Acoustic Analysis of Monophthongs in Monguor Language

Zhang Hanbin
Northwest Minzu University，Lanzhou, China
zhanghb06@163.com

Keywords: duration; formants; monophthong; Monguor

Abstract. The purpose of the study is to analyze the acoustic features of Monophthongs in Monguor language. On the base of acoustic parameter database of the Monguor speech, the study discusses the differences in duration and F2 difference between short and long monophthongs in monosyllabic words and the differences in duration and F2 difference in first syllable, middle syllable(s), and last syllable for short and long monophthongs. The results show that in Monguor language, short and long vowels have the opposed relation and therefore differentiate the meanings of words.

1. Introduction

Monguor is also called Tuzu language which was first recorded in 1885 by G. N. PoTanin[1], a Russian traveler. He recorded about 650 Monguor words and discussed the relation between Monguor and Mongolian language. Antoine Mostaert and Albrecht De Smedt，Mongolian linguists and Belgian missionaries, published Monguor Phonétique in Anthropos from 1929 to 1931[2,3,4]. Japanese researcher Jiaodaozhengjia (2008) studied Monguor phonetics in Huzhu dialect[5].

Chinese researchers have made great contributions to the description of Monguor phonology. Qinggeertai induced 11 monophthongs：a, i, e, o, u, ə, aa, ii, ee, oo, uu (1950s)[6], and later 10 monophthongs：5 short monophthongs --- a, i, e, o, u, and 5 long monophthongs --- a:, i:, e:, o:, u: (1991)[7]. Zhaonasitu (1981) concluded 12 monophthongs：a, aa, i, ii, e, ee, o, oo, u, uu, ə, ɘ[8]. Li Keyu (1979) put forward 10 monophthongs：a, i, e, o, u, a:, i:, e:, o:, u:[9]. Xi Yuanlin (1986) summed up 10 monophthongs：ɑ, ɑ̆, i, ĭ, ə, ə̆, o, ŏ, u, ŭ[10]. In this study, 10 monophthongs Li Keyu proposed will be described acoustically.

2. Method

2.1 Source of Corpus.

This study built an acoustic parameter database of the Monguor speech. The Monguor words recorded came from several Monguor dictionaries and other references on Monguor. The listing 839 words were composed of monosyllabic, trisyllabic, and quadrisyllabic words.

2.2 Speech Signal Collection.

The recording equipment for this study includes a Dell Notebook, a Behringer recording microphone, and a sound card of YAMAHA Steinberg. The recording sampling rate was 44.1 kHz with 16 bits resolution ratio. The sound was saved with *.wav file. The sound was recorded in a recording studio. Each word was read three times.

2.3 Speakers.

Two speakers, one male, aged 45, and one female, aged 48, read all the words. The two speakers served as teachers in a local primary school. Their speech belongs to Huzhu dialect.

2.4 Research Questions.

The questions addressed in this study are:

(1) What are the differences in duration and formants between short and long monophthongs in monosyllabic words?

(2) Are there significant differences in duration and formants in first syllable(FS), middle syllable(s)(MS), and last syllable(LS) for short and long monophthongs?

3. Results and Discussion

3.1 Acoustic Analysis of Monguor Monophthongs.

The differences in duration and formants between short and long monophthongs in monosyllabic words are shown in the following tables.

In this study, only the second formant (F2) is measured with the unit of Hz. The F2 difference is the absolute value between the onset value and the target value of the second formant. The mean and standard deviation (Std. D) differences between /u/ and /u:/ are the highest (198-550, 196-417), while between /i/ and /i:/ the lowest (168-187, 182-195)(in Table 1).

Table 1. The results of differences in formants

	F2 difference [Hz]			
	Mean	Std. D	T	P
/a/-/a:/	137-195	145-215	-2.883	0.004
/e/-/e:/	197-177	187-201	0.622	0.535
/i/-/i:/	168-187	182-195	-0.766	0.444
/o/-/o:/	247-296	212-246	-1.453	0.148
/u/-/u:/	198-550	196-417	-5.996	0.000

The T-test results show that the F2 difference between /a/-/a:/ ($P=0.004 < 0.05$) or /u/-/u:/ ($P=0.000 < 0.05$) is significantly different, while the F2 difference between /e/-/e:/ ($P=0.535 > 0.05$), /i/-/i:/ ($P=0.444 > 0.05$) or /o/-/o:/ ($P=0.148 > 0.05$) is the same statistically.

The unit of the duration of a sound is millisecond (ms). The mean difference between /i/ and /i:/ is the highest (121-265), while the Std. D difference between /u/ and /u:/ is the highest (66-95) , as shown in Table 2. The T-test results show that the duration between every pair of short and long monophthongs is significantly different ($P=0.000 < 0.05$).

Table 2. The results of differences in duration

	Duration (ms)			
	Mean	Std. D	T	P
/a/-/a:/	113-213	55-68	-15.224	0.000
/e/-/e:/	129-237	56-80	-8.473	0.000
/i/-/i:/	121-265	65-87	-12.389	0.000
/o/-/o:/	136-222	61-67	-9.177	0.000
/u/-/u:/	102-233	66-95	-9.671	0.000

3.2 Acoustic Analysis of Monguor Monophthongs in Different Syllables.

The acoustic data (F2 and duration) in this part come from the trisyllabic and quadrisyllabic words. The F2 and duration of each pair of monophthongs will be analyzed respectively.

/a/-/a:/

According to the result of ANOVA, there is no significant difference in F2 difference among the FS, MS, and LS for /a/ and /a:/ ($P=0.152 > 0.05$) in Monguor words. However, F2 difference between the short and long monophthongs is significantly different with $P=0.001 < 0.05$. The F2 difference between /a/ and /a:/ in the FS is the greatest. The F2 difference in the MS for /a:/ is the smallest.

The sound duration among the FS, MS, and LS for /a/ and /a:/ is significantly different with $P=0.000 < 0.05$. The multiple comparisons show that the duration between the FS (mean: 112ms) and MS (mean: 117ms) is the same with $P=0.908> 0.05$, while the duration between the FS and LS (mean: 179ms) ($P=0.000 < 0.05$), or between MS and LS ($P=0.000 < 0.05$), has the significant difference. The duration between the short (mean: 104ms) and long (mean: 188ms) monophthongs is significantly different with $P=0.000 < 0.05$. Duration between /a/ and /a:/ in the FS is the greatest, while in the LS the smallest.

/e/-/e:/

The result of ANOVA shows that the F2 difference for /e/ and /e:/ in Monguor words among the FS, MS, and LS is significantly different (P=0.013 < 0.05). The multiple comparisons show that the F2 difference between the FS (mean of F2 difference [MF2D]: 148Hz) and LS (MF2D: 225Hz) is different with P=0.046 < 0.05, while the F2 difference between the FS and MS (MF2D: 154Hz) (P=0.994 > 0.05), or between MS and LS (P=0.080 > 0.05) is the same. The F2 difference between the short (MF2D: 145Hz) and long (MF2D: 206Hz) monophthongs is significantly different with P=0.009 < 0.05. The F2 difference for /e:/ in each syllable is almost the same.

The sound duration between every two syllables among the FS (mean: 165ms), MS (mean: 121ms), and LS (mean: 212ms) for /e/ and /e:/ is significantly different with all P=0.000 < 0.05. The duration between the short (mean: 115ms) and long (mean: 220ms) monophthongs is significantly different with P=0.000 < 0.05. The sound duration in the MS for both /e/ and /e:/ is the smallest.

/i/-/i:/

According to the result of ANOVA, there is no significant difference in F2 difference among the FS, MS, and LS for /i/ and /i:/ (P=0.697 > 0.05) in Monguor words. However, F2 difference between the short (MF2D: 152Hz) and long (MF2D: 301Hz) monophthongs is significantly different with P=0.000 < 0.05. The F2 difference for /i:/ lowers from the FS to the LS, while the F2 difference for /i/ increases from the FS to the LS.

The sound duration among the FS, MS, and LS for /i/ and /i:/ is significantly different with P=0.000 < 0.05. The multiple comparisons show that the duration between the FS (mean: 138ms) and MS (mean: 113ms) is the same with P=0.296 > 0.05, while the duration between the FS and LS (mean: 203ms) (P=0.000 < 0.05), or between MS and LS (P=0.000 < 0.05), has the significant difference. The duration between the short (mean: 89ms) and long (mean: 257ms) monophthongs is significantly different with P=0.000 < 0.05. The sound duration in the FS and the MS for either /i/ or /i:/ is almost the same, but it is the longest in the LS for either /i/ or /i:/.

/o/-/o:/

The result of ANOVA demonstrates that there is no significant difference in F2 difference among the FS, MS, and LS for /o/ and /o:/ (P=0.899 > 0.05) in Monguor words. However, F2 difference between the short (MF2D: 197Hz) and long (MF2D: 315Hz) monophthongs is significantly different with P=0.000 < 0.05. The F2 difference is the greatest in the MS for /o/, while it is the greatest in the FS for /o:/. The F2 difference between the MS and LS for /o:/ is almost the same, while the F2 difference between the FS and LS for /o/ is almost the same.

The sound duration among the FS, MS, and LS for /o/ and /o:/ is significantly different with P=0.000 < 0.05. The multiple comparisons show that the duration between the FS (mean: 142ms) and MS (mean: 121ms) is the same with P=0.193 > 0.05, while the duration between the FS and LS (mean: 209ms) (P=0.000 < 0.05), or between MS and LS (P=0.000 < 0.05), has the significant difference. The duration between the short (mean: 103ms) and long (mean: 213ms) monophthongs is significantly different with P=0.000 < 0.05. The sound duration in the MS for both /e/ and /e:/ is the smallest, but in the LS for both /e/ and /e:/ the longest. The variation trend about the sound duration for both /e/ and /e:/ is the same.

/u/-/u:/

The result of ANOVA indicates that the F2 difference for /u/ and /u:/ in Monguor words among the FS, MS, and LS is significantly different (P=0.000 < 0.05). The multiple comparisons show that the F2 difference between every two syllables among the FS (MF2D:312 Hz), MS (MF2D: 206Hz), and LS (MF2D: 518Hz) for /u/ and /u:/ is significantly different with all P < 0.05. The F2 difference between the short (MF2D: 198Hz) and long (MF2D: 569Hz) monophthongs is significantly different with P=0.009 < 0.05. The variation trend about the F2 difference for both /e/ and /e:/ is the same. The F2 difference in the MS for both /u/ and /u:/ is the smallest, but in the LS for both /u/ and /u:/ the greatest.

The sound duration among the FS, MS, and LS for /u/ and /u:/ is significantly different with P=0.000 < 0.05. The multiple comparisons show that the duration between the FS (mean: 102ms) and MS (mean: 107ms) is the same with P=0.933 > 0.05, while the duration between the FS and LS (mean:

221ms) (P=0.000 < 0.05), or between MS and LS (P=0.000 < 0.05), has the significant difference. The duration between the short (mean: 86ms) and long (mean: 229ms) monophthongs is significantly different with P=0.000 < 0.05. The sound duration increases from the FS to the LS for both /u/ and /u:/.

4. Summary

On the base of acoustic parameter database of the Monguor speech, the study discusses the differences in duration and F2 difference between short and long monophthongs in monosyllabic words and the differences in duration and F2 difference in first syllable, middle syllable(s), and last syllable for short and long monophthongs. The study shows that the difference between the onset value and the target value of F2 for /a/-/a:/ and /u/-/u:/ is not the same, while the same for /e/-/e:/, /i/-/i:/ and /o/-/o:/. The duration between every pair of short and long monophthongs is significantly different. The result shows that, in Monguor language, short and long vowels have the opposed relation and therefore differentiate the meanings of words. The duration among first syllable, middle syllable(s), and last syllable for short and long monophthongs is significantly different. The result shows that the duration in the last syllable is longer than that in first syllable or middle syllable(s) for either short or long monophthong. There is no consistent result for F2 difference among first syllable, middle syllable(s), and last syllable for each pair of short and long monophthongs.

Language is the most important cultural resource for human beings. Minority language is a valuable cultural treasure of the Chinese nation, and also a component of the cultural diversity of the whole human race. At present, because of the economic, political, technological factors and social changes, the use of some Chinese minority languages decreases, and even some Chinese minority languages are disappearing or has disappeared. Therefore, the purpose of the study is for the preservation of the precious language resource. The results of the study can be applied in speech recognition and speech synthesis. The main significance of the study lies in language application and protection.

Acknowledgements

This work was financially supported by Fund Yxm2015014 to Northwest Minzu University.

References

[1] G.N. Potanin: The Tangut-Tibetan Borderlands of China and Central Mongolia, Royal Russian Geographical Society, Sankt-Petersburg Vol. 2 (1893), p. 1-567.

[2]A. de Smedt, and A. Mostaert: Le dialecte monguor parlé par les Mongols du Kansou occidental, Anthropos Vol. 24(1929), p. 145-166.

[3]A. de Smedt, and A. Mostaert: Le dialecte monguor parlé par les Mongols du Kansou occidental, Anthropos Vol. 25(1930), p. 657-669, 961-973.

[4]A. de Smedt, and A. Mostaert: Le dialecte monguor parlé par les Mongols du Kansou occidental, Anthropos Vol. 26(1931), p. 253.

[5]Jiaodaozhengjia: *The study on Huzhu Dialect of the Tu Language* (Nakanaishi Printing Co.， Ltd., 2008).

[6]Qinggeertai (Chinggeltai): Mongolian Language Family Dialects, Research Series Vol. 13(1957, 1958).

[7]Qinggeertai (Chinggeltai): *The Tu Language and the Mongolian Language* (The Inner Mongolia Peoples Press, Hohhot 1991).

[8]Zhaonasitu: *A Brief Account of the Tu Language* (Nationalities Press, Beijing 1981).

[9]Li Keyu: *Orthography draft of the Tu Language* (1980).

[10]Xi Yuanlin: *Phonological system in Tu Language. The Symposium for National Languages* (Sichuan National Press, Chengdu 1986).