

Comparative Spectral Analysis of Mongolian Sounds of Same Middle Age Male and Female

Chimedbaljir Narantungalag
School of Information and Communication Technology
Mongolian University of Science and Technology, Mangolia
Email: chnara [AT] gmail.com

ABSTRACT---- *This paper presents the results of spectral analysis of all vowels and consonants of Mongolian language using non-parametric and parametric methods of analysis. Early time for spectral analysis have been used the Period gram and it's versions, Fast Fourier Transform, MUSIC and PMUSIC methods, but today except them are used the automated speech signal processing software as PRAAT and AUDACITY. Spectral analysis of all sounds of Mongolian language was performed by these methods. Detailed analysis was performed for separately pronounced all sounds and pronounced at the beginning, middle and end of the words. The spectral analysis parameters as the formant frequencies, bandwidth of the formants, formant power densities of the vowels and consonants were analyzed. Done comparative evaluation depending on position of sounds in words (at beginning, middle and end of word) and for male or female. Results of comparative spectral analysis show the followings. Firstly, should be noted differences of formant frequencies of same age male and female. Secondly, spectral parameters (formant frequencies, bandwidth of the formants, formant power densities) of separately pronounced sounds vary according to their positions in the words. It's not observed deterministic or rational changes. First formant frequencies of female pronunciations predominantly exceed that of the male. But in next formants happened predominance of male formant frequencies depending on sound and position. Formant frequencies tend to increase at the beginning, middle, and end of the word.*

Keywords---- Formant, formant frequency, bandwihih, power spectrum density

1. INTRODUCTION

Spectral analysis has very wide application in different fields, including earthquake and vibration, economy, meteorology, astronomy and other fields.

Spectral models of acoustic signal are very useful in speech processing, particularly in speech synthesis and speech recognition and improves quality of intelligence in speech generation (production) process [6, 7, 8]. In radar and sonar systems spectral analysis provides location information of object in direct sight distance.

In the health sector electrocardiogram (ECG) and encephalogram (EEG) are widely used for patient diagnosis.

There are two approaches of spectral analysis. Spectral analysis classical or non-parametric method [2, 3, 6, 9]; Model based parametric method. This method makes spectra parametric (parameterize spectra) and reduces spectra evaluation transforming it into model parametric evaluation. The accuracy of parametric method is higher comparatively to that of non-parametric method. It's occurs necessity to make spectral analysis of deterministic signals.

Except abovementioned methods there are specialized software tools as Praat and Audacity for speech signal processing and analysis. Using these tools should be determined formant parameters of sounds of Mongolian language.

Spectral analysis was performed by non-parametric and parametric methods using computer. This is first detailed spectral analysis of Mongolian language.

Final results of the analysis are the determined formant parameters or excitation parameters of the Mongolian language for block "Vocal Tract System" of speech chain [1, 5].

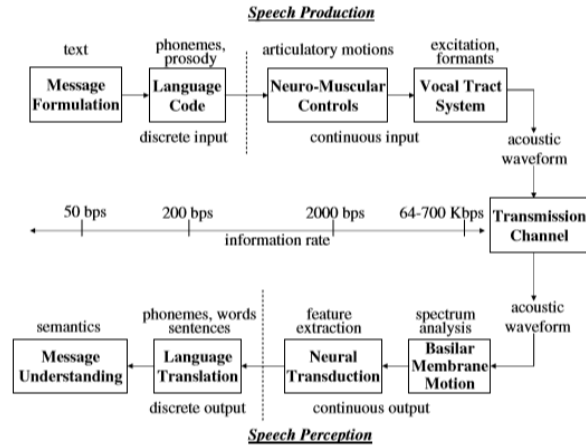


Figure 1. The Speech Chain

2. NON-PARAMETRIC AND PARAMETRIC METHOD SPECTRAL ANALYSIS

2.1 Acoustic signals used for analysis

For detailed spectral analysis and comparison of Mongolian language sounds have been used acoustic signals pronounced by male and female 4 persons' of 18 and 50 age. Analysis made by FFT and PMUSIC mehtods.

Acoustic signals of all vowels and consonants separately pronounced by male and female of 40 age were used for analysis with specialized software tool "Praat".

Also acoustic signals of all sounds (or context phonemes) extracted from selected words (according to position at beginning, middle, end of word), pronounced by male and female of 40 age were used for formant parameter analysis with specialized software tool "Praat".

Performed spectral analysis and comparison of formant parameters of separately pronounced vowels and consonants and context phonemes.

Using obtained results of research or formant parameters we suggest simplified formant model.

2.2 Results of non-parametric and parametric method spectral analysis

Inherent analysis have been used following formulas.

Furrier transform:

$$X(k) = \sum_{n=0}^{N-1} x(n)e^{-j2\pi kn/N}, \quad k = 0, 1, 2, \dots, N-1$$

Power spectrum density:

$$P(\omega) = \sum_{-\infty}^{\infty} r(k)e^{-i\omega k}$$

From the analysis were observed followings:

Parametric method uses model close to real signal and most adequate for determination of it's parameters.

MUSIC method is most suitable for determination of spectral component frequencies and their dependences.

According to specific features or frequencies vowels can be divided into 3 groups [4]:

- A. “A, O” vowels 625-950 Hz (male), 810-1085 Hz (female)
- B. “ y , θ , Y” vowels 300-480 Hz (male), 400-550 Hz (female)
- C. “ ə , И” vowels 93-148 Hz (male), 210-257 Hz (female).

Vowels have own carrier frequencies and upper frequency components. Maximum 4 from these frequencies determine distinguish attributes of each sound from others.

Also by the power difference of specific attribute frequency and carrier frequency vowels are distinguished each from others. Specific attribute frequencies are most higher frequencies than carrier one’s and more energy frequencies.

Correlation based PMUSIC method decreases occurrence of random components and emphasizes main components of signal. And permits to express the signal with a few frequencies and constant component.

From calculation by FFT, PMUSIC methods and comparison of obtained spectral analysis results of periodical consonants (M, H, P), fricatives (C, III), and stops (Б, Г, Д, И, Т) it is clear that PMUSIC method gives more explicit maximum meanings e.g. formant frequencies and their levels are more clear or understandable.

Duration and other parameters of sounds which compose word are changed due to co-articulation between them. Specially at the merging sections of sounds. In this case vowels have more changes. However if vowels are pronounced with pause in the mid of word parameters are similar to separately pronounced one’s.

Spectral analysis are made by extracted acoustic signals part which related to particular phonemes from Mongolian words as “MANAN”, “MINJ”, “TSOM”, “TUM”, “HAVAR” and performed with PMUSIC method.

All consonants included in this analysis have 3 formants.

3. SPECTRAL ANALYSIS MADE BY “PRAAT”

3.1 Analysis of parameters

In spectral analysis are widely used software tools. Spectral analysis presented in the paper was performed using Praat.

The Praat software tool is widely used for DSP, such as the processing and analysis of speech digital signals. Using Praat it is possible to extract out and save the desired part (sentence, word, signals of any sound or phonemes) from the continuous speech signal and process it. Praat permits definition of duration as time domain and spectrogram and formant frequencies as frequency domain parameters and illustrate them numerically or graphically.

In this research, the formant frequencies are calculated and plotted and analyzed. Furthermore, word and speech synthesis are possible using these formants.

In the research participated 40 age male and female, made pronunciation of separate sounds and selected words from which extracted acoustic signals of any sounds or phonemes and done analysis of formant parameters relatively to position of sounds in the words. So defined formant parameters as formant frequencies, frequency bandwidth of formants, power spectrum density comparatively evaluated.

The results obtained by analysis shown in tables. In Table 2.1 shown formant frequencies of vowels for male (M) and female (F).

Should be noted that 5th formant frequency is not available by Praat tool.

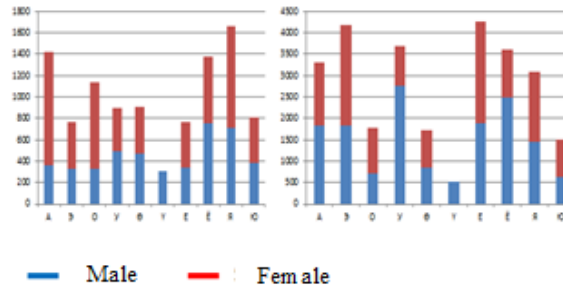
Table 2.1. Comparison of vowels formant frequencies for male and female

| | | |
|--|------------------|------------------|
| | Formant 1, F_1 | Formant 2, F_2 |
|--|------------------|------------------|

| | Female | Male | Female / Male | Female | Male | Female / Male |
|---|--------|------|---------------|--------|------|---------------|
| A | 1063 | 359 | 2.96 | 1479 | 1832 | 0.81 |
| Ә | 434 | 333 | 1.3 | 2354 | 1822 | 1.29 |
| O | 809 | 329 | 2.46 | 1049 | 724 | 1.45 |
| Y | 404 | 488 | 0.83 | 933 | 2759 | 0.39 |
| Ө | 434 | 471 | 0.92 | 894 | 835 | 1.07 |
| Ү | | 310 | | | 514 | |
| E | 423 | 339 | 1.25 | 2364 | 1894 | 1.25 |
| Ě | 622 | 757 | 0.82 | 1112 | 2493 | 0.45 |
| Я | 945 | 713 | 1.33 | 1625 | 1451 | 1.12 |
| Ю | 419 | 387 | 1.08 | 875 | 638 | 1.37 |

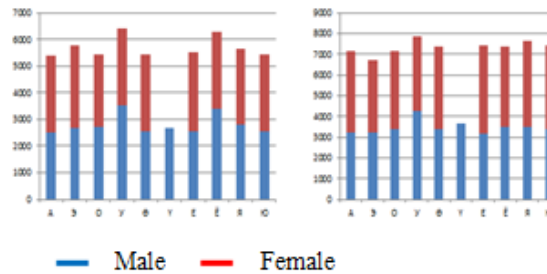
Continue of Tab.2.1.

| Formant 3, F_3 | | | Formant 4, F_4 | | |
|------------------|------|---------------|------------------|------|---------------|
| Female | Male | Female / Male | Female | Male | Female / Male |
| 2879 | 2519 | 1.14 | 3945 | 3216 | 1.23 |
| 3069 | 2689 | 1.14 | 3507 | 3206 | 1.09 |
| 2705 | 2727 | 0.99 | 3784 | 3373 | 1.12 |
| 2856 | 3538 | 0.81 | 3587 | 4276 | 0.84 |
| 2863 | 2564 | 1.12 | 3988 | 3369 | 1.18 |
| | 2687 | | | 3639 | |
| 2965 | 2567 | 1.16 | 4242 | 3165 | 1.34 |
| 2903 | 3389 | 0.86 | 3896 | 3488 | 1.12 |
| 2830 | 2798 | 1.01 | 4143 | 3504 | 1.18 |
| 2914 | 2528 | 1.15 | 4036 | 3409 | 1.18 |



Male Female

Fig. 2.1. 1st and 2nd formant frequencies of singular vowels



Male Female

Fig. 2.2. 3rd and 4th formant frequencies of singular vowels

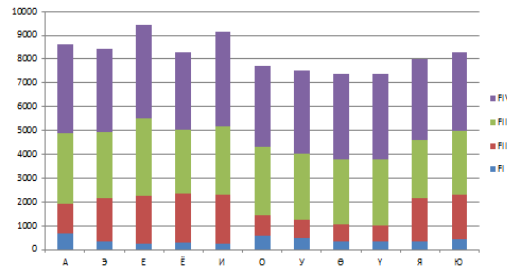


Fig. 2.3. Four formant frequencies singular vowels /Male/

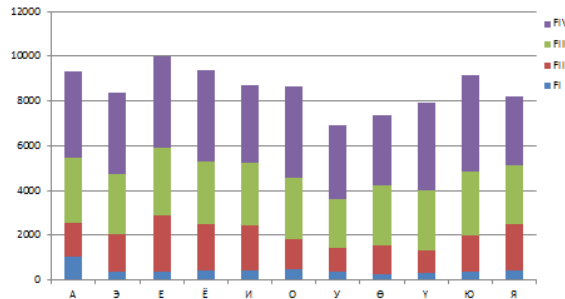


Fig.2.4. Four formant frequencies singular vowels /Female/

From Tab.2.1 and Fig. 2.1 -2.4 it is seen that first formant frequencies vary in the range 404-1063 Hz for female, and 310-757 Hz for male. Except some vowels female formant frequencies exceed male one's 8-30 %, even some formant frequencies 3 times. For remained 3 formant frequencies observed similar changes, but for some vowels male frequencies are predominant comparatively female one's. In figures 2.3 and 2.4 cyan color shows 1st formant frequency, red one - 2nd, green one – 3rd, purple blue one – 4th formant frequencies respectively.

3.2 Difference of formant frequencies of singular vowels and vowels at the different position of word

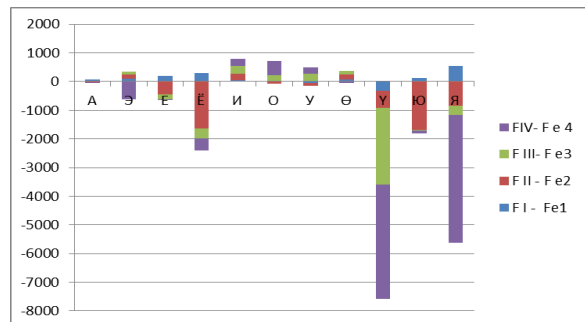


Fig. 2.5. Difference of formant frequencies of singular vowels and vowels at the beginning of word /Female/

$$F_I - F_{31}, \quad F_{II} - F_{32}, F_{III} - F_{33}, F_{IV} - F_{34}$$

From Fig. 2.5 it is seen that for some vowels, for example Y, Ю, Я, Е, difference of frequencies are negative (below zero). This means formant frequencies at the beginning large (increased) than singular vowels one's. For volwe A frequencies are same.

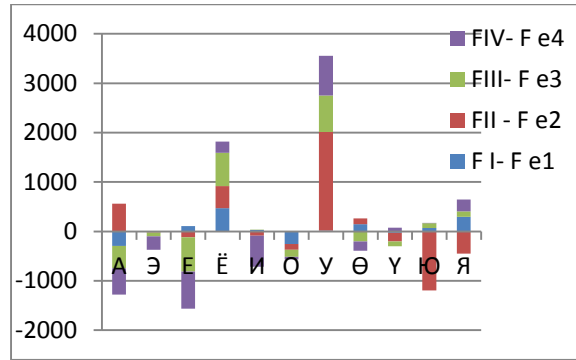


Fig. 2.6. Difference of formant frequencies of singular vowels and vowels at the beginning of word /Male/

$$F_I - F_{e1}, F_{II} - F_{e2}, F_{III} - F_{e3}, F_{IV} - F_{e4}$$

Fig. 2.6 shows specific that all formant frequencies of singular vowels E and Y are positive or large than at the beginning positions one's. For other vowels mostly negative.

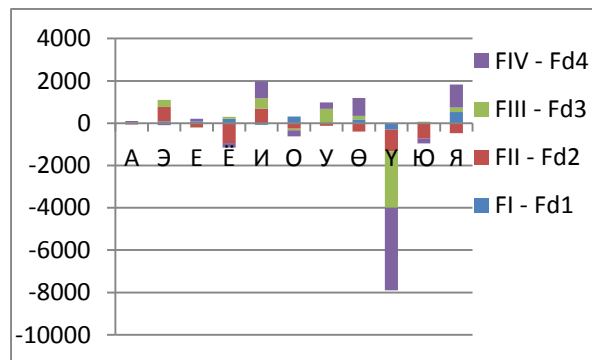


Fig. 2.7. Difference of formant frequencies of singular vowels and vowels at the middle of word /Female/

$$F_I - F_{d1}, F_{II} - F_{d2}, F_{III} - F_{d3}, F_{IV} - F_{d4}$$

From Fig. 2.7 we can see one specific that all formant frequencies of Y vowel pronounced in the middle of word considerably increased comparatively to separately pronounced one's.

For other cases of vowels positions differences of frequencies are positive and negative depending on vowels.

We can say that formant frequencies are changed according to the positions in word.

3.3 Formant parameters of double vowels and diphthons

Comparison of formant frequencies of double vowels shown in Tab. 2.2.

Table 2.2. Comparison formant frequencies of double vowels

| | Formant 1, F_1 | | | Formant 2, F_2 | | |
|----|------------------|------|--------------|------------------|------|--------------|
| | Female | Male | Female /male | Female | Male | Female /male |
| Aa | 1026 | 735 | 1,4 | 1490 | 1267 | 1,2 |
| Ээ | 409 | 376 | 1,08 | 2370 | 1883 | 1,26 |
| Ий | 283 | 268 | 1,05 | 2794 | 2091 | 1,34 |
| Оо | 652 | 743 | 0,9 | 1012 | 2721 | 0,37 |
| Уу | 393 | 493 | 0,8 | 1007 | 2621 | 0,38 |
| Өө | 389 | 430 | 0,9 | 950 | 735 | 1,3 |
| Үү | 398 | 323 | 1,23 | 1216 | 846 | 1,44 |

Continue of Tab.2.2.

| Formant 3, F_3 | | | Formant 4, F_4 | | |
|------------------|------|--------------|------------------|------|--------------|
| Female | Male | Female /male | Female | Male | Female /male |
| 2700 | 2688 | 1,00 | 3983 | 3364 | 1,2 |
| 3035 | 2677 | 1,13 | 4306 | 3160 | 1,36 |
| 3231 | 2940 | 1,1 | 4270 | 3417 | 1,25 |
| 2402 | 3618 | 0,66 | 3604 | 3824 | 0,94 |
| 2453 | 3158 | 0,78 | 3435 | 3541 | 0,97 |
| 2855 | 2490 | 1,15 | 4091 | 3162 | 1,3 |
| 2867 | 2734 | 1,5 | 4075 | 3499 | 1,16 |

From Tab. 2.2. it is seen that first formant frequencies vary in the range 283-1026 Hz for female, and 268-743 Hz for male. Most of female formant frequencies of double vowels are predominant in comparison of male one's respectively 5-40% (1st formant), 20-40% (II formant), 10% - 50% (III formant), 16% - 36% (IV formant). But for some double vowels male formants exceed female formants from 10% to 25%.

With increase of formant number 2nd and 3rd formant frequencies increased 3%-270% depending on double vowels.

Comparison of singular and double vowels shows that formant frequencies for double one's increased comparatively to singular one's.

3.4 Formant frequencies of diphtongs

Comparison of formant frequencies of diphtongs shown in Tab. 2.3.

Table 2.3. Comparison formant frequencies of diphtongs

| | Formant 1 | | | Formant 2 | | |
|----|-------------|------|-----------------|-------------|------|-----------------|
| | Fe- male | Male | Female /male | Fe- male | Male | Female /male |
| Ай | 824 | 529 | 1,56 | 2217 | 1803 | 1,23 |
| Эй | 420 | 348 | 1,2 | 2481 | 2020 | 1,23 |
| Ой | 615 | 442 | 1,4 | 1685 | 1750 | 0,96 |
| Уй | 324 | 309 | 1,05 | 2452 | 1709 | 1,43 |
| Үй | 259 | 280 | 0,9 | 2394 | 2055 | 1,16 |

Continue of Tab.2.3.

| Formant 3 | | | Formant 4 | | |
|-----------|------|--------------|-----------|------|--------------|
| Female | Male | Female /male | Female | Male | Female /male |
| 3078 | 2767 | 1,11 | 4431 | 3273 | 1,35 |
| 3119 | 2685 | 1,16 | 4387 | 3192 | 1,37 |
| 2954 | 2562 | 1,15 | 4112 | 3259 | 1,26 |
| 3112 | 2347 | 1,33 | 4337 | 3314 | 1,3 |
| 2966 | 2602 | 1,14 | 3948 | 3265 | 1,2 |

Most of female diphtong have more formant frequencies in comparison with male one's.

4. CONCLUSION

1. By means of Praat tool determined in detail formant parameters as formant frequencies, formant bandwidth, power spectrum density of vowels and consonants in singular pronunciation case and in different position case in the word. These parameters are the key parameters of vocal tract of speech chain and the purpose of research was the determination of them.
2. Depending on singular pronunciation case and context position pronunciation formant parameters of sounds are changed, but have not deterministic behavior. The changes depend on previous and next sound.
3. By the specific of formant parameters vowels can be divided into groups.
4. Depending on groups of consonants (fricatives, stops, noisy and so on) defined formant parameters of them

5. REFERENCES

1. P. B. Denes and Pinson. The speech chain. W.H. Freeman Company, 2nd Edition, 1993.
2. M. Ghil, C. Taricco. Advanced spectral analysis methods. 1999, 2019
3. Petre Stoica, Randolph Moses. Spectral Analysis of Signals. 2005, Prentice Hall, Inc

4. B. Damdinsuren, G. Demberel, Z. Buyankhishig, B. Sukhbat, CH. Narantungalag and others. Speech recognition technique and software. Report of research work. 2004, Ulaanbaatar.
5. B. Damdinsuren. Digital speech processing. Computerized speech synthesis. Automatic speech recognition. 2017, Ulaanbaatar. “Nom khur” LtD company.
6. L. R. Rabiner. Theory and application of Digital Signal Processing. Translation from English. Moscow, 1978.
7. Application of Digital Signal Processing. Edited by Oppenheim. Translated from English. Moscow, Communication, 1980.
8. L. R. Rabiner., Biing-Hwang Juang. Fundamentals of speech recognition. Prentice Hall PTR Englewood Cliffs. New Jersey
9. Advanced Spectral Analysis *Methods*. <https://www.researchgate.net/publication>. 2019.