



Diphthongized Vowels in the Yi County Hui Chinese Dialect

Fang Hu¹, Minghui Zhang²

¹Institute of Linguistics, Chinese Academy of Social Sciences

²Department of Linguistics, Graduate School, Chinese Academy of Social Sciences

hufang@cass.org.cn, 1136781591@qq.com

Abstract

This paper is an acoustic phonetic description of the vowels and diphthongs in the Yi county dialect and compares the diphthongized vowels with monophthongs and diphthongs in terms of their temporal structures and spectral characteristics. Results show that rising diphthongs have two targets while falling diphthongs only have one dynamic target. Diphthongized vowels occur as an intermediate category between monophthongs and diphthongs. Diphthongized vowels have distinctive onsets but a neutralized offset.

Index Terms: monophthongs, diphthongs, diphthongized vowels, Hui Chinese dialects, the Yi county dialect.

1. Introduction

Vowels are traditionally classified as monophthongs, diphthongs, triphthongs, and in some rare cases, tetraphthongs. Monophthongal and diphthongal vowels are commonly detected in the world's languages. However, there is a long debate on the definition of the diphthong in the literature. Some phoneticians view diphthongs as a single vowel with phonetically complex nucleus ([1], [2], [3]), while others view diphthongs as a sequence of two vowels or a combination of one vowel and one semivowel ([4], [5]). Even for English, for instance, there is controversy regarding what a diphthong is ([6], [7], [8]).

In the literature regarding Chinese vowels and diphthongs, [9] is a seminal work, where the author points out that only falling diphthongs are true diphthongs in Wu dialects, and rising diphthongs are not true diphthongs (pp. 65-66). [10] renewed the dichotomy between falling and rising diphthongs. On the basis of acoustic and lingual kinematic data from the Ningbo Wu dialect, it is argued that falling diphthongs have one dynamic target, while rising diphthongs are composed of two targets. In other words, the diphthong [ai], for instance, is not a sequence of [a] and [i], but a single dynamic event phonetically, and should be treated as being phonologically distinctive to the monophthongal vowel [a]; by contrast, the diphthong [ia] is simply a sequence of [i] and [a]. And the acoustic results from the Longchang dialect, a southwestern Mandarin dialect, also support the dichotomy between falling and rising diphthongs ([11]).

The diphthongized vowels in the Yi county dialect, a Hui Chinese dialect, occur as an intermediate category between the monophthong and diphthong dichotomy. In the Yi county dialect, there are 3 monophthongal vowels [ɿ u a], 2 rising diphthongs [ia ua], 3 falling diphthongs [ai au au], 1 level diphthong [iu], and 6 diphthongized vowels [i:ɐ u:ɐ y:ɐ ε:ɐ o:ɐ ɻ:ɐ]. The vowel inventory of the Yi county dialect is typologically uncommon in that (1) it only has 3 monophthongs, (2) it does not have [i], and (3) it has 6 diphthongized vowels. Following the tradition in the dialectological literature, the diphthongized vowels in Hui dialects were transcribed as a long vowel element followed by

the insertion of a coda-like vowel element, usually a schwa or other neutralized vowel element ([12], [13]). The long vowel symbol is to emphasize that the first element is the nucleus of the syllable; otherwise, the first element would be interpreted as a glide as in old literature ([14]).

This paper is an acoustic phonetic description of the vowels and diphthongs in the Yi county dialect. The production of the diphthongized vowels was compared with monophthongs and diphthongs in terms of their temporal structures and spectral characteristics. It is argued that the diphthongized vowels should phonologically be grouped with monophthongs, although they are phonetically diphthong-like. That is, the diphthongized vowels are an intermediate category between monophthongs and diphthongs.

2. Methodology

10 speakers, 5 male and 5 female, were recorded during the fieldwork trip in 2013 summer. They were all native adult speakers without any reported history of speech and hearing disorders.

Meaningful monosyllabic words were used as test words. Normally each target vowel has 2 test words, one is a zero-initial syllable V and the other is a labial stop-initial syllable CV. The Yi county dialect has 6 tones, and the 2 of them, *yang-qu* and *ru*, are phonetically short. Therefore, the long vs. short difference was also sampled in this study. Each test word was placed in a carrier sentence [X, ŋa³¹ t^hu³ X pi³ n³¹ t^hε^ɐ31] "X, I read X for you to listen". 5 repetitions were recorded.

The audio sound was recorded directly into a laptop PC through a TerraTec DMX 6Fire USB sound card with a SHURE SM86 microphone. The sampling rate is 11,025 Hz.

The target vowels were labeled in praat 5.3.48 ([15]). The first element, transition, and second element were annotated for each diphthongized vowel or diphthong. The lowest four formants were extracted in the mid point of each target vowel or vowel element. Additionally, F2 range and rate of change were also calculated for each diphthongized vowel or diphthong.

3. Results

3.1. Diphthongized vowels and the monophthongs

As mentioned above, there are 3 monophthongs [ɿ u a] and 6 diphthongized vowels [i:ɐ u:ɐ y:ɐ ε:ɐ o:ɐ ɻ:ɐ] in the Yi county dialect. Figure 1 shows the distribution of the first two formants for the 3 monophthongs and the 5 diphthongized vowels (the first element) from 5 male (left) and 5 female (right) speakers respectively in a two-dimensional acoustic space (F1 against F2) with the origin of the axes to the top right of the plot. The ordinates are bark scaled, but the values along the axes are labeled in Hertz. Each 2-sigma ellipse was based on 25 data points (5 repetitions × speakers; the citation position: the first X position in the carrier sentence). The first

element is about 17 ms longer than the offset element in male speakers, whereas the onset element is about 12 ms shorter than the offset element in female speakers.

In the diphthongized vowels [i:ɐ u:ɐ y:ɐ ɛ:ɐ o:ɐ ɤ:ɐ], the first element is overwhelmingly longer than the transition and second element. The first element is about 177 ms in both male and female speakers; the transition is about 85 ms in male speakers and 118 ms in female speakers; the second element is about 93 ms in male speakers and 112 ms in female speakers.

As summarized further in the bar charts in Figure 3, diphthongized vowels and different types of diphthongs differ in temporal organization. In diphthongized vowels, the first element is longest component; in rising diphthongs, the offset element is longest component; and falling and level diphthongs have a comparatively balanced temporal structure.

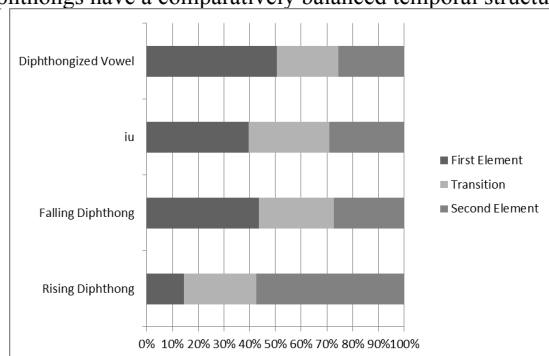


Figure 3: Temporal structure of rising diphthongs, falling diphthongs, level diphthong, and diphthongized vowels in the Yi county dialect: normalized time.

3.2.2. Spectral properties

Diphthongs could either be a dynamic articulatory event or a sequence of two events ([10]). It is thus of interest to examine if the two elements of diphthongs or diphthongized vowels have a steady spectral target, and if yes, what the nature of the target is. As mentioned earlier, the Yi county dialect only has 3 monophthongs, including the apical vowel [ɿ]. And as shown in Figure 1, the Yi county vowel system is a typical triangle pattern, if the diphthongized vowels are taken into account. In this study, monophthongs and the first elements of diphthongized vowels are both used as reference vowels, which whom the diphthong elements are compared.

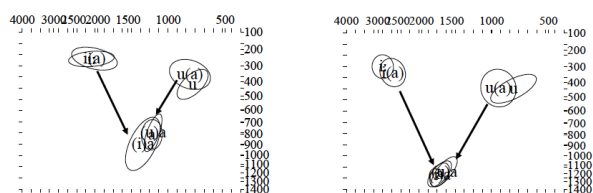


Figure 4: Ellipses for diphthong elements and the corresponding vowels in the Yi county dialect: male speakers (left) and female speakers (right). x-axis: F2; y-axis: F1.

First, examine the two rising diphthongs [ia ua]. Figure 4 compares the distribution of rising diphthong elements and the corresponding monophthongs [u a] and the first element of diphthongized vowel [i:] in the acoustic F1/F2 plane. From the figure we can see that although there are certain coarticulatory

effects, the ellipses of diphthong elements [i u a] heavily overlap with vowel elements [i: u a] respectively. This suggests both the onset and offset elements of rising diphthongs have steady frequency regions, i.e. acoustic targets.

Second, examine the three falling diphthongs [ai au au:]. Figure 5 shows the distribution of falling diphthong elements in the acoustic F1/F2 plane composed by the three reference vowel elements [i: u a]. Note that in the figure the diphthongs [ai au au:] were labeled as [ɛi au eu] respectively. From the figure we can see that the ellipses of the three onset elements [a]s in [ai au au:] heavily overlap with that of the monophthong [a] in both male and female speakers. And again, the onset elements are subject to certain coarticulation so that [a] in [ai] has a greater F2 value and [a] in [au] has a smaller F2 value than the monophthong [a].

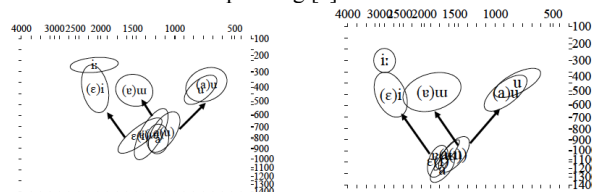


Figure 5: Ellipses for diphthong elements and the corresponding vowels in the Yi county dialect: male speakers (left) and female speakers (right). x-axis: F2; y-axis: F1.

Regarding the offset elements in falling diphthongs, the diphthong element [i] does not reach their target positions but ends up at the mid-high position in the acoustic F1/F2 plane; the diphthong element [u] does not have a comparable monophthong vowel, but it also ends up at the mid-high position. The ellipses of diphthong element and monophthong [u] overlap with each other; however, the diphthong element [u] stays at the mid-high position, and it is lower than the monophthong [u] in female speakers. In male speakers, both the diphthong element and monophthong [u] are located in a mid-high position. Another thing that should be noted is that, as evidenced by the size of ellipses, the offset diphthong elements in falling diphthongs are more variable than their corresponding vowel elements.

In summary, falling diphthongs have steady spectral specifications only on onset elements, which signifies an acoustic target, but not on offset elements. The production of falling diphthongs [ai au au:] all begin from a similar [a]-like spectral specification but end up with different offset positions, namely [+front] for [ai], [+back] for [au], and [-front, -back] for [au:], due to their different dynamic settings. In other words, [ai au au:] each has a distinctive dynamic target. Phonologically speaking, in addition to the static phoneme /a/, the monophthong, there are three more dynamic /a/-phonemes.

Third, the level diphthong [iu] is basically in parallel with rising diphthongs. As shown in Figure 6, the production of level diphthong [iu] generally begins from a formant setting similar to the vowel element [i:] and ends up at a frequency region similar to the vowel element [u]. That is, both onset and offset elements of [iu] seem to have acoustic targets.

Finally, examine the diphthongized vowels themselves. The distribution of the onset and offset elements of the diphthongized vowels [i:ɐ u:ɐ y:ɐ] was shown in Figure 7, and that of the diphthongized vowels [ɛ:ɐ o:ɐ y:ɐ] was shown in Figure 8. It's clear from the figures that the offset elements of diphthongized vowels are all located in a low vowel position,

although they are subject to apparent coarticulatory effect brought out by their onset elements as well. Also, it is observed that the offset elements are generally more variable than the onset elements in diphthongized elements. It seems that the production of diphthongized vowels tends to be released with a neutralized tongue position, which results in a somehow centralized low vowel specification, rather than having a similar acoustic offset target. Phonologically, the distinction of diphthongized vowels is attributable to the onset elements only, and the neutralized offset release is a redundant phonetic characteristic. The only exception is [u:ɐ], which contrasts with the monophthong [u].

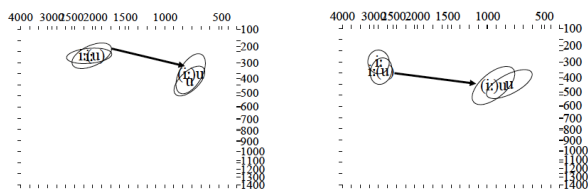


Figure 6: Ellipses for diphthong elements and the corresponding vowels in the Yi county dialect: male speakers (left) and female speakers (right). x-axis: F2; y-axis: F1.

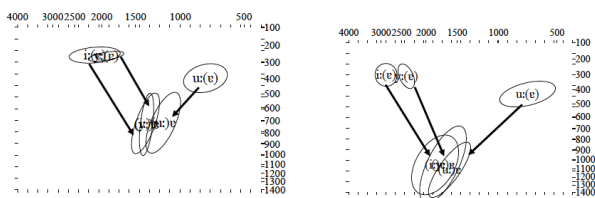


Figure 7: Ellipses for the onset and offset elements of the diphthongized vowels [i:ɐ, u:ɐ, y:ɐ] in the Yi county dialect: male speakers (left) and female speakers (right). x-axis: F2; y-axis: F1.

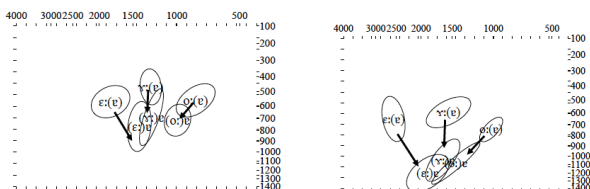


Figure 8: Ellipses for the onset and offset elements of the diphthongized vowels [ɛ:ɐ, o:ɐ, y:ɐ] in the Yi county dialect: male speakers (left) and female speakers (right). x-axis: F2; y-axis: F1.

3.2.3. Dynamic aspects

Diphthongs are composed of not only onset and offset elements but also the movement connecting them. Range and rate of formant change, especially the second formant (F2), is a useful measure to characterize the dynamic aspects of diphthong production ([19], [20], [21], [22]). Table 1 summarized mean F2 range and rate of change for diphthongs and diphthongized vowels in the Yi county dialect.

Table 1. Range ($\Delta F2$ in Hz) and rate (in Hz/ms) of F2 change for the diphthongs and diphthongized vowels in the Yi county dialect.

	male		female	
	$\Delta F2$	Rate	$\Delta F2$	Rate
ua	440	4.2	699	6.2
ia	772	9.6	869	9.1
au	218	2.5	252	2.2
au	344	3.6	535	4.5
ai	622	7.6	957	7.6
ɤ:ɐ	29	0.4	53	0.5
o:ɐ	174	2.6	669	4.3
u:ɐ	395	4.6	450	4.8
ɛ:ɐ	336	4.7	688	6.0
y:ɐ	551	5.4	804	6.5
i:ɐ	715	8.7	1034	7.2
iu	1326	10.8	1866	11.4

First, the two rising diphthongs have different ranges and rates of F2 change. The diphthong [ua] has a substantially smaller range and rate of F2 change than [ia].

Second, the three falling diphthongs have different ranges and rates of F2 change, too, in the order of [ai] > [au] > [au]. The results suggest that the production of falling diphthongs is associated with their own dynamic settings.

Third, it's difficult to distinguish diphthongized vowels from one another in terms of F2 range or rate of change, especially among [u:ɐ, ɛ:ɐ, y:ɐ]. This could further serve as evidence that dynamic properties are not important to characterize diphthongized vowels. Rather, they are basically in parallel with monophthongs.

Fourth, the level diphthong [iu] has the greatest F2 range and rate of change.

Finally, suffice it to conclude that it is not of great interest to compare dynamic aspects between different types of diphthongs, especially between the diphthongs with two targets and those with one dynamic target.

4. Conclusions

This paper gives an acoustic phonetic description of the vowels and diphthongs in the Yi county Hui Chiense dialect. Rising diphthongs have two targets, whereas falling diphthongs only have one dynamic target. The level diphthong resembles the rising diphthong in terms of formant pattern but resembles the falling diphthong in terms of temporal organization. Results suggest that diphthongized vowels are basically in parallel with the monophthong vowels. Each diphthongized vowel has a distinctive onset formant pattern, whereas it is released towards a neutralized formant frequency region. And the fact that the diphthongized vowels occur as an intermediate category between monophthongs and diphthongs provide us a better understanding of the phonetics and phonology of vowels and diphthongs in general.

5. Acknowledgements

This work is supported by CASS Innovation Project "Language universals and diversities: an experimental approach to the phonetics and grammar in Chinese dialects" (P.I.: Dr. HU Fang).

6. References

- [1] Malmberg, B. *Structural linguistics and human communication*, Berlin: Springer-Verlag, 1963.
- [2] Abercrombie, D. *Elements of general phonetics*, Edinburgh: Edinburgh University Press, 1967.
- [3] Catford, I. *Fundamental problems in phonetics*, Edinburgh: Edinburgh University Press, 1977.
- [4] Sweet, H. *A handbook of phonetics including a popular exposition of the principles of spelling reform*, Oxford: Clarendon Press, 1877.
- [5] Jones, D. *Outline of English phonetics (2nd Ed.)*, New York: E. P. Dutton, 1922.
- [6] Pike, K. L. "On the phonemic status of English diphthongs", *Language*, 23: 151-159, 1947.
- [7] Lehiste, I. & Peterson, G. E. "Transitions, glides, and diphthongs", *Journal of the Acoustical Society of America*, 33: 268-277, 1961.
- [8] Holbrook, A. and Fairbanks, G. "Diphthong formants and their movements", *Journal of Speech and Hearing Research*, 5: 38-58, 1962.
- [9] Chao, Y. R. *Studies in the Modern Wu Dialects*, Peking: Tsinghua University Research Institute Monograph, 4, 1928.
- [10] Hu, F. "Falling diphthongs have a dynamic target while rising 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: *Collected Papers in Language Studies*, 10: 12-37, Shanghai: Shanghai Lexicography Press, 2013.
- [11] Qiu, Y. and Hu, F. "Longchang vowels" [in Chinese], *Collected Papers in Language Studies*, 10: 38-51, Shanghai: Shanghai Lexicography Press, 2013.
- [12] Hirata, S. (ed.) *Studies in Hui dialects* [in Chinese], pp. 20, Tokyo: Kohbun, 1998.
- [13] Xie, L. and Shen, M. *The Yi county (Hongcun) dialect* [in Chinese], pp. 2-3, Beijing: China Social Sciences Press, 2008.
- [14] Wei, J. "Fieldwork on the Yi county dialect" [in Chinese], *Sinology Quarterly*, pp. 35-85, Beijing: Peking University Press.
- [15] Boersma, P. and Weenink, D. (2014). *Praat: doing phonetics by computer* [Computer program]. Version 5.3.68, retrieved 20 March 2014 from <http://www.praat.org/>.
- [16] Smith, J. O. and Abel, J. S., "Bark and ERB Bilinear Transforms", *IEEE Trans. Speech and Audio Proc.*, 7(6):697-708, 1999.
- [17] Soquet, A., Saerens, M. and Jospa, P., "Acoustic-articulatory inversion", in T. Kohonen [Ed], *Artificial Neural Networks*, 371-376, Elsevier, 1991.
- [18] Stone, H.S., "On the uniqueness of the convolution theorem for the Fourier transform", NEC Labs. Amer. Princeton, NJ. Online: <http://citeseer.ist.psu.edu/176038.html>, accessed on 19 Mar 2008.
- [19] Gay, T. "Effects of speaking rate on diphthong formant movements", *Journal of the Acoustical Society of America*, 44: 1570-3, 1968.
- [20] Manrique, A. "Acoustic analysis of the Spanish diphthongs", *Phonetica*, 36: 194-206, 1979.
- [21] Jha, S.K. "Acoustic analysis of the Maithili diphthongs", *Journal of Phonetics*, 13: 107-115, 1985.
- [22] Hu, F. "An acoustic analysis of diphthongs in Ningbo Chinese", In: *Proceedings of Eurospeech 2003*, pp. 801-804. Geneva, Switzerland, 2003.