Study on Phonation Features of Mongolian, Tibetan and Yi Language*

Chen Jiayou, Shen Mixia, Wu Xiyu, Kong Jiangping
Chinese Academy of Social Sciences, P.R. China
E-mail: jpkong@263.net , chenjy@nation.cass.net.cn

ABSTRACT

This paper is concerned with the study of phonation features of Mongolian, Tibetan and Yi language, which are languages in China. Among these languages, Tibetan and Yi are tone languages, which belong to Sino-Tibetan language family. Mongolian is not a tone language, which belongs to Altaic language family. The method of acoustical analysis includes: 1) inverse-filtering, which is used to extract glottal source from the speech sound; 2) spectrum tilt (ST) calculating, which is used to quantitate the spectrum feature by dB/oct; and 3) parameter extracting, which is used to extract the parameters of fundamental frequency (F0), speed quotient (SQ) and open quotient (OQ). The result shows: 1) the differences of phonation features exist among these three languages; 2) there are great differences between the phonations of male and female; 3) the correlation of the phonation parameters is studied, from which the nature of these parameters can be described and defined.

1. INTRODUCTION

Mongolian, Tibetan and Yi are languages in China. Among these languages, Tibetan and Yi are tone languages, which belong to Sino-Tibetan language family. Mongolian is not a tone language, which belongs to Altai language family. In Tibetan (Lhasa speech which is the standard speech of Tibetan), there are 4 basic tones, so the fundamental frequency is very important in speech communication. In Yi (Liangshan speech which is the standard speech of Yi language), there are 4 basic tones and two phonation types, which are usually called as tense voice (a kind of pressed voice) and lax voice (modal voice). The two phonation types are very important in distinguish meanings in Yi language. In Mongolian (Kerchin speech), the fundamental frequency is not so important as that in Tibetan and Yi, but has 2 different phonation types, which are traditionally called yin (cool) sound and yang (hot) sound. These 2 phonation types have been proved to be different in many respects acoustically and physiologically.

The purpose of this study is to find the differences of phonation features existing in these languages. In this study, sustained vowels of Mongolian, Tibetan and Yi are analyzed acoustically.

We have 30 speakers, 15 male and 15 female for each language. For each speaker, a sustained vowel /a/ are recorded twice. There are 60 samples for each language and 180 samples altogether in this study.

2. RESEARCH METHOD

The acoustical analysis used in this study has 3 steps, which are: 1) inverse-filtering these samples to extract speech glottal source; 2) calculating the spectrum, then calculating the spectrum tilt, and finally the spectrum tilt has been quantitated by dB/oct; 3) from the extracted speech glottal source (differential form), fundamental frequency, open quotient and speed quotient are extracted and calculated;

2.1 Inverse Filtering

The method of inverse filtering used in this study is the cascade inverse filtering (Gold and Rabiner,1968; Klatt and Klatt, 1990). See the following equation.

$$y(nT) = A'x(nT) + B'x(nT-T) + C'x(nT-2T)$$

where *x(nT-T)* and *x(nT-2T)* are the previous two samples of the output, *x(nT)*, and the constants A', B' and C' are defined by the following equations:

$$A' = \frac{1.0}{A} \qquad A = 1. - B - C$$

$$B' = \frac{-B}{A} \qquad B = e^{(-\pi BW \cdot T)}\cos(2\pi FT)$$

$$C' = \frac{-C}{A} \qquad C = e^{(-2\pi BW \cdot T)}$$

In order to inverse filtering the speech samples, an inverse filtering system has been establised. In this system, there are 5 inverse filters, which are used for the inverse filtering of 5 formants, and a pair of inverse filter and filter, which are designed for inverse filtering of nasal formant and anti-formant.There are 2 main steps in the inversed filtering. 1) initializing the formant central frequency and bandwidth by co-variance LPC; 2) these parameters can be modified acrording to the FFT spectrum, untill the speech glottal source are obtained. See figure 1.

ISBN 1-876346-48-5 © 2003 UAB

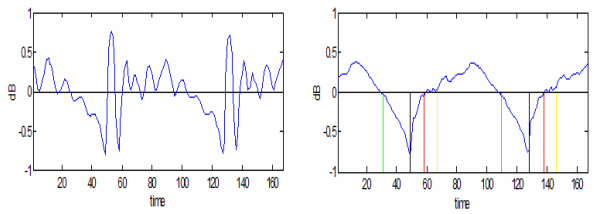

Figure 1: Speech waveform and the inverse-filtered glottal source with marks of glottal opening instant, peak of glottal airflow and glottal closure instant.

2.2 Spectrum Tilt Calculating

After the speech glottal source has been extracted, the spectrum of the speech glottal source is calculated by FFT. After that, the maximum of each harmonics is detected automatically. Then the polynomial curve fitting is used to obtain the contour of spectrum tilt with these data. Finally the value of spectrum tilt is quantitated by dB/oct from the polynomial curve fitted contour.

2.3 Calculating of OQ and SQ

The glottal period, opening phase and closing phase are defined in figure 2. See figure 2.

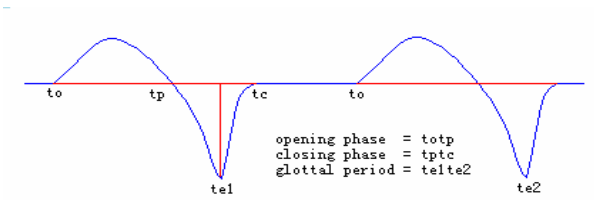

Figure 2: The definitions of glottal period, opening phase and closing phase.

The fundamental frequency, open quotient and speed quotient in this study are defined as follows:

F0=1/glottal period

OQ=((opening phase+closing phase)/period)× 100%

SQ=(opening phase/closing phase) × 100%

3. PHONATION FEATURES OF MONGOLIAN

Besides the parameters of F0, OQ and SQ, the parameters of 5 formants and bandwidths are also extracted during the inverse filtering. Before the analysis of these parameters, they have been first sorted on F0, and then treated by polynomial curve fitting, so the relationship among them can be easily found.

3.1 Relationship among Parameters

The relationships among F0, OQ, SQ and ST are displayed in figure 3, from which it can be seen that along with the increasing of F0, the OQ increases, then goes flat and falls a little at the end of the contour. The SQ increases all the time. The ST decreases, while the F0 increases. In general, F0, OQ and SQ have a direct ratio and F0, OQ and SQ have an inverse ratio with ST.

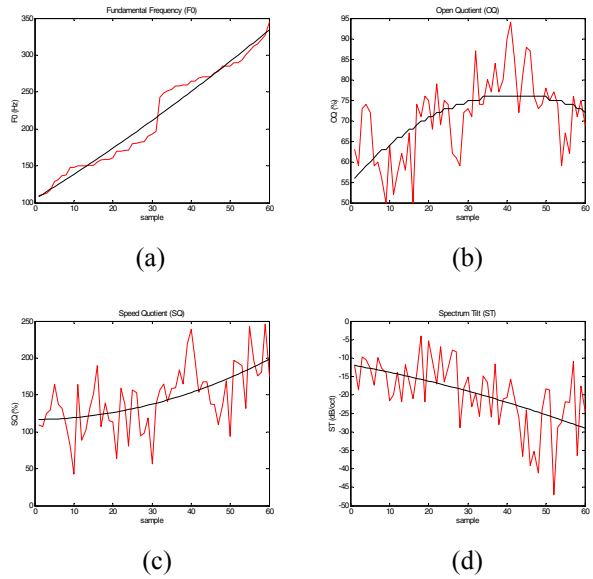

Figure 3: a) the parameters of F0 and the polynomial curve fitted contour; b) the parameters of OQ and the polynomial curve fitted contour; c) the parameters of SQ and the polynomial curve fitted contour; d) the parameters of ST and the polynomial curve fitted contour.

Figure 4 displays the parameters of 5 formants and the polynomial curve fitted contours, from which we can find that along with the increasing of F0, the 5 formants are all increasing. It means that F0 has a direct ratio with formants of speakers under the comfortable voice condition.

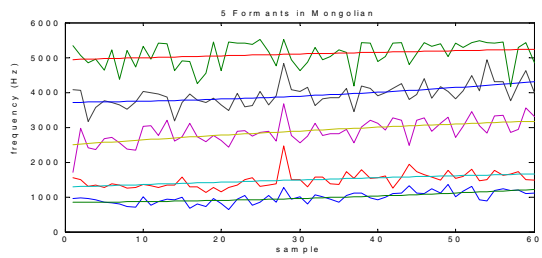

Figure 4: Parameters of 5 formants and the polynomial curve fitted contours.

3.2 Correlation Analysis

The correlation among F0, OQ, SQ and ST has been calculated. From these correlation, we can find that F0 has positive correlation with OQ and SQ, and has negative correlation with ST. OQ has correlation with SQ. The

correlations among F0 and the 5 formants is also calculated and from them, we can find that F0 and the first 4 formants have correlation and the formant 5 only has correlation with formant 3 and 4.

4. PHONATION FEATRUES OF TIBETAN

The parameters of F0, OQ, SQ and the 5 formants with their bandwidths are extracted and some of them are analized in this section. Before the analysis of these parameters, they have been first sorted on F0, and then treated by polynomial curve fitting, so the relationship among them can be easily found.

4.1 Relationship among Parameters

The relationship among F0, OQ, SQ and ST is displayed in figure 5. From the figure, it can be seen that along with the increasing of F0, the OQ increases quickly. The SQ and ST decrease, while F0 increases. In general, F0 and OQ have a direct ratio. F0 has an inverse ratio with SQ and ST. The relationship in Tibetan is not as same as that in Mongolian.

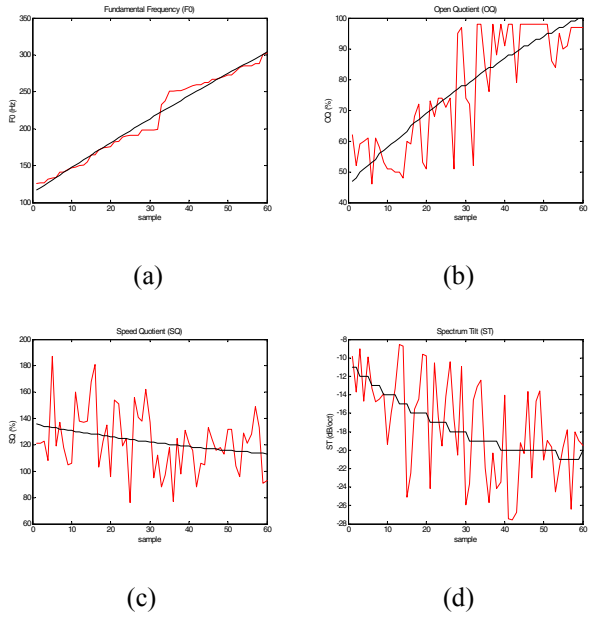

Figure 5: The parameters of F0, OQ, SQ and ST of Tibetan with their polynomial curve fitted contours. The figure arrangement is as same as that in figure 3.

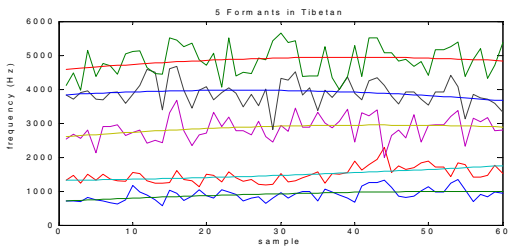

Figure 6: Parameters of 5 formants and their polynomial curve fitted contours.

Figure 6 displays the parameters of 5 formants and the polynomial curve fitted contours, from which we can find that along with the increasing of F0, the first 3 formants are all increasing. It means that F0 has a direct ratio with the first 3 formants. The 4th formant is flat and the 5th formant increases first and then decreases a little.

4.2 Correlation Analysis

The correlations among F0, OQ, SQ and ST have been calculated. From the correlation coefficients, we can find that F0 has positive correlations with OQ, and it has negative correlation with SQ and ST. OQ has positive correlation with ST. The correlation coefficients of F0 and the 5 formants are calculated and from them, we can see that F0 has correlation with the first 3 formants, but the correlation between the formant 3 and the formant 4 is also significant at the 0.01 level.

5. PHONATION FEATURES OF YI LANGUAGE

The parameters of F0, OQ, SQ and the parameters of the 5 formants and band widths are extracted through the inverse filtering. Before the analysis of these parameters, they have been first sorted on F0, and then treated by polynomial curve fitting, so the relationship among them can be easily found.

5.1 Relationship among Parameters

The relationships among F0, OQ, SQ and ST are displayed in figure 7. From the figure, it can be found that along with the increasing of F0, the OQ and SQ increase all the time. The ST decreases, while the F0 increases. In general, F0, OQ and SQ have a direct ratio, and F0, OQ and SQ have an inverse ratio with ST.

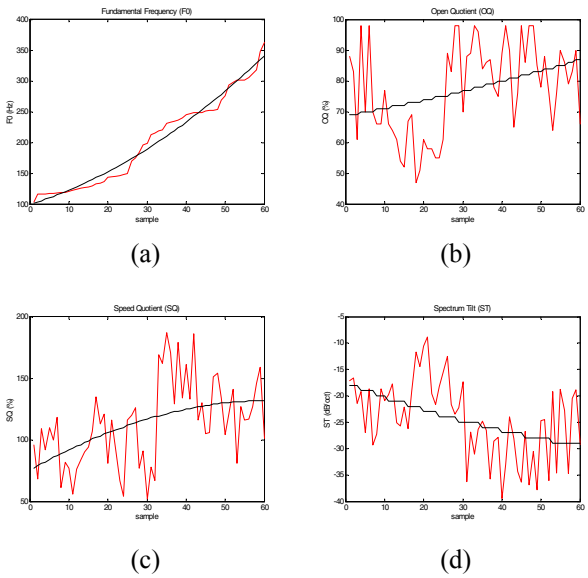

Figure 7: The parameters of F0, OQ, SQ and ST of Yi language with their polynomial curve fitted contours. The figure arrangement is as same as that in figure 3.

ISBN 1-876346-48-5 © 2003 UAB

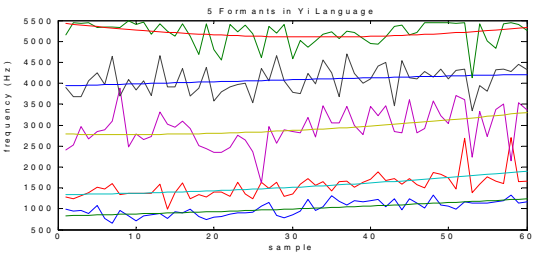

Figure 8: Parameters of the 5 formants and their polynomial curve fitted contours.

Figure 8 displays the parameters of 5 formants and the polynomial curve fitted contours, from which we can find that along with the increasing of F0, the first 4 formants are all increasing. It means that F0 has a direct ratio with these formants. The 5th formant decreases first and then goes up.

5.2 Correlation Analysis

The correlations among F0, OQ, SQ and ST have been calculated and from these coefficients, we can see that F0, OQ and SQ have positive correlations with each other and these 3 parameters have negative correlation with ST. The correlation coefficients of F0 and 3 formants are also calculate. Generally speaking, the correlation between F0 and the first 3 formants is significant at 0.01 level.

6. COMPARATIVE STUDY

In order to compare the phonation features of these languages, the average parameters of male and female are given in table 7. The parameters show that F0 of female is larger than that of male for each languages. The largest F0 is that of Mongolian female and samllest one is that of Yi male. The OQ of female is larger than that of male for each language. The largest OQ is that of Tibetan female and the smallest one is that of Tibetan male. The SQ of female is larger than that of male for Mongolian and Yi language, and the SQ of female is smaller than that of male for Tibetan. The ST of female is larger than that of male for all these 3 languages. In generally speaking, the phonation of Tibetan is quite different from the others.

	mf	mm	tf	tm	yf	ym
F0	278	154	263	163	266	137
OQ	77	66	94	60	85	70
SQ	171	119	116	130	133	93
ST	-25	-14	-20	-15	-29	-20

Table 7: The average parameters of male and female. In the first row, the first letter in each column stands for languages, which are Mongolian, Tibetan and Yi language, and the second letter stands for male and female.

	mf	mm	tf	tm	yf	ym
F1	1142	881	1085	885	973	831
F2	1712	1393	1584	1392	1617	1364
F3	3134	2733	3037	2720	2910	2809

The parameters of these 3 languages show that the first 3 formants of female are all larger than that of male. The largest F1 is that of Mongolian female and the smallest F1 is that of Yi male. The largest F2 is that of Mongolian female and the smallest F2 is that of Yi male. The largest F3 is that of Mongolian female and the smallest F3 is that of Tibetan male.

7. CONCLUSIONS

The research conclusions are: 1) there are close relationship among the phonation parameters and the first 3 formants for each language; 2) there are great differences between the phonation parameters and the first 3 formants of male and female; 3) the differences of phonation features do exist among these three languages, especially the that of Tibetan.

REFERENCES

[1] Caodao Bater and Kong Jiangping., 1999, EGG Study of 84 Mongolian Speakers. Proceedings of 4th National Conference on Modern Phonetics, Aug. Beijing, China, published by Jincheng Publishing House.

[2] Chen Jiayou and Kong Jiangping., 1999, EGG (electroglottography) Signal Study of 92 Yi Language Speakers, The Proceedings of 4th National Conference on Modern Phonetics, Aug. Beijing, China, published by Jincheng Publishing House.

[3] Caodao Bater, Chen Jiayou, Shen Mixia and Kong Jiangping., 2000, EGG (electroglottography) Signal Study of 100 Tibetan Speakers. A Selected Book for the Technology of man-machine Speech Communication edited by the Department of Electronic Engineering, Qing Hua University (Beijing).

[4] Gold B. And Rabiner L.R. "Analysis of Digital and Analog Formant Synthesizers," IEEE Trans. Audio Electoacoust. AU-16, 81-94, 1968.

[5] Klatt D.H. and Klatt L.C., "Analysis, Synthesis and Perception of Voice Quality Variations among Female and Male Talkers," J.Acoust. Soc. Am. 87,820-857, 1990.

* The research is funed by the National Natural Science Foundation of China (No.10274105).