

The Vowel Harmony in Two-syllabic Uyghur Words and its Acoustical Analysis

Gulnur Arkin¹, Seyyare Imam², Akbar Pattar¹ and Askar Hamdulla^{1*}

¹*Institute of Information Science and Engineering, Xinjiang University, China*

²*School of Politics and Public Administration, Xinjiang University, China*
askarhamdulla@sina.com

Abstract

A statistical analysis has been made on 969 two-syllabic words in Uyghur Acoustical Database. The major part of this paper concentrates on the basic acoustic features of two-syllabic harmonious words including formant frequency, resonance peak value, vowel duration, vowel pitch and the sound intensity. And we have summarized certain rules and conclusions of great importance for adjusting vowel harmony in parametric speech synthesis or waveform concatenated speech synthesis system.

Keywords: *Two-syllabic Word, Vowel Harmony, Acoustic Analysis, Uyghur Language*

1. Introduction

The acoustic feature analysis of vowel harmony in Uyghur words 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]. The Uyghur language belongs to Turkic branch of the Altaic language family. The Uyghur vowel harmony consists of the vowel harmony in stems and the vowel harmony between stem and affixes [2]. They obey similar rules and contain two basic harmony types such as the harmony in lips (consistency of being round or not round) and tongue position (consistency of being in front or behind).

Uyghur vowel harmony and Schwa phenomenon (neutral vowel phenomenon) have been researched on by many scholars (Alling 1999, Coorie 1997) [3] in different methods. For instance, the Uyghur phonetic harmony rules, including harmony in lips and tongue position and vowel deflection phenomenon, and the broadband formant pattern, formant value (F1, F2), and the basic acoustic features of vowel harmony according to the distribution of acoustic vowels [4]. However, due to the limitations of method, the amount of data, and the lack of experimental phonetics support, vowel harmony acoustic study has not yet achieved a distinctive progress. This has been seriously affected the teaching, researching and speech acoustic works about the Uyghur language.

There are thirty-two phonemes in the phonological system of the modern Uyghur standard language. Eight of are vowel phonemes [5]. They are: a [a], ä [ɛ], e [e], i [i], o [ɔ], ö [ø], u [u], ü [y]. Uyghur vowel harmony is usually expressed as following phenomena:

Consistency in tongue position: Vowels in stem and affixes are all the frontal vowels or the back vowels.

Consistency in shape of lips: Vowels in stem and affixes are all the vowels that lips will be round or not round. In addition, these neutral vowels (i, e) can emerge with front vowels simultaneously and sometimes with back vowels.

In recent years, with the development of phonetics and computer technology, speech synthesis technology of Uyghur language has made a tremendous progress. The speech

* Corresponding Author

synthesis technology can synthesize with higher clarity and better speech intelligibility. However, the naturalness of synthesized speech still needs a further study. In paper, based on the experimental phonetics and using the basic acoustic features of vowels the author investigated the acoustic Features of vowel harmony in two-syllabic Uyghur words. In this paper the author considers the acoustic performance and characteristics of vowel harmony, which includes: formant frequency, resonance peak value, vowel duration, vowel pitch and the sound intensity, and further research on the distribution pattern and regularities in Uyghur words, and the influence on the harmony of Two-syllabic words. This paper summarized certain rules which can be important reference to improving the naturalness of Uyghur speech synthesis technology simultaneously.

2. Experimental Data and Measurements

2.1. The Source of Two-syllabic Words

Uyghur is one of the main languages in Altaic language family with distinct agglutinative character. Uyghur has special vowel harmony in phonology [6]. This paper has done statistic on 969 two-syllabic words, for example, Q-Q type: e+ta=eta(tomorrow), H-H type: a+na=ana(mother), Z-H/Q type ik+ra=ikra(admit) / i+ga=iga (possess) respectively, and studied their basic acoustic features. The data of the Uyghur Language Acoustical Database has been built by voice 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.

2.2. Recording the Two-syllabic Words and Extracting the Acoustical Parameters

The author 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 author 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).

3. Analysis on the Distribution of Acoustic Parameters

In order to understand the acoustic characteristics of vowel harmony comprehensively, the author studied two different situations: front vowels and vowel harmony (denoted as Q-Q) and back vowels and vowel harmony (denoted as H-H).

In the case of front vowels and vowel harmony(denoted as Q-Q), the following four cases are studied separately: (1) front and not round –front and not round (abbreviated as QBY-QBY), (2) front and round – front and round (abbreviated as QY-QY), (3) front and not round - the front and round (abbreviated as QBY-QY), (4) front and round –front and not round acoustic characteristics (abbreviated as QY-QBY). For the case of back vowels and vowel harmony (abbreviated as H-H), the authors have studied another four different situations. They are: (1) back and not round –back and not round (abbreviated as HBY-HBY); (2) back and round – back and round (abbreviated as HY-HY); (3) back and not round –back and round (abbreviated as HBY-HY); (4) back and round –back and not

round (abbreviated as HY-HBY). The acoustic characteristics of neutral vowels and front/back vowels (abbreviated as Z-Q/H) are taken into consideration as well. Finally, the vowel harmonies of the acoustic features such as duration pitch and sound intensity distribution pattern and formant frequency values have been analyzed.

3.1. The Distribution Mode of Vowel Duration

Figures.1, 2 and 3 are the distribution diagram of average vowel duration of two-syllabic harmony words (black dash line for female, black solid line for male). Figures.1~3 showed that the duration of Q-Q vowel harmony or H-H vowel harmony distribution mode of the male or /and female. From the Figures we can see that the word-end vowel duration value is higher than the word-head vowel duration value. In each harmony forms, the average duration value of female is higher than that of male.

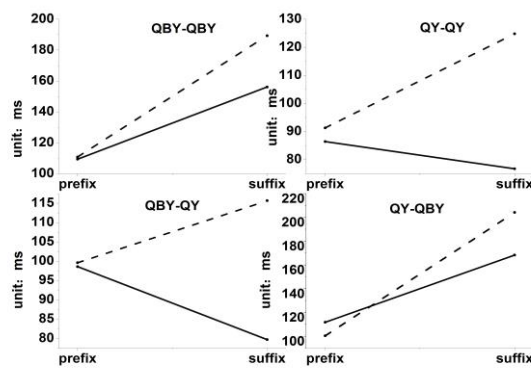


Figure 1. Q-Q Type Harmonious Word Length Distribution Mode

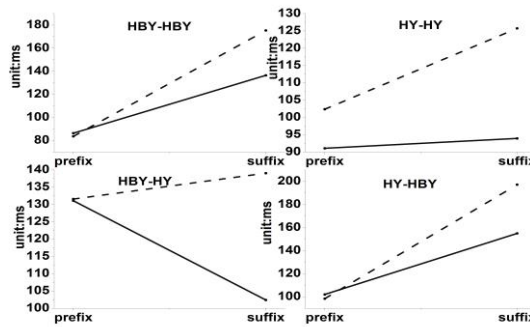


Figure 2. H-H Type Harmonious Word Duration Distribution Mode

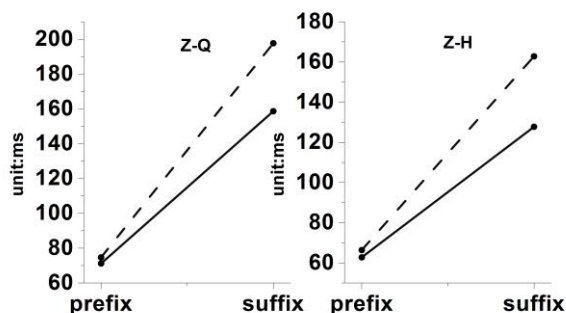


Figure 3. Z-Q/H Type Harmonious Word Duration Distribution Mode

For male in QBY-QY and HBY-HY harmony forms, the word-head vowel duration value is significantly higher than word-end vowel duration, and vowel duration distribution of QY-QY harmonious words is basically similar to the duration distribution of QBY-QY harmonious words. Characteristics of Z - Q and Z - H type harmonious words duration distribution patterns are almost the same. From above analyses we can conclude following rules:

- a) The first syllable vowel characteristics determine the characteristics of the last syllable vowel, and the vowel harmony curve extends in a fixed direction.
- b) Whether the stem vowel is a front vowel or a back vowel, neutral vowel has no alternation.

3.2. The Distribution Mode of Pitch

Figures.4, 5 and 6 are distribution diagram of average vowel pitch of two-syllabic harmonious words (black dash line for female, black solid for male). Either for female or male, the pitch distribution can be described as “lower at first then becoming higher”, denoted by "L - H model", and has high position of pitch curve. The same characteristics appeared in the QY-QBY or /and HY-HBY type harmonious word for the female. From the Figs.4, 5 and 6 we can see that, the curve of pitch characteristics varies from high at the beginning, to a lower position and have a bottom value. For female the Z-H characteristic of pitch distribution and QY-QBY or/and HY-HBY type of harmonious words are basically the same. From the above analyses we can also conclude following rules:

- a) Vowel harmony features are diffused in fixed direction, from the first syllable to ending syllable, also from down to up or from up to down.
- b) Neutral vowels neither promote nor block other vowel harmony process. Therefore, neutral vowel is transparent in the Uyghur language.

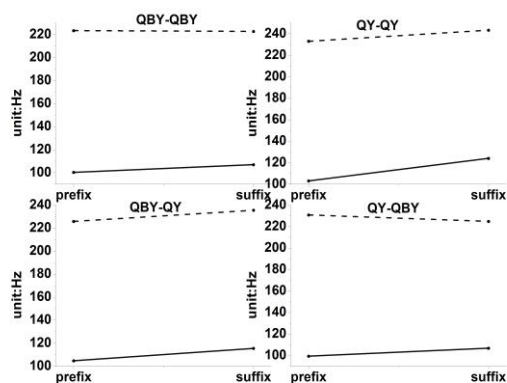


Figure 4. Q-Q Type Harmonious Word Pitch Distribution Mode

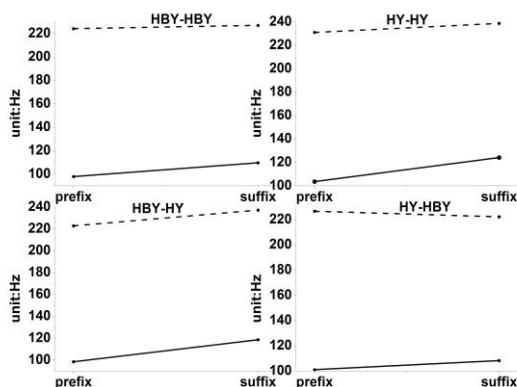


Figure 5. H-H Type Harmonious Word Pitch Distribution Mode

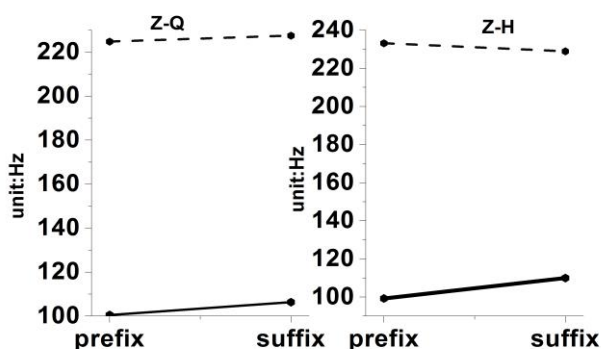


Figure 6. Z-Q/H Type Harmonious Word Pitch Distribution Mode

3.3. The Distribution Mode of Sound Intensity

The distribution diagram of average sound intensity of two-syllabic harmony words is shown in figures.7, 8 and 9 (black dash line for female, black solid line for male).

It can be seen from the figs.7, 8 and 9 that, whether it is front vowel or a back vowel harmony for male or/and female the word-head syllables sound intensity stronger than the word-end syllables. And the female has stronger sound intensity in the word-head syllables than male.

From the fig.7 we can find that the Q-Q type harmonious words, that is QBY-QBY, QBY-QY and QY-QBY, the male has slightly higher sound intensity of word-end syllables than female. From the fig.8 we can see that the sound intensity distribution characteristics of HBY-HBY or/and HY-HBY harmonious words are similar to the distribution characteristics of QBY-QBY harmonious words for male. The Z-Q / H type of harmonious word is shown in fig.9. From the fig.9 we can see that whether it is male or female the word-end vowel sound intensity is stronger than the word-head. However, female word-head sound intensity stronger than the male vice versa case, i.e., for the word-end syllables.

By the way we have also concluded following rules of harmonious words that:

- a) Both in front vowel harmony or back vowel harmony, lip harmonious phenomena will appear only after the building of tongue position harmony.
- b) Neutral vowels have no alternations

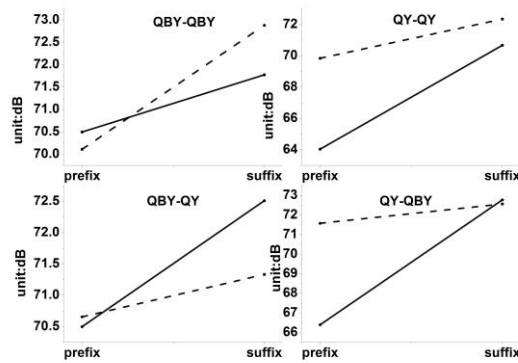


Figure 7. Q-Q Type Harmonious Word Intensity Distribution Mode

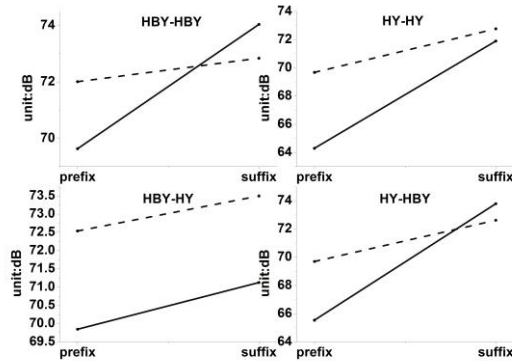


Figure 8. H-H Type Harmonious Word Intensity Distribution Mode

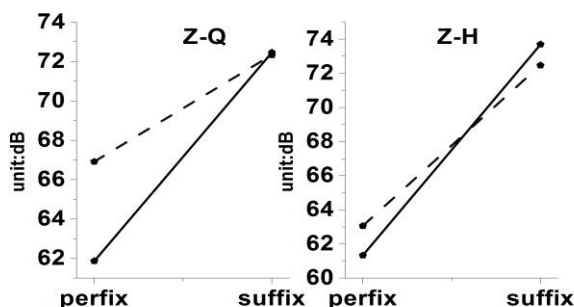


Figure 9. Z-Q/H Type Harmonious Word Intensity Distribution Mode

4. Discussions and Conclusions

Author names and affiliations are to be centered beneath the title and printed in Times New Roman 12-point, non-boldface type. (See example below)

Before giving a summary of this paper we can briefly discuss the vowel harmony about Uyghur languages. The vowel harmony almost exists all Uyghur words. However, there are some exceptional phenomena that out of those rules. For example:

il“hanging”+mAk=ilmaq“hang up”

il“hanging”+mAk=ilmaq“hook”

It can be said that the exceptions of Uyghur vowel harmony are mostly occurred in the combination of a few stems with suffixes. And the stems usually come with the vowels [i] and [e]. Why can be said so? That has much to do with the historical evolution of Uyghur vowels, Uyghur phonological system and the system of orthography.

In this paper the author only concentrated on the basic acoustic features of Two-syllabic Uyghur harmonious words, the vowel duration, sound intensity and pitch are the main acoustic features in vowel harmony analysis. Their distribution mode varies by language [8]. In order to get objective and accurate studying on acoustic features of two syllabic Uyghur harmonious words, the author measured first and second formant of testing word from all harmonious types and made statistical analysis on formant frequency, and all vowel formant data are extracted on the areas where vowels are relatively stable. The discrete degree is an important index (especially for the formant). Usually the ratio (σ / x) is used as the "dispersion" to measure the distance from the variables to the average value. (σ represents the standard deviation and x for the average value). If discrete degree is greater than ten percent, the vowel can be considered as another vowel [9]. The following table is the statistical results of the word-head syllables and word-end syllables formant frequency with different harmony type. The formant and position of the tongue are the basic theoretical issues on acoustic phonetics, many scholars studied the problem from different fields (physical, acoustic, auditory) [10].

From the Table.1 and 2 we can easily see that, whether it is Q-Q or H-H harmonious the formant form, formant values are concentrated and discrete degree is less than ten percent. Z-Q/H type harmony also has similar properties. By this, in different phonetic environment, we believe that altering tongue position does not cause the changing on sound quality of vowels in harmonious form.

Table 1. Formant Frequency Statistics (F1)

Type of harmony	Formant	F1			
	Syllable	The average (Hz)	Standard deviation	Sampling points (a)	Discrete degree (%)
QBY-QBY	first	672	57.5	82	8.5
	last	710	63.7	82	8.9
QY-QY	first	450	34.5	52	7.6
	last	476	21.0	52	4.4
QBY-QY	first	634	33.5	7	5.2
	last	466	24.3	7	5.2
QY-QBY	first	428	31.3	54	7.3
	last	695	45.5	54	6.5
Z-Q	first	406	32.4	32	7.9
	last	700	56.5	32	8.0
HBY-HBY	first	706	50.4	321	7.1
	last	775	66.6	321	8.5
HY-HY	first	488	30.4	91	6.2
	last	493	20.3	91	4.1
HBY-HY	first	715	63.4	27	8.8
	last	491	25.3	27	5.1
HY-HBY	first	473	38.3	167	8.0
	last	792	44.3	167	5.5
Z-H	first	502	43.2	38	8.6
	last	766	54.9	38	7.1

In a summary, in this paper, we have investigated the basic acoustic features of Two-syllabic Uyghur harmonious words. We consider the acoustic performance and characteristics of vowel harmony. Including the formant frequency, resonance peak value, vowel duration, vowel pitch and the sound intensity, and further research on their distribution pattern and regularities in Uyghur words, also their influence on the harmony of Two-syllabic words. And we draw the following conclusions:

- (1) The first syllable vowel characteristics determine the characteristics of last syllable vowel, vowel harmony features passes in fixed direction.
- (2) Whether the stem vowel is a front vowel or back vowel, neutral vowel has no alternation. We presented with some characteristics of harmonious words that: both in front vowel harmony or back vowel harmony, lip harmonious features appear only after the building of tongue position harmony.
- (3) Vowel harmony features are diffused in fixed direction, from the first syllable to ending syllables, also from down to up or from up to down.
- (4) Neutral vowels neither promote nor block other vowel harmony process; therefore, neutral vowel is transparent in the Uyghur language.
- (5) Formant parameters are more concentrated, discrete degree is less than ten percent.

We hope the summarized certain rules in this paper which could be useful for the improving the naturalness of Uyghur speech synthesis technology.

Table 2. Formant Frequency Statistics (F2)

Type of harmony	Formant Syllable	F2			
		The average (Hz)	Standard deviation	Sampling points (a)	Discrete degree (%)
QBY-QBY	first	2051	82.1	82	4.0
	last	2051	82.1	82	4.0
QY-QY	first	2044	89.4	82	4.3
	last	1990	69.6	52	3.4
QBY-QY	first	1947	77.3	52	3.9
	last	2027	99.7	7	4.9
QY-QBY	first	1972	75.2	7	3.8
	last	1968	79.4	54	4.0
Z-Q	first	2033	83.3	54	4.0
	last	2398	88.6	32	3.6
HBY-HBY	first	2059	92.5	32	4.5
	last	1431	90.5	321	6.3
HY-HY	first	1524	85.5	321	5.6
	last	1156	90.2	91	7.8
HBY-HY	first	1330	89.6	91	6.7
	last	1346	93.6	27	6.9
HY-HBY	first	1192	97.2	27	8.1
	last	1141	86.7	167	7.5
Z-H	first	1531	78.4	167	5.1
	last	2097	92.6	38	4.4

Acknowledgments

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

References

- [1] X. Ru. Zhao, Z. Zhu, "The Uyghur Jian Zhi", National press, (1985), pp.16-21.
- [2] B.Yi, "Forms and Characteristics of Modern Uyghur vowel harmony", Journal of national language, no.2, (2006), pp.51-53.
- [3] Y. Wei and J. Zhang, "Auto segmentation analysis of Uyghur vowel harmony", Language Sciences, no.10, (2011), pp.511-517.
- [4] B. Yi, "Experimental phonetics research of vowels in modern Uyghur", China Social Science Press, (2012), pp. 92-93.
- [5] M. Litip, "Acoustics parameters and prosodic features analysis of multilevel Uyghur speech units", Master's Thesis, Xinjiang University, (2009), pp.7-8.
- [6] D. Tursun, A.Hamdulla, "Experimental Phonetics Based Analysis of Devoiced Vowels in Uyghur Language", Journal of Chinese Information Processing, no.5, (2010), pp.117-123.
- [7] <http://www.fon.hum.uva.nl/praat/>
- [8] Zulpiya. Aman, Askar. Hamdulla, "Acoustic Analysis of the Prosodic Features of the Tri-syllabic Words in Uyghur language," Journal of Chinese information, no.5, (2009), pp.104-107.
- [9] H. qiao. B. Aximu, "Acoustic phonetics based Uyghur vowel analysis," Ethnic Chinese, no.5, (1988), pp.4-13.
- [10] Rana. Dawut, Askar. Hamdulla and Dlimurat. Tursun, "Experimental phonetics research of long vowels in Uyghur language," Communication technology, no.1, (2010), pp.5-7.

Authors



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 in Xinjiang University, China. Her research interests include Experimental Phonetics and Uyghur Linguistics.



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..



Akbar Pattar, He received his B.E. degree in radio electronics from Xinjiang University, China, in 1983. He has been working as a teacher in School of Information Science and Engineering, Xinjiang University since 1983. His research interests include experimental phonetics and natural language processing.



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, Xinjiang University. He has published more than 160 technical papers on speech synthesis, natural language processing and image processing. He is a senior member of CCF and an affiliate member of IEEE.