vowels occupy high central regions in the acoustic vowel plane, as they are historically related to high front vowels in general ([28]).

3.1.2. Monophthongs in CVN syllables

Less vowel contrasts are retained in CVN syllables than in CV syllables. It can be seen from Figure 2 that the monophthongs [a i ə o y] in CVN syllables, which are denoted by the nasal ending [-ŋ] in the following parenthesis, have a less peripheral distribution than the corresponding reference monophthongs [a i u y] in CV syllables in the acoustic F1/F2 vowel plane. Meanwhile it should also be noted that the ellipses for [a i o y] in CVN syllables extensively overlap with those for their counterparts [a i u y] in CV syllables, respectively. This means that vowels in CVN and CV syllables have comparable spectral targets. Additionally, there is a new monophthong [ə] in CVN syllables that is not detected in CV syllables. The schwa [ə] in CVN syllables could possibly have resulted from neutralized vowel contrasts in CV syllables; however, it is difficult to trace detailed procedures of vowel neutralization in current Hangzhou phonology.

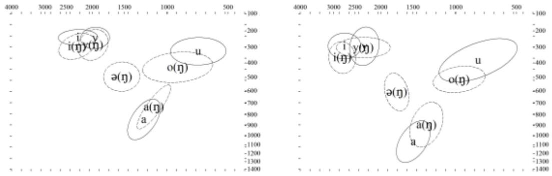

Figure 2: *2-sigma ellipses for the monophthongs [a i ə o y] in CVN and corresponding reference monophthongs [a i u y] in CV syllables in Hangzhou: male (left) and female speakers (right).*

3.1.3. Monophthongs in CVC syllables

CVC syllables have even less vowel contrasts in Hangzhou, and the vowels in checked syllables are about 53% shorter than in CV syllables. It was reported in the literature that there were three monophthongs [a o ə(e)] in CVC syllables in Hangzhou ([1], [2], [4], [29]). But only a two-way contrast [a o] was found in all the 12 speakers in this study.

Figure 3 shows the distribution of the two monophthongs [o a] in CVC syllables, which are denoted by the glottal stop [-ʔ] in the following parenthesis, and the corresponding reference monophthongs [a i u y] in CV syllables in the acoustic F1/F2 vowel plane. It can be seen from the figure that the ellipse for the short vowel [a] in CVC syllables extensively overlaps with the corresponding monophthong [a] in CV syllables. That is, they have comparable spectral targets. In other words, the short vowel in a checked syllable is not necessarily reduced in vowel quality.

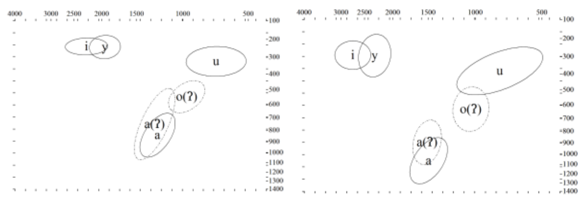

Figure 3: 2-sigma ellipses for the monophthongs [a o] in CVC and the corresponding reference monophthongs [a i u y] in CV syllables in Hangzhou: male (left) and female speakers (right).

3.2. Diphthongs

Hangzhou vowel phonology shares a commonality with other Chinese dialects in that a vowel element V and the preceding on-glide G form a rising diphthong, but differs in that in addition to the three high-vowel on-glides [i- u- y-], there is an apical one [ʮ-].

There are 9 rising diphthongs [ia iɔ iɛ ua uɛ uo ʮa ʮɛ ʮo] in CGV syllables, 5 rising diphthongs [ʮə(ŋ) uə(ŋ) ia(ŋ) ua(ŋ) ʮa(ŋ)] in CGVN syllables, and 5 rising diphthongs [io(ʔ) iɛ(ʔ)

ua(ʔ) yɛ(ʔ) ɥa(ʔ)] in CGVC syllables in Hangzhou. In addition, there are 2 falling diphthongs [ei ou], and a triphthong [uei]. Due to the space limit, discussion focuses on falling diphthongs and the rising diphthongs in CGV syllables.

3.2.1. Temporal organization

Figures 4 and 5 show mean durations in second and in percentage respectively for diphthong components in falling diphthongs and rising diphthongs in CGV syllables in Hangzhou.

First, both onset and offset components in both falling and rising diphthongs have steady states. Second, the onset has a longer duration than the offset in a falling diphthong. On average, the onset takes more than 44%, the offset takes about 32%, and the transition takes less than 24% of syllable duration in a falling diphthong. Third, rising diphthongs exhibit temporal organizations different to falling diphthongs in general. In a rising diphthong, the offset is longer than the onset. [iɛ uo] are exceptions, and they have temporal organizations similar to falling diphthongs.

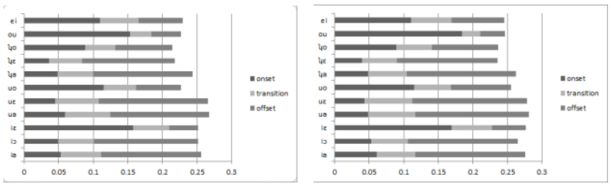

Figure 4: *Mean durations in second for the falling diphthongs and rising diphthongs in CGV syllables in Hangzhou: male (left) and female speakers (right).*

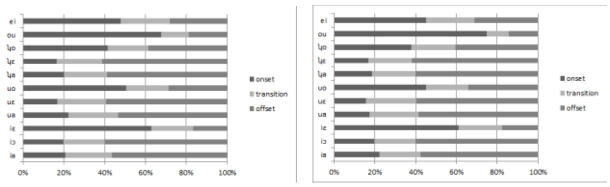

Figure 5: *Mean durations in percentage for the falling diphthongs and rising diphthongs in CGV syllables in Hangzhou: male (left) and female speakers (right).*

3.2.2. Spectral properties

Figures 6 to 9 compare 2-sigma ellipses for onset and offset components in rising diphthongs [ia iɔ iɛ ua uɛ uo ɥa ɥɛ ɥo ou ei] with their monophthongal counterparts in CV syllables. The IPAs with parentheses denote diphthong components.

It can be observed from the figures that both onset and offset components in rising diphthongs have comparable spectral distributions to their monophthongal counterparts in general. First, the offset components in rising diphthongs overlap with their monophthongal counterparts respectively, suggesting that rising diphthongs have similar spectral offset targets as their corresponding monophthongs do. Second, although there are certain coarticulatory effects, the onset components [i] and [u] in rising diphthongs overlap with their monophthongal counterparts in general. Third, the onset component [ɥ] in rising diphthongs shows more variability, suggesting more effects of coarticulation. In summary, it seems that the onset is subject to anticipatory coarticulation, but suffice it to conclude that rising diphthongs have both onset and offset targets.

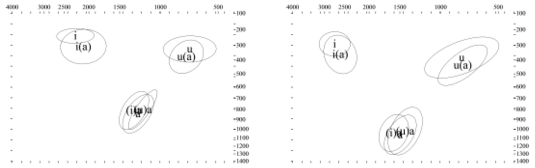

Figure 6: *2-sigma ellipses for the onset and offset components of [ia ua] and their monophthongal counterparts in CV syllables in Hangzhou: male (left) and female speakers (right).*

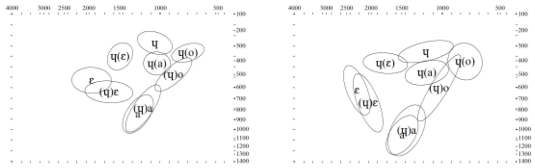

Figure 7: *2-sigma ellipses for the onset and offset components of [ua ɥɛ ɥo] and their monophthongal counterparts in CV syllables in Hangzhou: male (left) and female speakers (right).*

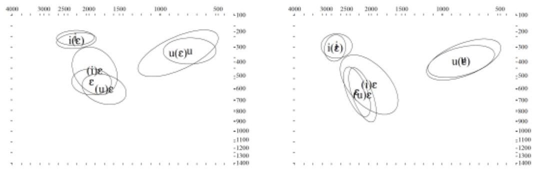

Figure 8: *2-sigma ellipses for the onset and offset components of [iɛ uɛ] and their monophthongal counterparts in CV syllables in Hangzhou: male (left) and female speakers (right).*

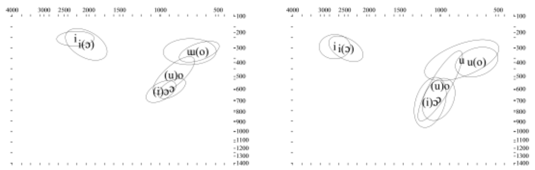

Figure 9: *2-sigma ellipses for the onset and offset components of [uo iɔ] and their monophthongal counterparts in CV syllables in Hangzhou: male (left) and female speakers (right).*

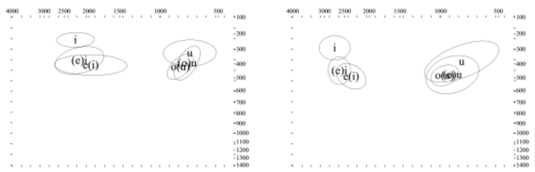

Figure 10: *2-sigma ellipse for the onset and offset parts of [ou ei] and their monophthongal counterparts in CV syllables in Hangzhou: male (left) and female speakers (right).*

Figure 10 shows 2-sigma ellipses for the components of falling diphthongs [ei ou] in CV syllables in acoustic F1/F2 vowel plane. And the ellipses for the monophthongs [i u] in CV syllables were superimposed for reference. First, both [ei] and [ou] have short spectral movements, as the ellipses for the onset and offset extensively overlap with each other in both diphthongs. Second, the offset in [ei] has a spectral distribution separate to its monophthongal counterpart [i] in the acoustic vowel plane, whereas the ellipses for not only the offset but also the onset in [ou] overlap with the corresponding monophthong [u]. But this does not mean that the two falling diphthongs are different. Rather, the observed difference between [ei] and [ou] suggests that it is not important to have specific spectral targets for the production of the two falling diphthongs. What is essential is being spectrally dynamic for [ei] and [ou]. Remind that Hangzhou does not have mid-high monophthongs but has mid-low vowels [ɛ ɔ]. It seems that spectral dynamics helps distinguish mid-high vowels from their neighboring high and mid-low vowels, namely [i ei ɛ] and [u ou ɔ] respectively.

A piece of independent evidence comes from the “-er” suffixation. The diminutive suffix is phonetically realized as [əl] in Hangzhou, and there is no resyllabification as in Beijing Mandarin ([30]). That is, [əl] stays as an individual syllable during the morphological process of er-suffixation. It is however intriguing to observe that er-suffixation triggers monophthongization for the falling diphthongs in Hangzhou. That is, the two falling diphthongs [ei ou] are realized as monophthongs [e o] respectively after er-suffixation in Hangzhou. Figure 11 plots [e o] with the other monophthongs in CV syllables in Hangzhou. That is, the Hangzhou monophthongs actually distinguish four levels of vowel height, as the falling diphthongs [ei ou] should be treated as being distinctive to other monophthongs in CV syllables. It could be observed from the figure that the ellipse for [e] overlaps with that for [ɛ] in female speakers and the ellipse for [o] overlaps heavily with that for [u] in both male and female speakers. This explains why [e o] are usually realized as dynamic vowels, i.e. falling diphthongs. That is, as mentioned above, being spectrally dynamic helps make sufficient distinctions among neighboring vowels.

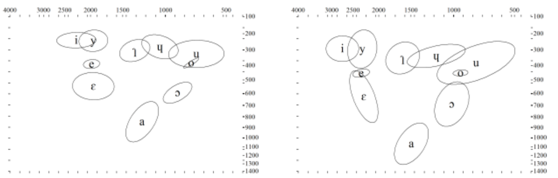

Figure 11: *Hangzhou vowels with the monophthongized diphthongs [e<ei o<ou]: male (left) and female speakers (right).*

3.2.3. Dynamic aspects

Dynamic aspects were expressed as range and rate of spectral change for the second formant in diphthong production ([31], [32]). Table 2 summarized mean range and rate of F2 change for the diphthongs in CGV syllables in Hangzhou.

It can be seen from the table that diphthongs are characterized by F2 range and rate of change, if they are accounted within their own series. First, [ei] has greater F2 range and rate of change than [ou] in falling diphthongs. Second, [iɔ] > [ia] > [ie] in i-series rising diphthongs. Third, [uɛ] > [ua] > [uo] in u-series rising diphthongs. Fourth, [ɥɛ] > [ɥa]/[ɥo] in ɥ-series rising diphthongs. What is general is that a greater F2 range of change is correlated with a greater F2 rate of change. And the data is consistent across male and female speakers. However, there is no general pattern, if all diphthongs are taken into account together. The results corroborate previous studies that the F2 range and rate of change could characterize diphthongs in languages with a simple inventory of diphthongs, but would have problems in languages with a complex inventory of diphthongs ([31], [32], [33], [34], [35], [36]). And in the latter cases, it is better to classify diphthongs into different groups and then check if there is a general pattern of spectral dynamics in diphthong production.

Table 2. *Mean range in Hz and rate in Hz/ms of the F2 change for the diphthongs in CV syllables.*

Diphthongs		Male		Female	
		Δ F2	rate	Δ F2	rate
Rising	ia	834.2	14.4	1018.2	18.3
	iɔ	1093.8	21.3	1364.7	25.5
	iɛ	373.2	7.2	724.1	12.2
	ua	462.8	7.1	638.2	9.4
	uɛ	931.9	14.8	1342.8	19.6
	uo	228.1	4.9	346.5	6.5
	ɥa	179	3.5	255	4.6
	ɥɛ	209.5	4.4	310	6.2
	ɥo	145.5	3.4	258.4	5.1
falling	ou	51.6	1.7	71.7	2.7
	ei	212.8	3.9	263.7	4.5

4. Conclusions

This paper gives an outline of vowel phonology in Hangzhou on the basis of an acoustic phonetic description of vowel production. And the results are of theoretical interest to vowel phonetics and phonology in general.

Hangzhou monophthongs in CV syllables have a triangular distribution in acoustic F1/F2 vowel planes, and as shown by the 2-sigma ellipse, each vowel has a spectral region that serves as a static acoustic target. Less vowel oppositions are found in CVN syllables, and only a two-way vowel contrast [a o] is detected in CVC syllables.

Falling diphthongs are grouped with monophthongs, as they have a dynamic spectral target ([19]). The two falling diphthongs [ei ou] in Hangzhou both have a small range of spectral change, and occupy a mid-high position in the acoustic F1/F2 vowel plane. And evidence from er-suffixation supports to conclude that in addition to static spectral targets, dynamic properties play an important role in distinguishing falling diphthongs from neighboring monophthongal vowels, namely [i ei ɛ] and [u ou ɔ] respectively.

Rising diphthongs are totally different from falling diphthongs in terms of spectral properties. Rising diphthongs have two spectral targets as their monophthongal counterparts do, and they are vowel sequences.

5. Acknowledgements

This work is supported by Chinese National Social Science Foundation (Project number: 15BYY073; P.I.: Dr. HU Fang).

6. References

[1] Simmons, R. V. *The Hangzhou Dialect*. Ph.D. dissertation, The University of Washington, Seattle, 1992.

[2] Li, R. and Bao, S. *The Hangzhou dialect Dictionary* [in Chinese]. Jiangsu: Jiangsu Education Publishing House. 1998.

[3] Qian, N. *Studies in the contemporary Wu dialects* [in Chinese]. Shanghai: Shanghai Education Publishing house. 1992.

[4] Wang, Q. "The Hangzhou phonology" [in Chinese]. *Journal of Tsinghua University (Philosophy and Social Sciences)* (01), 57-67. 1999.

[5] Jin, J. "Perception studies in Hangzhou Dialect Rising Tones" [in Chinese]. *Journal of Sun Yat-Sen University (Social Science Edition)*, 55(1), 70-82. 1992.

[6] Li, J. "Experimental Study on Hangzhou dialect tone and tone sandhi" [in Chinese]. *Journal of Chinese Phonetics*, 2, 108-112. 2009.

[7] Hu, W., Hu, F. & Jin, J. "Diphthongization of the Nuclear Vowel and the Emergence of the Tetraphthong in the Hetang Cantonese", *Proceedings of Interspeech 2016*, 983-987, San Francisco, USA, 2016.

[8] Hartman, L. M. "The segmental phonemes of the Peiping dialect", *Language*, 20: 28-42, 1944.

[9] Luo, C. and Wang, J. *An Outline of General Phonetics* [in Chinese], Beijing: China Science Press, 1957.

[10] Cheng, R. L. "Mandarin phonological structure", *Journal of Linguistics*, 2, 135-158, 1966.

[11] Cheng, C. C. "A synchronic phonology of Mandarin Chinese", *Monograph on Linguistic Analysis*, No. 4, The Hague: Mouton, 1973.

[12] Chao, Y.-R. *A grammar of spoken Chinese*, Berkeley: University of California Press, 1968.

[13] You, R., Qian, N. and Gao, Z. "On the phonemic system of Putonghua" [in Chinese], *Chinese Language*, 5, 328-334, 1980.

[14] Duanmu, S. *The Phonology of Standard Chinese* (2nd ed.), Oxford: Oxford University Press, 2007.

[15] Duanmu, S. *Syllable Structure: The Limits of Variation*. Oxford University Press, 2008.

[16] Duanmu, S. "Syllable and syllable structure in Chinese", in *Encyclopedia of Chinese Language and Linguistics*, ed. by R. Sybesma (editor-in-chief), W. Behr, Y. Gu, Z. Handel, C.-T. J. Huang, and J. Myers, Leiden: Brill, 2016.

[17] Chao, Y.-R. "The non-uniqueness of phonemic solutions of phonetic systems", *Bulletin of the Institute of History and Philology, Academia Sinica*, 4, 363-397, 1934.

[18] Chao, Y.-R. *Studies in the Modern Wu Dialects*, Peking: Tsinghua University Research Institute Monograph, 4, 1928.

[19] Hu, F. "Falling diphthongs have two targets: Acoustics and Articulation of the diphthong production in Ningbo Chinese", *Journal of the Acoustical Society of America*, Vol. 134, No. 5, Pt. 2, 4199, Nov 2013; full paper in Chinese in *Bulletin of Linguistic Studies*, 10: 12-37. Shanghai: Shanghai Lexicography Press, 2013.

[20] Qiu, Y. and Hu F. "Longchang vowels" [in Chinese], Bulletin of Linguistic Studies, 10: 38-51, Shanghai: Shanghai Lexicography Press, 2013.

[21] Xia, L. and Hu, F. "Vowels and diphthongs in the Taiyuan Jin Chinese dialect", *Proceedings of Interspeech 2016*, 993-997, San Francisco, 2016.

[22] Hu, F. and Zhang, M. "Diphthongized vowels in the Yi county Hui Chinese dialect", *Proceedings of Interspeech 2014*, 1703-1707, Singapore, 2014.

[23] Hu, F. and Zhang, M. "On the diphthongized vowel in Qimen Hui Chinese dialect", *Proceedings of the 18th International Congress of Phonetic Science*, paper number 0100, Glasgow, U.K.: the University of Glasgow, 2015.

[24] Hu, F. and Ge, C. "Vowels and diphthongs in Cangnan Southern Min Chinese dialect", *Proceedings of Interspeech 2016*, 978-982, San Francisco, 2016.

[25] Karlgren, B. *Études sur la phonologie chinoise*. Archives d' Études Orientales, Vol. 15, Leyde et Stockholm, 1915-26.

[26] Boersma, P & Weenink, D. "Praat: doing phonetics by computer" [Computer program]. Version 6.0.19, 2016.

[27] Smith, J. O. and Abel, J. S. "Bark and ERB Bilinear Transforms", *IEEE Trans. Speech and Audio Proc.*, 7(6): 697-708, 1999.

[28] Hu, F. "On the distinctive features for the high front vowels in Ningbo and Suzhou Wu Chinese – with reference to sound changes of high vowels" [in Chinese], *Zhongguo Yuwen (Chinese Language)*, 320 (5), 455-465, 2007.

[29] Xu, Y. *Phonetic studies in Zhejiang north Hangjiahu dialect* [in Chinese]. Beijing: China Social Science Press. 1: 40-41, 2007.

[30] Chao, Yuen-ren. *Mandarin Primer*. Cambridge, MA: Harvard University Press, 1948.

[31] Gay, T. "Effect of speaking rate on diphthong formant movements", *Journal of the Acoustical Society of America*, 44: 1570-3, 1968.

[32] Gay, T. "A perceptual study of American English diphthongs", *Language and speech*, 13: 65-88, 1970.

[33] Manrique, A. "Acoustic analysis of American English diphthongs", *Language and Speech*, 13: 223-230, 1986.

[34] Jha, S. K. "Acoustic analysis of diphthongs", *Journal of phonetics*, 13: 223-230, 1986.

[35] Bladon, A. "Diphthongs: A case study of dynamic auditory proceeding", *Speech Communication*, 4: 145-154, 1985.

[36] Hu, F. An acoustic analysis of diphthongs in Ningbo Chinese", *Proceedings of Eurospeech 2003*, pp. 801-804, Geneva, Switzerland, 2003.