On the Diphthongized vowels in Qimen Hui Chinese

[1]Fang Hu, [2]Minghui Zhang
[1]Institute of Linguistics, Chinese Academy of Social Sciences
[2]Department of Linguistics, Graduate School, Chinese Academy of Social Sciences
[1]hufang@cass.org.cn, [2]goldenzmh@hotmail.com

ABSTRACT

This paper gives an acoustic phonetic description of the diphthongized vowels in the Qimen Hui dialect. The results suggest that the diphthongized vowels in Qimen have similar acoustic phonetic properties with those in Yi County Hui, namely they belong to an intermediate category between monophthong and diphthong. However, the diphthongized vowels in Qimen are typologically closer to plain diphthongs in the continuum of monophthong and diphthong.

Keywords: diphthongized vowel, monophthong, diphthong, Qimen Hui Chinese dialect.

1. INTRODUCTION

Monophthongs and diphthongs are the most common vowel categories in the world's languages. Phonetically, the former refers to a pure vocalic sound whose articulation is relatively fixed in the total duration; the later refers to a vocalic sound starting at one articulatory position gliding up or down towards another position. However, it has been controversial whether the diphthongs should be viewed as a sequence of two vowels ([1], [2]), or defined as a single dynamic articulation event with complex nucleus ([3], [4], [5]). Diphthongs are traditionally subdivided into two main categories: rising diphthongs and falling diphthongs ([6]). It is reported that falling diphthongs are easier to alternate with monophthongal vowels than rising diphthongs ([7]). Chao points out that in Wu Chinese dialects, only falling diphthongs are true diphthongs ([8]). On the basis of the acoustic as well as articulatory data from Chinese dialects, recent studies argue that rising diphthongs are composed of two targets while falling diphthongs have only one dynamic target ([9], [10]).

In addition to monophthongs and diphthongs, [11] reports that diphthongized vowels occur as an intermediate category in the Yi county Hui dialect. The diphthongized vowels are described as long glides or vocalic elements being followed by short open vowels in dialectological works ([12], [13], [14]). Meanwhile it is admitted that Hui dialects do not contrast in long versus short vowels ([15]).

Diphthongized vowels are phonetically similar to plain diphthongs in that they both have a changing vowel quality in production. Diphthongization is a common phonetic and/or phonological process, both synchronically and diachronically ([16]). And the emergence of diphthongized vowels seems to be a shared innovation among a number of Hui dialects in their historical development.

This paper gives an acoustic phonetic description of the diphthongized vowels as well as monophthongs and plain diphthongs in Qimen dialect. The vowel inventory in Qimen is shown in table 1. Monophthongs, rising diphthongs, and diphthongized vowels occur in open CV syllables as well as CVN syllables, i.e. those having a nasalized rhyme or a nasal ending; the falling and level diphthongs only occur in CV syllables.

Table 1: The vowel inventory of Qimen dialect

		CV	CVN
Monophthong		ɿ i u y a e o ɚ	æ̃ ã õ æn ɐŋ
Diphthong	Rising	ie ia io ua	iã iõ uæ̃ uã iæn iɐŋ yæn
	Falling	ei	/
	Level	ui	/
Diphthongized vowel		iːɐ uːɐ yːɐ	ĩːɐ ũːɐ ỹːɐ

2. METHODOLOGY

10 native adult speakers with a balanced gender ratio provided the speech data and none of them were reported speech disorders or hearing impairments.

Two meaningful monosyllabic words were used as test words for each target vowel in the vowel inventory. Each test word was placed in the carrier sentence [X, ɳ^{22} tʰu^{33} X fæ̃^{33} ʂa^{41} tæ̃^{23}] (X, you read X for me to listen) and 5 repetitions were recorded. The audio sound was recorded into a laptop PC through a TerraTec DMX 6Fire USB sound card with a SHURE SM86 microphone. The sampling rate is 11,025 Hz.

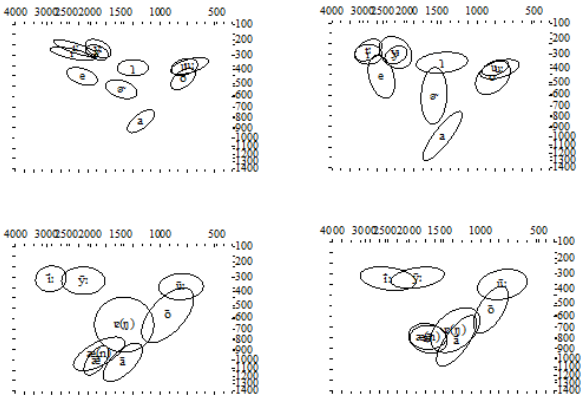
Sampled data were annotated in praat 5.3.48 [17]. The target vocalic element was annotated for each monophthong. And three vocalic elements, namely onset, transition and offset, were annotated for each diphthong or diphthongized vowel. The lowest four formants were extracted in the mid-point of the target vocalic element. The temporal structure was calculated for each diphthong and diphthongized vowel. And the range and rate of the second formant (F2) change were also measured.

3. RESULTS

3.1. Diphthongized vowels and monophthongs

Figure 1 plots the first two formants of the monophthongs and the onset element of diphthongized vowels in CV syllables (upper) and CVN syllables (lower) from 5 male speakers (left) and 5 female speakers (right). The acoustic vowel space was determined by the first formant (F1: y-axis) and the second formant (F2: x-axis) with the origin of the axes to the top right. The ordinates are Bark scaled ([18]), but the values along the axes are still labeled in Hertz. Each two-sigma ellipse, which intuitively manifests the vowel distribution, is based on 25 sampled data points.

Figure 1: Vowel ellipses for the monophthongs and the onset elements of diphthongized vowels CV syllables (upper) and CVN syllables (lower): male (left) and female speakers (right).



In CV syllables, the 8 monophthongs have a triangular distribution and can be grouped into 3 levels of vowel height and 3 levels of vowel backness. [i u y ɿ] are high vowels; [e ɚ o] are mid vowels; [a] is a low vowel which doesn't contrast in backness. [i y e] are front vowels; [u o] are back vowels; [ɿ ɚ a] are central vowels. It can be seen from the figure that the onset elements of diphthongized vowels [iːɐ uːɐ yːɐ] heavily overlap with the corresponding monophthongal counterparts [i u y], respectively. This means that the onset elements of diphthongized vowels themselves do not contrast with monophthongal vowels. Rather, the dynamic aspects of diphthongized vowels contribute to the distinctiveness.

The diphthongized vowels [ĩːɐ ũːɐ ỹːɐ] also occur in CVN syllables. As shown in the lower panel of Figure 1, in general, lesser vowel contrasts are detected in CVN syllables than in CV syllables. First, there is no apical vowel [ɿ] and monophthongal [i u y e]. Second, the mid central vowel is [ɐ], lower than [ɚ] in CV syllables. Last but most interestingly, the vowel space is quadrangular because the low vowels [æ] in [æ̃ æn] and [a] in [ã] contrast in backness.

3.2. Diphthongized vowels and diphthongs

Diphthongized vowels and diphthongs are structurally similar in that they are composed of two elements and a transition; but they differ in temporal organization, spectral properties, and dynamic aspects. Due to the space limit, discussion focuses on the 6 diphthongized vowels and diphthongs in CV syllables.

3.2.1. Temporal organization

Figure 2 gives mean durations of the onset, transition and offset for the 6 diphthongized vowels and diphthongs in millisecond (upper) and percentage (lower).

Durations of the diphthongized vowels and diphthongs range from 360 ms to 460 ms, and [io] is the longest both in male and female speakers. In diphthongized vowels, the onset elements are about 200 ms in both male and female speakers; the transitions are about 92 ms in males and 113 ms in females; the offset elements are about 100 ms in males and 80 ms in females. In rising diphthongs, the onset elements are quite short: about 50 ms in males and 43 ms in females; the transitions are about 90 ms in males and 110 ms in females; the offset elements are the longest: 260 ms in males and 250 ms in females. In the falling diphthong [ei], the onset elememnt is 78 ms in males and 63 ms in females; the transition is about 85 ms both in males and females; the offset element is the longest: 234 ms in males and 222 ms in females. In the level diphthong [ui], the onset element is about 53 ms in both males and females; the transition is about 113 ms in males and 125 ms in females; the offset is the longest: 220 ms in males and 246 ms in females.

In summary, different diphthongal categories have different temporal organization in general, which is better demonstrated by the percentage data in the lower panel of Figure 2. And Figure 3 gives the pooled mean data for different categories.

First, the diphthongized vowel differs from the other categories in that the onset element has an overwhelmingly longest duration in diphthongized vowels, about 51% of the total duration, whereas in other categories, the offset element is the longest, about 60% of the total duration. Among diphthongs, the falling diphthong has the longest duration for the onset element, 18% of the total duration, the rising diphthongs have the shortest duration, 11% of the tonal duration, and the level diphthong occurs in-between, 13% of the total duration. The results are consistent with the Yi county data in that the onset element is the dominating segment in a diphthongized vowel ([10]).

Figure 2: Temporal organizations for the diphthongized vowels and diphthongs in millisecond (upper) and percentage (lower): male (left) and female speakers (right).

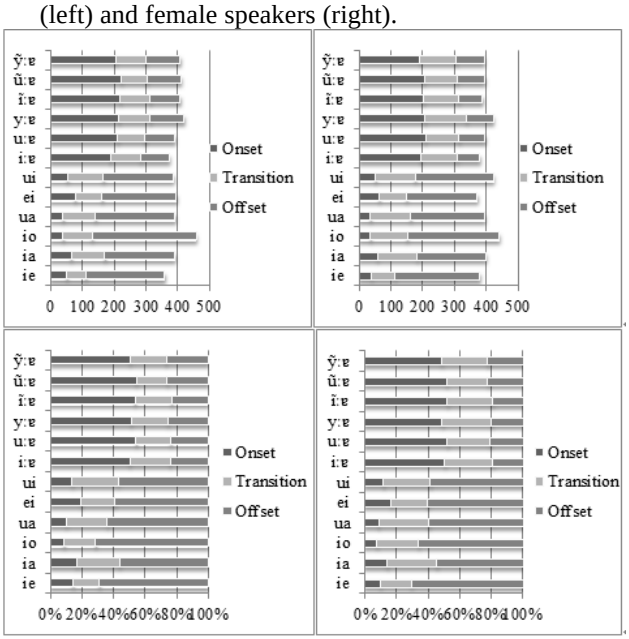
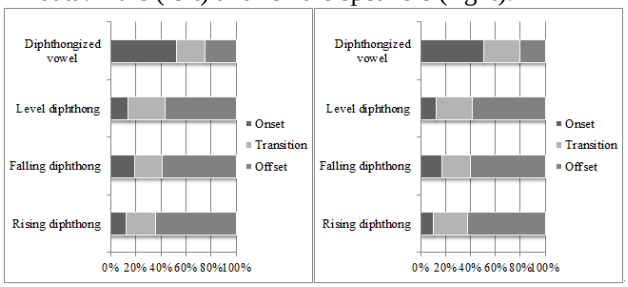


Figure 3: Temporal organizations for the pooled data: male (left) and female speakers (right).



3.2.2. Spectral properties

It is of interest to examine whether the onset and onset elements have steady spectral targets in different diphthongized vowels and diphthongs. And the corresponding monophthongs serve as references for comparison.

Figure 4: Vowel ellipses for the onset and offset elements of diphthongized vowels [iːɐ uːɐ yːɐ] (upper) and [ĩːɐ ũːɐ ỹːɐ] (lower) and their corresponding monophthongs: male (left) and female speakers (right).

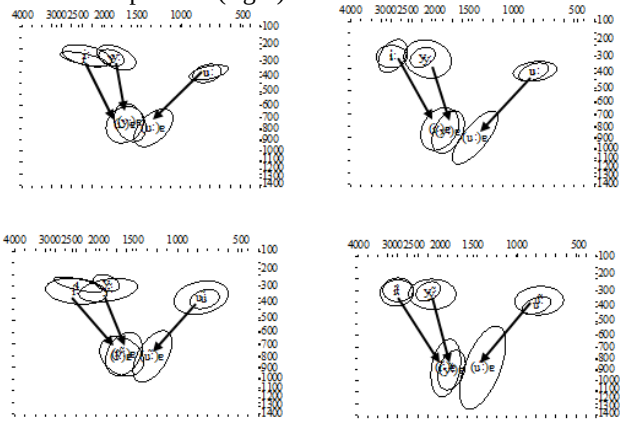


Figure 5: Vowel ellipses for the onset and offset elements of rising diphthongs [ia ua] (upper) and [ie io] (lower) and their corresponding monophthongs: male (left) and female speakers (right).

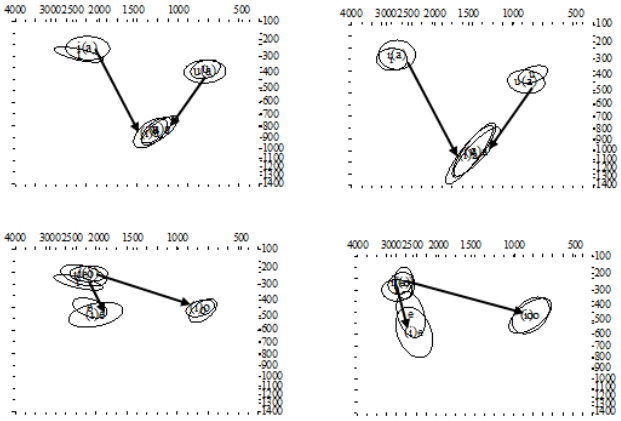
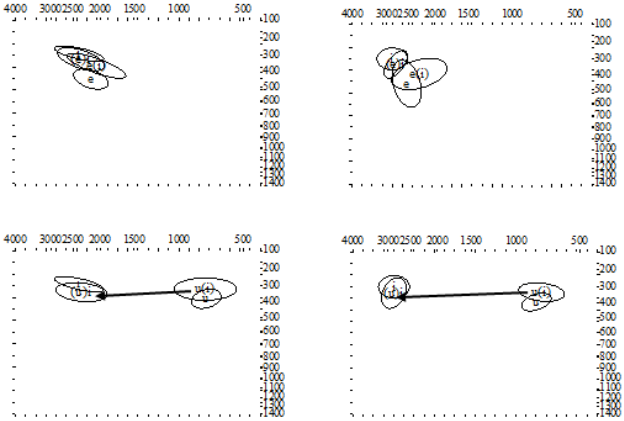


Figure 6: Vowel ellipses for the onset and offset elements of the falling diphthong [ei] (upper) and the level diphthong [ui] (lower) and their corresponding monophthongs: male (left) and female speakers (right).



As shown in Figure 4, the ellipses for the onset elements in diphthongized vowels intensively overlap with those of corresponding monophthongs; the ellipses for the offset elements in [iːɐ yːɐ] overlap with each other while the ellipse for the offset element in [uːɐ] has a separate distribution. And these are also true for the nasalized [ĩːɐ ũːɐ ỹːɐ]. It seems that the production of diphthongized vowels begins from a distinctive spectral specification and ends up with a somehow neutralized spectral region towards an [ɐ]-like configuration.

Figure 5 compares the diphthong elements of [ia ua ie io] with their corresponding monophthongs in acoustic vowel planes. The onset elements in rising diphthongs have a similar distribution with their monophthongal counterparts. The ellipses for offset elements in [ia ua io] overlap with those for their monophthongal counterparts, too. The results suggest that both onset and offset elements in a rising diphthong have specific spectral regions, i.e., targets. And even articulatory overshoot is detected in the production of the offset element of [ie], since

[e] in [ie] has a lower and posterior distribution than the corresponding monophthong [e] in the acoustic vowel plane in both male and female speakers.

Ellipses for diphthong elements in the falling diphthong [ei] and the level diphthong [ui] and their corresponding monophthongal vowels are shown in Figure 6. The [e] in [ei] has a different distribution with the monophthong [e] in the acoustic vowel plane, while [i] in [ei] and the monophthong [i] have a similar distribution. This means that the production of [i] in [ei] is better controlled than that of [e] in [ei]. In other words, the onset element [e] in [ei] behaves like a glide while the offset element [i] becomes the target nucleus in [ei]. This is quite different from the behaviors of falling diphthongs in other Chinese dialects, where the onset in a falling diphthong has a well-controlled acoustic target and the offset is so variable that it does not have a specific spectral target. In other words, the production of diphthong offset is usually constrained by the dynamic aspects ([9], [10], [11]).

The production of [ui] is in parallel with that of rising diphthongs. That is, both onset and offset elements seem to have their own acoustic targets.

3.2.3. Dynamic aspects

The range and rate of the second formant (F2) change are also of importance for characterizing diphthongs ([19], [20], [21], [22]). Table 2 summarizes the mean data of F2 range and rate of change for the production of the diphthongized vowels and diphthongs in Qimen.

Table 2: Range (ΔF2 in Hz) and rate (in Hz/ms) of F2 change for diphthongs and diphthongized vowels in Qimen dialect.

Vowel types		**Male**		**Female**	
		ΔF2	Rate	ΔF2	Rate
Rising Diph.	ie	59	1.0	166	2.3
	ua	421	4.1	520	4.1
	ia	936	8.8	1148	9.2
	io	1234	13.5	1564	13.4
Falling Diph.	ei	274	3.2	488	5.7
Level Diph.	ui	1489	13.2	1883	15.1
Diphthongized Vowel	yːɐ	206	2.1	312	2.4
	ỹːɐ	242	2.6	337	3.0
	ũːɐ	468	5.8	584	5.6
	uːɐ	544	6.4	611	5.8
	iːɐ	652	6.8	915	7.9
	ĩːɐ	798	8.4	986	9.0

First, the 4 rising diphthongs have different ranges and rates of F2 change. That is, each rising diphthong is characterized by a distinctive F2 range and rate of change. Second, the F2 range and rate of change in the falling diphthong [ei] are apparently greater than those in its rising counterpart [ie]. Third, the F2 range of change for the level diphthong [ui] is the greatest among all the diphthongs and diphthongized vowels in both male and female speakers; the F2 rate is the greatest in females, too, and is the second greatest in males.

Finally, F2 range and rate of change can characterize the three diphthongized vowels in both CV syllables [iːɐ uːɐ yːɐ] and nasalized syllables [ĩːɐ ũːɐ ỹːɐ]. However, the diphthongized vowels in CV syllables [iːɐ uːɐ yːɐ] have similar ranges and rates of F2 change with their nasalized counterparts [ĩːɐ ũːɐ ỹːɐ], respectively.

4. CONCLUSIONS

This paper gives an acoustic phonetic description of the diphthongized vowels in Qimen dialect by comparing them with monophthongs and diphthongs.

Results suggest that the diphthongized vowels in Qimen have similar acoustic phonetic properties with those in Yi County ([11]). That is, the diphthongized vowels belong to an intermediate category between monophthong and diphthong. The onset elements [i u y] in diphthongized vowels [iːɐ uːɐ yːɐ] or [ĩːɐ ũːɐ ỹːɐ] have the longest durations and steady spectral specifications, i.e. targets, respectively, whereas the offset elements are neutralized to an [ɐ]-like configuration.

What should be emphasized is that the onset elements themselves are generally distinctive and consequently the offset elements are phonologically redundant for the diphthongized vowels in Yi County. But that is not the case in Qimen. In Qimen, the diphthongized vowels [iːɐ uːɐ yːɐ] contrast with the monophthongs [i u y]. That is, in addition to the onset element, which is the dominating segment in terms of spectral properties as well as temporal organization, the dynamic aspects contribute to the characterization of diphthongized vowels, too. The diphthongized vowels in Qimen are thus typologically one-step closer to plain diphthongs in the continuum of monophthong and diphthong. And the nasalized diphthongized vowels [ĩːɐ ũːɐ ỹːɐ] add complexity to the vowel phonology of the Qimen Hui Chinese dialect.

5. ACKNOWLEDGEMENT

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] Sweet, H. 1877. *A handbook of phonetics including a popular exposition of the principles of spelling reform*. Oxford: Clarendon Press.

[2] Jones, D. 1992. *Outline of English phonetics (2nd Ed.)*. New York: E. P. Dutton.

[3] Malmberg, B. 1963. *Structural linguistics and human communication.* Berlin: Springer-Verlag.

[4] Abercrombie, D. 1967. *Elements of general phonetics*. Edinburgh: Edinburgh University Press.

[5] Catford, I. 1977. *Fundamental problems in phonetics*. Edinburgh: Edinburgh University Press.

[6] Ren, H. 1986. *On the Acoustic Structure of Diphthongal Syllables*. Ph.D. dissertation, UCLA. Published in *UCLA Working Papers in Phonetics* 65.

[7] Lipski, J. M. 1979. On vowel-diphthong transitions. *Current Issues in the Phonetic Sciences*, Vol. 9, Part I, 551-561, Amsterdam: John Benjamins.

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

[9] Hu, F. 2013. 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. Full paper in Chinese published in *Languages Studies*, 10. Shanghai: Shanghai Lexicography Press, 12-37.

[10] Qiu, Y., Hu, F. 2013. Longchang vowels. *Languages Studies*, 10. Shanghai: Shanghai Lexicography Press, 38-51.

[11] Hu, F., Zhang, M. H. 2014. Diphthongized Vowels in the Yi County Hui Chinese Dialect. *Proceedings of Interspeech 2014*, Singapore, 1703-1707.

[12] Wei, J. G. 1935. Fieldwork on the Yi county dialect. *Sinology Quarterly.* Peking: Peking University Press, 35-85.

[13] Chen, Y. 2006. *A Comparative Study on Dialects in Qimen County*. MA thesis, Guizhou University.

[14] Xie, L., Shen, M. 2008. *The Yi county (Hongcun) dialect*. Peking: China Social Sciences Press.

[15] Hirata, S. (ed.). 1998. *Studies in Hui dialects*. Tokyo: Kohbun.

[16] Endersen, H. 1972. Diphthongization. *Language*, Vol. 48, No.1, 11-50.

[17] Boersma, P. and Weenink, D. Praat: doing phonetics by computer [Computer program]. Version 5.3.68, retrieved 20 March 2014 from http://www.praat.org/.

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

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

[20] Manrique, A. 1979. Acoustic analysis of the Spanish diphthongs. *Phonetica*, 36, 194-206.

[21] Jha, S.K. 1985. Acoustic analysis of the Maithili diphthongs". *Journal of Phonetics*, 13, 107-115.

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