International Journal of Multimedia and Ubiquitous Engineering
Vol.11, No.1 (2016), pp.179-190
http://dx.doi.org/10.14257/ijmue.2016.11.1.18

The Acoustical Analysis of Vowel Harmony Issues in Uyghur Words

Seyyare Imam, Gulnur Arkin and Askar Hamdulla*

Xinjiang University, Urumqi, China
**askarhamdulla@gmail.com*

Abstract

The vowel harmony phenomena in Uyghur words, its representation forms and main types are introduced first in this paper. Then, from the experimental phonetics point of view, the feasibility and affectivity of acoustical analysis method for harmony analysis are verified. finally, as the application part, the acoustic features such as formant frequency, resonance peak value, vowel duration, vowel pitch and the sound intensity etc. for every disyllabic, trisyllabic and quad syllabic Uyghur words are collected from the "Acoustical Database for Uyghur Language" established by Laboratory of Institute of Ethnology and Humanities of Chinese Academy of Social Sciences and Xinjiang University. And the statistical analysis carried out on those acoustical feature values and some conclusions and rules are also summarized.

Keywords*: Vowel Harmony, Experimental Phonetics; Acoustic Parameters, Uyghur Language*

1. Introduction

The acoustic feature analysis of vowel harmony is not only an important part of Phonetics, but also a basic research of speech synthesis. Vowel harmony refers to the relationship among vowels on phonetic structure [1]. There are still existing differences on theoretical understanding and explanation for vowel harmony. One view [2] says that vowel harmony is kind of assimilation phenomenon, and vowel harmony is defined as a highly mixing phenomenon between the vowels in all inherited words of a language. On contrast, another view differ the concepts of vowel harmony and mixing phenomenon, for assimilation theory still can't satisfy some questions about vowel harmony. This view believes that new vowel harmony can be found in each vowel contained syllables within a word. Also, the so-called vowel harmony can be defined as the collocation relationship of vowels on phonetic structure [3]. Harmony based on position of tongue and shapes of lip (except for Lop dialect) are the essentials of vowel harmony in Uyghur language, and the characteristics of vowel in word-stem are attributed to the decisive factor in Uyghur vowel harmony [4]. The variants of additional parts which usually are added before or after the word (stem) are relied on the roundness or flatness of lips and the height of tongue position of the first vowel. Generally speaking, the affix with similar vowel characteristic to the vowel in stem is selected to connect the stem. Following is schematic diagram of vowel harmony.

* Corresponding Author

ISSN: 1975-0080 IJMUE

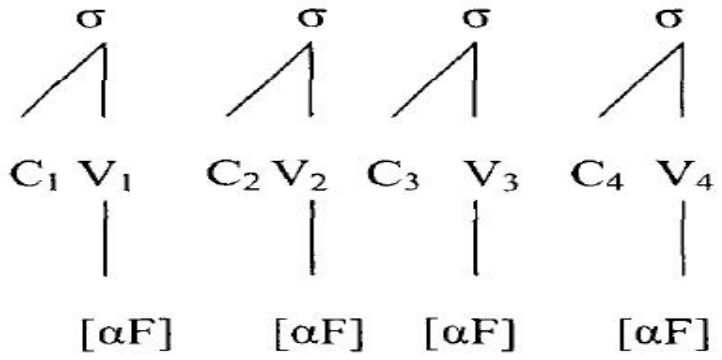

Figure 1. Schematic Diagram of Vowel Harmony

One of the main languages in Altaic language system—Uyghur is an adhesive language, and has its unique rules on acoustic vowel harmony. Foreign scholars (Alling 1999) [5] analyzed Uyghur vowel harmony and neutral vowel phenomenon from different aspects and using different methods. For example, they discussed the harmony rules in Uyghur including vowel harmony, harmonious lips, and vowel displacement phenomenon in their papers. Also, the broadband formant mode, resonance peak (F1, F2), dynamic pronunciation characteristics and acoustic distribution of vowel harmony were considered to analyze the basic acoustic characteristics of vowel harmony. However, the insufficient data, limited nature of former methods and lacking support from Experimental Phonetics has made many acoustic problems in Uyghur vowel harmony research remaining unsolved until today.

Some more intensive research on the harmony phenomenon of Uyghur was made by researchers on this field. Abdukerim Baki's "The modern Uyghur language" systematically demonstrated the Uyghur harmony rules, and discussed the harmony phenomenon between Uyghur vowels and between vowels and consonants [6]. Sulayman Seper made explanation about the harmony phenomenon of vowels and consonants [7]. These research activities were conducted only on the speech characteristics of harmonious words. But, the research on the acoustic properties of Uyghur harmony has still been staying behind. So, the clarity and naturalness of Uyghur speech synthesis and speech recognition has been directly influenced.

Since the establishment of "Uyghur acoustic parameters Database", the statistical analysis on Uyghur acoustic parameters has received its further progress and a number of research papers on Uyghur acoustic features have been published. For example, "Experimental phonetics study of long-vowels in Uyghur language "(Askar Hamdulla), "Formant features of vowel harmonious words in four-syllabic Uyghur words"(Gulnur Arkin, Askar Hamdulla), "Research on the vowel patterns of disyllabic words in Uyghur language "(Aynur Nurtay, Askar Hamdulla), "Experimental Phonetics based Analysis of Semivowels in Uyghur Language" (Hankiz Yilahun, Askar Hamdulla) *etc*. However, the contents of those studies are only a little portion of Uyghur phonetic research, and little research has been conducted on Uyghur acoustic characteristics. Therefore, the study on acoustic features of Uyghur standard vowel harmony has great research significance and potential value.

The research on the acoustic space of Uyghur vowel harmony has limited results. The acoustic study on Uyghur vowel harmony is still subtle and lacks the domestic and foreign research results. Although the research on Uyghur speech synthesis and speech

recognition have got admirable progress, but the efficiency and accuracy problem still couldn't get ideal solution. One main reason for accuracy problem is believed that the understanding of Uyghur acoustic characteristics is not satisfactory enough. So, whether it is considered from the basic research aspect of social needs or from the point of information technology, an intensive study on Uyghur acoustic features can't be delayed.

2. Representation of Uyghur Vowel Harmony

The structure of vowel system decides the types of vowel harmony. According to the characteristics of vowel harmony, harmony phenomenon can be found in two types, which are the harmony between the vowels in stem, and the harmony between the vowels in stem and vowels in affixes. Before explaining these types of harmony phenomenon, observing the phonetic system of Uyghur vowels is advisable.

Every language has its unique phonetic system. For example, one of the very much related languages to Uyghur, Uzbek has phonetic system which consisted of six vowels; While Uyghur has eight vowels in its phonetic system. Uyghur is a popular language for a number of ethnic groups in Xinjiang, also showing important significance in communication for Uyghur people. There are 32 phonemes in modern Uyghur phonetic system [8], including eight vowel phonemes below:

ئا(a) , ئە(e) , ئې(é) , ئى(i) , ئو(o) , ئۆ(ö) , ئۇ(u) , ئۈ(ü)

（a[ɑ] , ä[ɛ] , e[e] , i[i] , o[ɔ] , ö[ø] , u[u] , ü[y]）

Usually, Uyghur vowels are divided into the Front vowels and Back vowels. Front vowel includes a[ɑ], o[ɔ], u[u], while back vowels refer ä[ɛ], ö[ø], ü[y]. Current studies are taking the Central vowel into consideration, too. Therefore, based on the horizontal position of tougne when vowel are pronunced, Uyghur vowels can be analyzed in three categories including Front vowels, Back Vowels and Central vowels. Table 1 shows the vowels in different categories above.

Table 1. Vowel Categories Based on Tongue Position

Front Vowels	Middle Vowels	Back Vowels
ئە(e) , ئۆ(ö) , ئۈ(ü)	ئى(i)	ئا(a), ئې(é), ئو(o), ئۇ(u)

Lip-shape harmony phenomenon also can be found in Uyghur vowel harmony. According to the formed shape of lips when vowels are pronunced, either round or flat, Uyghur vowels can be observed in Round vowel and Flat vowel categories.

Table 2. Rounded/Flattened Vowels

Rounded Vowel	Flattened Vowel
ئو(o), ئۇ(u), ئۆ(ö) , ئۈ(ü)	ئا(a), ئې(é), ئە(e) ,ئى(i)

Summing up the mentioned properties above, Uyghur vowels should be analyzed in multi-dimensional categories shown in Table 3.

Table 3. Overall Uyghur Vowel Categories

Vowel types / Tongue position	Front		Middle		Back	
	Flattened	Rounded	Flattened	Rounded	Flattened	Rounded
High		ﯛ(ü)	ﯩ(i)			ﯘ(u)
-High		ﯚ(ö)			ﯥ(é)	ﯗ(o)
-Low	ﻪ(e)					
Low					ﺎ(a)	

The actual representation of Uyghur vowel harmony usually shows the similarity of vowel characteristics in stem and affix syllables. For example, all vowels in stem and affixes are Front vowels, or all vowels are Back vowels *etc*. In short, vowels are found in same vowel categories in vowel harmony phenomenon. As for Central vowels, they can appear with both of Front and Back vowels together. For example:

tiräk （Pillar） biraq （However）

qeza （Horsemeat sausage） nimä （What）

ilan （Snake）

3. Basic Types of Vowel Harmony

Tongue position harmony is the essential harmony phenomenon of Uyghur vowels. But Lip-shape harmony is usually achieved based on tongue position harmony and Round-lip harmony. A brief analysis of Uyghur vowel harmony is given in following section.

3.1 Vowel Harmony in Stems

In modern Uyghur, the harmonious relationship between the vowels in word-stem is called stem-vowel harmony. If the vowel in the first syllable in the stems is Back vowel (such as ‘a’, ‘o’, ‘u’), and then the vowel in following syllables will also be Back vowel (such as ‘a’ or ‘u’). Similarly, Front vowel (‘ɛ’, ‘ø’, ’y’) in the first syllable of stem is also followed by Front vowel contained syllable.

Front vowel——Front vowel:

ɛtɛ （Tomorrow） yɛttɛ (Seven)

ɛzgy (Opression) dyglɛk (Round)

Yzym （Grapes） øtyk （Boots）

Back vowel——Back vowel:

Alma （Apple） paqa （Frog）

Qosaq （Belly） buqa （Bull）

Oruq （Thin） bulut （Cloud）

In above examples,' ɛ-ɛ, y-y, i-i, a-a, u-u' vowel groups keep strict harmony relationship due to the same vowel characteristics, which is usually called full-harmony. Other harmony groups can't satisfy the full-harmony requirements, so it is called semi-harmony. Semi-harmony relationship is formed between the vowels of similar characteristics. The above harmony relationships are also applied to the relationship between the vowels in stem and affixes.

3.2 Vowel Harmony between the Vowels in Stem and Affixes

The harmonious relationship—the mutual adaptation and assimilation between the vowels in stem and affixes is called the vowel harmony between stems and affixes. Vowel harmony plays an important role in appropriate connecting of affixes to word-stem. The vowel in affix is always constrained by the vowel in the last syllable of stem. For example, as for the two variants of name plural forms "lar, lɛr", if the vowel in the last syllable of stem is 'a' which is back vowel, 'lar' is chosen, because the vowel 'a' in plural form is Back vowel. In the same way, if Front vowel 'ɛ' is in the last syllable of stem, the Front vowel contained variant of plural 'lɛr' is chosen. For example:

Kit**a**b + lAr = kitabl**a**r (Book--Books)
[Back vowel] [Back vowel]

ɛynɛk + lAr = ɛynɛklɛr （Glass---Glasses）
[Front vowel] [Front vowel]

These harmony rules are found in conditional tone affixes sA （-sa, -sɛ） and the verb negativity affix mA (-ma, -mɛ) too. The affixes which keep the tongue position harmony rules are quite popular in Uyghur language. There are more examples:

Y**a**z + mA = yazm**a** （Write----Don't write）
[Back vowel] [Back vowel]

K**ɛ**l + mA = kɛlm**ɛ** （Come----Don't come）
[Front vowel] [Front vowel]

m**a**ŋ（walk） + sA = maŋs**a** （Walk--- If walk）
[Back vowel] [Back vowel]

ɛt（do） + sA = ɛts**ɛ** （Do--- If do）
[Front vowel] [Front vowel]

In addition to above affixes types, the position affixes, the direction affixes and inferiority affixes are also keep the tongue position harmony rules. Above examples show the fact that affixes in Uyghur always have two types of variants due to tongue position vowel harmony. However, in actual speaking and writing, vowel harmony in affixes and stem is often found unstable.

In modern Uyghur, some affixes only carry one form and can be connected to stem without considering its vowel characteristics. For example, objective affix '-ni ', word formation functional affixes 'tʃi', 'siz' *etc*.

3.3 The Acoustic Parameters of Vowel

The acoustic parameters of vowel (Duration, pitch, intensity and formant) are of great significance in speech studies. The extraction of vowel acoustic parameters is also basic task in speech research. Due to the complexity of voice parameter representation and lacking the consensus on acoustic parameter extraction criteria, the extraction standard of acoustic parameters of one specific language has been quite difficult. The non-unified criteria make referencing between languages is un-applicable and can't prevent the extraction of acoustic parameters by configuration of self constructed local systems. This is not good for sharing of knowledge, and in-depth comparative study between languages.

3.3.1 Formants of Vowel

Study on the formant feature of vowels refers to the timber characteristics of the voice. Vowel is usually measured by five formant parameters. Among them, pramameter value F1 and F2 are the main inlfuencing factors for voice timber, and F3 has the clarification (sharpening) role for high Front vowel /i/, /y/. Effect of F4 and F5 has not been clarified yet. But, combined experiments obtained the fact that F4 and F5 value has contribution to the polishing the timber of voices. Usually, the formant feature is found at the poles of the vocal tract transmission frequency response, as shown in Figure 2.

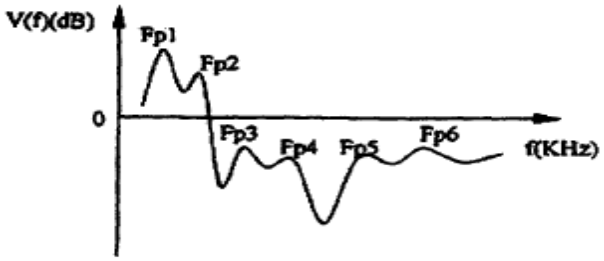

Figure 2. Frequency Responses

Vowel articulation is produced by vibrating of vocal cord, and has some regular patterns. Ups and downs are generated when it goes through the vocal tract. The frequencies of same or similar values to the inherent frequency of vocal tract will get strengthened, while other frequencies will weakened and even diminished at last. On the three dimensional spectrogram, the strengthened frequency components are represented as darker 'formant bar'. The frequecy components are arranged in frequency axis in increasing order. For example, Formant parametes are found as the lowest frequency (the first formant) F1 → F2 and F3 *etc*.

Formant parameters and tone quality of vowel are in continuous changing mode. So, measuring formant parameters of vowel requires a target point. The following points are noted from our experience:

- Vowel /a/ produces the highest frequency component as it holds the biggest amplitude of mouth opening. The moment at the highest value on the continuous

changing first formant curve is chosen for target point, and then we can get F2, F3, and F4 easily.

- Since /i/ is highest Front vowel, the first formant value of /i/ is the lowest one while the second formant value is the highest. So, the target point is taken at the point after two longest distances on the changing formant curve.
- The highest Back vowel /u/ produces the lowest F1 value and the minimum distance between F1 and F2. So, F1 is used as the lowest formant curve, and the nearest point around F1 and F2 is chosen for target point.

In acoustics, size of physiological organs such as nose, throat and mouth will influence the formants regional responses. Distiguishing of different pronunciations and vowels is mainly relied on distributed formant values. Red line in Figure 3 shows the central formant.

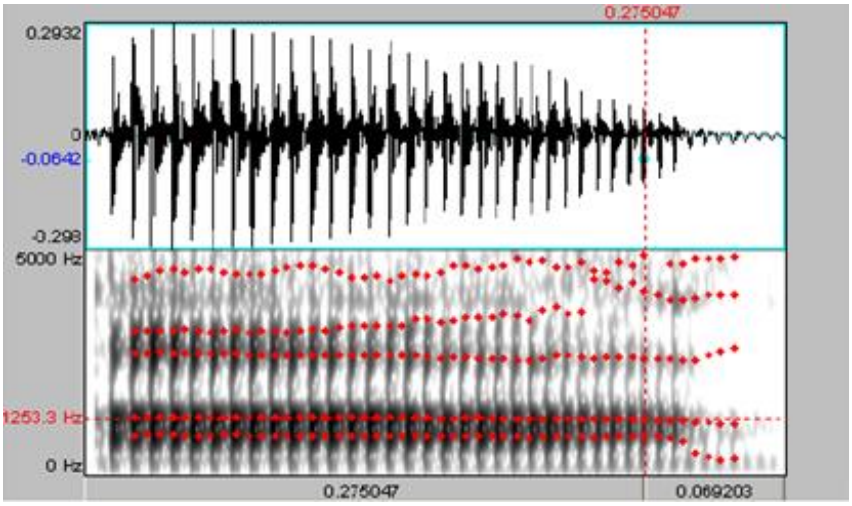

Figure 3. Central Formant

3.3.2 Duration of Vowel

Duration is measured by the length of each waveform (Unit: MS). Vowel Duration (VD) always plays the feature distinguishing role in vowel acoustic study. In modern Uyghur acoustics, the time length of vowels in each syllable determine the vowel duration.

3.3.3 Sound Intensity

The so-called sound intensity (Intensity) refers to the strength of dominant tone in speech signal, also called Short-time Average Energy.

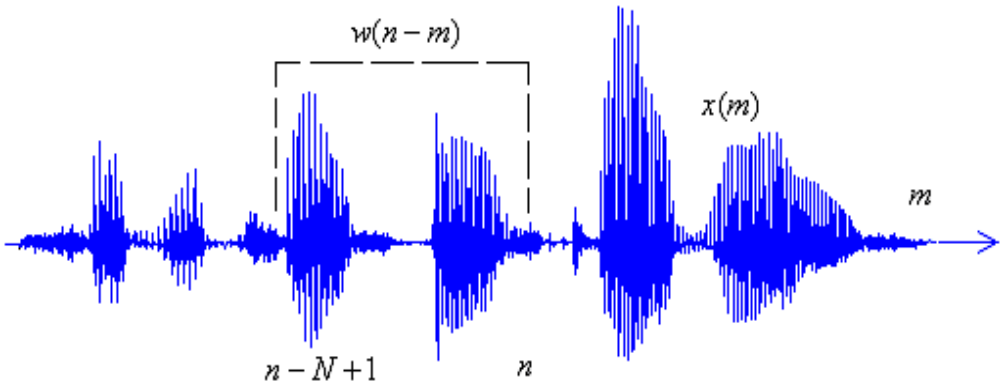

Figure 4. Speech Intensity

Intensity represents the strength of voice, also called "Level of intensity" or "Energy", its unit is dB, it is measured by the strength at vowel target point.

Intensity characterized by the properties as follows:

- As for one corresponding target point, the influence from Microphone is big.
- In order to prevent the error DC bias of speech signal, the average value of speech signal should be subtracted before measuring.

3.3.4 Pitch

Pitch is another very important acoustical parameter of speech signal and basic characteristic of voice. Pitch is a perceptual property of sounds that allows their ordering on a frequency-related scale or more commonly, pitch is the quality that makes it possible to judge sounds as "higher" and "lower" in the sense associated with musical melodies. Pitch can only be determined in sounds that have a frequency that is clear and stable enough to be distinguishable from noise. Pitch is a major auditory attribute of musical tones, along with duration, loudness, and timbre.[9]

4. Experimental Data and Measurements

4.1. The Source of Words

The data of the Uyghur Language Acoustical Database has been built by Speech Laboratory of Institute of ethnology and humanities of Chinese Academy of Social Sciences and Xinjiang University. The research of Uyghur Language Acoustical Database is an information technology project organized by the National Languages Committee and Ministry of Education. We collect the total of 1395 harmony words composed of two syllabic words subtotal of 969, three syllabic words subtotal of 333 and four syllabic words subtotal of 93. And extract the acoustical features such as duration, formant, intensity, and pitch values *etc*. for each word of them. The averaged acoustic values are given in Table 1 below.

According to the average values of those acoustic feature values, we have conduct the statistical analysis on for vowels positioned in front syllable, second syllable, third syllable, or last syllable of the words. Then, we have summarized the basic acoustical distributions, and obtain some rules and conclusions.

4.2. Recording the Words to Speeches and Extracting the Acoustical Parameters

We recorded the voices of one male and one female speaker, from 35 to 40 years old and both of them are professional announcer of The Central People's Broadcasting Station, with IBM R51 type notebook PCs and a creative sound blaster card in the standard recording studio of Chinese Academy of Social Sciences Institute of Ethnology and Anthropology. They read each single word twice. The authors extracted acoustic parameters of harmony words with speech analysis software --- Pratt. Such as for each syllable and the length of vowel (unit: Ms); for each syllable the author took three points: the starting point, turning point and end point of the pitch as the target value (unit: Hz), And for each syllable the acquisition of the strongest point is the intensity target value (unit: dB).The first four formant vowels in vowel F1-F4 measurements are on the vowel target position. The target position is typical, strong energy part of the vowel formant pattern (unit: Hz).

Table 1. Averaged Acoustical Feature Values of Vowels in Uyghur Words

Words / Parameters		Two syllabic words			three syllabic words			Four syllabic words		
		duration (ms)	intensity （dB）	pitch （Hz）	duration (ms)	intensity （dB）	pitch （Hz）	duration (ms)	intensity （dB）	pitch （Hz）
First syllable	F	103	70	223	88	70	225	93	70	220
	M	100	72	101	82	65	91	97	71	98
Second syllable	F	146	72	227	95	71	234	66	69	231
	M	120	73	114	65	68	111	65	72	122
Third syllable	F				149	72	224	74	69	224
	M				108	70	109	65	70	109
Forth syllable	F							145	68	205
	M							122	68	88

5. Discussions and Conclusions

In this paper the author only concentrated on the basic acoustic features of Uyghur harmonious words, and the vowel duration, sound intensity and pitch are the main acoustic features in vowel harmony analysis. Their distribution mode varies by language [8]. And we draw the following conclusions:

(1) Vowel harmony features are diffused in fixed direction, from the first syllable to ending syllables, also from bottom to top or from top to bottom. The vowel characteristics in first syllable determine the characteristics of the vowel in last syllable.

(2) From the distribution of duration values of disyllabic, two syllabic and quad syllabic words, we can conclude that the duration time of first syllable in two syllabic harmony words always longer than others. Meanwhile, the duration time of middle syllable and last syllable in tri-syllabic harmony words always longer than others.

(3) From the distribution of intensity values of disyllabic and two syllabic words, their peak values positioned at last syllable whether female or male.

(4) The pith values distributed in range of from 205Hz to 234Hz for female speaker and from 88Hz to 122Hz to male speaker. This further verified the existence of

vowel harmony feature and one directional diffusion characteristics in Uyghur words.

Acknowledgment

This work was supported by the National Social Science Foundation of China (13BYY062) and Natural Science Foundation of China (61063023).

References

[1] X. R. Zhao and Z. N. Zhu, "The Uyghur Jian Zhi", National press, **(1985)**, pp.16-21.
[2] "Ci Hai", Shanghai Lexicographical Publishing House, **(1980)**.
[3] K. A. Reyiyi, A. Litip and Y. Ebeidulla, "Modern Uyghur Language", Urumqi: Xinjiang People's Publishing House, **(2003)**, pp. 92-110.
[4] Y. Wei, "A Phonological Study of Harmony Phenomena in Uyghur", Ph.D. Thesis, East China Normal University, **(2012)**.
[5] A. Emily, "Uyghur vowel harmony", Manuscript, Harvard University, **(1999)**.
[6] A. Baki, "Modern Uyghur", Ethnic Publishing House, **(1983)**.
[7] S. Seper, "Contemporary Uyghur", Urumqi: Xinjiang People's Publishing House, **(2000)**.
[8] H. Niyaz, O. Hapiz, G. Islam and Bahargul, "Basic Uyghur", Urumqi: Xinjiang University Publishing House, **(1997)**, pp:86-88.
[9] https://en.wikipedia.org/wiki/Pitch_(music)
[10] B. Yi, "Forms and Characteristics of Modern Uyghur vowel harmony", Journal of national language, no. 2, **(2006)**, pp. 51-53.
[11] Y. Wei and J. Zhang, "Auto segmentation analysis of Uyghur vowel harmony", Language Sciences, vol. 10, **(2011)**, pp. 511-517.
[12] B. Yi, "Experimental phonetics research of vowels in modern Uyghur", China Social Science Press, **(2012)**, pp. 92-93.
[13] M. Litip, "Acoustics parameters and prosodic features analysis of multilevel Uyghur speech units", Master's Thesis, Xinjiang University, **(2009)**, pp.7-8.
[14] D. Tursun and A. Hamdulla, "Experimental Phonetics Based Analysis of Devoiced Vowels in Uyghur Language", Journal of Chinese Information Processing, no. 5, **(2010)**, pp. 117-123.
[15] http://www.fon.hum.uva.n1/praat/
[16] Z. Aman and A. Hamdulla, "Acoustic Analysis of the Prosodic Features of the Tri-syllabic Words in Uyghur language," Journal of Chinese information, vol. 5, **(2009)**, pp. 104-107.
[17] H. Qiao and B. Aximu, "Acoustic phonetics based Uyghur vowel analysis", Ethnic Chinese, vol. 5, pp. **(1988)**, pp. 4-13.
[18] R. Dawut, A. Hamdulla and D. Tursun, "Experimental phonetics research of long vowels in Uyghur language", Communication technology, vol. 1, **(2010)**, pp. 5-7.

Authors

Seyyare Imam, She received her B.E. and M.S. degree in politics and philosophy, from Xinjiang University and Xinjiang Normal University of China, in 1995 and 2008, respectively. Since 1996, she has been working as a teacher in politics and public administration, Xinjiang University. Her research interests include Uyghur Phonetics and Linguistics, social language, and related issues about Uyghur ontology construction *etc*.

Gulnur Arkin, She has received her B.E. and M.S. degree in Communication Engineering from Xinjiang University, China, in 2011 and 2014, respectively. Currently, she is pursuing her PhD degree in Computer Applications Technology in Xinjiang University, China. Her research interests include Experimental Phonetics and Uyghur Linguistics.

Askar Hamdulla, He received B.E. in 1996, M.E. in 1999, and Ph.D. in 2003, all in Information Science and Engineering, from University of Electronic Science and Technology of China. In 2010, he was a visiting scholar at Center for Signal and Image Processing, Georgia Institute of Technology, GA, USA. Currently, he is a professor in the School of Software Engineering, Xinjiang University. He has published more than 140 technical papers on speech synthesis, natural language processing and image processing. He is a senior member of CCF and an affiliate member of IEEE.

