Statistical Survey of Monophthong Formants in Mandarin for Students Being Trained as Broadcasters

Zihou MENG, Yudong CHEN, Xiaohua LI

Communication Acoustics Laboratory, Communication University of China, Beijing P.R.C.

Tel/Fax:+86-10-65783400, E-mail: mzh@cuc.edu.cn

Abstract: To study the group characteristics of Mandarin monophthong formants, the monophthongs of 212 female students and 126 male students from the department of broadcasters were recorded. The monophthong formants were measured using LPC method. The mean value and deviation of formants were given with statistical meaning. The results show the difference from the previous measurements by other researchers decades ago.

Key words: Mandarin, monophthong, formants, statistical survey

1. Introduction

There have been researchers on the monophthong formants characteristics of Mandarin (Putonghua) since 1960's[1][2][3]. Because the number of the speech samples was limited, the results were insignificant in statistics and can not be used to describe the group characteristics of the monophthong formants of Mandarin.

The computer aided Putonghua Shuiping Ceshi (Mandarin level testing) technique has been studied in Communication University of China since 1996. To develop computer aided PSC technique, it needs a statistical survey of the Mandarin speech characteristics with a large number of samples.

Using the Mandarin speech database recorded in the Communication University of China, the monophthong formants of 212 trained female students and 126 trained male students were measured and analyzed. The result is meaningful in statistics. It should be helpful to the study of the group characteristics of Mandarin monophthong formants. It is also helpful for the study of parametric speech standard of Mandarin.

2. Measurement method

The statistical survey is based on the Mandarin speech database in the Communication University of China. All speakers are students from the broadcaster's department, aged from 19 to 21 years old and have been trained for 1 to 3 years. 212 female students and 126 male students were measured. The dialect atlas (hometown) of the students is listed in Table 1.

Table 1. Dialect atlas of the students

Dialect Area	Female	Male
Anhui	6	0
Beijing	19	17
Chongqin	6	6
Fujian	2	0
Gansu	3	3
Guangdong	3	1
Guangxi	0	1
Guizhou	1	0

Dialect Area	Female	Male
Hebei	7	5
Henan	14	8
Heilongjiang	24	14
Hubei	4	8
Hunan	6	4
Jilin	10	7
Jiangsu	13	5
Jiangxi	3	0
Liaonin	21	11
Neimenggu	2	3
Ningxia	2	0
Shandong	20	8
Shanxi( )	5	0
Shannxi( )	9	5
Shanghai	6	2
Sichuan	9	3
Tianjin	4	9
Xizang(Tibet)	1	0
Xinjiang	4	2
Yunnan	1	2
Zhejiang	7	2
Total	212	126

The 10 monophthongs in the test are listed in Table 2. These 10 monophthongs of Mandarin are selected according to reference [4], and it is not completely same with the monophthong catalogue used by other researchers. Usually, ê[ɛ] is only used as the mood word " ( )",and its voice is unstable. In phonetics, there is dispute of whether the vowel should be put into the Mandarin speech frame. In this paper, considering the special application background, ê[ɛ] is dealt with as an independent monophthong.

The software used to measure the formants and is SFSwin. The LPC technique is used in SFSwin with a formant trajectory estimation dynamic program to get the optimum formants estimation[5].

Table 2. Monophthongs in the test

No.	tag	International phonetic symbol	Chinese character with the monophthong
1	i	[i]	
2	u	[u]	
3	ü	[y]	
4	o	[o]	
5	e	[½]	
6	a	[a]	
7	ï	[ɿ]	final of " "
8	î	[ʅ]	final of " "
9	ë	[ə]	final of " "
10	ê	[ɛ]	

Among the 10 monophthongs, [ɿ],[ʯ],and [ə] are very difficult to speak for untrained speakers. Since the speakers in this test are being trained as professional broadcaster, they are required to speak the vowels independently. If it is really too difficult to speak the monophthongs independently, then they can speak the Mandarin character with the monophthong as the final. The monophthong is then cut from the sound of the character in signal processing.

For signal processing, all recorded monophthong signals are checked by listening. The data with deficiencies are repaired or cut into the stable pieces. The processed sound files are then put into the SFSwin to measure the first five formants and the fundamental frequencies. Based on the analysis, it found that tones has no meaningful affection on formants, therefore all sound samples in this paper are with the first tone (plain flat).

3. Results and discussion

The results of the statistical survey of the formants are listed in Table 3 and Table 4. The mean frequency of the first five formants for each monophthongs is listed in Table 3 for female and male students separately. For each formant, the ratio of standard deviation to the mean frequency of the formant is listed in Table 4.

Although 212 female students and 126 male students were recorded, the number of effective sound samples is not the same as the number of students because some sound samples are not correct or with deficiency. In signal processing, those sound samples with deficiency were deleted. In the work of this paper, the number of effective sound samples is large enough to guarantee the measurement data to be meaningful in statistics..

Table 3.　Mean frequency of the first five formants
for 10 monophthongs in Mandarin

speaker	formant	i[i]	u[u]	ü[y]	o[o]	e[½]	a[a]	ï[ɿ]	î[ʯ]	ë[ə]	ê[ɛ]
female	F1	311	354	305	654	622	952	399	416	656	856
	F2	2871	762	2411	947	1334	1371	1762	2092	1370	2119
	F3	3882	3432	2813	3410	3358	3444	3271	2805	3376	3159
	F4	4613	4370	4061	4276	4452	4422	4599	4045	4474	4541
	F5	6642	7100	6098	6026	6118	5943	6792	5760	6091	6229
male	F1	279	342	280	532	501	795	355	351	525	705
	F2	2240	701	1992	817	1163	1168	1410	1719	1179	1789
	F3	3168	2800	2331	2949	2736	2945	2702	2316	2735	2656
	F4	4117	4552	3352	3679	3608	3727	4324	3194	3583	3844
	F5	6413	6587	6358	6280	6300	6056	6500	6170	6310	6206

Table 4.　Ratio of the standard deviation to the mean frequency of formants

speaker	formant	i[i]	u[u]	ü[y]	o[o]	e[½]	a[a]	ï[ɿ]	î[ʯ]	ë[ə]	ê[ɛ]
female	F1	0.09	0.13	0.08	0.12	0.12	0.13	0.18	0.17	0.15	0.12
	F2	0.06	0.12	0.05	0.10	0.09	0.10	0.07	0.08	0.10	0.08
	F3	0.06	0.12	0.04	0.05	0.06	0.07	0.06	0.18	0.06	0.08
	F4	0.04	0.06	0.05	0.05	0.05	0.05	0.05	0.11	0.05	0.05

	F5	0.08	0.04	0.15	0.12	0.12	0.13	0.07	0.16	0.12	0.10
male	**F1**	0.11	0.10	0.10	0.12	0.13	0.07	0.17	0.13	0.10	0.08
	F2	0.05	0.09	0.05	0.09	0.06	0.06	0.09	0.07	0.07	0.06
	F3	0.06	0.12	0.05	0.07	0.05	0.06	0.06	0.15	0.07	0.06
	F4	0.18	0.23	0.06	0.08	0.06	0.06	0.22	0.10	0.05	0.07
	F5	0.07	0.07	0.08	0.06	0.06	0.07	0.07	0.06	0.07	0.11

Figure 1 is the probability density distribution function of F2 for female [i] which has 173 effective samples, while Figure 2 is the probability density distribution function of F4 for male [ɛ] which has only 40 effective samples. Both figures show the typical normal distribution pattern. In the work of this paper, the number of male students is just more than half of the number of female students, but based on the observation of the probability density distribution function of each formant for each monophthong, even the minimum effective number, as showed in Figure 2, is enough to guarantee the measurement to be meaningful in statistics.
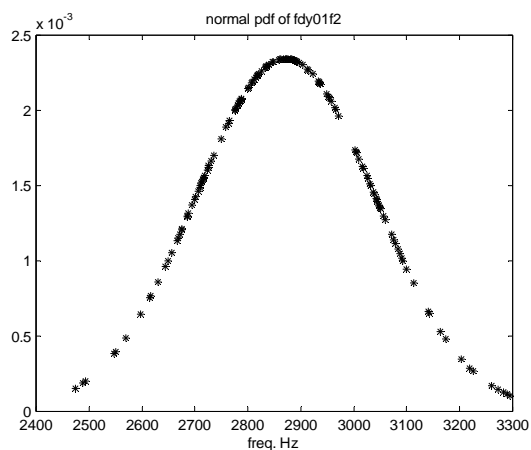

Figure 1. Probability density distribution function of F2
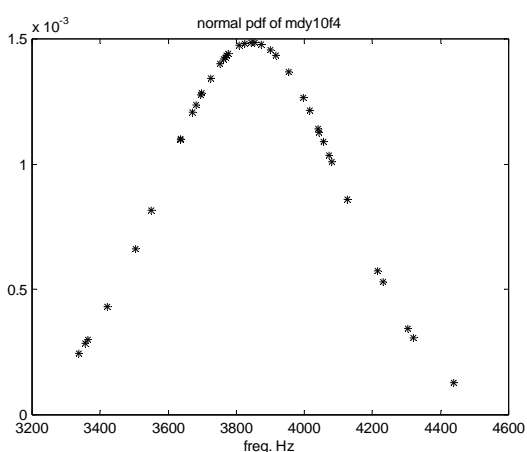for female [i] with 173 effective samples

Figure 2. Probability density distribution function of F4
for male [ɛ] with 40 effective samples

Figure 3 is the spreading of the first three formants of female students group for the ten monophthongs. Figure 4 is for male students. The vertical width in the figure is double of the standard deviation of the corresponding formant. This represents the spreading range of the formant.

Figure 5 and Figure 6 are the acoustic monophthong plots for female and male students. The frequency of F1 and F2 are labeled in mel to be consistent with that used in phonetics. Compared with Fig. 2-13 in Ref. [4], it can be seen that in Figure 5 and Figure 6, [o] is outside the vowel triangle composed of [i], [u], and [a], while in Ref. [4], [o] is inside the vowel triangle. It is not clear whether the Fig. 2-13 in Ref. [4] is meaningful in statistics, and whether it is the male data or female data. The fact that [o] is outside the vowel triangle in Figure 5 and Figure 6 is from the practical statistical survey. The difference from Ref. [4] possibly means that speaking style of the students has been changed with the time, at least the speaking style of the trained students group has been different from that decades ago[7][8], or different from the subjects used in Ref. [4].

Compared Figure 5 with Figure 6, it can be seen that the triangle in Figure 5 for female students is similar to the triangle in Figure 6 for male students, but it is larger than the male triangle. This may imply that besides the absolute value of formants, there are some other attributes in the listening perception of sound that has not been completely revealed by current studies.
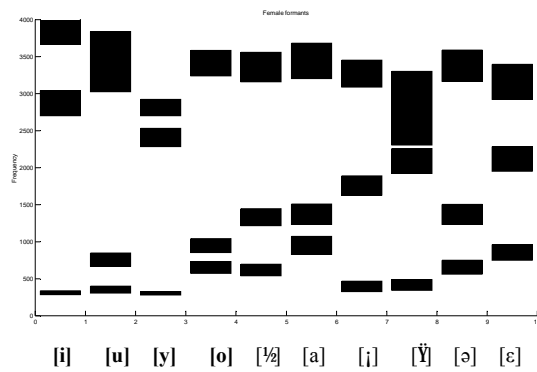

Figure 3. Spreading of the first three formants of female
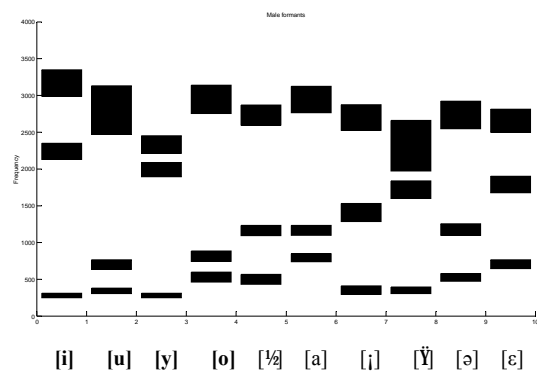students for ten monophthongs

Figure 4. Spreading of the first three formants of male
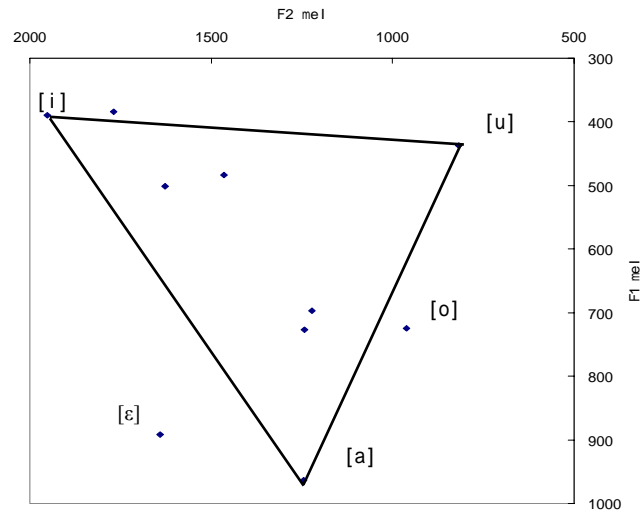students for ten monophthongs

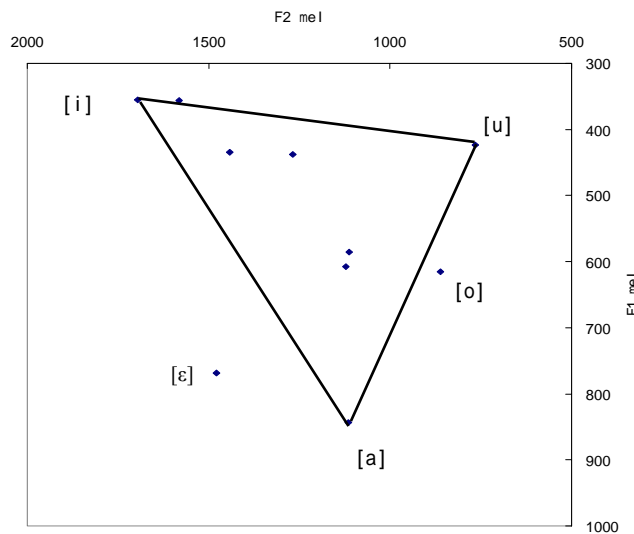Figure 5. Acoustic monophthong plot for female students

Figure 6. Acoustic monophthong plot for male students

Compared the measured data with the data measured by other researchers[1][2][3], most of the data measured by other researchers are not within the spreading range of this work, as mentioned in the previous paper[7]. The test subjects in Ref. [1] are 4 male and 4 female. The test subjects in Ref. [2] are 19 male and 19 female. The subjects in Ref. [3] are 7 male and 5 female. For all these works, the number of the subjects is too small to produce the result with statistical meaning, and the age range of the subjects in these work is very wide. The subjects in the work of this paper are all with the age from 19 to 21 years old and are all university students. The number of subjects is large enough to guarantee the survey with statistical meaning and produce the typical formants data representing this subject group. It is not meaningful to comment whether the data in this work is more precise or more correct than the data measured by other researchers, because the work in this paper has been the first time to survey the

Mandarin monophthongs with very large number of subjects for a specified group.

4. Summary

The work of this paper is the first time to conduct a statistical survey of Mandarin monophthong formants for the students being trained as broadcasters in Communication University of China. The result for female group is different from that of male group. Generally, a formant frequency of female is higher than the corresponding frequency of male formant. This is believed to be caused by acoustic model of monophthongs. Different subject should possess different formants. But, different subject can make the same listening perception of a specified monophthong. Therefore, although the statistical survey of monophthong formants is very useful for monophthong synthesizing and the study of voice source mechanism, more important study should be conducted to explore the listening perception mechanism of monophthongs.

Reference

[1] WU Zhongji, "Spectral Analysis and Formants Measurement for Vowels and Consonants in Standard Chinese", Journal of the Acoustics Society of China, Jan. 1964; **1**(1): p33-39

[2] XU Huangzhang, "The Mean Spectral Characteristics of Ten Vowels on the Standard Chinese (Putonghua)", Journal of the Acoustics Society of China, 1965; **2(**1): p1-7

[3] ZHANG Jialu, "A Semi-Dynamic Method for Spectrographic Analysis of Vowels", ACTA ACUSTICA, 1979; **4**(1): p23-29

[4] LIN Tao, ***Course of Phonetics***, Peking University Press, 1992: p47-52

[5] D.Talkin, "Speech Formant Trajectory Estimation Using Dynamic Programming with Modulated Transition Costs" JASA, **S1**,1987: p. S55

[6] ZHANG Jialu, "On the Development of Speech Technology", ACTA ACUSTICA, 2004; **29**(3): p193-199

[7] MENG Zihou, "Statistical Survey of Female Monophthong Formants in Mandarin", ACTA Acoustica, 2006, Vol.25 No.2

[8] MENG Zihou, CHEN Yudong, LI Xiaohua, "Fundamental Frequency Survey of Mandarin Monophthongs", WESPAC IX, June 26-28, 2006, Seoul Korea