EXTRACTION OF THE ACOUSTIC FEATURES OF SEMI-VOWELS IN THE KURDISH LANGUAGE

Özlem BATUR DİNLER
Siirt University, Department of Computer Engineering, Siirt-Turkey
o.b.dinler@siirt.edu.tr

Nizamettin AYDIN
Yildiz Technical University, Department of Computer Engineering, Istanbul-Turkey
nizamettin@ce.yildiz.edu.tr

Abstract: In this study, it was aimed to extract the acoustic features of semi-vowels in the Kurdish language spoken in Turkey at the phoneme level. For this purpose, voice recordings were collected from 42 adult speakers who know how to read and write in Kurdish. The isolated voice signals of the data set generated with these recordings were obtained by Praat software. The acoustic features of the isolated semi-vowels obtained were acquired by extracting the F1, F2, F3 and F4 formant frequencies in the Matlab environment using the LPC method. Semi-vowels, the phonetic analysis of which was examined, were then addressed with vowels at the end of this study, and their similarities and differences in terms of the phonetic feature were revealed according to the examination results of the analyzed data. This is the first study in which the acoustic analysis of Kurdish semi-vowels was performed. This study will contribute to the researchers who will carry out studies on Kurdish voice processing systems in the future.
Keywords: Speech processing, Kurdish, Phonetic feature, Formant frequency analysis, Lpc, Praat.

Introduction

Kurdish is an Iranian language of the Indo-Iranian language group of the Indo-European language family in terms of the language family and is an inflected language when examined in terms of the structure. Kurdish communities speak many different dialects, including Kurmanji, Sorani, Lori, Gorani, and Zazaki, according to the regions they live in and the tribes and principalities they depend on (Ciwan., 1992). In this study, the Kurmanji dialect, the most spoken dialect of the Kurdish language, is taken as a basis. Kurdish is the most spoken language in Turkey among the living languages following the Turkish language. Therefore, in this study, the acoustic structure of the phonemes in Kurdish was examined in order to develop the voice processing applications such as speech recognition, speaker recognition, voice to text, and text to voice, in the Kurdish language using biometric data consisting of Kurdish sound signals for the determination of crimes committed in Turkey. The examinations were carried out with the sorting of the formant frequencies of the isolated semivowel phonemes in Kurdish.

Formant frequencies are the resonance frequencies of the vocal tract that are expressed as the peak amplitude in the frequency spectrum of the sound. With these frequencies, the most important acoustic properties of semivowels were obtained.

Formation Process of Semivowels and Vowels

The speech signal is produced as a result of the vibrating of airflow coming from the lungs in vocal cords and passing of the vibrations through the throat, oral cavity, and nasal cavity and forming in the tongue, teeth, jaws, palates, and lips, the articulator structures. The smallest sound units that make up a spoken language consist of vowels and consonants. Vowels are the voiced sounds produced by the oscillation as a result of the vibrating of vocal cords. Consonants are the sounds produced by the obstacles against the airflow in the vocal tract. For this reason, the vocal cord vibrations do not have any significance in the production process of consonants. However, this is not the case for the production of every consonant. Thus, vowels and some of the consonants are voiced. This is also observed in the semivowels (/w/ and /y/) of Kurdish. When the semivowels are pronounced, they have both the consonant and the voiced sound properties. When examined, while the voiced sounds exhibit a periodic structure, voiceless sounds exhibit a non-periodic structure (Artuner., 1994).

Semivowels in Kurdish

There are two semivowels in Kurdish, being /w/ and /y/. In Kurdish, semivowels are classified into 2 main groups as labial and palatal semivowels by their phonation places as in Table 1.

Table 1: Semivowels in Kurdish.

	Labials	Prepalatals
Semivowels	/w/	/y/

When the semivowels in Kurdish are analysed phonetically, they demonstrate the following properties:

- /w/ sound is a continuous voiced sound pronounced with two lips (Ciwan., 1992, Khan., & Lescot., 1971).
- /y/ sound is a voiced sound pronounced in the prepalate (Ciwan., 1992, Khan., & Lescot., 1971).

Vowels in Kurdish

There are 8 vowels in Kurdish, being /a/, /e/, /ê/, /i/, /î/, /o/, /u/, and /û/. Vowels in Kurdish are classified into 3 main groups as long and short vowels by their pronunciation time when being pronounced, as prepalatal and postpalatal vowels by the motions the tongue has in the prepalate and postpalate in the oral cavity, and as high, mid, and low vowels by the proximity of the tongue to the mandible and palate as in Table 2.

Table 2: Vowels in Kurdish.

	Long Vowels			Short Vowels		
	High	Mid	Low	High	Mid	Low
Prepalate	/î/, /û/	/ê/				/e/
Postpalate			/a/, /o/	/i/, /u/		

The remaining parts of the article are arranged as follows: Materials and methods used are presented in Chapter 2, the formant frequency analysis results of the isolated semivowels are examined in Chapter 3, and a general deduction is made from the obtained results in Chapter 4.

Materials and Methods

In this study, the acoustic properties of Kurdish semivowels were sorted by the formant analysis method. The formant analysis is one of the most important methods used to determine the distinctive properties or significant frequency combinations from the speech data. They are demonstrated with F1, F2, F3, and F4. In this study, the formant analysis values were obtained in Matlab environment with the Linear Predictive Coding (LPC) method, and sound analysis process was performed with Praat software.

The Data Set Used and Its Properties

The properties of the data set formed for the acoustic analysis of Kurdish semivowels are presented in Table 3. The recordings were obtained in a silent room environment with the Easy Voice Recorder Pro voice recording program and General Mobile Discovery II + brand mobile phone.

Table 3: Properties of the database used.

Isolated semivowels	The number of male speakers	The number of female speakers	The age of male speakers	The age of female speakers
/w/	42	8	19-35	21-30
/y/	42	8	19-35	21-30

LPC Method

LPC is a technique used to characterize spectral properties from the sound signals. In this technique, the prediction of the n sample of the speech signal s(n) is expressed as the linear combination of the previous p samples. This is shown in equation 1 and the a_i values in Equation 1 refer to the LPC coefficients (Sel., 2013).

$$S(n) = -\sum_{i=1}^{p} a_i s(n-i) \qquad (1)$$

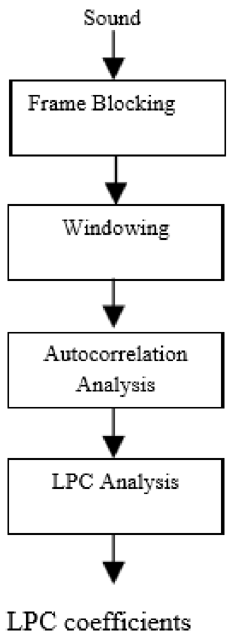
In the LPC method, the difference between the actual signal value s(n) and the predicted signal value S(n) of the speech equals to the error signal. This is expressed in Equation 2. The LPC coefficients in this method make production to calculate the error signal to be minimum (Sel., 2013).

$$e(n) = s(n) - S(n) = s(n) + \sum_{i=1}^{p} a_i s(n-i) \qquad (2)$$

With the z-transfer transformation, a filter with a finite impulse response, which is based on the human vocal tract and its properties, is obtained with the A(z) LPC analysis filter in Equation 3,

$$H(z) = 1 - \sum_{i=1}^{p} a^i z^{-i} = A(z) \qquad (3)$$

Steps of the LPC method
Steps constituting the LPC method are presented in Figure 1.



Şekil 1: The block diagram of the steps constituting the LPC method.

Frame Blocking: It is the processing of the speech signal in a small period of time, especially between 20 and 30 ms, and obtaining more stable characteristics of the speech signal (Karasartova., 2011).
Windowing: It is used to block discontinuous areas at the beginning and end of each framed sound signal. The examples of windowing are Hamming, Hanning, Rectangular, Barlett, and Kaiser. The most commonly used windowing example in voice processing applications is the Hamming windowing. This windowing is expressed by Equation 4 [7]. (The symbol "N" in this equation indicates the number of samples in each frame).

$$w(n) = 0.54 - 0.46\cos\left(\frac{2n\pi}{N-1}\right) \qquad (4)$$

Autocorrelation Analysis: The autocorrelation analysis is performed to each windowed frame signal. The autocorrelation analysis equation is expressed by Equation 5. In this equation, the symbol p indicates the analysis degree that can range from 8 to 16 (Eray., 2008).

$$r_l(m) = \sum_{n=0}^{N-1-m} \hat{x}_l(n) . \hat{x}_l(n+m), m = 0,1,2 \ldots\ldots, p \qquad (5)$$

The LPC Analysis: In this operation, p+1 LPC coefficient values are calculated from each frame applied with the autocorrelation analysis. These calculated values are called the LPC parameters.

Results and Discussion

The mean and standard deviation values of the frequencies F1, F2, F3 and F4 of the semivowels in the data set are presented in Table 4 for male speakers and in Table 5 for female speakers.

Table 4: The mean and standard deviation values of semivowels for males.

	Mean Formant Values			
	F1	F2	F3	F4
/w/	378.305	1056.323	2848.916	4124.831
/y/	299.084	2263.655	3309.581	4300.899
	Standard Deviation Formant Values			
	F1	F2	F3	F4
/w/	38.697	444.588	421.887	655.171
/y/	35.827	395.661	376.336	675.381

When the mean formant frequency values of the male speakers in Table 4 are examined, the format analysis of the semivowels of males is as follows:

- Since the phoneme with the highest F1 value is the /w/ sound, the /w/ phoneme is pronounced more clearly than the /y/ phoneme. In this case, the /y/ phoneme is pronounced more closed than the /w/ phoneme.

- Since the phoneme with the highest F2 value is the /y/ sound, the /y/ phoneme is pronounced more closely to the prepalate than the /w/ phoneme by the position of the tongue.

Table 5: The mean and standard deviation values of semivowels for females.

	Mean Formant Values			
	F1	F2	F3	F4
/w/	405.272	1069.249	3259.623	4447.568
/y/	323.411	2768.787	3835.873	4837.657
	Standard Deviation Formant Values			
	F1	F2	F3	F4
/w/	74.102	290.413	641.514	588.197
/y/	59.256	242.791	344.807	503.110

When the mean formant frequency values of female speakers in Table 5 are examined, it is observed that the format analysis of the semivowels of females shows the same conclusions as male speakers and the mean formant frequency values of female speakers have higher frequency values than male speakers. This is because the female vocal tract is longer than the male vocal tract.

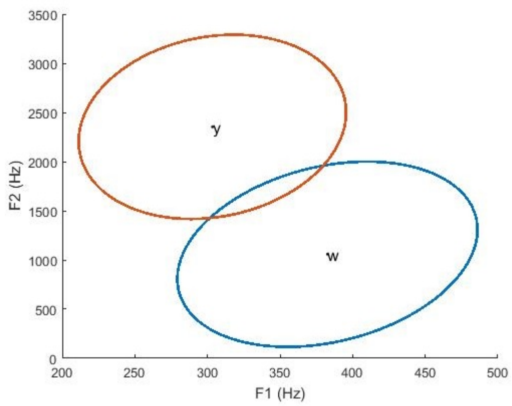


Figure 2: The scatter graph of the mean F1 and F2 frequency values of all speakers for semivowels.

In Figure 2, which is drawn according to the mean and standard deviation results of F1 and F2 frequency values of all speakers for semivowels, it is observed that the F2 formant frequency values of the /y/ semivowel are in areas with high values and the F1 formant frequency values of the /w/ semivowel are in areas with lower values, and the

F1 formant frequency values of the /y/ semivowel are in areas with low values and the F1 formant frequency values of the /w/ semivowel are in areas with higher values.

According to the classification of the vowels (Table 2), it can be said by the shape of the tongue, jaw, and lips when pronouncing semivowels [/w/, /y/] and vowels [/a/,/e/,/ê/,/i/,/î/,/o/,/u/ ,/û/] that;

- The /w/ phoneme shows similar phonetic properties with /a/, /o/, /e/, and /ê/ vowels compared to the /y/ phoneme in terms of the size of the jaw angle and it is different form /î/, /u/, /ı/, and /û/ high vowels.

- The /y/ phoneme shows similar phonetic properties with /e/, /ê/, /î/, and /û/ pre-tongue vowels compared to the /w/ phoneme and it is different form /a/, /o/, /i/, and /u/ post-tongue vowels.

Conclusion

Sorting the acoustic properties of the vowel and semivowel phonemes for the development of the speech processing applications, such as the voice to text and text to voice in Kurdish is of great importance. For this reason, in this study, the vowel and semivowel phonemes in Kurdish were examined, and their characteristic features were sorted. At the same time, the phonetic properties of the isolated semivowels and vowels in Kurdish were introduced, and their pronunciation relations were revealed.

References

Ciwan. M. (1992). *Türkçe açıklamalı Kürtçe dilbilgisi*, Sweden, MA: Jîna Nû Press.

Artuner. H. (1994). *The Design and Implementation of a Turkish Speech Phoneme Clustering System*, Ankara, MA: Ankara University, Phd Thesis.

Khan, E.D.B & Lescot, R. (1971). *Grammaire Kürde (Dialecte Kurmandji)*, Paris, Avesta Press.

http://www.mathworks.com/products/signal/

http://www.fon.hum.uva.nl/praat/

Sel, İ. (2013). *Türkçe metinler için hece tabanlı metinden konuşma sentezleme sistemi*, Elazığ, MA: Fırat University, Master Thesis.

Karasartova, S. (2011). *Metinden bağımsız konuşmacı tanıma sistemlerinin incelenmesi ve gerçekleştirilmesi*, Ankara, MA: Ankara University, Master Thesis.

Eray, O. (2008). *Destek vektör makineleri ile ses tanıma uygulaması*, Denizli, MA: Pamukkale University, Master Thesis.

Neel, J. (2005). *Cluster analysis methods for speech recognition*, Paris, MA: Centre for Speech Technology.

