

TÜRKÇE SESLERİN İSTATİSTİKSEL ANALİZİ

Pakize ERDOGMUS ⁽¹⁾

Ali ÖZTÜRK ⁽²⁾

Abant İzzet Baysal University
Technical Education Faculty
Electronic and Comp. Education Dept.
Asit. Prof.

Abant İzzet Baysal Üniversitesi
Technical Education Faculty
Electrical Education Dept.
Lecturer

ÖZET

Ses işleme, sesin zaman ve frekans özelliklerini temel alır. Ses gerçekte analog bir işaret iken, sayısal işarete çevrilerek bilgisayara aktarılan ses bir sayı dizisi haline gelir. Bu çalışmada, sayısal hale getirilmiş seslerin(Çalışmada harfler kullanılmıştır.) zaman ve frekans domenindeki yapıları Matlab ortamında istatistiksel olarak incelenmiş ve sonuçları sunulmuştur.

Anahtar Kelimeler: Sayısal İşaret İşleme; Ses işleme; Ses Tanıma; Ses Sentezi; İstatistiksel Analiz

THE STATISTICAL ANALYSIS OF TURKISH LETTERS

ABSTRACT

Speech processing is based on time and frequency properties of signal. Originally, speech is an analog signal, but it is converted to digital signal for processing in computer. And it is an array for computer. In this study, it has been studied Turkish letters' time and frequency domain analysis as statistically in Matlab and Excel. And the interesting results has been submitted.

Key Words: Digital Signal Processing, Speech Processing; Speech Recognition; Speech Synthesis; Statistical Analysis

1. INTRODUCTION

Sound waves are variations in air pressure. And two key features of sound wave are its *amplitude* which is a measure of air pressure at a point in time and *frequency* which is the rate at which this amplitude varies.[1]

speech is also a sound wave produced by human via vocal fold, velum , tongue and e.t.c. Speech is an effortless and highly efficient form of communication that doesn't require any special equipment. Speech Processing has the following sub titles. [2]

- Coding: To store a speech using as few bits as possible while retaining high quality
- Synthesis: To convert a text string into a speech waveform
- Recognition: To convert a speech waveform into text
- Identity Verification
- Enhancement

Speech is an analog signal that can be converted to digital signal easily. One of the speech feature is its amplitude stated before. This can be obtained easily from the recorded digital speech signal. The other speech feature,frequency are usually obtained via Fourier transforms (FTs), Short time fourier transforms (STFTs) or Linear Predictive Coding techniques . In this study we obtained this feature via FFT.

The most popular titles in speech processing are speech recognition and speech synthesis because of the fact that human-man interaction is very important especially for both disabilities and children. Even though many speech/voice processing tasks, like speech and

word recognition, reached satisfactory performance levels on specific applications, and although a variety of commercial products were launched in the last decade, many problems remain an open research area, and absolute solutions were not found out, yet.[3] The main obstacles slowing down these studies are the difference speaking style and grammar.

2 SPEECH PROCESSING

In this study, firstly 18 Turkish speech letters were recorded by microphone with Windows Sound Recorder. The sampling rate was used to as 22.050 kHz with 8-bit for all recorded letters.

Table 1. Turkish Letters Used in This Study

Consonants	B	C	D	F	G	H	P	S	T	Z
Vowels	A	E	I	İ	O	Ö	U	Ü		

Then it can be possible to filter and denoise these letters' signals. But when it was applied some signals, it has been seen that it is redundant for this study. But it is quite easy to denoise signals with Matlab Wavelet Toolbox.

It could be possible to record letters spoken by different peoples. As it has been known that signal can change according to the person. Child, female and male voices may have a bit different features. But since this study is not a pure speech recognition study, letters has been spoken by a person.

Secondly, it has been applied FFT in order to get frequency features of these signals. Vowels and consonants time representation has been given in Figure 1 and Figure 2 respectively.

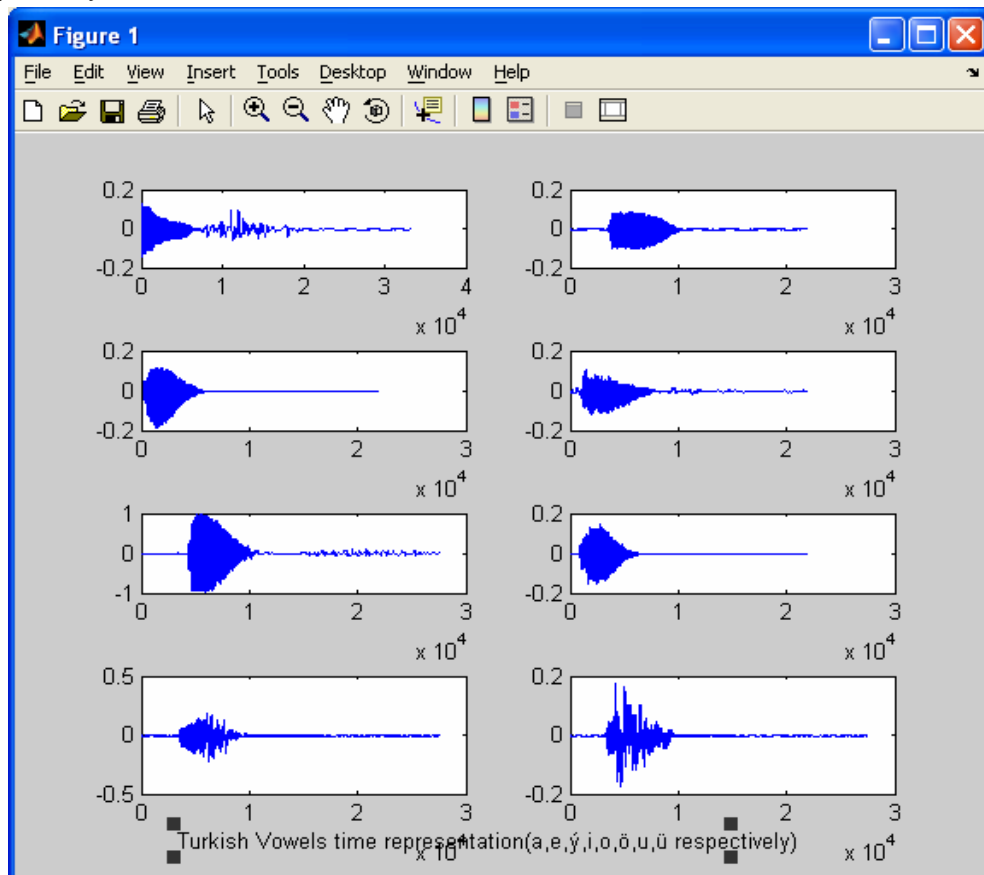


Figure 1. Turkish Vowels Time Representation(a,e,ı,i,o,ö,u,,ü)

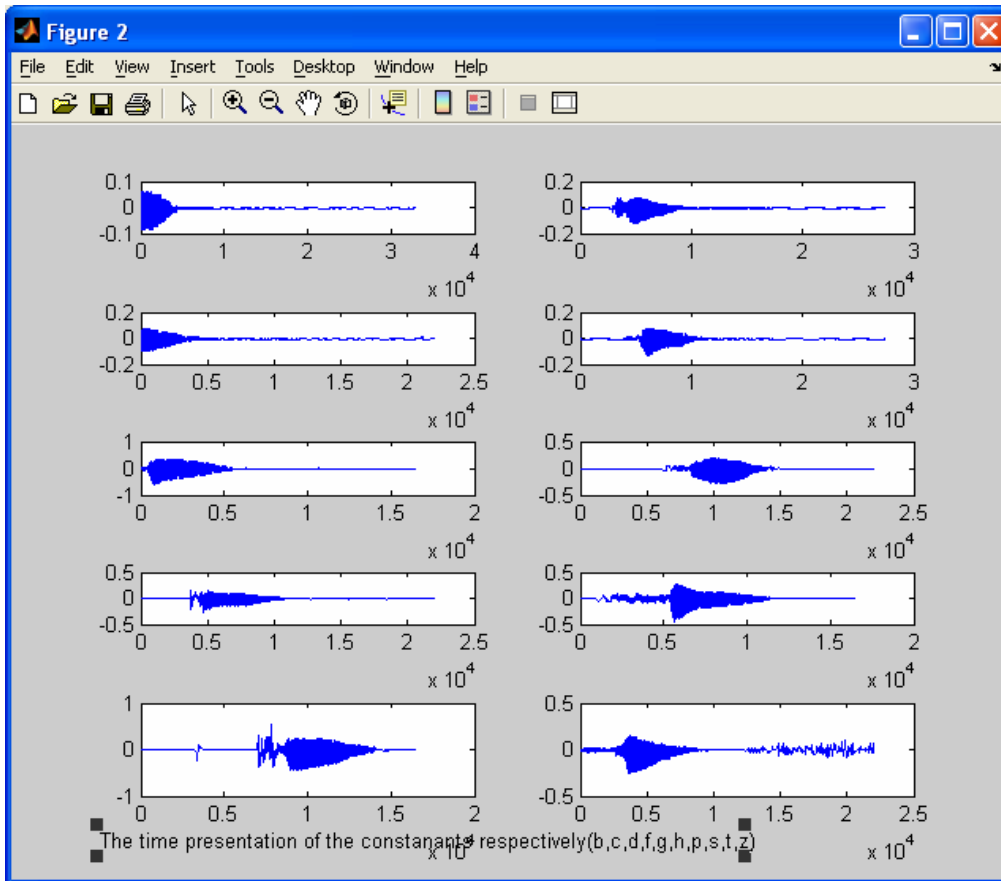


Figure 2. Turkish Constanants Time Representation(b,c,d,f,g,h,p,s,t,z)

And when it was observed some of these letters time presentations are quite similar. So it has been taken FFT in Matlab and some of these letter pairs frequency representation is as follows. It has been taken FFT for *i,o,e,i* vowels and *s,z,c,f* constanants.

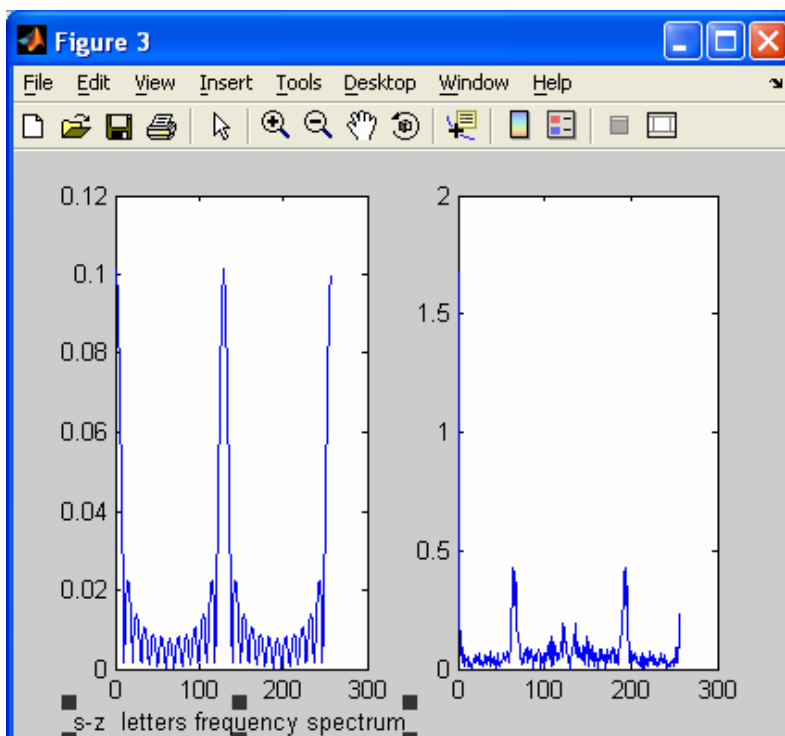


Figure 3. s-z Frequency Spectrum

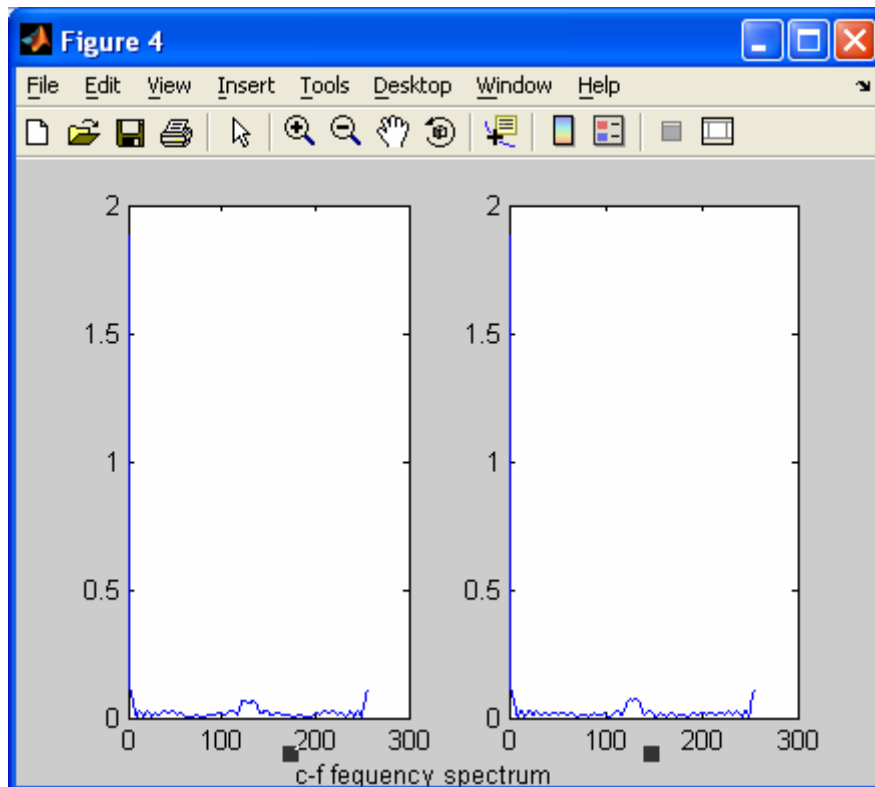


Figure 4. c-f Frequency Spectrum

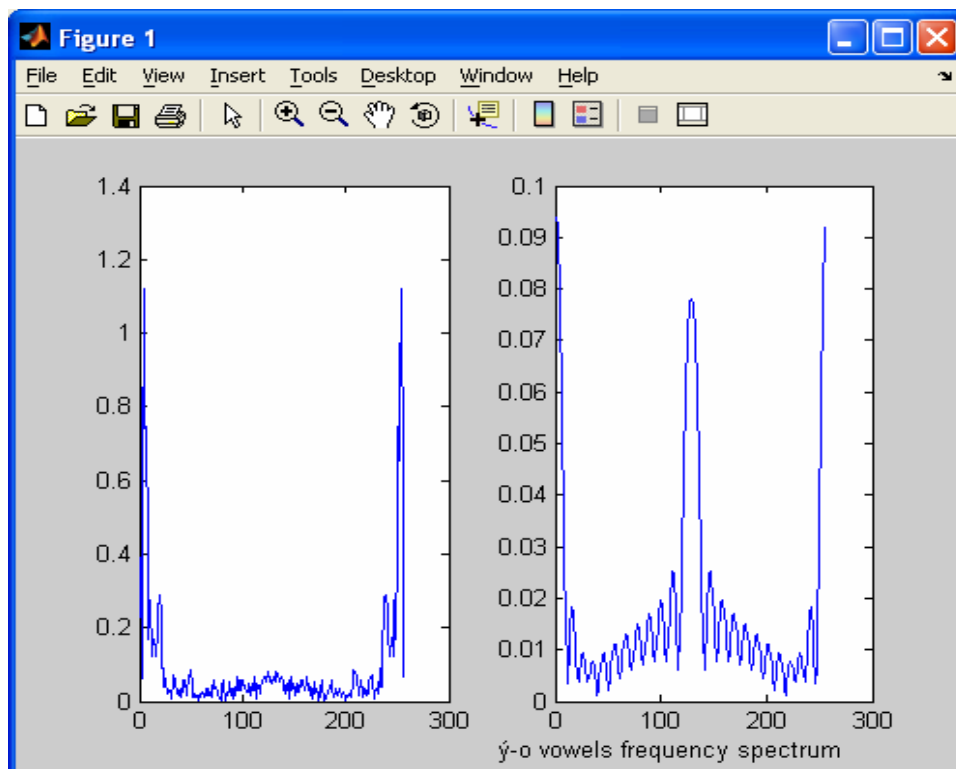


Figure 5. i-o Frequency Spectrum

3 STATISTICAL ANALYSIS

Recorded signals have been read with Matlab as a number array with Wavread Command. It has been calculated mean, standart deviation and modes of these arrays in time domain. It has been shown in Table 2. These statistics can be obtained from the Matlab figure window easily. But it has been calculated by Excel.

Table 2. The statistics of Letter signals

Letters	Mean	Absolute Mean	Std. Deviation
A	-0,00395	0,012367	0,020343
E	-0,00369	0,009688	0,018192
I	-0,0039	0,011522	0,02428
i	-0,00382	0,009856	0,016072
O	-0,00396	0,074739	0,165849
ö	-0,0039	0,011594	0,024669
U	-0,00407	0,0129	0,028433
ü	-0,00387	0,011119	0,023834
B	-0,00413	0,005926	0,009666
C	-0,00385	0,007506	0,013891
D	-0,00394	0,006431	0,010865
F	-0,00413	0,007601	0,013222
G	-0,00389	0,037777	0,081246
H	-0,00406	0,019653	0,040311
P	-0,00384	0,015176	0,03035
S	-0,00389	0,030842	0,054286
T	-0,00393	0,045256	0,083597
Z	-0,00353	0,021545	0,034016

As it has been seen in the table the highest standart deviations of the letters are belong to g and t. And the lowest standart deviation is belong to i.

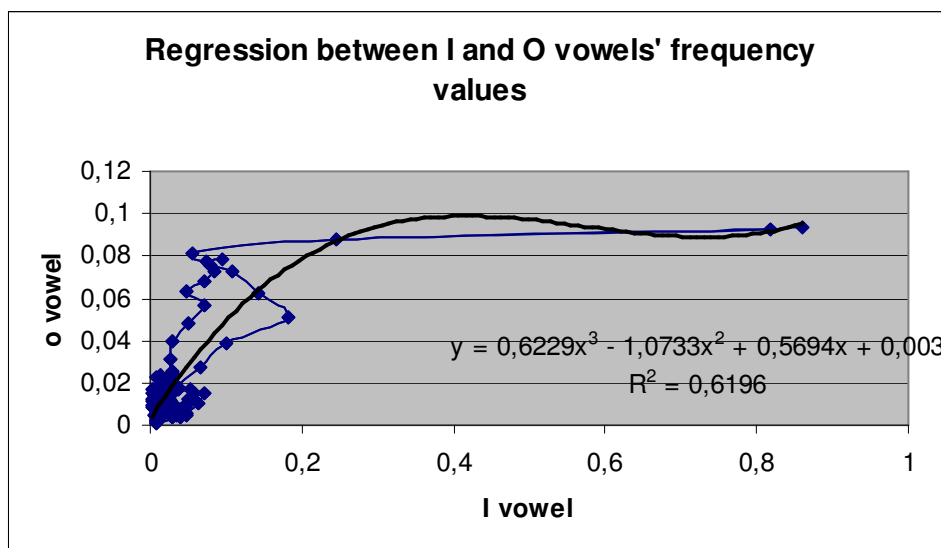


Figure 6. The Regression between the i-o vowels

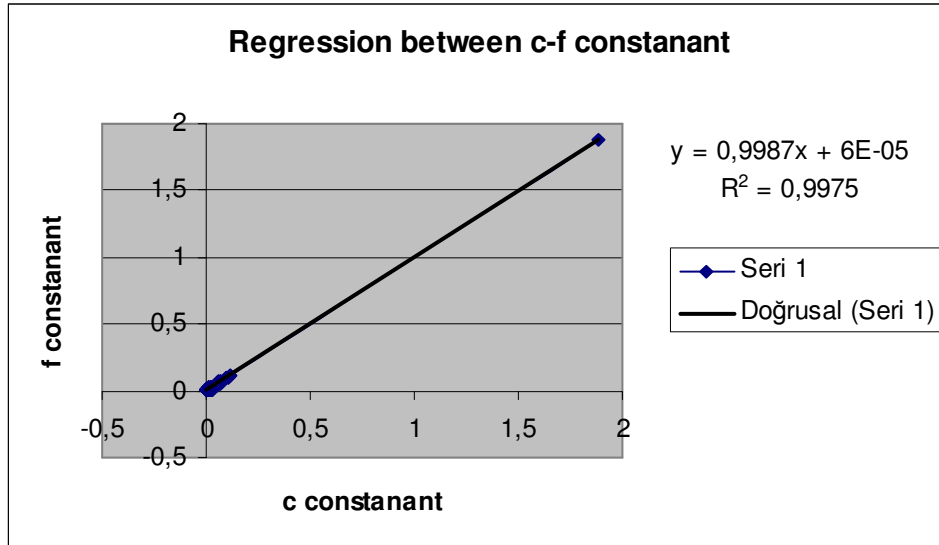


Figure 7. The Regression between the c-f letter.

4 RESULTS

In this study, it has been aimed to find some useful information using for speech recognition of the letters. And as it has been seen both in tables and figures there are a lot of useful information to use. Especially statistics can be used for recognition parts parameters. And as it has been in the figures some letters are very similar. The correlation coefficient is quite meaningful for these letters. So the process for these letters' recognition will be much more hard. These study will continue as Turkish letter recognition .

5 REFERENCES

- [1] CALLAN R., "Artificial Intelligence", Palgrave Mcmillan, 2003.
- [2] <http://www.ee.ic.ac.uk/hp/staff/dmb/courses/speech/overview.pdf>
- [3] Avcı, E., Akpolat, Z. H., "Speech Recognition Using A Wavelet Packet Adaptive Network Based Fuzzy Inference System", Expert System with Applications, Vol 31, 2006.