

**TÜRKÇE METİNLER İÇİN KONUŞMA MOTORU
GELİŞTİRİLMESİ**

**2014
YÜKSEK LİSANS TEZİ
BİLGİSAYAR MÜHENDİSLİĞİ**

Erdem ERKAN

TÜRKÇE METİNLER İÇİN KONUŞMA MOTORU GELİŞTİRİLMESİ

Erdem ERKAN

**Karabük Üniversitesi
Fen Bilimleri Enstitüsü
Bilgisayar Mühendisliği Anabilim Dalında
Yüksek Lisans Tezi
Olarak Hazırlanmıştır**

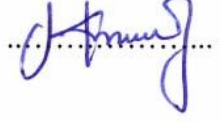
KARABÜK

Ocak 2014

Erdem ERKAN tarafından hazırlanan “TÜRKÇE METİNLER İÇİN KONUŞMA MOTORU GELİŞTİRİLMESİ” başlıklı bu tezin Yüksek Lisans Tezi olarak uygun olduğunu onaylarım.

Yrd. Doç. Dr. İsmail KURNAZ

Tez Danışmanı, Bilgisayar Mühendisliği Anabilim Dalı



Bu çalışma, jürimiz tarafından oy birliği ile Bilgisayar Mühendisliği Anabilim Dalında Yüksek Lisans tezi olarak kabul edilmiştir. 16/ 01/ 2014

Ünvanı, Adı SOYADI (Kurumu)

İmzası

Başkan : Doç. Dr. İsmail Rakıp KARAS (KBÜ)

Üye : Yrd. Doç. Dr. Hüseyin DEMİREL (KBÜ)

Üye : Yrd. Doç. Dr. İsmail KURNAZ (KBÜ)



31./01./2014

KBÜ Fen Bilimleri Enstitüsü Yönetim Kurulu, bu tez ile, Yüksek Lisans derecesini onamıştır.

Prof. Dr. Mustafa BOZ

Fen Bilimleri Enstitüsü Müdürü



“Bu tezdeki tüm bilgilerin akademik kurallara ve etik ilkelere uygun olarak elde edildiğini ve sunulduğunu; ayrıca bu kuralların ve ilkelerin gerektirdiği şekilde, bu çalışmadan kaynaklanmayan bütün atıfları yaptığımı beyan ederim.”

Erdem ERKAN

ÖZET

Yüksek Lisans Tezi

TÜRKÇE METİNLER İÇİN KONUŞMA MOTORU GELİŞTİRİLMESİ

Erdem ERKAN

Karabük Üniversitesi

Fen Bilimleri Enstitüsü

Bilgisayar Mühendisliği Anabilim Dalı

Tez Danışmanı:

Yrd. Doç. Dr. İsmail KURNAZ

Ocak 2014, 64 sayfa

İnsansı yeteneklere sahip bilgisayarlar, en başta engelliler olmak üzere tüm insanların yaşam kalitesini artırmaktadır. Bir bilgisayarın metni seslendirme, diğer bir deyişle konuşabilme özelliği de bu yetenekler arasında ilk sıralarda yer almaktadır. Bu tez çalışmasında, dijital ortamdaki Türkçe metinleri ses sinyallerine dönüştürebilen akıcı, anlaşılır bir Türkçe metin sentezleyici yazılım (MKS) geliştirilmiştir. Yazılım farklı ses veri tabanlarının uyarlanmasına uygun bir yapıda olmakla beraber, şu an sadece erkek sesi olarak seslendirme yapabilmektedir. Ayrıca uygulama genişletilebilir bir telaffuz modülüne sahiptir. Uygulamada metin sentezleme tekniği olarak ikili ses (difon) sentezleme tekniği kullanılmıştır. Uygulama MOS tipi bir değerlendirme testinden 5 tam puan üzerinden 3,72 puan almıştır.

Anahtar Sözcükler : Türkçe, metinden konuşma sentezleme (MKS), difon, metin seslendirme.

Bilim Kodu : 902

ABSTRACT

M. Sc. Thesis

DEVELOPING SPEECH ENGINE FOR TURKISH TEXT

Erdem ERKAN

Karabük University

Graduate School of Natural and Applied Sciences

Department of Computer Engineering

Thesis Advisor:

Assist. Prof. Dr. İsmail KURNAZ

January 2014, 64 pages

Computers with human-like capabilities, enhance the quality of life of all people, including people with disabilities in the first place. In other words, a computer voice to speak text feature of these capabilities is in the first place. In this thesis, digital media fluent in Turkish that can convert text to audio signals, understandable Turkish text to speech software (TTS) has been developed. The software is currently only capable of voice as the voice of Turkish men. But it has suitable structure for other voices. In addition, applications have expanded pronunciation module. Diphone synthesis technique has been used as a technique to synthesize the text. Application of MOS type assessment test score was 3.72 points out of 5.

Key Word : Turkish, text-to-speech synthesis (TTS), diphone, speech synthesis.

Science Code : 902

TEŐEKKÜR

Bu tez alıőmasının planlanmasında, araőtırılmasında, yürütülmesinde ve oluşumunda ilgi ve desteęini esirgemeyen, bilgi ve tecrübelerinden yararlandıęım, yönlendirme ve bilgilendirmeleriyle alıőmamı bilimsel temeller ışığında őekillendiren sayın hocam Yrd. Do. Dr. İsmail KURNAZ'a teőekkürlerimi sunarım.

Baőta eőim Yasemin ERKAN olmak üzere tüm aileme manevi hibir yardımı esirgemedен yanımda oldukları için tüm kalbimle teőekkür ederim.

İÇİNDEKİLER

	<u>Sayfa</u>
KABUL.....	ii
ÖZET.....	iv
ABSTRACT.....	v
TEŞEKKÜR.....	vi
İÇİNDEKİLER	vii
ŞEKİLLER DİZİNİ.....	x
ÇİZELGELER DİZİNİ	xii
SİMGELER VE KISALTMALAR DİZİNİ	xiii
BÖLÜM 1.	1
GİRİŞ.....	1
BÖLÜM 2.	3
KONUŞMA MOTORLARI.....	3
2.1. KONUŞMA MOTORLARININ KULLANIM ALANLARI	3
2.2. GEÇMİŞTEN GÜNÜMÜZE KONUŞMA MOTORU SİSTEMLERİ.....	4
2.3. KONUŞMA SENTEZLEMEDE KARŞILAŞILAN ZORLUKLAR	9
2.3.1. Veri Tabanı Hazırlamada Karşılaşılan Zorluklar	9
2.3.2. Metin Ön İşlemede Karşılaşılan Zorluklar	10
2.3.3. Bazı Kelimelere Has Özel Okunuşlar.....	10
2.3.4. Aruz (Tonlama, Vurgu ve Süre Hesap Analizi) Zorlukları	11
2.4. ÖRNEK KONUŞMA MOTORU SİSTEMLERİ.....	11
2.4.1. MITalk	11
2.4.2. Infovox.....	11
2.4.3. Bell Labs TTS System	12
2.4.4. Cnet Psola	13
2.4.5. Festival.....	13
2.4.6. MBROLA	13
2.5. TÜRKÇE TEMELLİ KONUŞMA MOTORLARI	14

	<u>Sayfa</u>
2.6. TÜRKÇE İÇİN YAPILAN ÇALIŞMALAR.....	14
2.7. FARKLI DİLLER İÇİN YAPILAN ÇALIŞMA ÖRNEKLERİ.....	16
BÖLÜM 3.	18
KONUŞMA SENTEZLEME.....	18
3.1. KONUŞMA SENTEZLEME TEKNİKLERİ	18
3.1.1. Söyleyiş Sentezleyiciler.....	18
3.1.2. Formant Sentezleyiciler.....	19
3.1.3. Parça Sentezleyiciler.....	21
3.1.3.1. TD-PSOLA	22
3.1.3.2. MBR-PSOLA.....	24
BÖLÜM 4.	26
TURS TÜRKÇE KONUŞMA MOTORU.....	26
4.1. SES VERİ TABANI OLUŞTURULMASI.....	27
4.1.1. Sistem İçin Gerekli Tüm İkili Seslerin Belirlenmesi.....	27
4.1.2. Kayıt Ortamı	30
4.1.3. Ses Formatı	30
4.1.3. Seslerin Taşıyıcı Kelimelerden Ayrılması.....	31
4.1.4. Ses Veri Tabanı Normalizasyonu	32
4.2. METİN ÖN İŞLEMLERİ.....	33
4.3. İKİLİ SES BİRLEŞTİRME ALGORİTMASI.....	34
4.4. SİSTEMİN OPTİMİZASYONU.....	38
4.4.1. Türkçe Heceleme Algoritması	38
4.4.2. Ses Çıkışını Düzeltme Çalışmaları	40
4.4.2.1. Ayrık Zamanlı Fourier Transformu (DFT).....	40
4.4.2.2. DFT'nin Koda Uyarlanması	40
4.5. EZGİSEL ÇALIŞMALAR.....	41
4.6. TELAFUZ ÜNİTESİ.....	44
4.7. TAŞINABİLİR SÜRÜM DENEMESİ	45

	<u>Sayfa</u>
BÖLÜM 5.	47
SİSTEMİN TEST EDİLMESİ.....	47
5.1. MOS (MEAN OPINION SCORE).....	47
5.2. SİSTEM MOS PUANININ BELİRLENMESİ.....	47
BÖLÜM 6.	50
SONUÇLAR.....	50
KAYNAKLAR.....	52
EK AÇIKLAMALAR A. DİFON VE TAŞIYICI KELİME TABLOSU.....	55
ÖZGEÇMİŞ.....	64

ŞEKİLLER DİZİNİ

Sayfa

Şekil 2.1.	Konuşma motoru genel yapısı.....	3
Şekil 2.2.	Charles Wheatstone tarafından geliştirilen konuşma makinesi.....	4
Şekil 2.3.	Riesz tarafından geliştirilen konuşma makinesi.....	5
Şekil 2.4.	Riesz tarafından geliştirilen konuşma makinesi kontrol tuşları	6
Şekil 2.5.	Riesz ve Watkins'in VODER şeması	7
Şekil 2.6.	Konuşma motorlarının tarihsel gelişimi.....	9
Şekil 2.7.	Bell Laboratuvarları TTS Sistemi modülleri.....	12
Şekil 3.1.	Sırasıyla alçak geçiren, yüksek geçiren, bant geçirmeyen ve bant geçiren filtre çıkışları.....	19
Şekil 3.2.	Bant geçiren filtrelerin seri bağlanması	20
Şekil 3.3.	Bant geçiren filtrelerin seri bağlanması sonucu giriş ve çıkış sinyali	20
Şekil 3.4.	Bant geçiren filtrelerin paralel bağlanması ve çıkış sinyali	20
Şekil 3.5.	TD-PSOLA zaman domeninde genişletme	23
Şekil 3.6.	MBROLA ile “bonjour” kelimesinin seslendirilmesi	24
Şekil 4.1.	TURS Türkçe seslendirme sistemi ekran görüntüsü	26
Şekil 4.2.	Türkçe ikili seslerin hesaplanması için kullanılan fonksiyon	28
Şekil 4.3.	Örnek ikili sesler ve taşıyıcı kelimeleri.....	30
Şekil 4.4.	İkili seslerin (difonların) taşıyıcıdan çıkarılması	32
Şekil 4.5.	Normalize edilmiş ab ikili sesi ab.wav.....	33
Şekil 4.6.	Birleştirme bayt tamsayı dönüşümü	34
Şekil 4.7.	Oku kelimesi için kullanılacak ikili ses dizisi.....	34
Şekil 4.8.	İkili ses birleştirme algoritması	35
Şekil 4.9.	Oku için ikili ses dizisi dalga işaretleri	36
Şekil 4.10.	Birleştirme sonucu elde edilen oku.wav dosyası işareti.....	36
Şekil 4.11.	Ses sinyali perde frekansı artırılması	38
Şekil 4.12.	Türkçe heceleme algoritması akış diyagramı.....	39
Şekil 4.13.	Ayrık zamanlı Fourier Transformu (DFT)	41
Şekil 4.14.	Ayrık zamanlı Fourier Transformunun tersi (IDFT).....	41

	<u>Sayfa</u>
Şekil 4.15. Cümle sınıfı kod bloğu örneği.....	42
Şekil 4.16. Kelime sınıfı kod bloğu örneği	43
Şekil 4.17. İkili ses sınıfı kod bloğu örneği.....	43
Şekil 4.18. TURS telaffuz veri tabanı dosyası	45
Şekil 4.19. TURS mobil test ekranı.....	46
Şekil 5.1. Cümle bazında MOS puan grafiği	49

ÇİZELGELER DİZİNİ

	<u>Sayfa</u>
Çizelge 3.1. Türkçe sesli harflerin okunuş biçimleri	21
Çizelge 4.1. Hesaplanan Türkçe ikili sesler	29
Çizelge 4.2. WAV dosya formatının yapısı	31
Çizelge 4.3. MBR-PSOLA veri tabanı kayıt yapısı (bardaktan).....	37
Çizelge 5.1. TURS tarafından seslendirilen cümleler	48
Çizelge 5.2. MOS derecelendirme çizelgesi	48
Çizelge 5.3. Katılımcı puanları çizelgesi.	49

SİMGELER VE KISALTMALAR DİZİNİ

SİMGELER

C#	: C Sharp
P(t)	: giriş ses sinyali perde periyodu
$\tilde{P}(\tilde{t})$: çıkış ses sinyali perde periyodu
t_i	: giriş sinyali anlık zamanı
\tilde{t}_k	: çıkış sinyali anlık zamanı
α	: ses sinyali germe faktörü
$X_1(n)$: pencerelemiş ses sinyali parçası

KISALTMALAR

dB	: desibel
DECTalk	: Digital Equipment Corporation Talk System
DFT	: Discrete Fourier Transform
Hz	: Hertz
IDFT	: Inverse Discrete Fourier Transform
KHz	: Kilo Hertz
LPC	: Lineer Predictive Coding
MBR-PSOLA	: Multi-Band Resynthesis Pitch Synchronous OverLap Add
MIT	: Massachusetts Institute of Technology
MKS	: Metinden Konuşma Sentezleme
MOS	: Mean Opinion Score
ms	: mili saniye
OVE	: Orator Verbis Electricis
PAT	: Parametric Artificial Talker
PSOLA	: Pitch Synchronous OverLap Add
SGK	: Sosyal Güvenlik Kurumu

TASA : Turkish Automatic Spelling Algorithm
TD-PSOLA : Time Domain Pitch Synchronous OverLap Add
TTS : Text to Speech
TURS : Türkçe Seslendirme Sistemi
VODER : Voice Operating Demonstrator
WAV : Waveform Audio File Format

BÖLÜM 1

GİRİŞ

Konuşma motorları, diğer bir deyişle metinden konuşma sentezleme (MKS) sistemleri yazılı metinleri girdi olarak alıp, bunları konuşma diline çeviren sistemlerdir. Türkçe için yeterli sayıda olmasa da başta İngilizce olmak üzere, diğer tüm dillerde konuşma motorları konusunda birçok çalışma yapılmıştır. Ülkemizde henüz olmamakla beraber yurtdışında birçok üniversitede konuşma motoru, TTS laboratuvarları kurulmuştur.

Konuşma motorlarının kullanım alanları en başta görme engelli bireyler olmak üzere, çağrı merkezi otomasyonları, navigasyon sistemleri, toplu taşıma araçları sesli uyarı sistemleri, zihinsel özürlü bireylerin eğitimi olarak sıralanabilir. Ancak konuşma motoru sistemleri, teknolojinin hızla geliştiği dünyamızda her gün yeni kullanım alanlarıyla karşımıza çıkmaktadır. Konuşma motorları sayesinde en başta görme engelliler, dış dünya ile hiç olmadığı kadar sıkı bağlar kurma fırsatını yakalamışlardır.

Konuşma motorları, diğer bir deyişle konuşma sentezleyiciler, genel olarak, söyleyiş sentezleyiciler, formant sentezleyiciler ve parça sentezleyiciler olmak üzere üç grupta incelenebilir.

Söyleyiş sentezleyiciler insan akustik organlarını model alıp, ses çıktısını bu modele uygun olarak üretmeye çalışan sentezleyicilerdir. İnsanın akustik organlarının modellenmesi zorluğu nedeniyle, uygulamada az karşılaşılan sentezleyici tipidir.

Formant sentezleyiciler, konuşmayı bir sinyal üretici ve bu sinyali şekillendiren seri-paralel filtre yapılarının, üretilecek sese göre tasarlanması temeline dayanmaktadır.

Son olarak parça sentezleyiciler de, ilgili parçaları bünyesinde barındıran bir ses veri tabanı ve bu veri tabanındaki parçaları, gerçek zamanlı olarak, istenilen konuşma metnine göre birleştirme esasına göre çalışan sentezleyici tipidir.

Konuşma motorunun ses kalitesini, sistemin doğallığı ve anlaşılabilir olması belirler. Bu iki özellik birlikte konuşma motorunun insan sesine yakınlığını temsil eder. Bu bağlamda bu tez çalışmasının amacı Türkçe tabanlı, parça sentezleyici tipte bir konuşma motoru uygulaması geliştirilmesidir. Parça sentezleyici tipte bir konuşma motorunun geliştirilmesinde karşılaşılan zorluklar ve bu uygulamanın doğallığını artırmaya yönelik yapılan çalışmaların detaylı bir şekilde anlatılması çalışmanın amaçları arasındadır. Ayrıca uygulama ücretsiz olarak görme engellilerin kullanımına sunulacaktır.

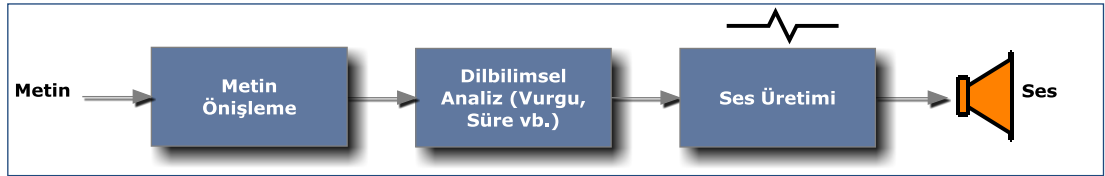
Hazırlanan bu çalışma altı bölümden oluşmaktadır. Bunlardan ilki giriş bölümü olup burada çalışmanın kısa bir özeti verilmiştir. İkinci bölümde konuşma motorlarının kullanım alanları, geçmişten günümüze kadar konuşma motorlarının gelişimi, geliştirilen konuşma motoru örnekleri ve Türkçe tabanlı konuşma motorları hakkında literatür çalışmalarına yer verilmiştir. Üçüncü bölümde konuşma sentezleme teknikleri genel bir çerçevede anlatılmış, seçilen sentezleme tekniğinin popüler yöntemleri üzerinde durulmuş, dördüncü bölümde çalışma kapsamında geliştirilen konuşma motoru TURS'un geliştirme aşamaları algoritmalar ve kod bloğu örnekleriyle detaylı bir şekilde anlatılmaya çalışılmıştır. Beşinci bölümde sistemin test edilmesi ve değerlendirilmesi üzerinde durulmuştur. Geliştirilen sistem MOS tipi bir değerlendirme yöntemiyle test edilmiş ve 5 tam puan üzerinden 3,72 puan almıştır. Son olarak altıncı bölümde, geliştirilen konuşma motorunun artı ve eksileri değerlendirilmiş, geliştirme sürecinde edinilen tecrübeler, iyileştirmeler hakkında tespitlerde bulunulmuş, bu tez çalışmasının devamı niteliğindeki ilerideki amaç ve hedefler belirtilmiştir.

BÖLÜM 2

KONUŞMA MOTORLARI

2.1. KONUŞMA MOTORLARININ KULLANIM ALANLARI

Konuşma motorları bilgisayar veya benzeri ortamlarda bulunan metinleri konuşmaya çeviren yazılımlardır. Konuşma motorlarının çok çeşitli kullanım alanları vardır. Bunların en başında görme engelli insanlar gelmektedir. Günümüzdeki konuşma motorları sayesinde görme engelli insanlar da en az görme engelli olmayan insanlar kadar bilgisayarlardan faydalanabilmektedirler. Bunun yanında konuşma motorları çağrı merkezi otomasyonlarında, navigasyon sistemlerinde, toplu taşıma araçları sesli uyarı sistemlerinde, okuma sorunu olmayan bilgisayar kullanıcıları, robot teknolojilerinde ve hatta zihinsel özürli bireylerin eğitimi gibi birçok alanda kullanılmaktadır. Şekil 2.1’de konuşma motoru genel yapısı verilmiştir.



Şekil 2.1. Konuşma motoru genel yapısı.

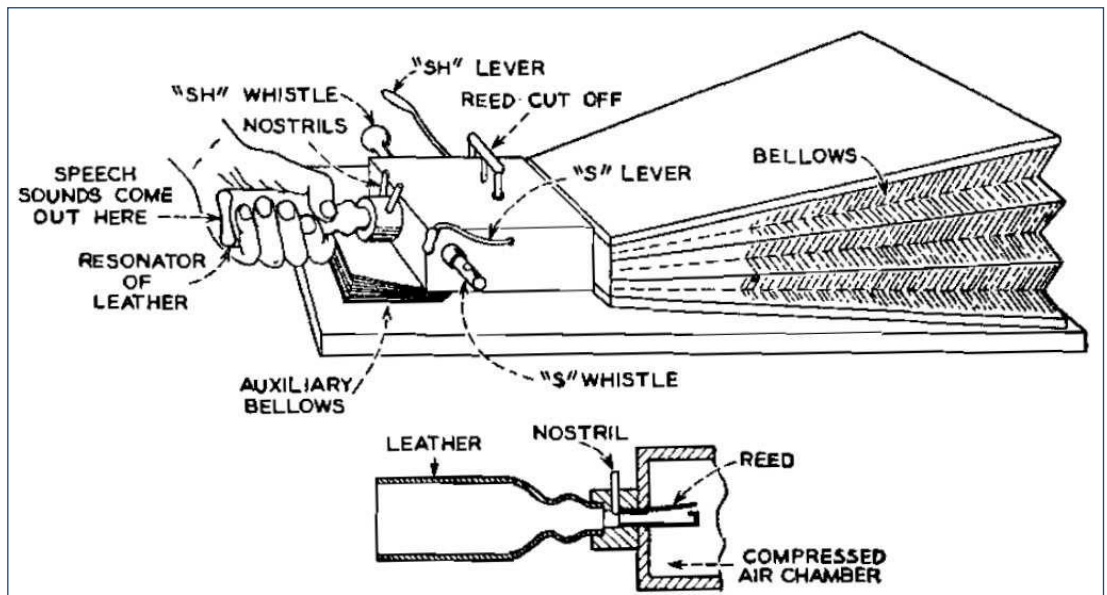
Şekil 2.1’de konuşma motoruna girdi olarak alınan metin ilk olarak ilgili metni standartlaştırmaya yönelik bazı ön işlemlere girmekte, sonrasında da metnin dilbilimsel analizi yapılmaktadır. En son olarak da sinyal işleme teknikleriyle çıkış sesinin üretilmesi gerçekleşmektedir.

2.2. GEÇMİŞTEN GÜNÜMÜZE KONUŞMA MOTORU SİSTEMLERİ

Günümüz konuşma motorlarının temeli 1779 yılında Kratzenstein tarafından atılmıştır. Kratzenstein insan ses sistemine benzer bir mekanik yapı geliştirmiş, bu mekanik sistemdeki rezonatör görevi gören kamışların titreşimi ile a, e, i, o, u sesli harflerinin fizyolojik farklılıklarını göstermiştir [1].

Birkaç yıl sonra 1791 de Wolfgang von Kempelen Kratzenstein'in makinesinden daha ayrıntılı bir makine geliştirmiştir. Bu makine Akustik Mekaniksel Konuşma Makinesi olarak adlandırılmıştır. Konuşma makinesi tek elle tutulan körük yardımıyla çalışmakta, diğer el ile mekanik aletin dört ayrı geçidi kontrol edilmekteydi. Körük tutan el yardımıyla sesli harfler, diğer el yardımıyla da bazı sessiz harfleri üretilebilmekteydi. Aslında Kempelen Konuşma Makinesi çalışmalarına Kratzenstein'den çok daha önce 1769 yılında başlamıştı [1-3].

1800'lerin ortalarında Charles Wheatstone, Şekil 2.2'de gösterilen Kempelen'in konuşma makinesinin daha da iyileştirilmiş bir sürümünü geliştirmiştir. Bu haliyle konuşma makinesi biraz daha karmaşık bir yapıya sahip olmakla beraber ünlü ve ünsüz sesleri, bazı ses öbeklerini, hatta bazı kelimeleri de üretebilir bir yapıya sahip kavuşmuş oldu.

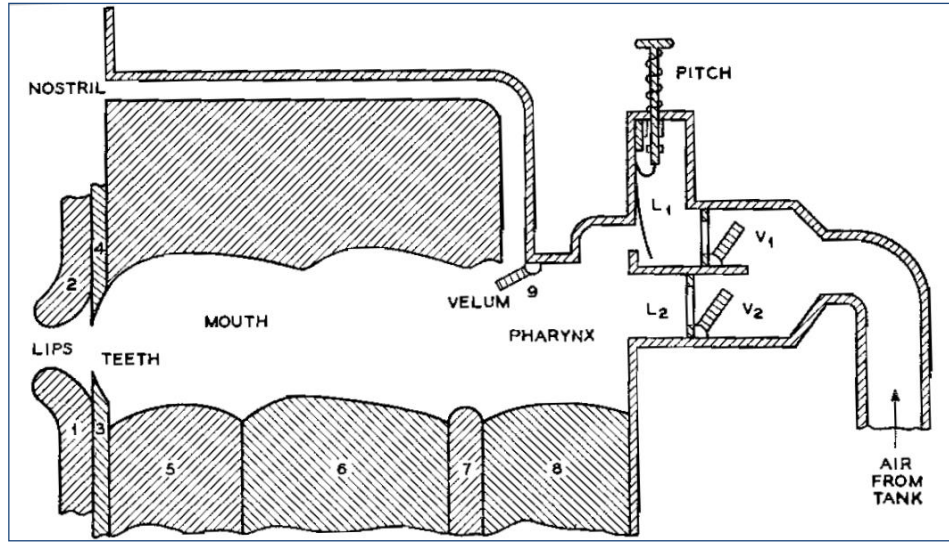


Şekil 2.2. Charles Wheatstone tarafından geliştirilen konuşma makinesi [1].

1838 yılında Willis belirli bir ses ve bu sesin geometrisi arasındaki ilişkiyi tespit etti. Willis farklı ünlü sesleri, organların borsal yapısına benzer tüp rezenatörler yardımıyla üretmişti. Çalışmalarının sonucunda ürettiği ünlü seslerin kalitesinin tüp rezenatörün uzunluğuna, aynı zamanda da çapına bağlı olduğunu keşfetti [3].

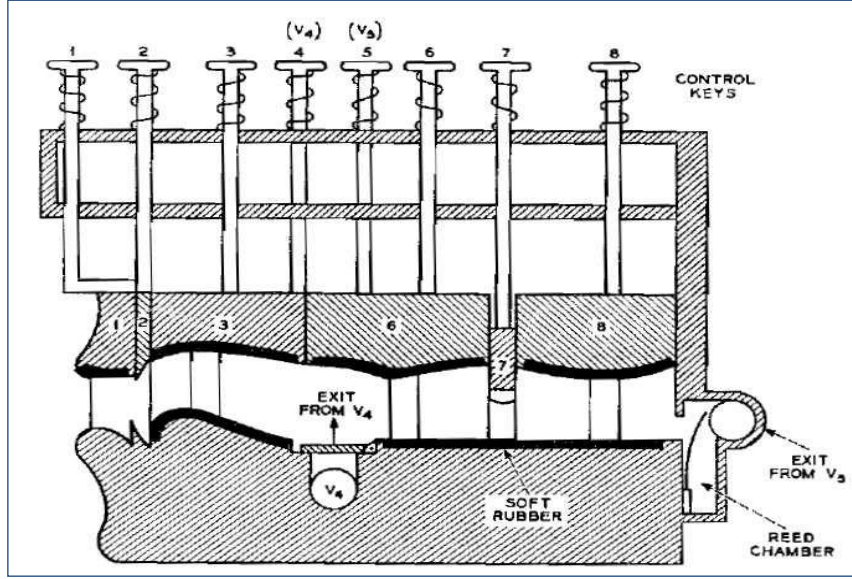
1800'lü yılların sonlarında Alexander Graham Bell ve babası Wheatstone'un konuşma makinesinden ilham alarak bir konuşma makinesi yaptılar [1,2].

1937'de Riesz Şekil 2.3'de gösterilen mekanik konuşma makinesini geliştirdi.



Şekil 2.3. Riesz tarafından geliştirilen konuşma makinesi [1].

Şekil 2.4'de Riesz tarafından geliştirilen konuşma makinesi kontrol tuşları gösterilmektedir. Ünsüz sesler V2 valfi tarafından kontrol edilmektedir. 1 ve 2 tuşları dudakları, 3 ve 4 tuşları dişleri, 5, 6 ve 7 tuşları dili, 8 tuşu yutağı temsil etmekteydi [1].

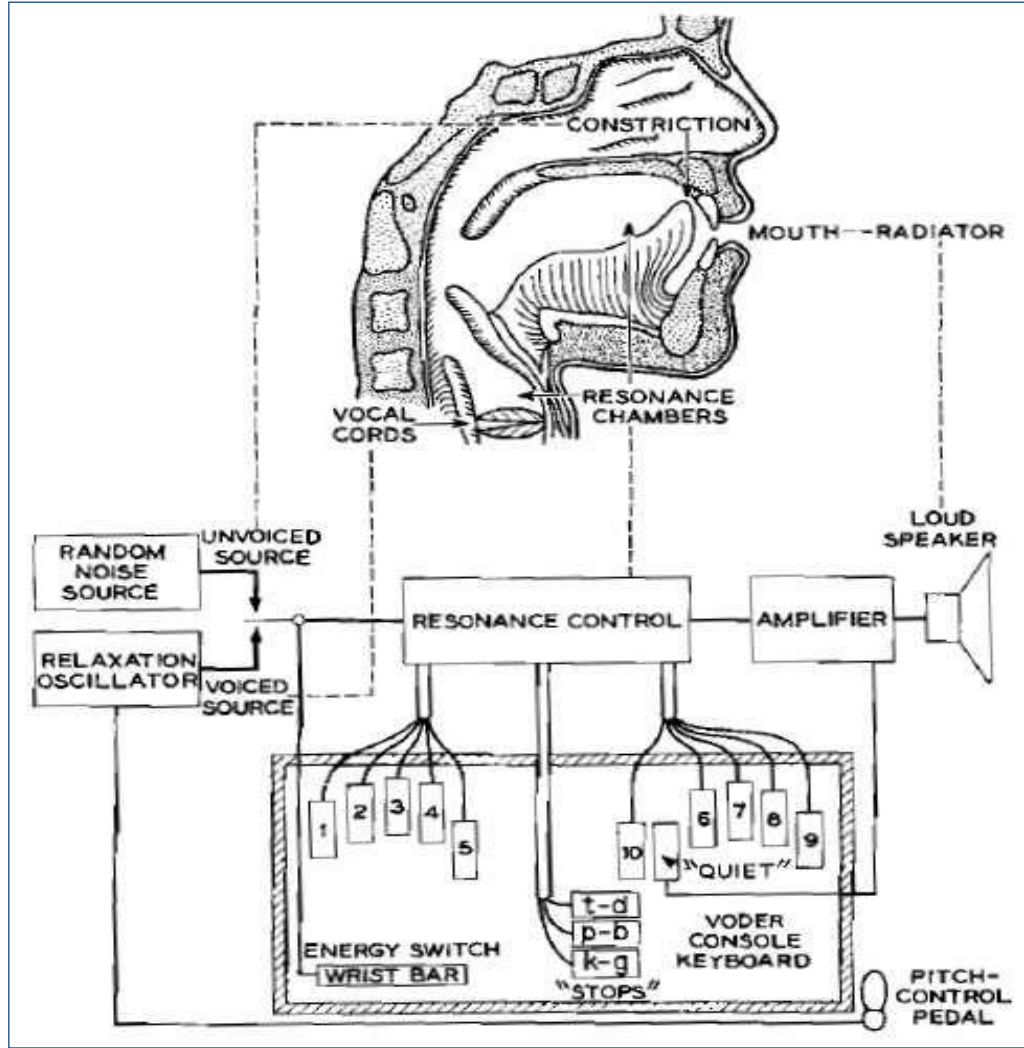


Şekil 2.4. Riesz tarafından geliştirilen konuşma makinesi kontrol tuşları [1].

Mekanik ve yarı elektrikli sistemlerle ses üretme araştırmaları 1960'lı yıllara kadar devam etmiş, ancak pek başarılı sonuçlar elde edilememiştir [1,2].

Bütünüyle elektrikli olarak çalışan konuşma motoru, ilk olarak Stewart tarafından 1922 yılında geliştirilmiştir [4]. Bu konuşma motoru bir zil ve iki titreşim devresinden oluşmaktaydı. Ancak konuşma motoru sadece tek ünlü sesleri üretebiliyor, buna karşılık ünsüz ve birleştirilmiş sözcükleri üretmiyordu [3].

1939 yılında Riesz ve Watkins tarafından elektrikli olarak tasarlanmış ve konuşma motoru olarak dikkate alınması gereken ilk cihaz VODER (Voice Operating Demonstrator) geliştirilmiştir [1]. VODER'in şematik gösterimi Şekil 2.5'de sunulmuştur.



Şekil 2.5. Riesz ve Watkins'in VODER şeması [1].

Cihazın titreşim kontrol kutusu 10 adet bant geçiren filtre içermekte, bu filtreler üretim esnasında paralel olarak çalışmaktaydı. Tüm bu filtreler bilek barı (Wrist Bar) denilen yapıyla, seçilen rastlantısal gürültü kaynağı (Random Noise Source) veya vızıltı üretici (Relaxation Oscillator) tarafından uyarılmaktaydı. Ayak pedalı ise sinyal üreticinin perde, diğer bir deyişle genlik değerini belirlemekteydi. Bant geçiren filtre çıkışları potansiyometreye gönderilmekteydi. Bu potansiyometrelerde yine 10 adet tuş yardımıyla kontrol edilmekteydi [1].

VODER'den sonra konuşma motorları bilim dünyası için çok daha ilgi çekici bir alan haline gelmiştir [3].

İlk formant sentezleyici konuşma motoru PAT (Parametric Artificial Talker) Walter Lawrence tarafından 1953 yılında tanıtılmıştır [4].

Aynı sıralarda Gunnar Fant ilk ardışık bağlı formant sentezleyicili konuşma motoru OVE I (Orator Verbis Electricis)'i tanıtmıştı. Sonralarında bu proje OVE II ve OVE III olarak devam etti. [3].

İlk söyleyiş sentezleyici konuşma motoru ise George Rosen tarafından M.I.T. (Massachusetts Institute of Technology)'de tanıtılmıştı [4].

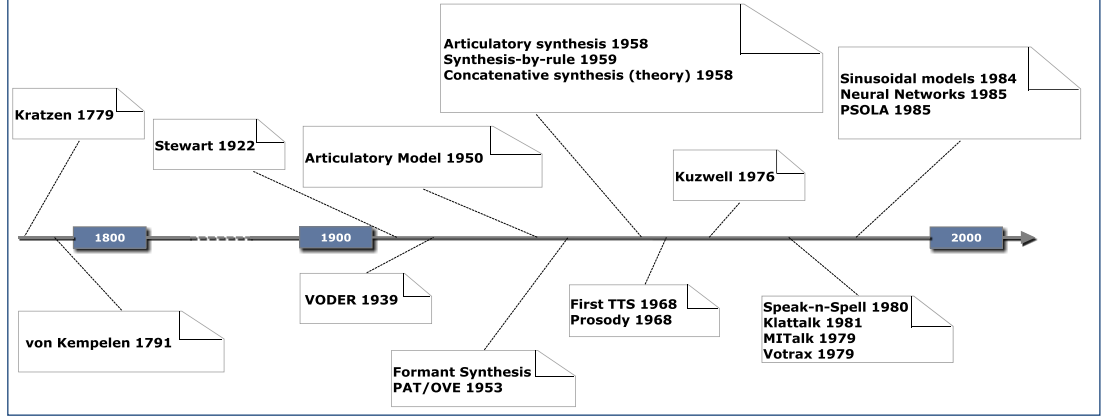
İngilizce için ilk tam metinsel konuşma motoru 1968 yılında Japonyada bulunan Electrotechnical laboratuvarında Noriko Umeda ve arkadaşları tarafından geliştirilmiştir [4]. Sistemin temeli söyleyiş sentezleyici modeline dayanmaktaydı. Sistem anlaşılır bir sentezleme yapabilmekle beraber, günümüz konuşma motorlarının seslendirme kalitesiyle kıyaslanamaz bir ses çıktısı üretmekteydi [3].

1979 yılında Allen, Hunnicutt, ve Klatt MIT'de geliştirilen MITalk laboratuvarı metinden konuşma sistemini tanıtmışlardır [3].

1981 yılında Dennis Klatt daha gelişmiş bir konuşma motoru olan Klattalk sistemini tanıtmıştır. 1970'lerin sonları ve 1980'li yılların başlarında ticari konuşma motorlarında artış olmuştur [4].

Konuşma sentezi için ilk tümleşik devre, muhtemelen kademeli formant synthesizer ve basit alçak geçiren yumuşatma devrelerinden oluşan Votrax yonga (çip) olmuştur. 1978 yılında Richard Gagnon ucuz bir Votrax tabanlı Tip-n-Talk sistemi tanıtmıştır. İki yıl sonra, 1980 yılında, Texas Instruments doğrusal tahmini kodlama (LPC) tabanlı Speak-n-Spell sentezleyici yongası TMS-5100 tanıtmıştır. TMS-5100 çocuklar için okuma yardımcısı olarak kullanılmıştır. 1982 yılında Street Electronics TMS-5100 yongasının gelişmiş versiyonu olan TMS-5220 yongasıyla geliştirilmiş Echo düşük maliyetli ikili ses birleştirici (Echo low-cost diphone synthesizer) konuşma motorunu tanıtmıştır. Aynı tarihlerde Speech Plus şirketi Prose-2000 metinden konuşma sentezleyici sistemini, bundan bir yıl sonrada meşhur konuşma

motoru DECtalk ve Infovox SA-101 konuşma motorları tanıtılmıştır [3,4]. Şekil 2.6'da Konuşma Motorlarının tarihsel gelişimi verilmiştir.



Şekil 2.6. Konuşma motorlarının tarihsel gelişimi [3].

2.3. KONUŞMA SENTEZLEMEDE KARŞILAŞILAN ZORLUKLAR

Konuşma sentezlemenin, seçilen sentezleme tekniği ile ilgili zorluklarının olmasının yanında, doğrudan içerikle alakalı zorlukları da mevcuttur. Örnek olarak formant sentezleme tekniğinin temeli, insan vokal organlarının modellenmesine dayanır. Vokal organlardaki değişken parametreler göz önüne alındığında ilgili tekniğin karmaşıklığı açıkça görülmektedir. Çalışmanın konusu olan eklemeli sentezleme tekniğini kullanan, konuşma motorunun geliştirme sürecinde ise, ses veri tabanının hazırlanması, metin ön işleme problemleri, özel okunuşa sahip kelime yapıları (seslendirme) gibi zorluklarla karşılaşmış, cümle içindeki kelimenin tonlama, vurgu, süre hesap analizlerinin yapılması konularına ilişkin model çözümler önerilmiştir.

2.3.1. Veri Tabanı Hazırlamada Karşılaşılan Zorluklar

Eklemeli sentezleme yapan konuşma motorları, kendi veri tabanlarına sahip konuşma motorlarıdır. Teknikte kullanılan en küçük ses birimi çeşidine göre veri tabanı boyutu değişmektedir. Birleştirmede tek ses elemanları birleştirme birimi olarak seçilmiş ise ilgili dilde mevcut tek seslerin büyüklüğü oranında bir veri tabanına ihtiyaç olacaktır. Türk alfabesinde 29 harf bulunmaktadır. Bu 29 harfin her biri bir

tekli sesi ifade etmektedir. Bunlara ek olarak yazılışlarında ve okunuşlarında ufak farklılık gösteren sesler mevcuttur. Örneğin a ve â gibi. İşte ilgili dildeki tüm tekli seslerden kasıt budur. Birleştirme tekniğinde birleştirme birimi olarak ikili, üçlü, dörtlü vb. seslerin yanı sıra kelimeler dahi kullanılabilir. Birleştirmede kullanılacak ses birimi elemanları büyüdükçe veri tabanı boyutu da artmaktadır. Örnek olarak ikili seslerin (difon) birleştirme birimi olarak kullanıldığı bir Türkçe sentezleyicide Türkçe 'deki tüm ikili sesler 900 ile 1000 arasında değişmektedir.

2.3.2. Metin Ön İşlemede Karşılaşılan Zorluklar

Metin ön işleme tüm metnin veri tabanı ile uyumlu bir hale getirilme işlemidir. Bu işlem metnin büyük küçük harf uyumundan, rakamların yazı dilindeki ifadesine dönüştürülmesine, para birimleri, tarih, saat vb. tüm ifadelerin normalizasyona tabi tutulmasıdır. Tüm dillerde de olduğu gibi Türkçe'de de cümlenin yapısına göre okunuşu değişen ifadeler bulunmaktadır. Örnek olarak "S.G.K." kısaltması cümle içinde "SGK" veya "sosyal güvenlik kurumu" şeklinde okunabilmektedir. Yine aynı şekilde "/" işareti cümle içinde "bölü" şeklinde okunurken adres içeren metinlerde "taksim" şeklinde okunmaktadır [5]. Konuşma motorunun, ifadenin doğru seslendirme şekline karar verip, ona göre seslendirmesi gerekmektedir.

2.3.3. Bazı Kelimelere Has Özel Okunuşlar

Genel olarak Türkçe yazıldığı gibi okunan bir dil olsa da aynı özellikteki dillerde olduğu gibi özel durumlar içerebilir. Örnek olarak "fedakarlık" kelimesinin okunurken yazıldığından farklı bir formata seslendirilmesi gerekir. İlgili kelime okunurken "fedaakearlık" şeklinde okunmalıdır. Ayrıca dile sonradan giren kelimelerin okunuşlarının belirlenmesi de ayrı bir problem olarak karşımıza çıkmaktadır. Konuşma motorunun, özel okunuş gerektiren bu ve benzer kelimeleri belirlemesi ve bunların doğru telaffuzunu yapması gerekmektedir.

2.3.4. Aruz (Tonlama, Vurgu ve Süre vb.) da Karşılaşılan Zorlukları

Konuşmadaki tonlama, vurgu ve süre ile yakından alakalı olan aruz kavramının soyut ve somut olmak üzere iki farklı yaklaşımı mevcuttur. Somut olarak aruz yaklaşımı, akustik parametrelerden olan ses perdesi, süresi ve ses yoğunluğu kavramlarıyla ifade edilir. Bu yaklaşım aruzun somut anlamlar ve özel akustik şekiller arasındaki ilişkisini tanımlar. Soyut olarak aruz yaklaşımı ise sesin dilbilimsel yapısından ziyade fonetik doğasını, sınıflandırılması vb. konularını ifade eder [6]. Görüldüğü gibi aruzun somut yaklaşımlarına sinyal işleme teknikleriyle kısmen cevap verilebilirken, soyut kavramlarının (konuşmacının ruh hali, kızgınlık vb.) aktarılmasında zorluklarla karşılaşmaktadır.

2.4. ÖRNEK KONUŞMA MOTORU SİSTEMLERİ

2.4.1. MITalk

MITalk konuşma motoru Allen, Hunnicutt, ve Klatt tarafından 1970'li yıllarda MIT' de geliştirilmeye başlanmıştır. 1979 yılında da proje son halini almıştır [3,4]. MITalk sadece PASCAL ve C gibi genel amaçlı dillerde kodlanmıştır. Günümüzde konuşma motorlarının bir dilden diğer bir dile taşınabilir yapıda olması yaygın bir kanı olarak karşımıza çıkmaktadır. MITalk en bilinen konuşma motorları arasındadır [7].

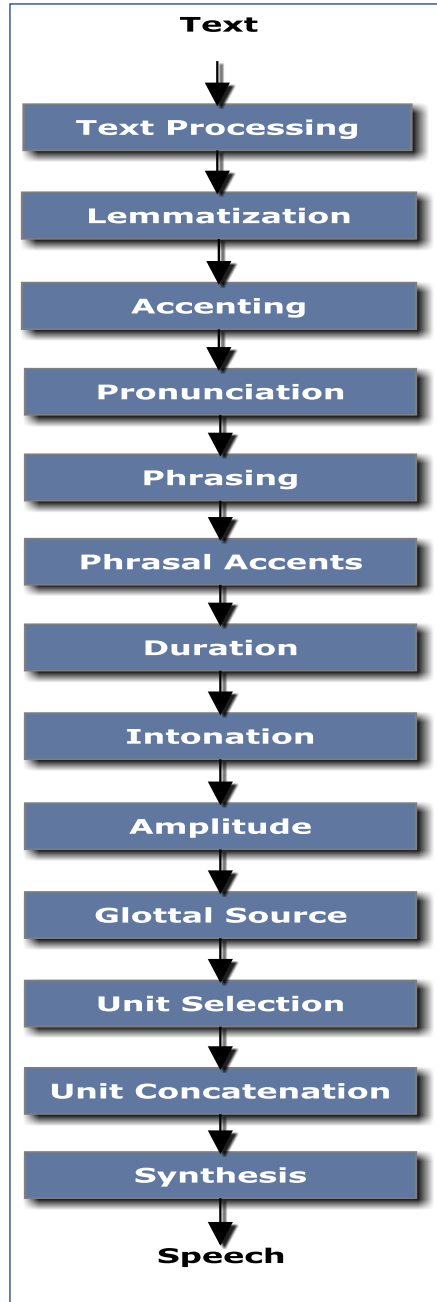
2.4.2. Infovox

Infovox ticari amaçlı geliştirilmiş, çoklu dil desteği olan, en ünlü konuşma motorlarından biridir. İsviçre de geliştirilmiştir. İlk geliştirilen sürümü ardışık formant sentezleyici tipindedir [3,8].

2003 yılında Infovox (İsviçre), Babel Teknoloji (Belçika), Elan Speech (Fransa) şirketleri Acapela grup şirketi adı altında birleşmiştir. Günümüzde bu grup şirketi halen konuşma motoru yazılım ve hizmetleri üretmeye devam etmektedir. Şu an yüz farklı ses ile otuz farklı dilde konuşma motoru hizmeti verebilmektedir [9,10].

2.4.3. Bell Laboratuvarları TTS Sistemi

AT&T Bell laboratuvarlarında geliştirilen, çoklu dil desteği olan, ikili üçlü vb. şekilde ifade edilen doğal konuşma parçalarının birleştirilmesi temeline dayanan konuşma sistemidir. Sistem mimarisi modüler bir yapıda inşa edilmiştir. Bu modüler yapı sistemli ve hesaplama kolaylığı avantajı sunan bir mimaridir. Sistem 13 farklı modülden oluşmaktadır [11]. Şekil 2.7’de ilgili şekiller gösterilmiştir.



Şekil 2.7. Bell Laboratuvarları TTS Sistemi modülleri [11].

2.4.4. Cnet Psola

1980'lerde France Telecom-CNET tarafından tanıtımı yapılmış ikili ses (difon) tabanlı birleştirmeli yapıda bir konuşma motorudur. Birleştirme tekniği olarak, kendi patenti altındaki PSOLA (Pitch Synchronous OverLap Add) tekniğini kullanmaktadır. Gerçek zamanlı, çoklu kanal sentezleme yapmaktadır. İlk sürümü Almanca olarak geliştirilmekle beraber, karma sistemler için İngilizce, İtalyanca ve Fransızca diğer laboratuvarlarla ortaklaşa aynı konuşma motoru kullanılarak geliştirilmiştir. Konuşma motorunun şu an için İngilizce, Almanca, Fransızca ve İspanyolca desteği bulunmaktadır [12].

2.4.5. Festival

Sistem Edinburg Üniversitesi Konuşma Teknolojisi Araştırma Merkezinde (Edinburgh University Centre of Speech Technology Research) C++ programlama dilinde geliştirilmiştir. Çok dilli bir konuşma motorudur. İngiliz İngilizcesi, Amerikan İngilizcesi ve İspanyolca desteği mevcuttur. Proje "festvox" adı altında halen devam etmektedir [13].

2.4.6. Mbrola

İkili ses (difon) birleştirme temeline dayanan bir konuşma motorudur. Birleştirmede PSOLA tekniğine benzer yapıdaki MBR-PSOLA (Multi-Band Rcsynthesis Pitch Synchronous OverLap Add) tekniğini kullanmaktadır. Belçika'daki Mons Üniversitesinin TCTS laboratuvarında geliştirilmiştir. Sistem askeri ve ticari olmayan uygulamalar için ücretsiz olarak kullanılabilir. MBROLA şu an Türkçe de dahil olmak üzere en az bir erkek ve bir bayan sesi ile 33 dilde destek vermektedir. MBROLA TTSBOX adı altında bir Matlab toolbox sunarak da lisans ve lisansüstü öğrencileri TTS sistemlerinin anlaşılabilirliğini artırmayı amaçlamaktadır [14,15].

2.5. TÜRKÇE TEMELLİ KONUŞMA MOTORLARI

İncelenen konuşma motorları çoğunlukla İngilizce temelli konuşma motorlarıdır. Çok dilli konuşma motorları da hemen bütün dillerde destek vermektedir. Geliştirilen projelerin çoğu da üniversite destekli araştırma merkezlerinde yürütülmektedir. Yapılan araştırmalarda ticari veya kısa soluklu projelerin dışında, Türkçeyi temel alan, bitmiş veya devam eden bir konuşma motoru sistemine rastlanmıştır. İlgili konuda yapılan araştırma ve çalışmaların bahsedilen eksiklikleri giderme noktasında katkıları olacağı kanaatine varılmıştır.

2.6. TÜRKÇE İÇİN YAPILAN ÇALIŞMALAR

Eker çalışmasında Türkçe için ikili ses birleştirme tekniğini kullanan bir metin seslendirme sistemi anlatmaktadır. Sistem girdi olarak bir metin almakta ve çıktı - olarak bu metne karşılık gelen Türkçe sesleri üretmektedir. Çalışmada Türkçe'nin fonetik bir dil olduğu, bu sistemin ufak değişikliklerle benzer fonetik diller içinde kullanılabilir olduğu vurgulanmıştır [24].

Sak çalışmasında kulağa doğal insan sesi gibi gelen, birleştirmeli bir konuşma sentezleme sistemi üzerinde çalışılmış ve sentetik konuşma kalitesini değerlendirmek için öznel dinleme testleri yapılmıştır. Sonuç olarak MOS benzeri bir derecelendirmede 4,2 puan alan bir Türkçe konuşma sentezleme sistemi geliştirilmiştir [16].

Ünaldı çalışmasında mevcut konuşma motoru sistemleri ve bu sistemlerde kullanılan konuşma sentezleme yöntemleri ile taşınabilir cihazlar için var olan uygulama geliştirme platformlarını incelemiştir. Bu inceleme sonucunda difonları birleştirerek sentezleyen TDPSOLA yöntemi ile Java 2 Platformunda gelen kısa mesajları konuşmaya çeviren bir taşınabilir uygulama üzerinde çalışılmıştır [17].

Aşlıyan ve Günel çalışmalarında, Türkçe metinlerin insan sesine dönüştüren hece birleştirmeli bir konuşma sentezleme sistemi geliştirilmiştir. Çalışmanın ilk aşamasında Türkçe metni heceleyen bir algoritma TASA (Turkish Automatic

Spelling Algorithm) tasarlanmıştır. Algoritma C++ dili kullanılarak test edilmiştir. Hatalı heceleme oranı yaklaşık %0 olarak gözlemlenmiştir. Böylece, Türkçe 'de var olan heceler tespit edilmiştir. Çalışmanın ikinci aşamasında her bir hece bir konuşmacı tarafından seslendirilerek ve ön işlemden geçirilerek wav dosyası olarak hece ses veri tabanına dâhil edilmiştir. Çalışmanın en son aşamasında, sistemin kullanıcı arabirimine eklenen metin TD-SOLA algoritması kullanılarak seslendirilmiştir. Bu metindeki sözcükler ön işlemden geçirilerek hecelere ayrılmıştır ve her bir heceye karşılık gelen ses dosyaları birbirine eklenerek sözcük ses dosyası elde edilmiştir. Sözcük ses dosyaları da eklenerek cümle ses dosyaları oluşturulup seslendirilmiştir [18].

Uslu ve İlk çalışmalarında, ikili sesbirim (difon) birleştiren eklemeli bir Türkçe Metinden Konuşma Sentezleme sistemi için geliştirilen ezgi modelini anlatmaktadır. Bu bağlamda pek çok dilde perde frekansı değişimini başarıyla temsil eden Fujisaki ezgi modeli, Türkçe'ye uygulanmıştır. Geliştirilen model, sentezlenen konuşmaya doğallık kazandırmış, sentez perde frekansı eğrisini, özgün perde frekansı eğrisine oldukça yaklaştırmıştır [19].

Canal, Kurnaz ve Yılmaz çalışmalarında Türkçe için ünlüden-ünlüye ses sentezleme yöntemini kullanan bir metin seslendirme sistemi anlatılmıştır. Sistem, girdi olarak yazılı metni almakta, standartlaştırmakta ve çıktı olarak bu metne karşılık gelen Türkçe sesleri üretmektedir. Geliştirilen uygulama, metni sadece seslendirmekle kalmayıp anlamsal olarak da incelemekte ve bu sayede metnin Türkçe'deki doğru okunuşunu belirleyebilmektedir. Sonuç olarak, Türkçe metinden konuşma sentezleme konusunda yaşanan sıkıntılar belirlenmiş, bu sıkıntılardan noktalama işaretleri, kısaltmalar, nümerik ifadeler ve ses değişimleri konularında çeşitli çözüm yöntemleri geliştirilmiş ve insan sesine yakın bir metinden konuşma sentezleme sistemi geliştirilmiştir [5].

Erdemir çalışmasında Matlab ortamında Türkçe metinleri seslendiren bir metin seslendirme sistemi sunmuştur. Bu sistem dalga şekli birleştirme yöntemiyle ve “en çok iki harfli hece” parçacık birimi temel alınarak oluşturulmuştur [20].

Arık çalışmasında eklemeli sentezleme tekniđi kullanılan bir hece tabanlı Türkçe metinden konuşma sentezleme yazılımı geliştirilmiştir. Çalışmada girdi metni, yazılım içine entegre edilen Türkçe heceleme algoritması ile hecelere ayrılmış ve ses sinyallerine dönüştürülmüştür. Yapılan doğrulama ve geçerleme çalışmaları sonucunda yazılımda üretilen konuşmanın kalitesi MOS benzeri bir derecelendirmeden 5 üzerinden 3,7 puan almıştır [21].

Tekindal ve Arık çalışmalarında görme engellilerin bilgisayar kullanmalarını etkinleştirmeyi amaçlamışlardır. Bu amaca yönelik olarak, daha kolay ve daha pratik parçalama ve etiketleme yapısı sunan, parça tabanlı ardışık eklemeli konuşma sentezleme tekniđi kullanan, hece tabanlı Türkçe metinden konuşma sentezleme yapabilen “Okuyucu TTS” isimli bir metin okuyucu yazılımı geliştirilmiştir. Yazılım Visual C#.NET programlama dili ile kodlanmıştır. Geliştirme ortamı olarak Visual Studio 2008 Professional geliştirme ortamı kullanılmıştır. Yazılım tamamlandıktan sonra görme engellilerin ihtiyaçlarına cevap verip vermediđine yönelik doğrulama ve geçerleme çalışmaları 10 görme engelli kişinin katılımı ile yapılmıştır. Yapılan doğrulama ve geçerleme çalışmaları sonucunda yazılımda üretilen konuşmanın kalitesi MOS benzeri bir derecelendirmeden 5 üzerinden 3,7 puan almıştır [22].

2.7. FARKLI DİLLER İÇİN YAPILAN ÇALIŞMA ÖRNEKLERİ

Chappell ve Hansen çalışmalarında sınırlı bir veri tabanı kullanarak sentezleme yapan eklemeli tipteki konuşma motorlarında, ekleme yapılan parçalar arasındaki geçişleri ve pürüzsüzlüđü sağlamak için yöntemler önermektedirler. Çalışmada önerilen farklı yöntem ve algoritmalar karşılaştırılmış artı ve eksilerinin üzerinde durulmuş, doğru kullanılan yumuşatma algoritmalar ile doğrudan birleştirme tekniklerinden daha iyi sonuç verdiđi görülmüştür. [23].

Vinet çalışmasında konuşmadaki vurgu ve doğallıđı artırmaya hedef alan bir formant sentezleyici konuşma motoru olan GLOVE’u geliştirmiştir. GLOVE bir kural oluşturma uygulaması olan RULSYS ile kontrol edilmektedir. Projenin genel amacı GLOVE tarafından üretilen ünsüz seslerin veri odaklı bir model kullanılarak deđiştirilmesidir. Bu üç işlem basamađından oluşan bir yapıdır. Birinci basamakta

değiştirilecek sesler tespit edilmekte, ikinci basamakta bu ses yeniden üretilmekte ve son olarak da ilgili sese yumuşatma teknikleri uygulanmaktadır [24].

Schröder ve Trouvain çalışmalarında Almanca için geliştirilmiş Mary adlı XML tabanlı konuşma motorunu tanıtmışlardır. Uygulamalarına sentezlenen konuşmanın duygusallığını artırma amaçlı bir duygu modülü de eklemişlerdir. [25].

Thomas çalışmasında Hint dilleri için Festival adlı konuşma motorunun omurgasını kullanan hece tabanlı bir konuşma motoru geliştirmiş, ikili ses tabanlı konuşma motorlarına oranla daha iyi sonuçlar elde etmişlerdir [26].

Bu çalışmada da, dijital ortamdaki Türkçe metinleri ses sinyallerine dönüştürebilen, farklı platformlara uyarlanabilirliği yüksek, geliştirmeye açık, akıcı, anlaşılır, veri tabanı boyutu düşük, Türkçe tabanlı eklemeli tipte bir konuşma motoru geliştirilmesi amaçlanmıştır.

BÖLÜM 3

KONUŞMA SENTEZLEME

3.1. KONUŞMA SENTEZLEME TEKNİKLERİ

3.1.1. Söyleyiş Sentezleyiciler (Articulatory Synthesis)

Konuşma motorları ilk geliştirilmeye başlandığında kullanılan tekniktir. İlk söyleyiş sentezleyici konuşma motoru 1958 yılında George Rosen tarafından M.I.T de tanıtılmıştır [4].

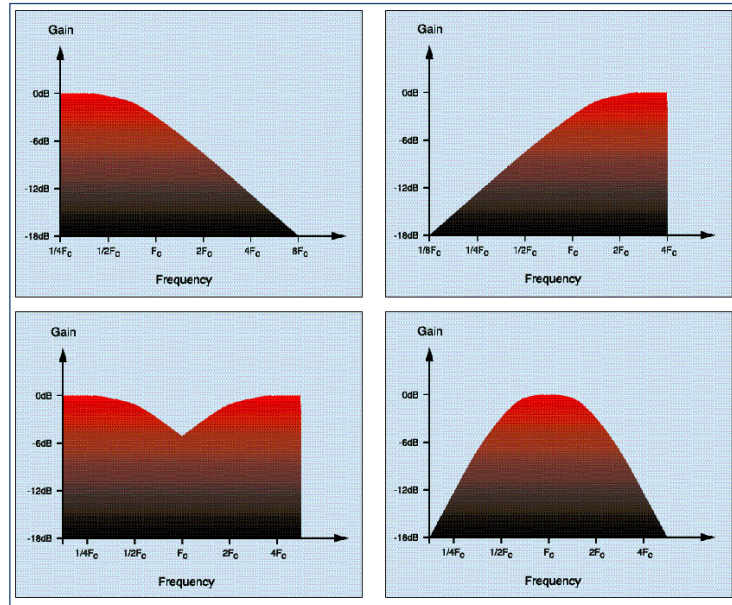
Söyleyiş Sentezleyiciler insan konuşma sisteminin nörofizyolojik ve biyomekaniksel açıdan modellenmesini temel alır. Bu teknik diğer teknikler arasında daha az gelişmiş bir teknik olsa da, detaylı ses dalgası üretimi için diğerlerinden daha doğru bir seçim olduğu izlenimini vermektedir. Söyleyiş sentezleme belki de insanın vokal organlarını bütünüyle modellenmesinin zorluğu hesaba katılmadan kesin bir sentezleme olarak tanıtılmıştır. Ses bilimciler henüz tam anlamıyla Söyleyiş sentezleyiciyi destekleyen bir model bulunmadığı konusunda hemfikirdir [27].

Söyleyiş sentezleyici modelleri insan konuşma organlarının oluşturan kasların ve diğer yapıların konuşma anındaki hareket değişimlerini baz alır. Bu veriler genel olarak 2 veya 3 boyutlu X-Ray analizleriyle elde edilir. Bu yüzden modellenmesi en zor tekniktir. Gerek modellenme zorluğu, gerekse çalışma anında bu modelin gerektirdiği hesaplama yoğunluğu sebebiyle günümüz sistemlerinde nadir kullanım alanı vardır [24].

3.1.2. Formant Sentezleyiciler (Formant Synthesis)

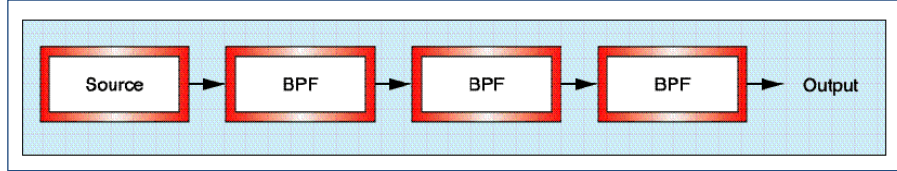
Formant sentezleme tekniđi son zamanlarda en çok kullanılan sentezleme tekniđi olarak karřımıza çıkmaktadır. Kural temelli sentezleme olarak da adlandırılmaktadır. Teknik yapısı itibariyle doğadaki tüm seslerin üretilmesine olanak sağlamaktadır. Diğer tekniklerde olduđu gibi ses örneklerinden oluşan bir veri tabanına ihtiyaç duymaz. Dolayısıyla formant sentezleyiciler daha az hafıza gereksinimi duyan aynı zamanda da çok fazla hesaplama yükü taşıyan sistemlerdir. Ayrıca Formant sentezleme tekniđiyle üretilmiş pek çok uygulama, insan sesine benzemeyen robotsu bir ses olarak karřımıza çıkmaktadır. DECTalk, MITalk, Infovox formant sentezleme tekniđi kullanılarak geliştirilmiş ilk konuşma motorlarıdır. Formant sentezleyiciler kendi aralarında paralel ve seri formant sentezleyiciler olarak ikiye ayrılrsa da bu iki yapının beraberce kullanılması daha doğal sonuçlar vermektedir [27-29].

Formant sentezleyiciler Yapısı basitçe “Kaynak-Filtre” temeline dayanmaktadır. Bu temel yapıda kaynak olarak bir osilatör veya benzeri bir sinyal üretici, filtre olarak ise bilinen temel filtreler kullanılmaktadır. Şekil 3.1’de 4 temel filtre çıkış sinyali gösterilmektedir. Bunlar sırasıyla alçak geçiren, yüksek geçiren, bant geçirmeyen ve bant geçiren filtre çıkış sinyalleridir [28,29].

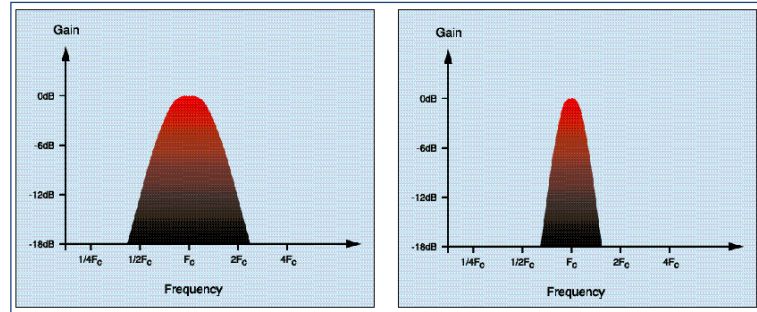


Şekil 3.1. Sırasıyla alçak geçiren, yüksek geçiren, bant geçirmeyen ve bant geçiren filtre çıkışları [29].

Şekil 3.1’de 4. Sırada çıkışı verilen bant geçiren filtre Şekil 3.2’de gösterildiği gibi seri olarak bağlandığında sistemin çıkışında Şekil 3.3’dekine benzer çıkışlar elde edilebilir.

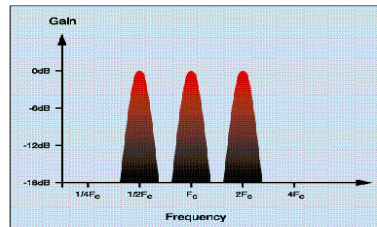
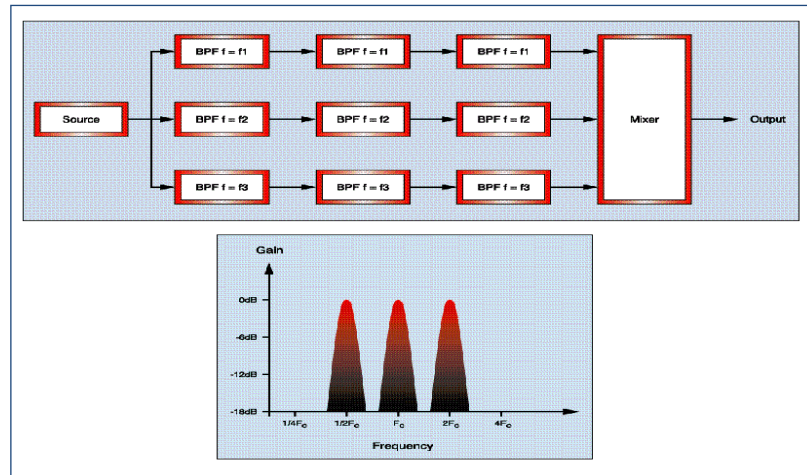


Şekil 3.2. Bant geçiren filtrelerin seri bağlanması [29].



Şekil 3.3. Bant geçiren filtrelerin seri bağlanması sonucu giriş ve çıkış sinyali [29].

Bant geçiren filtrenin paralel bağlanması ve sonucunda oluşan çıkış sinyali de Şekil 3. 4’de verilmiştir.



Şekil 3.4. Bant geçiren filtrelerin paralel bağlanması ve çıkış sinyali [29].

Şekil 3.4’den görüldüğü gibi seri ve paralel bağlama sonucu giriş sinyaline istenilen şekil verilebilmekte ve bu şekilde formant sentezleyici modellenilebilmektedir.

Modelleme sonucunda üretilen çıktıya, Latincedeki şekil anlamına gelen ‘formare’ yani formant denilmektedir [29].

3.1.3. Parça Sentezleyiciler (Concatenative Synthesis)

İngilizceden çeviride tam olarak “parça sentezleyici” kavramı elde edilmese de anlam olarak kastedilen karşılığı budur. Parça sentezleyiciler bahsedilen parçaların oluşturduğu veri tabanlarına sahiptirler. Bu ses parçaları ne kadar kısa seçilirse veri tabanı o oranda küçük olur. Ancak üretilen ses de aynı oranda doğallıktan uzaklaşır. Ses parçaları ne kadar büyük seçilirse doğallık o oranda artmakla beraber, veri tabanı büyük boyutlara ulaşabilir. Genel olarak parça sentezleyicilerde seçilen parça büyüklüğü, tek ses (phoneme) ile kelime arasında seçilir.

Tek ses sentezleyicilerde veri tabanı büyüklüğü tek bir dil için ilgili dildeki tüm tek seslerin büyüklüğü kadardır. Örneğin Türkçe de 29 harf bulunur. Bu 29 harfin her biri bir tek sesi ifade eder. Ancak Türkçe’deki tüm tekli sesler her bir 29 harfin sesiyle sınırlı değildir. Türkçede her bir harfin de uzun, kısa, kalın, ince vb. farklı seslendiriliş biçimleri mevcuttur. Buna örnek olarak Çizelge 3.1’de Türkçe’deki sesli harflerin okunuşları gösterilmiştir.

Çizelge 3.1. Türkçe sesli harflerin okunuş biçimleri [5].

Okunuşlar	Harfler							
	A	E	I	İ	O	Ö	U	Ü
Normal	a	e	ı	i	o	ö	u	ü
Uzun	Â	Ê	-	Î	Ô	-	Û	Û
	â	ê	-	î	ô	-	û	ÿ
İnce	Á	-	-	-	Ó	-	Ú	-
	á	-	-	-	ó	-	ú	-
Uzun-İnce	Ă	-	-	-	-	-	Ý	-
	ă	-	-	-	-	-	ý	-
Yumuşak	À	-	-	-	Ò	-	Û	-
	à	-	-	-	ò	-	ù	-
Geniş	-	Ë	-	-	-	-	-	-
	-	ë	-	-	-	-	-	-

Bir dildeki tüm seslerin doğru bir şekilde veri tabanında yer alması, sentezleyici diğer bir deyişle konuşma motoru etkinliğini belirleyen en önemli faktörlerden biridir.

Parça sentezleyiciler tek ses (fonem), ikili ses (difone), üçlü ses (trifone) veya çoklu ses (polifone) olmakla birlikte kelime sentezleyici olarak da adlandırılmaktadırlar. Kelime sentezleyici tarzındaki sistemler daha çok toplu taşıma araçlarındaki sınırlı, belirli kalıplardaki cümleleri sentezlemede kullanılırken, diğer yapılar daha çok ilgili dil veya dillere ait tüm metinlerin sentezlenmesi, bilgisayar tarafından okutulması amacıyla geliştirilir.

Parça sentezleyicilerde veri tabanından sonra sentezlemeyi etkileyecek diğer bir faktörde ilgili parçaların sentezlenme yani birleştirilme tekniğidir. Bilinen en yaygın birleştirme teknikleri PSOLA tekniğini temel alan TD-PSOLA, MBR-PSOLA (MBROLA) teknikleridir.

3.1.3.1. TD-PSOLA

TD-PSOLA (Time Domain Pitch Synchronous Overlap and Add), yani zaman domeni senkron perdelerini üst üste ekleme yöntemidir. Yöntem eklemelerdeki süreksizliği önlemek için, birbirine eklenmesi istenen ses parçalarının perdelerini doğru bir şekilde algılayarak, yine doğru bir şekilde senkronize birleştirilmesini amaçlar. Tekniğin en önemli sorunlarından biri sinyalin perde periyodlarının belirlenmesidir.

Yöntem iki kısma ayrılır. Bunlardan ilki giriş sinyallerini analiz eder ve parçalara ayırır, diğer kısım ise bu parçaları örtüştürür ve ekler. Analiz kısmında ilgili sesin periyodik parçalarının maksimum genlikleri, yani perde periyodu maksimumu (pitch mark) tespit edilir. Perde periyodu zaman aralıkları sabit kabul edilirse $P(t)$, eşitlik 3.1'den elde edilebilir.

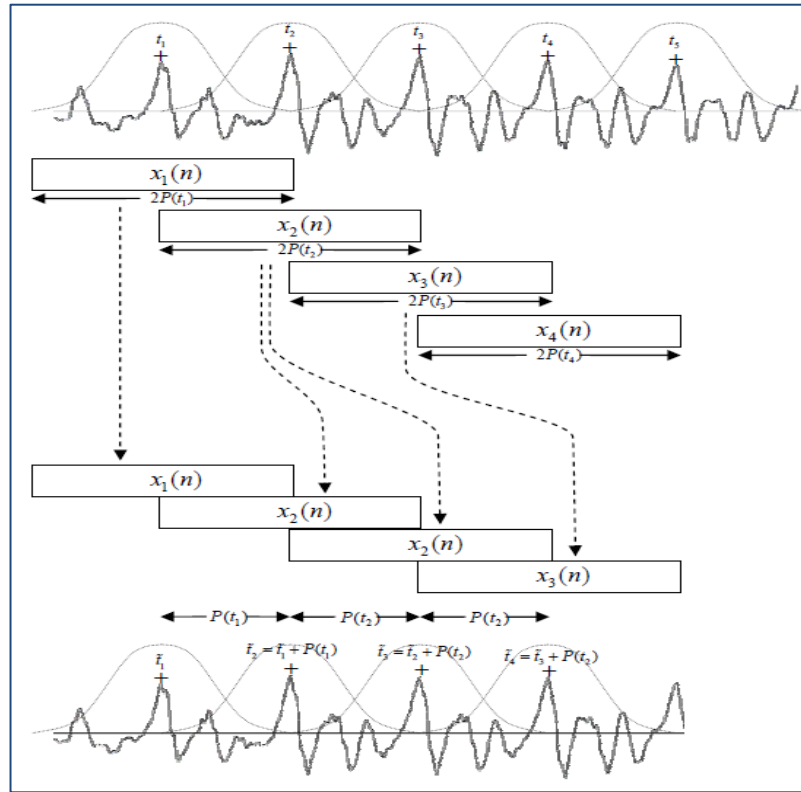
$$P(t_i) = t_{i+1} - t_i \quad (3.1)$$

Perde periyodu bulunan işaret bu periyoda göre ortalanarak parçalara ayrılır. Her parça iki perde periyodu uzunluğunda bir Hanning penceresiyle zarflanır. Bu Hanning pencereler üzerinden de örtüştürme işlemi yapılır. Hanning penceresi genel olarak ilgili sinyali zarflayan bir kosinüs işareti olarak ifade edilebilir.[30]

Germe faktörü α çıkış sinyali periyod uzunluğunu, diğer bir deyişle çıkış sinyali periyodunun giriş sinyaline oranla kaç kat artacağını belirler. Çıkış sinyali perde periyodunu $\tilde{P}(\tilde{t})$ ile ifade edersek bu periyod eşitlik 3.2'deki gibi yazılabilir. Çıkışın anlık zamanı \tilde{t}_k , "i" ilgili parçanın indeksi olmak üzere perde periyodu $P(t)$ 'den eşitlik 3.3'de ifade edildiği gibi elde edilebilir. Her birleştirmede çıkış perde periyodu \tilde{t}_k üstü üste getirilerek birleştirme işlemi yapılır. Şekil 3.5'de TD_PSOLA yöntemi ile yapılan genişletme işlemi grafiksel olarak ifade edilmiştir [30].

$$\tilde{P}(\tilde{t}) = \tilde{P}(\alpha t) = P(t) \quad (3.2)$$

$$\tilde{t}_{k+1} = \tilde{t}_k + \tilde{P}(\tilde{t}_k) = \tilde{t}_k + P(t_i) \quad (3.3)$$



Şekil 3.5. TD-PSOLA zaman domeninde genişletme [30].

3.1.3.2. MBR-PSOLA (Multi Band Resynthesis OverLap Add)

MBR_PSOLA diğerk bir deyişle MBROLA örtüşürme yöntemi Belçika'daki Mons üniversitesi Politeknik Fakültesinde geliştirilmeye başlanan MBROLA projesinde temelleri atılan örtüşürme yöntemidir [31].

MBR_PSOLA yöntemi TD-PSOLA yönteminin yenilenmiş bir halidir. Sonuçta her iki yöntemde PSOLA tabanlı yöntemlerdir.

Bu yöntemin en önemli özelliğı “pitch mark” diye adlandırılan tepe noktalarının tespit edilmesinin otomatikleşmesidir. Yöntem, birleştirilecek ses parçalarının sesli kısımları arasında ekleme yapmayı kolaylaştırmakta, bunun sonucunda da yüksek kalitede birleştirme sağlamaktadır [15].

Yöntem, zaman domeninde ses pencerelerini doğrudan örtüşürmekte ve pürüzsüz bir birleşme sağlamakta ve tüm bunları yaparken de bilgisayar tarafından yapılan hesaplama karmaşışını da azaltmaktadır. Yöntem, 16 KHz örnekleme frekansında intel486 işlemcili bir bilgisayarda test edilmiş diğerk yöntemlere oranla örnek başına ortalama 7 çalışma yükü daha az hesaplama sunmaktadır. Yöntem sayesinde örtüşürme sonrasında herhangi düzeltme işlemi gerekmemektedir [31].

MBROLA Şekil 3.6'da gösterildiğı gibi seslendirilmesi istenen seslerle beraber ilgili ses dosyasına ait perde ve süre gibi aruz bilgilerini içeren bilgileri de almaktadır.

```
; bonjour
_ 51 25 114
b 62
o 127 48 170
Z 110 53 116
u 211
R 150 50 91
_ 9
```

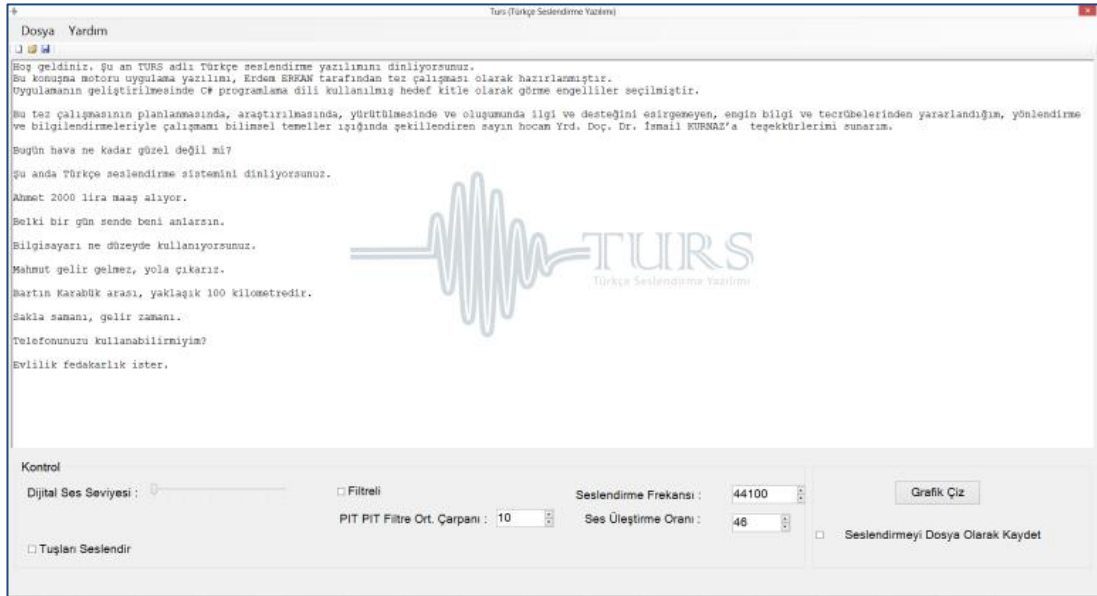
Şekil 3.6. MBROLA ile “bonjour” kelimesinin seslendirilmesi [31].

Şekil 3.6’da görüldüğü gibi “bonjour” bazı sesler ve bu seslere ait birtakım parametre değerleri ile sentezleyiciye gönderilmektedir. Örnek olarak ilgili kelime 51 mili saniyelik bir sessizlikle başlamakta, ilk perde periyodunun tepe noktası (pitch mark) 51 mili saniyenin %25’inden başlamakta ve 114 Hz frekansla devam etmektedir [31].

BÖLÜM 4

TURS TÜRKÇE KONUŞMA MOTORU

TURS Türkçe konuşma motoru, girilen Türkçe metinlerin Türkçe konuşma dilinde ses sinyalleri olarak çıktısını verebilen Microsoft Windows tabanlı bir ikili ses sentezleyicisi olarak geliştirilmiştir. Uygulama .Net Framework 4.5 üzerinde C# programlama dilinde, Microsoft Visual Studio 2012 geliştirme ortamı kullanılarak hazırlanmıştır. Ancak sistemin konuşmayı gerçekleştiren kodu tek bir sınıf halinde tasarlandığından, farklı yazılım dilleri ve işletim sistemlerine uyarlanması nispeten daha kolay olmaktadır. Uygulama metin dosyalarını açma, değişiklik yapma, kaydetme, yeni bir metin dosyası oluşturmanın yanı sıra tüm bu işlemleri ve ilgili metin dosyaların anlık olarak seslendirebilen bir yapıdadır. Şekil 4.1’de TURS konuşma motoru masaüstü uygulamasının ekran görüntüsü verilmiştir.



Şekil 4.1. TURS Türkçe seslendirme sistemi ekran görüntüsü.

Uygulamanın masaüstü versiyonunda geliştirme esnasında kullanılan “*Dijital Ses Seviyesi*”, “*Filtreli*”, “*PIT PIT Filtre Oralama Çarpanı*”, “*Seslendirme Frekansı*”,

“*Ses Üleştirme Oranı*”, “*Grafik Çiz*”, “*Seslendirmeyi Dosya Olarak Kaydet*” gibi değiştirilebilir parametre kontrolleri bulunmaktadır. Bu değişken parametre kontrolleri yardımıyla ses seviyesi programsal olarak artırılıp azaltılabilmekte, geliştirilen birkaç farklı filtre ilgili sese uygulanabilmekte, ses yükseltme alçaltma sırasında sese oluşan bozulmalar elimine edilebilmekte, seslendirme frekansı değiştirilebilmekte, seslendirme grafik olarak gösterilebilmekte ve uygulama tarafından seslendirilen metin wav dosyası olarak kaydedilebilmektedir.

4.1. SES VERİ TABANI OLUŞTURULMASI

Konuşma motoru tasarımında en fazla iş yükü veri tabanı hazırlama aşamasında karşımıza çıkmaktadır. Uygulama eklemeli tipte bir ikili ses sentezleyicisi olarak geliştirilmiştir. Bir metni konuşma sinyallerine çevirecek sistemin ilgili dil veya dillerdeki muhtemel tüm sesleri seslendirebilmesi için öncelikle uygulamanın kullanacağı tüm ikili seslerin doğru bir şekilde tespit edilmesi ve standart bir halde veri tabanında toplanması gerekmektedir.

4.1.1. Sistem İçin Gerekli Tüm İkili Seslerin Belirlenmesi

TURS Türkçe seslendirme sistemi için öncelikle Türkçede kullanılabilmesi muhtemel tüm ikili sesler belirlenmiştir. Bu bağlamda basit bir bilgisayar kodu yazılmıştır. Bu kod ile Türk alfabesindeki 8 sesli ve 21 sessiz harf, bunlara ek olarak da boşluk karakteri de bir harf gibi düşünülerek toplam 30 harfli bir alfabenin yan yana gelmesi muhtemel ikili sesler genel olarak belirlenmiştir. Matematiksel olarak belirlenen bu sesler tam olarak Türkçedeki ikili seslerin hepsine karşılık gelmemektedir. Çünkü Türkçe dâhil pek çok dilde yazımı aynı, okunuşu farklı olan sesler mevcuttur. Yani bir sesin uzun, kısa, ince, kalın vb. farklı okunuş biçimleri olabilmektedir. Örnek olarak “katil” ve “kartal” kelimelerinin her ikisinde de “a” sesi bulunmasına karşılık “katil” kelimesindeki “a” sesi uzun “kartal” kelimesindeki “a” sesleri ise kısa süreli seslendirilmelidir. Bu nedenle sistemin bir telaffuz ünitesi eklenmiş ve karşılaşılan eksiklikler bu şekilde giderilmeye çalışılmıştır. Belirlenen her ikili ses için anlamlı veya anlamsız bir taşıyıcı kelime belirlenmiştir. Taşıyıcı kelimeler barındırdıkları ikili seslerin doğal okunuşları ile kayıt altına alınmasını

sağlamışlardır. Örnek olarak “bd” ikili sesi için “abdi” taşıyıcı kelimesi belirlenmiştir. Bu şekilde “bd” ikili sesi “abdi” kelimesi içinde nasıl bir ses sinyaline karşılık geliyorsa veri tabanında da o şekilde kayıt altına alınmıştır. Şekil 4.2’de genel olarak ikili seslerin hesaplandığı kod bloğu görülmektedir.

```
char[] alfabe = {'_', 'a', 'b', 'c', 'ç', 'd', 'e', 'f', 'g', 'ğ', 'h', 'ı', 'i', 'j',  
                'k', 'l', 'm', 'n', 'o', 'ö', 'p', 'r', 's', 'ş', 't', 'u', 'ü', 'v', 'y', 'z'};  
  
private void Form1_Load(object sender, EventArgs e)  
{  
    string difone = "";  
    int row_index = 0;  
    for (int i = 0; i < alfabe.Length; i++)  
    {  
        char baz_harf = alfabe[i];  
  
        for (int t = 0; t < alfabe.Length; t++)  
        {  
            difone = baz_harf + alfabe[t].ToString();  
            dataGridView1.Rows.Add(1);  
            dataGridView1.Rows[row_index].Cells[0].Value = row_index;  
            dataGridView1.Rows[row_index].Cells[1].Value = difone;  
            row_index++;  
            difone = "";  
        }  
    }  
}
```

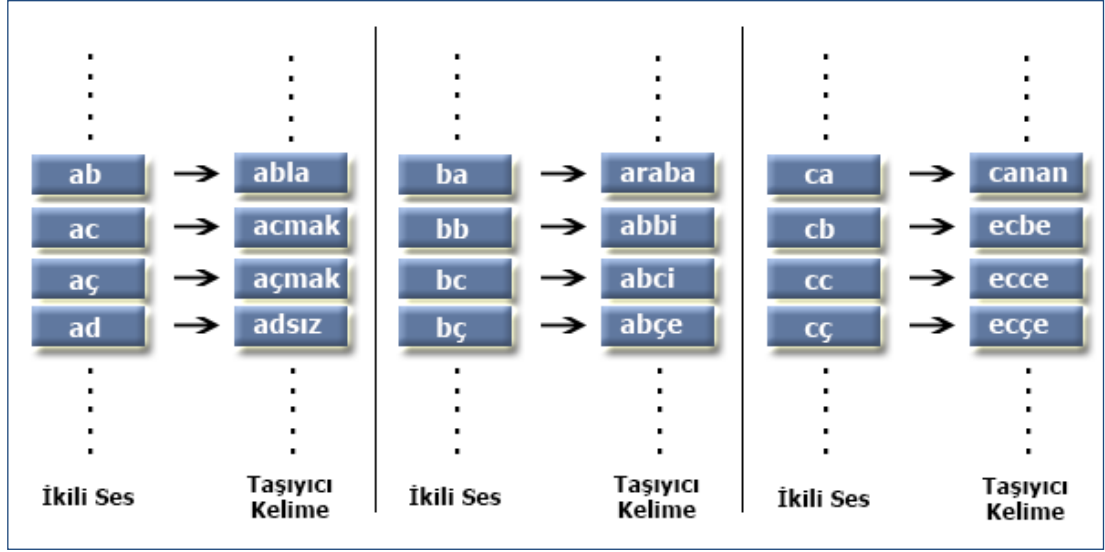
Şekil 4.2. Türkçe ikili seslerin hesaplanması için kullanılan fonksiyon.

Hesaplama sonucunda 900 adet ikili ses belirlenmiştir. Bazı seslere özgü okunuşların da sisteme eklenmesiyle 60 adet daha ikili sesin eklenmesi öngörülmektedir. Seslendirmede, belirlenen bu ikili sesler ile beraber üçlü ve dörtlü seslerinde kullanılmasının seslendirme kalitesini artıracığı, buna karşılık veri tabanı boyutunda yaklaşık 9000 adetlik fazladan ses parçası (değişken boyutta) yükü getireceği düşünülmektedir. Genel olarak hesaplanan tüm ikili sesler Çizelge 4.1’de sunulmuştur. Boşluk sesi “_” ile tanımlanmıştır. Uygulama için kullanılan tüm ikili sesler ve taşıyıcı kelimeler ayrıntılı olarak Ek Açıklamalar A.’da sunulmuştur.

Çizelge 4.1. Hesaplanan Türkçe ikili sesler.

Ses	a	b	c	ç	d	e	f	g	ğ	h	i	j	k	l	m	n	o	ö	p	r	s	ş	t	u	ü	v	y	z
—	a_	b_	c_	ç_	d_	e_	f_	g_	ğ_	h_	i_	j_	k_	l_	m_	n_	o_	ö_	p_	r_	s_	ş_	t_	u_	ü_	v_	y_	z_
_a	aa	ab	ac	aç	ad	ae	af	ag	ağ	ah	ai	aj	ak	al	am	an	ao	aö	ap	ar	as	aş	at	au	au	av	ay	az
_b	ba	bb	bc	bç	bd	be	bf	bg	bğ	bh	bi	bj	bk	bl	bm	bn	bo	bö	bp	br	bs	bş	bt	bu	bü	bv	by	bz
_c	ca	cb	cc	cç	cd	ce	cf	cg	cğ	ch	ci	cj	ck	cl	cm	cn	co	cö	cp	cr	cs	cş	ct	cu	cü	cv	cy	cz
_ç	ça	çb	çc	çç	çd	çe	çf	çg	çğ	çh	çi	çj	çk	çl	çm	çn	ço	çö	çp	çr	çs	çş	çt	çu	çü	çv	çy	çz
_d	da	db	dc	dç	dd	de	df	dğ	dğ	dh	di	dj	dk	dl	dm	dn	do	dö	dö	dö	dr	ds	dş	dt	du	dü	dv	dz
_e	ea	eb	ec	eç	ed	ee	ef	eg	eğ	eh	ei	ej	ek	el	em	en	eo	eö	ep	er	es	eş	et	eu	eü	ev	ey	ez
_f	fa	fb	fc	fç	fd	fe	ff	fg	fğ	fh	fi	fj	fk	fl	fm	fn	fo	fö	fp	fr	fs	fş	ft	fu	fü	fv	fy	fz
_g	ga	gb	gc	gç	gd	ge	gf	gg	gğ	gh	gi	gj	gk	gl	gm	gn	go	gö	gp	gr	gs	gş	gt	gu	gü	gv	gy	gz
_ğ	ğa	ğb	ğc	ğç	ğd	ğe	ğf	ğg	ğğ	ğh	ği	ğj	ğk	ğl	ğm	ğn	ğo	ğö	ğp	ğr	ğs	ğş	ğt	ğü	ğü	ğv	ğy	ğz
_h	ha	hb	hc	hç	hd	he	hf	hg	hğ	hh	hi	hj	hk	hl	hm	hn	ho	hö	hp	hr	hs	hş	ht	hu	hü	hv	hy	hz
_i	ia	ib	ic	iç	id	ie	if	ig	iğ	ih	ii	ij	ik	il	im	in	io	iö	ip	ir	is	iş	it	iu	iü	iv	iy	iz
_j	ja	jb	jc	jç	jd	je	jf	jg	jğ	jh	ji	jj	jk	jl	jm	jn	jo	jö	jp	jr	js	jş	jt	ju	jü	ju	iy	jz
_k	ka	kb	kc	kç	kd	ke	kf	kg	kğ	kh	ki	kj	kk	kl	km	kn	ko	kö	kp	kr	ks	kş	kt	ku	kü	kv	ky	kz
_l	la	lb	lc	lç	ld	le	lf	lg	lğ	lh	li	lj	lk	ll	lm	ln	lo	lö	lp	lr	ls	lş	lt	lu	lü	lv	ly	lz
_m	ma	mb	mc	mç	md	me	mf	mg	mğ	mh	mi	mj	mk	ml	mm	mn	mo	mö	mp	mr	ms	mş	mt	mu	mü	mv	my	mz
_n	na	nb	nc	nç	nd	ne	nf	ng	nğ	nh	ni	nj	nk	nl	nm	nn	no	nö	np	nr	ns	nş	nt	nu	nü	nv	ny	nz
_o	oa	ob	oc	oç	od	oe	of	og	oğ	oh	oi	oj	ok	ol	om	on	oo	öo	öp	or	os	oş	ot	ou	ou	ov	oy	oz
_ö	öa	öb	öc	öç	öd	öe	öf	ög	ög	öh	öi	öj	ök	öl	öm	ön	öo	öö	öp	ör	ös	öş	öt	öu	öü	öv	öy	öz
_p	pa	pb	pc	pç	pd	pe	pf	pg	pğ	ph	pi	pj	pk	pl	pm	pn	po	pö	pp	pr	ps	pş	pt	pu	pü	pv	py	pz
_r	ra	rb	rc	rç	rd	re	rf	rg	rğ	rh	ri	rj	rk	rl	rm	rn	ro	rö	rp	rr	rs	rş	rt	ru	rü	rv	ry	rz
_s	sa	sb	sc	sç	sd	se	sf	sg	sğ	sh	si	sj	sk	sl	sm	sn	so	sö	sp	sr	ss	sş	st	su	sü	sv	sy	sz
_ş	şa	şb	şc	şç	şd	şe	şf	şg	şğ	şh	şı	şj	şk	şl	şm	şn	şo	şö	şp	şr	şs	şş	şt	şu	şü	şv	şy	şz
_t	ta	tb	tc	tç	td	te	tf	tg	tğ	th	ti	tj	tk	tl	tm	tn	to	tö	tp	tr	ts	tş	tt	tu	tü	tv	ty	tz
_u	ua	ub	uc	uç	ud	ue	uf	ug	uğ	uh	ui	uj	uk	ul	um	un	uo	üo	üp	ur	us	uş	ut	uu	üü	uv	uy	uz
_ü	üa	üb	üc	üç	üd	üe	üf	üg	üğ	üh	üi	üj	ük	ül	üm	ün	üo	öö	üp	ür	üs	üş	üt	üu	üü	üv	üy	üz
_v	va	vb	vc	vç	vd	ve	vf	vg	vğ	vh	vi	vj	vk	vl	vm	vn	vo	vö	vp	vr	vs	vş	vt	vu	vü	vv	vy	vz
_y	ya	yb	yc	yç	yd	ye	yf	yg	yğ	yh	yi	yj	yk	yl	ym	yn	yo	yö	yp	yr	ys	yş	yt	yu	yü	yv	yy	yz
_z	za	zb	zc	zç	zd	ze	zf	zg	zğ	zh	zi	zj	zk	zl	zm	zn	zo	zö	zp	zr	zs	zş	zt	zu	zü	zv	zy	zz

Şekil 4.3’de ise taşıyıcı bazı kelimeler ve taşıdıkları ikili sesler gösterilmiştir.



Şekil 4.3. Örnek ikili sesler ve taşıyıcı kelimeleri.

4.1.2. Kayıt Ortamı

Taşıyıcı kelimeler içine gömülü ikili sesler, ses yalıtımı sağlanmış odada bir konuşmacı tarafından seslendirilmiştir. Seslendirmede kolaylık açısından ikili sesler başlangıcında bulunan sese göre a, b, c vb. şeklinde gruplara ayrılmış ve gruplar halinde seslendirmeye tabi tutulmuştur. Çekimlerde masaüstü bilgisayar fan seslerinin dezavantajlı olacağı düşünülerek, gürültü seviyesini en az seviyede tutmak için Intel i3 330M 2.13 GHZ işlemcili 4 GB ram belleğe sahip bir dizüstü bilgisayar ve Pirhana marka bir yaka mikrofonu kullanılmıştır. Ses kaybının olmaması ve kullanım kolaylığı açısından 44100 Hz. Tek kanal WAV (Waveform Audio File Format) dosya yapısı seçilmiştir. Kayıt için Audacity adlı yazılım kullanılmıştır.

4.1.3. Ses Formatı

Tüm cihazlar tarafından desteklenmesi, yaygın kullanım alanı olması, yapısının basit olması ve seslendirmede kullanılacak ses dosyaları üzerinde ekleme, kesme, güncelleme gibi işlemlerin yapılacak olması ses kayıt formatı olarak WAV dosya yapısının seçilmesinde etkili olmuştur. Bu dosya tipi temel ses dosyası tipi

olduğundan belirtilen işlemlerden önce herhangi bir dönüşüm gerektirmemektedir. Çizelge 4.2’de WAV dosya formatının yapısı ayrıntılı bir şekilde anlatılmıştır. Bu çizelgeden de anlaşılacağı gibi ilgili formatın ilk 44 baytlık kısmı dosyanın ana başlığını, 44. bayttan sonraki kısmı ise ses dosyasının ses örnek verisini oluşturmaktadır. Ses örneklerinde yapılan değişiklikler doğrudan dosya başlığı bölümünü de etkilemektedir. Örnek olarak ses verisi kısmında yapılan bir artma veya azalma başlık kısmındaki ses “Verisi Boyutunu” değişmesini gerektirmektedir. Bu doğrultuda seslerin birbirine eklenmesi, kesilmesi, uzatılması vb. tüm sinyal işleme işlemleri sonuçta oluşacak ses dosyasının hem başlık kısmını hem de ses örneklerini içeren kısmını doğrudan etkilemektedir.

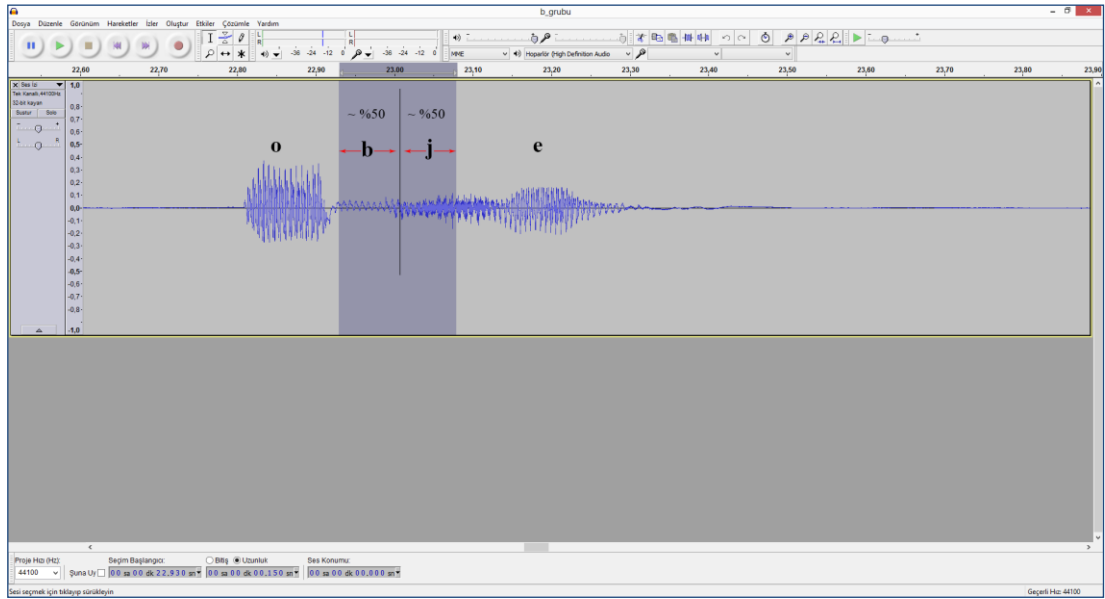
Çizelge 4.2. WAV dosya formatının yapısı.

Başlangıç	Boyut (Byte)	Adı	Açıklama
0	4	Zincir Kimliği (ChunkID)	ASCII biçiminde "RIFF" harflerini içerir
4	4	Zincir Boyutu (ChunkSize)	ChunkID ve ChunkSize 'ın boyutları hariç tüm dosya boyutunu temsil eder.
8	4	Format	ASCII biçiminde "WAVE" harflerini içerir
"WAVE" format 2 farklı alt zinciri temsil eder. Birinci alt zincir "fmt" dosyanın başlığını, ikinci alt zincir "data" dosyanın veri kısmını temsil eder			
12	4	1. Alt zincir Kimliği (SubChunk1ID)	ASCII biçiminde "fmt" harflerini içerir
16	4	1. Alt zincir Boyutu (SubChunk1Size)	
20	2	Ses Formatı (Audio Format)	Genel olarak 1 değerini alır. 1 Dışındaki değerler sıkıştırma oranını gösterir.
22	2	Kanal Sayısı (NumChannels)	Tek Kanal(Mono)=1, Çift(Stereo)=2
24	4	Örnekleme Oranı (SampleRate)	8000,44100 gibi değerler alır
28	4	Byte Oranı(ByteRate)	Örnekleme Oranı * Kanal Sayısı * Örnek Başına Düşen Bit Sayısı / 8 değerini alır.
32	2	Blok Ortalama (BlockAlign)	Kanal Sayısı * Örnek Başına Düşen Bit Sayısı / 8 değerini alır.
34	2	Örnek Başına Düşen Bit Sayısı (BitsPerSample)	8 bit için 8, 16 bit için 16 değerini alır.
36	4	2. Alt zincir Kimliği (SubChunk2ID)	ASCII biçiminde "data" harflerini içerir
40	4	2. Alt zincir Boyutu (SubChunk2Size)	Ses Verisi boyutu
44	...	Ses Verisi (DATA)	Ses Verisi

4.1.4. Seslerin Taşıyıcı Kelimelerden Ayrılması

İkili ses birimleri (difonlar) yine Audacity adlı yazılım sayesinde kendilerine ait taşıyıcı kelimelerinden çıkarılmış, yine kendi isimlerini taşıyan WAV dosyaları olarak veri tabanının bir parçası olarak kaydedilmiştir. Kesme işleminde her bir ikili ses uzunluğu için 50, 100 ve 150 ms seçenekleri değerlendirilmiş ve yapılan değerlendirme sonucunda da 150 ms’lik ikili seslerin en iyi sonucu verdiği tespit edilmiştir. İkili sesler için karar verilen uzunluk olan 150 ms, yaklaşık olarak 6-7 bin civarında ses örneği içermektedir. İkili sesi oluşturan seslerin yapısından kaynaklanan sebeplerden dolayı her bir ikili ses için istenen uzunluk sağlanamasa da uzunluk yakın değerlerde tutulmaya çalışılmıştır. Ayrıca kesme işleminde birleştirmede azami uyumluluk sağlanması için ikili sesler üzerindeki her bir tek

sesin eşit dağılımda olmasına ve kesme noktalarının periyodikliği bozmayacak şekilde seçilmesine özen gösterilmiştir. Dağılıma dikkat edilmeyen seslerde hatalı birleştirmelerle karşılaşmıştır. Şekil 4.4’de “obje” taşıyıcı kelimesinde “bj” ikili sesinin çıkartılması gösterilmektedir. “bj” ikili sesini “obje” ses dosyası içinden kesmek için otomatik bir yol olmadığı için ilgili ikili ses Audacity programı ile dikkatlice dinlenmiş “b” ve “j” sesinin eşit dağılımda olduğu yaklaşık 150 ms lik “bj” sesi “obje” sesinden kesilerek bj.wav olarak kaydedilmiştir. Tespit edilen tüm ikili sesler için aynı yöntem birer birer uygulanmıştır.

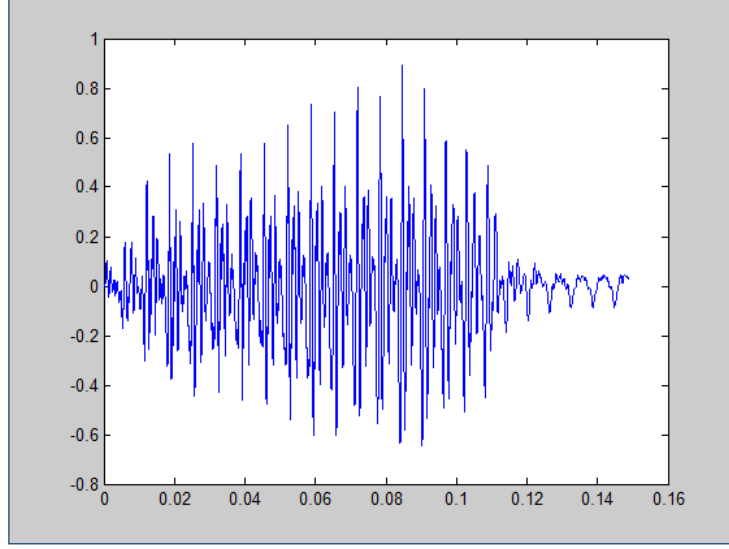


Şekil 4.4. İkili seslerin (difonların) taşıyıcıdan çıkarılması.

4.1.5. Ses Veri Tabanı Normalizasyonu

İkili seslerde birleştirme sonrasında oluşacak, ses yükselme ve alçalmalarını engellemek için, kaydedilen tüm seslere -1dB,+1dB aralığında normalize işlemi tabii uygulanmıştır. Şekil 4.5’de normalize edilmiş “ab” ikili sesi görülmektedir. Bu sayede normalize edilecek ikili ses işaretinin pozitif ve negatif (yukarı ve aşağı) yönlerdeki en büyük tepe değerleri +1dB ve -1dB değerlerine çekilmiş, diğer değerleri de bu oranda ölçeklendirilmiştir. Bu sayede veritabındaki tüm ikili sesler için standart bir ses yüksekliği seviyesi yakalanmıştır. Kayıt için kullanılan Audacity yazılımı istenilen ses dosyalarını istenilen aralıkta normalize eden bir fonksiyonu

bünyesinde barındırmaktadır. Bu fonksiyon sayesinde taşıyıcı kelimelerin kaydı sırasında konuşmacı ve donanımdan kaynaklanan ses alçalma ve yükselmeleri, bunlara bağlı olarak da birleştirme sonrasında meydana gelebilecek bozulmaların önlenmesi amaçlanmıştır.



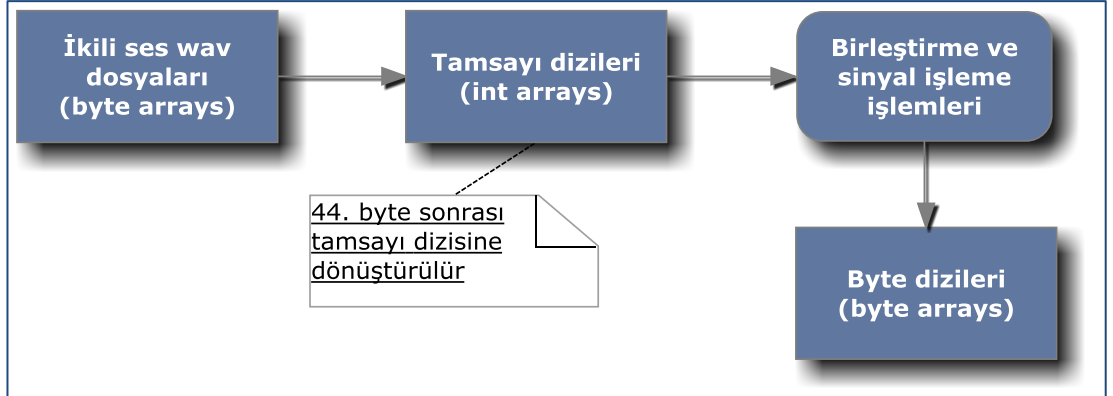
Şekil 4.5. Normalize edilmiş ab ikili sesi ab.wav.

4.2. METİN ÖN İŞLEMLERİ

Seslendirilecek metnin ön bir veya birkaç işlemden geçirilmesi genel olarak tüm dillerde aynı olmasına rağmen seslendirme diline ait özellikler de içermektedir. Ön işlemlerde noktalama işaretleri, rakamlar, tarih, para birimleri ve buna benzer özel işaretler ayıklanmakta, okunabilecek bir yapıya kavuşturulmaktadır. Sistem içerisindeki sayı-metin dönüştürücü sınıf sayesinde metin içinde sayısal değeri olan ifadeler metin karşılıklarına dönüştürülmesi ve tüm metnin veri tabanı ile eşleşir bir hale getirilmesi buna örnek olarak verilebilir. Sayı-metin dönüştürme ile metin içinde karşılaşılan “1568 Lira” ifadesi “binbeşyüzaltmışsekiz lira” olarak metin ön işleminden çıkmaktadır. TURS sisteminde de öncelikle girilen metin ön bir işlemden geçirilmiş ve tüm metin küçük harflere çevrilmiştir. Örnek olarak; seslendirilecek metine içindeki virgül ve benzeri noktalama işaretleri için yarım ses (75 ms), nokta da ise tam seslik (150 ms) sessizlik eklenmiştir.

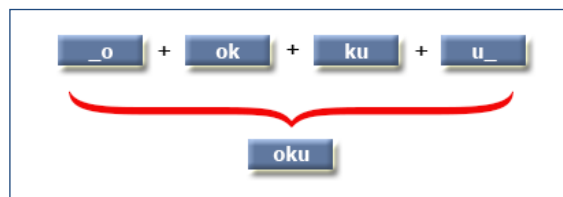
4.3. İKİLİ SES BİRLEŞTİRME ALGORİTMASI

Sistem, ikili ses (difon) sentezleyicisi olarak tasarlanmıştır. İkili ses sentezleyici sistemler daha küçük veri tabanları ile çalışmakta olup, seslendirme dilin hemen hemen bütün ihtiyaçlarına cevap verebilmektedir. Sistem ikili ses sentezleyici olduğundan seslendirme algoritması ikili seslerin belli kurallar çerçevesinde birbiri ardına eklenmesi mantığıyla çalışmaktadır. Birbiri ardına eklenecek her bir ikili ses veri tabanında bir wav dosyası olarak tutulmaktadır. Birleştirmede kullanılacak ikili seslere ait wav dosyaları öncelikle birer bayt dizisi olarak birleştirme işlemine girmekte, bu bayt dizileri birleştirme işleminde öncelikle tamsayı dizisi haline dönüştürülmektedir. Birleştirme de yapılan sinyal işleme dair tüm işlemler bu tamsayı dizileri üzerinden yürütülmekte, birleştirme işlemi sonucunda elde edilen sonuç tamsayı dizisi, tekrar bayt dizisine dönüştürülmektedir. Şekil 4.6 ilgili dönüşümü ifade etmektedir.



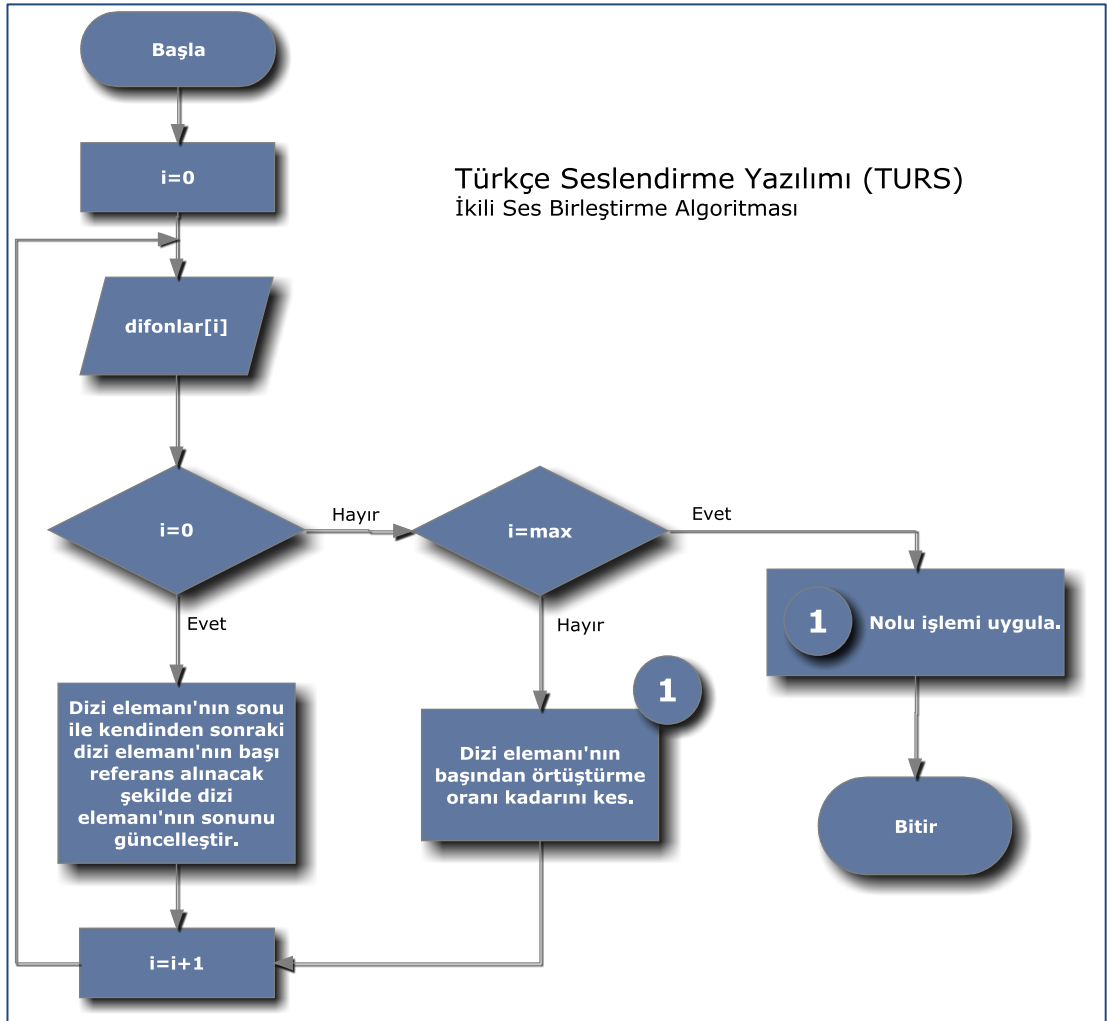
Şekil 4.6. Birleştirme bayt tamsayı dönüşümü.

Şekil 4.7’de örnek bir “oku” kelimesini elde etmek konuşma motorunun kullanılması gereken tüm ikili sesler gösterilmiştir.



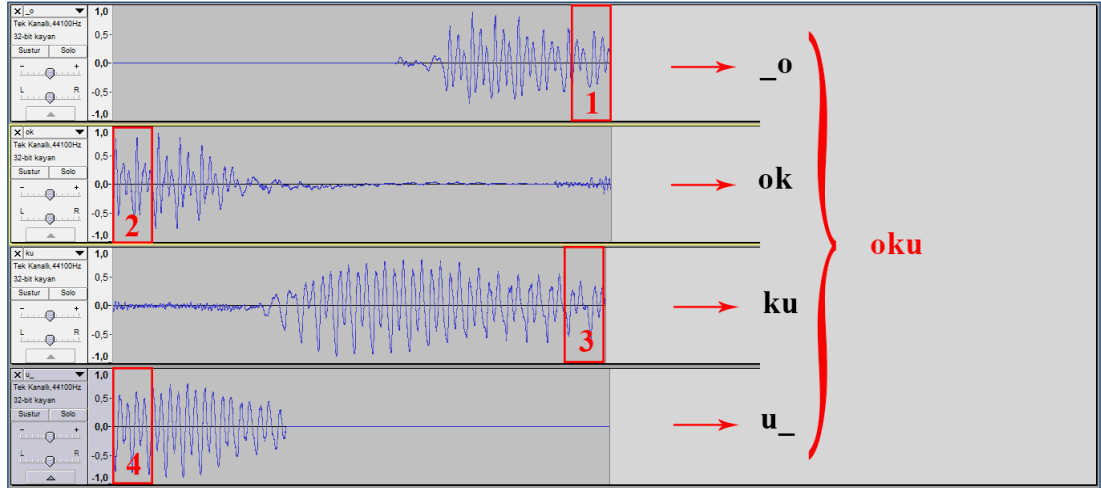
Şekil 4.7. Oku kelimesi için kullanılacak ikili ses dizisi.

İkili seslerin art arda eklenmesi TD- PSOLA yöntemine benzer bir şekilde, eklenecek ilk ikili sesin sonundaki belli bir kısmın, ikinci ikili sesin başındaki aynı orandaki kısım ile ortalaması alınarak güncelleştirilmesi ve bu güncelleştirmeye katkısı olan ikinci ikili sesin başındaki bu kısmın kesilerek atılması şeklinde yapılmaktadır. Sistemde ikili seslerin birleştirmesinde geliştirilen algoritma da Şekil 4.8'de gösterilmiştir.



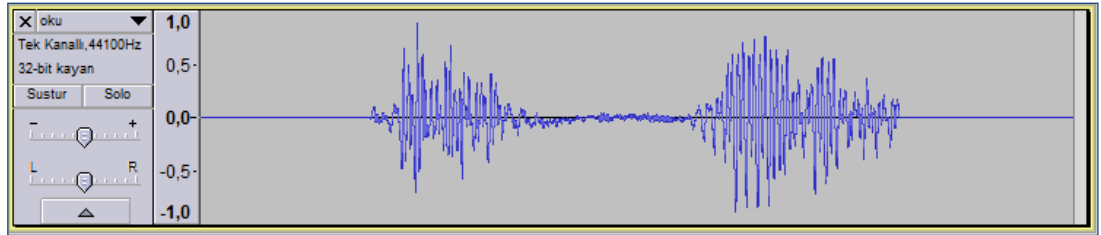
Şekil 4.8. İkili ses birleştirme algoritması.

Şekil 4.9'da oku kelimesini seslendirmek için birleştirilen ikili ses dalga şekli ve sembolik birleşme oranları görülmektedir.



Şekil 4.9. Oku için ikili ses dizisi dalga işaretleri.

Şekil 4.10'da birleştirme sonucu ortaya çıkan oku.wav ses dosyasının dalga şekli görülmektedir.



Şekil 4.10. Birleştirme sonucu elde edilen oku.wav dosyası işareti.

Yapılan dinleme testleri sonucunda 150 ms'lik (6000-7000 örnek) ikili seslerden oluşan bir veri tabanında %40 - %50 arasında bir birleştirme oranı anlaşılabilir sonuçlar vermektedir.

Geliştirilen uygulamada ayrıca örtüştürme tekniği olarak MBR-PSOLA tekniğine benzer bir örtüştürme tekniği de denenmiş ve bu farklı iki örtüştürmenin avantaj ve dezavantajları gözlemlenmiştir. MBR-PSOLA tekniğinde veri tabanı hazırlanması esnasındaki iş yükü fazlalığına karşılık daha esnek ve daha etkin bir örtüştürme sağladığı gözlemlenmiştir. MBR_PSOLA yöntemine benzer yapıda kullanılan veri tabanında ilgili ikili ses parçası baştan ve sondan 8'er parçaya ayrılmış olup, her parçada ilgili sese ait 8 perdelik periyodik işaretin uzunlukları tutulmuştur.

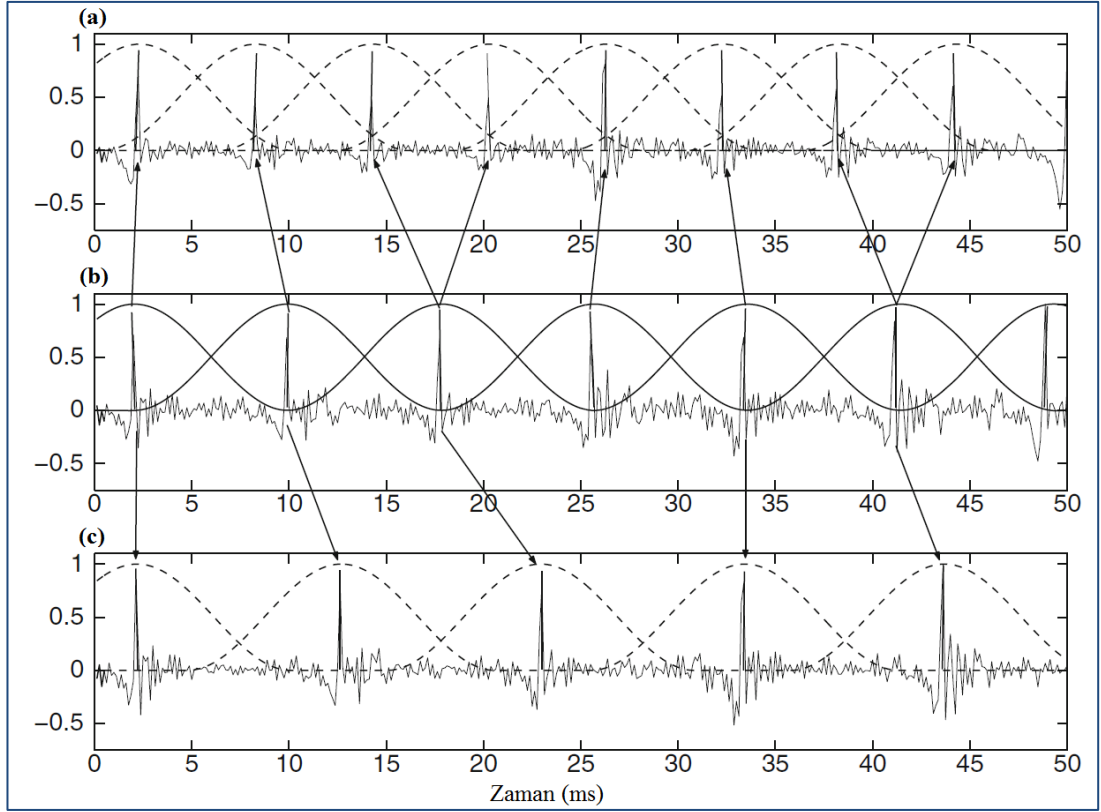
Örtüştürme işlemleri bu perdeler dikkate alınarak, perde-perde yapılmış, küçük perdeler büyük perdelerin içinde homojen dağılımı sağlanmıştır.

Çizelge 4.3’de “bardaktan” sözcüğüne ait ikili ses parçaları ve bunların perdelerine ait örnek bilgilerinin veri tabanındaki kayıtlar gösterilmektedir. Örneğin, çizelgenin ikinci satırında “ar” ikili sesine ait perde bilgileri yer almaktadır. Burada “ar” ikili sesini oluşturan “a” sesi 294 örneklilik bir perde periyodu ile başlamakta ve 277 örnekten oluşan perde periyodu ile sonlanmaktadır. Yine aynı şekilde “r” sesi 280 örneklilik bir perde periyodu ile başlamakta, 296 örnekten oluşan perde ile de sonlanmaktadır.

Çizelge 4.3. MBR-PSOLA veri tabanı yapısı (bardaktan).

Difon	bas1	bas2	bas3	bas4	bas5	bas6	bas7	bas8	son1	son2	son3	son4	son5	son6	son7	son8	Toplam	Ortalama
ba	0	0	0	0	0	0	0	3849	367	342	346	370	366	381	378	374	6773	365,5
ar	294	292	287	284	280	277	276	277	280	284	287	289	295	291	302	296	4591	286,9375
rd	323	326	359	357	331	329	346	359	351	341	335	336	332	332	332	347	5436	339,75
da	0	0	0	0	0	0	0	3450	286	280	276	279	280	280	288	285	5704	281,75
ak	264	267	265	272	279	286	288	317	3589	0	0	0	0	0	0	0	5827	279,75
kt	0	0	0	0	0	0	0	3662	2961	0	0	0	0	0	0	0	6623	
ta	0	0	0	0	0	0	0	1726	244	262	272	271	278	285	291	291	3920	274,25
an	262	259	260	255	253	253	251	247	247	281	241	240	237	235	236	235	3992	249,5
n_	302	311	309	312	311	310	303	217	2781	0	0	0	0	0	0	0	5156	296,875
																		296,7891

Ayrıca MBR-PSOLA veya benzeri tekniklerde ilgili sese ait perde frekanslarının uzunluklarının yerleri önceden bilindiğinden bu frekansları artırma (zamanda yavaşlama) ve azaltma (zamanda yavaşlama) gibi sinyal işlemlerindeki hesaplama zorluklarını ortadan kaldırmaktadır. Şekil 4.11 (a), (b) ve (c) de sırasıyla artan frekans değerleri detaylı bir şekilde sunulmuştur. Şekil sondan başa doğru incelendiğinde ise azalan frekanslar gözlemlenmektedir.



Şekil 4.11. Ses sinyali perde frekansı artırılması. a)Orijinal ses işareti, b)Perde frekansı iki perde artırılmış işaret ve c) Perde frekansı üç perde artırılmış işaret [32].

4.4. SİSTEMİN OPTİMİZASYONU

4.4.1. Türkçe Heceleme Algoritması

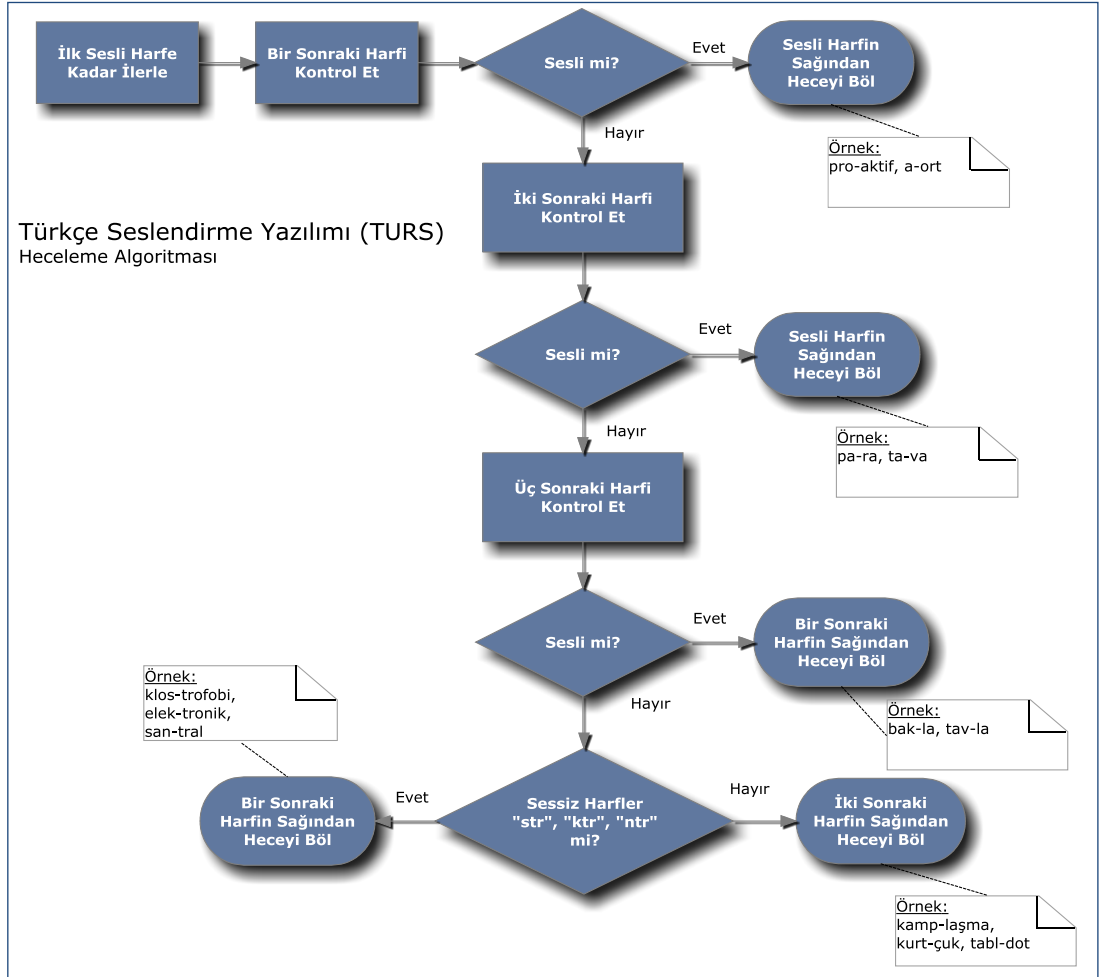
Uygulama için Türkçe kelimeleri hecelerine ayıran bir heceleme algoritması kullanılmıştır. Heceleme algoritması ile seslendirilmesi istenen metindeki kelimelerin hece kesim yerleri hesaplanmıştır. Hece kesim yerleri belirlenen kelimelerde ikili ses birleştirilmesi yapılırken hece kesim noktaları dikkate alınmış olup, bu kısımlardaki birleştirme oranı belli bir yüzde oranında azaltılmıştır. Bu sayede kelimenin doğal heceleme vurgusuna yakın bir vurgu elde edilmeye çalışılmıştır. Yapılan dinleme testleri sonucunda hece kesim noktalarında ikili ses birleştirme oranlarında yapılan %40 oranındaki azaltma çalışmaları yeterli düzeyde bir heceleme vurgusu sağlamıştır. Şekil 4.12’de TURS uygulamasında kullanılan

heceleme algoritması açıklanmaktadır. Algoritma TDK'nın Türkçe kelimelerin hecelemesine ilişkin bazı kurallar çerçevesinde oluşturulmuştur.

Bu kurallar;

- Türkçede tek harften oluşan hece olabilir ancak bu sessiz harf olmalıdır.
- Sessiz harfler bir araya gelerek hece oluşturamazlar.
- En fazla iki sessiz harf birlikte yazılabilir şekilde özetlenebilir.

Türkçe olmayan ve Türkçe'ye sonradan giren "elektronik", "kontrol" ve klostrfobi gibi kelimeler için ise fazladan bir kontrol yapılarak doğru bir heceleme yapılması sağlanmıştır.



Şekil 4.12. Türkçe heceleme algoritması akış diyagramı [33].

4.4.2. Ses Çıkışı Düzeltme Çalışmaları

Çıkış sesinde, gerek sesin genliğini arttırmaya yönelik gerekse birleştirme sonrasında ortaya çıkan birleştirme kusurlarını ortadan kaldırma amaçlı çeşitli filtreleme çalışmaları yapılmıştır. Bu çalışmalar kapsamında ilgili sesin Ayrık Zamanlı Fourier Transformu (DFT) ile istenmeyen bileşenlerin en aza indirilmesi amaçlanmıştır.

4.4.2.1. Ayrık Zamanlı Fourier Transformu (DFT)

Fourier Transformu bir sinyali zaman domeninden frekans domenine aktarmak için kullanılan en temel yöntemdir. Fourier Transformu sürekli işaretler için frekans dönüşümü sağlarken DFT ise mutlak toplanabilir diziler için frekans dönüşümünü sağlamaktadır. Matematiksel olarak $x(i)$ ayrık zaman işareti olmak üzere eşitlik (4.1a) ve (4.1b) de verildiği gibi DFT reel ve imajiner bileşenlerden oluşur.

$$ReX[k] = \sum_{i=0}^{N-1} x(i) \cos(2\pi ki/N) \quad (4.1a)$$

$$ImX[k] = - \sum_{i=0}^{N-1} x(i) \sin(2\pi ki/N) \quad (4.1b)$$

Bu bileşenlerin vektörel toplamı da frekans yoğunluğunu göstermektedir.

4.4.2.2. DFT'nin Koda Uyarlanması

Şekil 4.13'de DFT'nin Şekil 4.14'de de Ayrık Zamanlı Fourier Transformu'nun tersinin (IDFT) C# programlama dilindeki karşılıkları görülmektedir. Şekil 4.13'de "output" adlı dizi frekans yoğunluklarını, Şekil 4.14'deki "en_son" adlı dizi, istenmeyen frekansları yok edilmiş çıkış işaretini ve her iki kod bloğundaki "k" da ses sinyalinin örnekleme indeksini temsil etmektedir. İlgili program parçalarıyla önce çıkış sinyalinin DFT'si alınmıştır. Son olarak tespit edilen istenmeyen frekanslar yok edilerek, tekrar tersine dönüşüm (IDFT) ile çıkış sinyali elde edilmiştir. DFT gibi testi yapılan tüm filtreler masaüstü uygulamasındaki "Filtreli" parametresiyle ilişkilendirilmiş ve bu sayede testi yapılan tüm filtreleme işlemlerinin

yalın çıkış ile karşılaştırma fırsatı yakalanmıştır. Çıkış sinyalinde birtakım iyileşmeler gözlemlense de dönüşüm işlemlerinin uzun zaman almasından ötürü konuşma motorunda gözle görülür bir yavaşlama tespit edilmiştir. Bu bağlamda sonradan çıkış sinyali üzerinde yapılacak düzeltmelerin getireceği hesaplama yükü düşünülerek ses dosyaları üzerindeki gürültü gibi istenmeyen durumların taşıyıcı kelimelerin kaydı yapılırken engellenmesi tavsiye edilmektedir.

```
//dft ayrık zamanlı fourier transformu
for (int k = 0; k < N; k++)
{
    twoPikOnN = twoPiOnN * k;

    for (int j = 0; j < N; j++)
    {
        twoPijkOnN = twoPikOnN * j;
        reel_data[k] += ses_array[j] * Math.Cos(twoPijkOnN);
        im_data[k] -= ses_array[j] * Math.Sin(twoPijkOnN);
    }
    reel_data[k] /= N;
    im_data[k] /= N;

    output[k] = (int)Math.Sqrt(Math.Pow(reel_data[k], 2) + Math.Pow(im_data[k], 2)); // Magnitude

    if (output[k] > 800)
    {
        reel_data[k] = 0;
        im_data[k] = 0;
    }
}
}
```

Şekil 4.13. Ayrık Zamanlı Fourier Transformu (DFT).

```
//idft ayrık zamanlı fourier transformunun tersi
for (int k = 0; k < N; k++)
{
    twoPikOnN = twoPiOnN * k;

    for (int j = 0; j < N; j++)
    {
        twoPijkOnN = twoPikOnN * j;
        en_son[k] += (int)(reel_data[j] * Math.Cos(twoPijkOnN) - im_data[j] * Math.Sin(twoPijkOnN));
    }
}
}
```

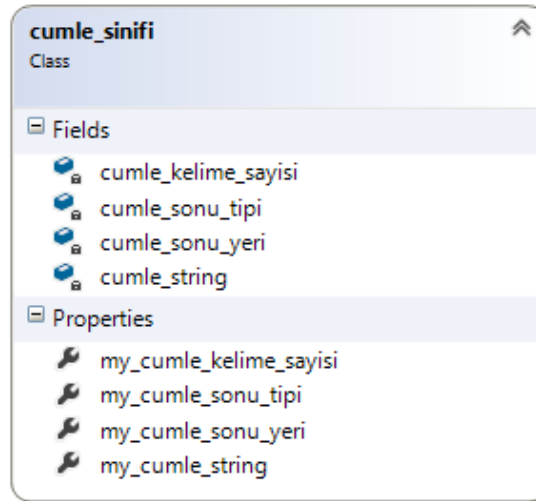
Şekil 4.14. Ayrık Zamanlı Fourier Transformunun tersi (IDFT).

4.5. EZGİSEL ÇALIŞMALAR

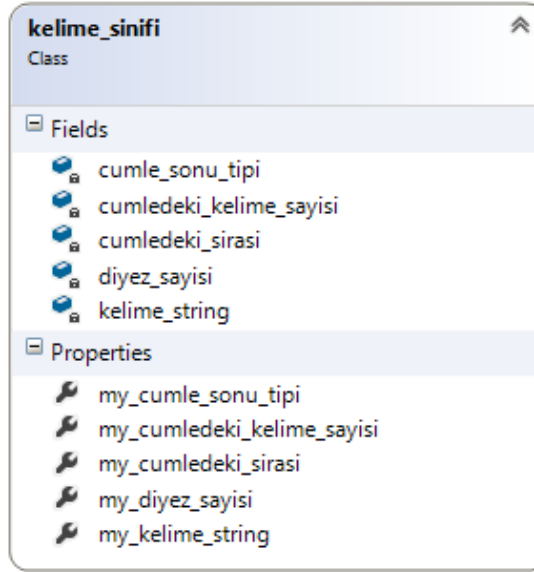
Bir konuşma motorunun doğallığını artıracak en önemli faktör seslendirmeye eklenen ezgidir. Ezgi simgelerden oluşan metinsel ifadeyi doğallığa bir adım daha

yaklaştırır. Bir cümlenin sonundaki ünlem, soru işareti ve nokta gibi sembollerin o cümleye kattığı anlam ile beraber cümlenin karşı tarafa aktarmak istediği sevgi, öfke, acıma vb. tüm duygular ezgisel yaklaşımlarla ifade edilebilir.

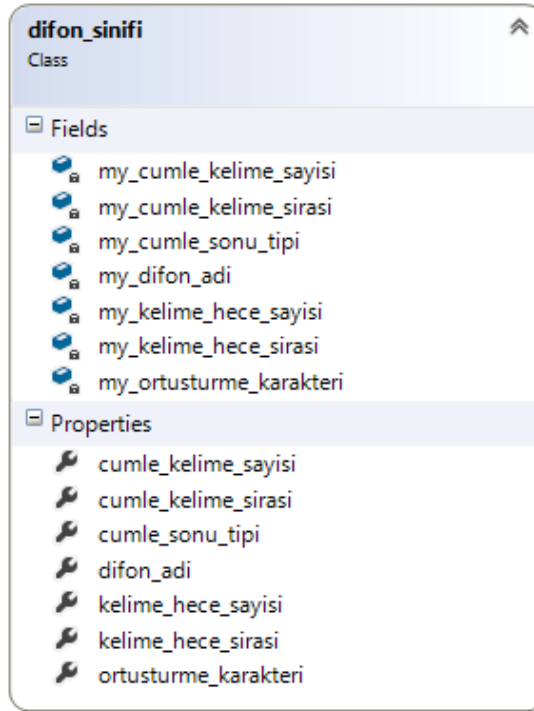
TURS konuşma motoru her ne kadar bir ikili ses sentezleyici olsa da, ilk olarak seslendirmesi istenen metni noktalama işaretlerine göre cümlelere ayırmaktadır. Sistemdeki cümle, kelime, ikili ses sınıfları vasıtasıyla da bazı kalıtsal özellikleri cümleden başlayarak bir alt sınıfa taşımaktadır. Bu sayede sistemin seslendirmede kullandığı en küçük ses birimi ikili ses elde edildiğinde bu ikili sesin hangi kelimeye ait olduğu, bu kelimenin hangi hecesinde yer aldığı, ilgili kelimenin kaç heceli olduğu, ait olduğu cümle yapısı vb. özellikler tespit edilebilmektedir. Şekil 4.15, Şekil 4.16 ve Şekil 4.17’de ilgili sınıfların diyagramları sunulmaktadır.



Şekil 4.15. Cümle sınıfı kod bloğu örneği.



Şekil 4.16. Kelime sınıfı kod bloğu örneği.



Şekil 4.17. İkili ses sınıfı kod bloğu örneği.

Belirlenen tüm bu özellikler sayesinde konuşma motoruna bazı ezgisel özellikler kazandırılmıştır. Örnek olarak “gidiyor musun.” ve “gidiyor musun?” kelimelerinin okunuşları sonlarındaki noktalama işaretlerine göre bazı farklılıklar kazanmışlardır. Uygulamanın vurgu modülü sayesinde sonu noktalı olan “gidiyor musun?” kelimesi düz bir tonlamada, seri bir şekilde seslendirilirken, sonunda soru işareti olan “gidiyor

musun?” kelimesinin diğ erinden %60 daha vurgulu ve “musun” parçasının sesinin %50 oranında alçaltılarak seslendirilebilmektedir.

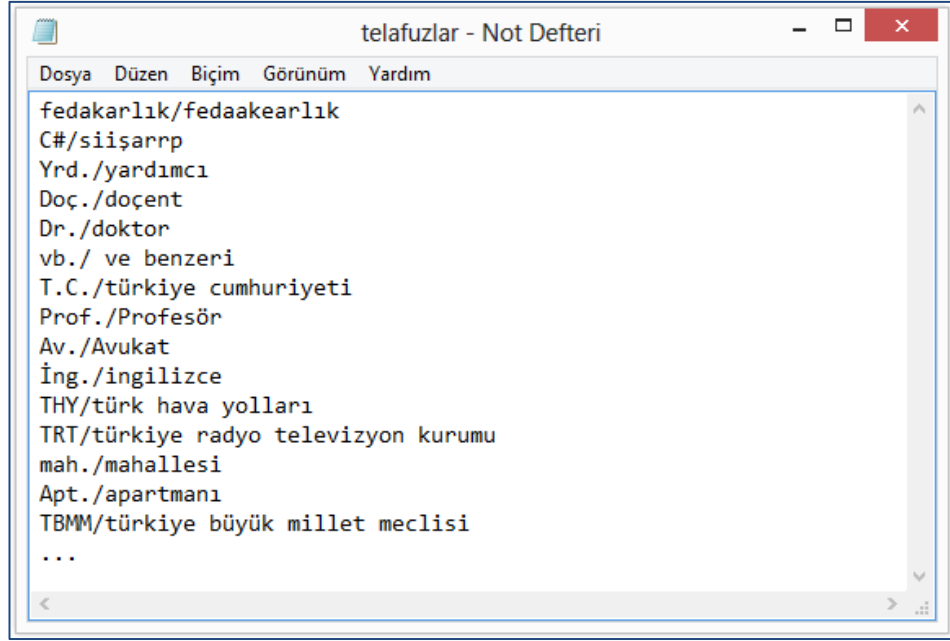
Gidiyor → *musun* → .

Gidiyor ↑ *musun* ↓ ?

Detaylı dil bilgisel analizler sonucunda bu ezgi sınıfının etkinliğinin daha da artırılması amaçlanmaktadır.

4.6. TELAFUZ ÜNİTESİ

Sisteme bir telaffuz ünitesi eklenerek farklı okunuş biçimleri gerektiren kelimelerin doğru bir şekilde seslendirilmesi sağlanmıştır. Telaffuz ünitesi kendisine özgü bir text veri tabanı ile çalışmaktadır. Taşınabilirlik ve kullanım sadeliği açısından text bir veri tabanı seçilmiştir. Sistem tarafından seslendirilecek kelimeler öncelikle telaffuz veri tabanındaki kelimelerle karşılaştırılmakta, ilgili veri tabanında karşılaşılan kelimelerin seslendirilmesi belirtilen okunuş biçimine göre yapılmaktadır. Telaffuz veri tabanının uygulamanın hızına olumsuz etkisinin ortadan kaldırılması doğrultusunda telaffuz veri tabanındaki kelimeler uygulamanın çalıştırılması esnasında bir diziye alınmakta ve bu sayede ilgili kelimelere çalışma sırasında daha hızlı erişim sağlanabilmektedir. Telaffuz veri tabanına telaffuzu farklı şekilde yapılması istenen kelimeler belirlenen bir formatta yapılmakta olup, ilk önce telaffuzu yapılacak kelime arkasından ilgili telaffuzun ne şekilde yapılacağı text dosyaya kaydedilmektedir. Telaffuz ünitesi farklı telaffuzu olması gereken kelimelerin yanı sıra, dilde kullanılan kısaltmaların seslendirilmesinde de kullanılmaktadır. Şekil 4.18’de konuşma motoru tarafından kullanılan telaffuz veri tabanı ’nın bir kısmı Microsoft Windows not defteri uygulamasıyla gösterilmektedir.



Şekil 4.18. TURS telaffuz veri tabanı dosyası.

4.7. TAŞINABİLİR SÜRÜM DENEMESİ

Daha önceden de belirtildiği gibi TURS konuşma motoru tek bir sınıf dosyası halinde diğer üçüncü parti yazılımlardan bağımsız olarak geliştirildiğinden, geliştirildiği dil haricindeki dil veya platformlara kolay uyarlanabilir yapıdadır. Bu doğrultuda ilgili uygulama Microsoft Windows Phone 7.1 Emülatör üzerinde de test edilmiş, uyarlamadan başarılı bir sonuç alınmıştır. Geliştirme ortamı olarak Microsoft Visual Studio 2012 ve .Net Framework 4.5 platformu kullanılmıştır. Sistem her ne kadar .Net Framework tabanlı bir mobil uygulamada test edilse de yapısal olarak Java ve benzeri diğer dillere taşınabilirliğinin kolay olacağı düşünülmektedir. Şekil 4.19’da mobil uygulama emülatör ekranı gösterilmektedir. TURS seslendirme sınıfı mobil uygulama projesine eklenmiş, çıkışa gönderilen bayt dizisinin seslendirme formatında yapılan değişiklikle sınıfın mobil uygulama üzerinde çalışabilirliği sağlanmıştır. Karşılaşılan Türkçe karakter seslendirme sorunu emülatörün yerel (local) ayarlarında yapılan değişiklikle giderilmiştir.



Şekil 4.19. TURS mobil uygulama test ekranı.

BÖLÜM 5

SİSTEMİN TEST EDİLMESİ

TURS konuşma motorunun sentetik konuşmasının değerlendirilmesi amaçlı dinleme testleri yapılmış bu testlerin sonuç ve grafikleri değerlendirilmiştir.

5.1. MOS (MEAN OPINION SCORE)

MOS, ses iletişimde özellikle internet telefonlarında iletilen ses kalitesi hakkında bir karara varmak için kullanılan bir test yöntemidir. MOS iletilen sesin kalitesi hakkında sayısal bir ölçü sunmaktadır. Testin uygulanmasında dinleyicilere birtakım sesler dinletilir ve kendilerinden bu seslerin Çizelge 4.5. de verilen puanlamalar çerçevesinde puanlamaları istenir. Tüm puanların aritmetik ortalaması sistemin genel puanını verir.

5.2. SİSTEM MOS PUANININ BELİRLENMESİ

Sistem MOS (Mean Opinion Score) değerlendirme sistemiyle değerlendirilen bir teste tabi tutulmuştur. Görme engelli ve görme engelli olmayan toplam 25 kişilik bir grup üzerinde anlaşılabilirlik noktasında yapılan değerlendirme sonucunda TURS konuşma motoru 5 tam puan üzerinden 3,72 puan almıştır. Derecelendirmede sistemin konuşma kalitesi ve anlaşılabilirliği test edilmiştir. Dinleyici guruba öncelikle Çizelge 5.1. deki Türkçe cümleler sistem tarafından seslendirilmiş ve Çizelge 5.2. deki değerlendirme sistemine göre puan vermeleri istenmiştir.

Çizelge 5.1. TURS tarafından seslendirilen cümleler.

MOS Değerlendirme Soruları	
1.	Bugün hava ne kadar güzel değil mi?
2.	Şu an Türkçe seslendirme sistemini dinliyorsunuz.
3.	Ahmet 2000TL maaş alıyor.
4.	Belki bir gün sende beni anlarsın.
5.	Bilgisayarı ne düzeyde kullanıyorsunuz?
6.	Mahmut gelir gelmez, yola çıkarız.
7.	Bartın Karabük arası yaklaşık 100 kilometredir.
8.	Sakla samanı, gelir zamanı.
9.	Telefonunuzu kullanabiliyormuyum?
10.	Evlilik, fedakarlık ister.

Çizelge 5.2. MOS derecelendirme çizelgesi.

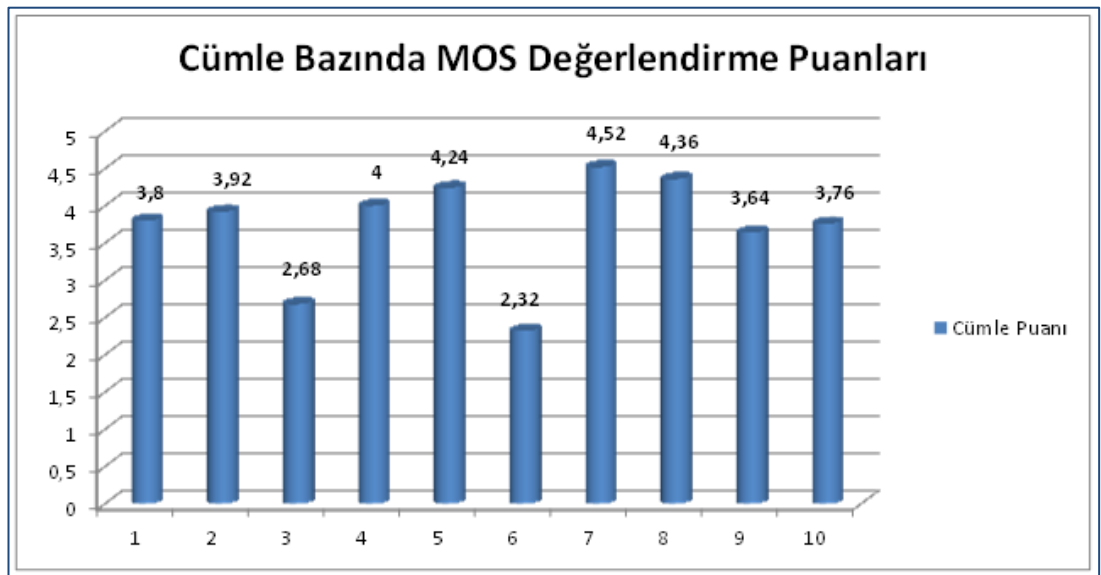
Puan	Değerlendirme
5	Mükemmel
4	İyi
3	Orta
2	Kötü
1	Berbat

Çizelge 5.3’de katılımcıların sorulara verdiği puanlar ve ortalamaları hem cümle hem de katılımcı bazında ayrı ayrı görülmektedir. Çizelgenin en sonunda ise genel ortalama belirtilmiştir.

Çizelge 5.3. Katılımcı puanları çizelgesi.

Katılımcı No	1. cümle	2. cümle	3. cümle	4. cümle	5. cümle	6. cümle	7. cümle	8. cümle	9. cümle	10. cümle	Gen. Ortalamalar
1	4	4	3	5	4	3	4	4	4	5	4,00
2	5	5	3	5	5	3	5	5	4	4	4,40
3	5	5	4	5	4	4	5	5	4	5	4,60
4	4	4	3	4	5	2	5	5	5	4	4,10
5	3	5	1	1	3	1	5	5	2	1	2,70
6	4	4	3	4	5	3	5	5	3	4	4,00
7	5	5	4	5	5	3	5	5	5	5	4,70
8	5	4	4	4	5	2	5	5	4	5	4,30
9	5	4	4	5	5	3	5	5	4	4	4,40
10	2	2	2	3	3	1	3	3	2	3	2,40
11	3	3	3	4	4	3	4	4	4	4	3,60
12	1	3	2	4	3	2	5	4	5	4	3,30
13	5	5	3	4	5	2	5	5	5	4	4,30
14	4	3	3	4	4	2	4	4	3	2	3,30
15	2	3	2	3	4	1	4	4	2	4	2,90
16	5	5	3	4	4	2	5	5	5	5	4,30
17	5	4	3	5	5	3	5	5	5	5	4,50
18	4	4	2	4	5	2	5	4	3	4	3,70
19	3	2	1	3	4	1	4	4	3	3	2,80
20	4	5	1	4	5	2	5	5	4	5	4,00
21	3	3	2	4	3	2	4	4	3	2	3,00
22	3	4	3	5	4	2	4	3	2	3	3,30
23	5	5	3	5	5	4	5	5	4	4	4,50
24	3	4	2	3	4	2	4	3	3	2	3,00
25	3	3	3	3	3	3	3	3	3	3	3,00
Cümle Ort.	3,8	3,92	2,68	4	4,24	2,32	4,52	4,36	3,64	3,76	3,72

Şekil 5.1. de ise bu puanların grafik olarak gösterimi verilmiştir. Puanların cümleden cümleye farklılık göstermesi, ilgili cümleyi oluşturan en küçük ses birimi ikili seslerin kalitesiyle alakalıdır. Bu kalite ikili sesin taşıyıcı kelimesinin kaydı, taşıyıcıdan çıkarılması, konuşmacı, donanım ve hatta katılımcının duyu algılaması gibi faktörlerle doğrudan ilişkilidir.



Şekil 5.1. Cümle bazında MOS puan grafiği.

BÖLÜM 6

SONUÇLAR

Bu çalışmada Türkçe metinleri ses sinyallerine çevirebilen, farklı platformlarda çalışabilirliği yüksek, geliştirmeye açık, eklemeli tipte TURS adlı bir konuşma motoru kütüphanesi geliştirilmiştir.

TURS uygulaması tek bir sınıf olarak geliştirilmiş ve bu sınıfta herhangi bir kütüphane (.Net vb.) kullanılmamıştır. Bu sayede TURS farklı platformlarda çalıştırılabilir yapıdadır. Bu çalışmada TURS biri masaüstü, diğeri mobil olmak üzere iki farklı uygulama üzerinde başarıyla çalıştırılmıştır.

Konuşma motorlarının geliştirme sürecinde karşılaşılan zorluklar veri tabanı hazırlamada karşılaşılan zorluklar, metin ön işleme zorlukları, özel telaffuz gerektiren kelimelerdeki seslendirme zorlukları ve ezgisel yaklaşımlardaki zorluklar olarak sıralanabilir. Bu çalışmada bahsi geçen zorluklar ayrıntılı olarak incelenmiş ve bunlara yönelik çözümler üretilmiştir.

Öncelikle eklemeli tipteki bir konuşma motorunun geliştirilmesindeki basamaklar detaylı bir şekilde ele alınmıştır. Bu bağlamda Türkçe için tüm sesleri ifade edebilecek eklemeli tipteki konuşma motorunun kullanacağı bir veri tabanı oluşturulmuştur. Konuşmanın sentezlenmesinde TD-PSOLA ve MBR-PSOLA gibi farklı birleştirme teknikleri denenmiştir. Bu tekniklerin olumlu ve olumsuz yönleri tespit edilmiştir. Sonuç olarak TD-PSOLA benzeri bir yöntemle birleştirme gerçekleştirilmiştir.

TURS farklı telaffuza sahip Türkçe kelime ve kısaltmaların okunabilirliğini artırmak için bir telaffuz modülüne sahiptir. Bu modül ile fedakarlık gibi kelimeler ile Dr., Yrd., TBMM gibi kısaltmalar doğru bir şekilde seslendirilebilmektedir.

Sistemin ürettiđi ses kalitesini artırmaya yönelik gürültü filtreleme amaçlı Fourier Transformu denenmiş, sinyal işleme çerçevesinde çıkış sesine yönelik yumuşatma ve filtreleme fonksiyonları geliştirilmiştir.

Sesin doğallığını artırmaya yönelik ise seslerin genliği ve birleştirme oranlarındaki deđişikliği temel alan vurgu ve ezgisel çalışmalar yapılmıştır.

Sonuç olarak Türkçe seslendirme yapabilen akıcı, anlaşılır, geliştirmeye açık ikili ses tabanlı eklemeli tipte bir konuşma motoru sistemi geliştirilmiştir. Çalışmada 150 ms boyutunda ikili sesler kullanılmış, yapılan denemeler sonucunda %46 'lık birleştirme oranının en iyi sonucu verdiği gözlemlenmiştir.

Sistemin ses kalitesi MOS benzeri bir yöntemle deđerlendirilmiştir. Bu yöntemde kullanılan parametreler Bölüm 5'te verilmiştir. Deđerlendirme sonucunda TURS, 5 tam puan üzerinden 3,72 puan almıştır.

Eklemeli tip sentezleme kullanılarak geliştirilen konuşma motorlarında daha iyi sonuçlar alabilmek için kullanılacak veri tabanının hazırlanmasında seçilecek mikrofon, kayıt ortamı, kayıt cihazı vb. tüm teknik donanımın hassasiyetle seçilmesi gerekmektedir. Kayıtların özenle yapılması, kayıt sırasında oluşan gürültü gibi olumsuz etkilerin tespit edilip giderilmesi daha sonradan oluşacak fazladan iş yükünü de önleyecektir.

Eklemeli sistemlerde ikili seslerin yanı sıra üçlü ve dörtlü seslerin beraberce kullanılmasının sentezleme sırasında yapılacak birleştirme sayısını azaltacağı, dolayısıyla daha az birleştirme ile daha fazla doğallığın sağlanacağı sonucuna varılmıştır. Bu bağlamda birleştirmede ikili, üçlü ve dörtlü seslerin beraberce kullanılarak, hece tabanlı bir birleştirme yapılmasının daha iyi sonuçlar vereceđi, bununla birlikte veri tabanı boyutunun da kabul edilebilir bir oranda artacaktır.

Sistemin iyileştirilmesi için daha ayrıntılı dil bilgisel analizler yapılarak ezgisel yaklaşımların üzerinde durulmalıdır. Bunun sonucunda ezgisel çalışmaların seslendirmedeki doğallığa olan etkisi artırılabilir.

KAYNAKLAR

1. Flanagan, J. L., Allen, J. B. and Hasegawa-Johnson, M. A., "Speech Analysis, Synthesis and Perception, 3rd Edition", *Springer-Verlag*, Berlin, 269-275 (1972).
2. Schroeder, M. R., "A brief history of synthetic speech", *Speech Communication*, North Holland, 13: 231-237 (1993).
3. İnternet: Helsinki University of Technology Laboratory, "History and Development of Speech Synthesis", http://www.acoustics.hut.fi/publications/files/theses/lemmetty_mst/chap2.html (2013).
4. Klatt, D., "Review of text-to-speech conversion for English", *Journal of the Acoustical Society of America*, 82 (3): 737-793 (1987).
5. Canal, Ş. M., Kurnaz, S. ve Yılmaz, E., "Türkçe metinden konuşma sentezlemede yaşanan sıkıntılar ve çözüm yöntemleri", *Havacılık ve Uzay Teknolojileri Dergisi*, 4 (10): 47-55 (2010).
6. Cutler, A., and Ladd, D. R., "Prosody: Models and Measurements", *Springer-Verlag*, New York, Tokyo, 1-3 (1983).
7. Allen, J., Hunnicutt, S. and Klatt, D., "From Text to Speech: The MITalk System", *Cambridge University Press*, Cambridge, 269-270 (1987).
8. Ljungqvist, M., Lindström, A. and Gustafson, K., "A new system for text-to-speech and its applications to swedish", *ICSLP*, 94 (4): 1779-1782 (1994).
9. İnternet: Acapela Group Web Page, "Acapela Group Profile", <http://www.acapela-group.com/company/profile/> (2013).
10. İnternet: Acapela Group Web Page, "Acapela Group", <http://www.acapela-group.com/company-information-acapela-group.html> (2013).
11. Möbius, B., Schroeter, J., Santen, J.V., Sproat, R. and Olive, J., "Recent Advances in Multilingual Text-to-Speech Synthesis", *In Fortschritte der Akustik*, Germany (1996).
12. Bigorgne, D., Boeffard, O., Cherbonne, B., Emerard, F., Larreur, D., Le Saint-Milon, J.L., Mètayer, I., Sorin, C. and White, S., "Multilingual PSOLA text-to-speech system", *IEEE*, 2: 187-190 (1993).

13. İnternet: Edinburgh University The Centre of Technology Research Laboratory , “The Festival Speech Synthesis System”, <http://www.cstr.ed.ac.uk/projects/festival/> (2013).
14. İnternet: The EMBROLA Project Home Page, “MBROLA”, <http://tcts.fpms.ac.be/synthesis/> (2013).
15. Dutoit, T. and Leich, H., “MBR-PSOLA: Text-To-Speech synthesis re-synthesis of the segments database”, *Speech Communication* **13**, North Holland, 435-440 (1993).
16. Sak, H., “A corpus-based concatenative speech synthesis system for Turkish”, Yüksek Lisans Tezi, *Boğaziçi Üniversitesi Bilişim Enstitüsü*, İstanbul, 10-22 (2004).
17. Ünalı, İ., “Taşınabilir cihazlar için Türkçe metinden konuşma sentezleme sistemi”, Yüksek Lisans Tezi, *Hacettepe Üniversitesi Fen Bilimleri Enstitüsü*, Ankara, 21-31 (2007).
18. Aşlıyan, R., ve Günel K., “Türkçe metinler için hece tabanlı konuşma sentezleme sistemi”, *Akademik Bilişim*, Çanakkale, 31-38 (2008).
19. Uslu B., ve İlk H.G., “Türkçe metinden konuşma sentezlemede Fujisaki ezgi modeli”, *IEEE SIU 2009*, Side, Antalya, 1-2 (2000).
20. Erdemir C., “Türkçe metin seslendirme için doğal konuşma sentezleme”, Yüksek Lisans Tezi, *İstanbul Üniversitesi Fen Bilimleri Enstitüsü*, İstanbul, 1-5 (2010)
21. Arık G., “Görme engelliler için bilgisayar kullanımının etkinleştirilmesi, erişilebilirlik ve bir Türkçe hece tabanlı konuşma sentezleme sisteminin geliştirilmesi”, Yüksek Lisans Tezi, *Gazi Üniversitesi Bilişim Enstitüsü*, Ankara, 1-10 (2011).
22. Tekindal, B. ve Arık G., “Görme engelliler için Türkçe metinden konuşma sentezleme yazılımı geliştirilmesi”, *Bilişim Teknolojileri Dergisi*, Ankara, 3: 9-18 (2012).
23. Cappell, D.T. and Hansen J.H.L., “A comparison of spectral smoothing methods for segment concatenation based speech synthesis”, *Speech Communication* **36**, Germany, 343-373 (2002).
24. Vinet, R., “Enhancing rule-based synthesizer using concatenative synthesis”, Master of Science, *Department of Speech, Music and Hearing Royal Institute of Technology*, Stockholm, (2003).
25. Schröder, M, and Trouvain, J., “The German text-to-speech synthesis system MARY: A tool for research, development and teaching”, *International Journal of Speech Technology*, Netherlands, 6: 365-377 (2003).

26. Thomas, S, “Natural sounding text-to-speech synthesis based on SyllableE-Like units”, Master of Science, *Department of Computer Science and Engineering Indian Institute of Technology Madras*, India, 13-17 (2007).
27. Tatham, M. and Morton, K., “Developments in Speech Synthesis”, *John Wiley & Sons Ltd*, England, 23-25 (2005).
28. Eker, B., “Turkish text to speech system”, Yüksek Lisans Tezi, *Bilkent Üniversitesi Mühendislik ve Fen Bilimleri Enstitüsü*, Ankara, 19-20 (2002).
29. İnternet: Sound on Sound Home Page , “Synth Secrets, Part 23: Formant Synthesis”, <http://www.soundonsound.com/sos/mar01/articles/synthsec.asp> (2013).
30. Sanjaume, J.B., “Audio time-scale modificationin the context of professional audio post-production”, Ph. D. Program, *Universitat Pompeu Fabra*, Barcelona-Spain, 13-17 (2002).
31. Dutoit, T., Pagel, V., Pierret, N., Bataille, F. and Van Der Vrecken, O., “The Mbrola Project: towards a set of high quality speech synthesizers free of use for non commercial purposes”, *ICSPL96*, Philadelphia-USA (1996)
32. Rao, K. S. “Predicting Prosody from Text for Text-to-Speech Synthesis”, *Springer*, New York, Heidelberg, Dordrecht, London, 19-21 (2012).
33. Yılmaz, A.E., “Türkçe metinden konuşma sentezleme uygulamaları için bir veri sözlük seti ve yazılım çerçevesi”, *Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi*, 4 (24): 735-744 (2009).

EK AÇIKLAMALAR A.

DİFON VE TAŞIYICI KELİME TABLOSU

Boşluk Grubu		A Grubu		B Grubu	
Difon	Taşıyıcı	Difon	Taşıyıcı	Difon	Taşıyıcı
—		a_	makara	b_	kab
_a	ankara	aa	paar	ba	araba
_b	bekar	ab	abla	bb	abbi
_c	cemil	ac	acmak	bc	abci
_ç	çeri	aç	açmak	bç	abçe
_d	demir	ad	adsız	bd	abdi
_e	efe	ae		be	bebek
_f	fedai	af	afgan	bf	abfe
_g	geri	ag	aglamak	bg	ebge
_ğ	ağlamak	ağ	ağlamak	bğ	abğa
_h	hemen	ah	ahlamak	bh	abhe
_ı	ırak	aı	paır	bı	bıçak
_i	iran	ai	pair	bi	biber
_j	jel	aj	ajmak	bj	obje
_k	kel	ak	akmak	bk	obka
_l	leman	al	almak	bl	abla
_m	mersin	am	amca	bm	imbe
_n	nerde	an	ankara	bn	ibne
_o	ordu	ao	paor	bo	boru
_ö	ömer	aö	paör	bö	börek
_p	pelit	ap	aptal	bp	abpa
_r	reyhan	ar	arsız	br	bravo
_s	sema	as	asmak	bs	absi
_ş	şevkat	aş	aşmak	bş	abşi
_t	terlik	at	atmak	bt	abti
_u	uyuz	au	kaut	bu	bulut
_ü	üzüm	aü	kaüt	bü	büyük
_v	veda	av	avrupa	bv	ebve
_y	yemek	ay	aymak	by	ebye
_z	zeybek	az	azmak	bz	abzi

C Grubu		Ç Grubu		D Grubu	
Difon	Taşıyıcı	Difon	Taşıyıcı	Difon	Taşıyıcı
c_	sac	ç_	saç	d_	murad
ca	canan	ça	çayır	da	dana
cb	ecbe	çb	eçbe	db	edbe
cc	ecce	çc	eçce	dc	edce
cç	ecçe	çç	eççe	dç	edçe
cd	ecdi	çd	eçde	dd	edde
ce	cemil	çe	çelebi	de	demir
cf	ecfe	çf	eçfe	df	edfe

cg	ecge	çg	eçge	dg	edge
cğ	ecği	çğ	eçğe	dğ	edğe
ch	eche	çh	eçhe	dh	edhe
cı	cızık	çı	çırak	dı	dıman
ci	ciğer	çi	çiçek	di	dibek
cj	ecje	çj	eçje	dj	edje
ck	ecke	çk	eçke	dk	edke
cl	ecle	çl	eçle	dl	edle
cm	ecme	çm	eçme	dm	edme
cn	ecnebi	çn	eçne	dn	edne
co	ecoma	ço	çocuk	do	dolap
cö	cömert	çö	çörek	dö	dönek
cp	ecpi	çp	eçpe	dp	edpe
cr	ecri	çr	eçre	dr	edre
cs	ecsi	çs	eçse	ds	edse
cş	ecşi	çş	eçşe	dş	esşe
ct	ecti	çt	eçti	dt	edte
cu	Cuma	çu	çuval	du	durak
cü	cüce	çü	çürük	dü	dümen
cv	ecve	çv	eçve	dv	edve
cy	ecye	çy	eçye	dy	edye
cz	ecze	çz	eçze	dz	edze

E Grubu

Difon	Taşıyıcı
e_	kete
ea	peap
eb	ebru
ec	ecdad
eç	eçmek
ed	edmek
ee	zeer
ef	eftal
eg	egmek
eğ	eğmek
eh	ehmek
eı	teır
ei	teir
ej	ejmek
ek	ekmek
el	eldiven
em	emsal
en	endam

F Grubu

Difon	Taşıyıcı
f_	kaf
fa	falaka
fb	efbe
fc	efce
fç	efçe
fd	efde
fe	fedaf
ff	effe
fg	efge
fğ	efğe
fh	efhe
fi	fitır
fi	fitil
fj	efje
fk	efke
fl	efle
fm	efme
fn	efne

G Grubu

Difon	Taşıyıcı
g_	mag
ga	gala
gb	egbe
gc	egce
gç	egçe
gd	egde
ge	gemi
gf	egfe
gg	egge
gğ	egğe
gh	eghe
gı	agımak
gi	gider
gj	agja
gk	agka
gl	agla
gm	agma
gn	agna

eo	peot	fo	foto	go	gomalak
eö	peöt	fö	fötör	gö	görenek
ep	eptek	fp	efpe	gp	egpe
er	erkek	fr	efre	gr	grup
es	eski	fs	efse	gs	egse
eş	eşsiz	fş	efşe	gş	egşe
et	etli	ft	efte	gt	egte
eu	teuk	fu	fular	gu	gurur
eü	teük	fü	füme	gü	güve
ev	evli	fv	efve	gv	egve
ey	eylem	fy	efye	gy	egye
ez	ezgi	fz	efze	gz	egze

Ğ Grubu

Difon	Taşıyıcı
ğ_	bağ
ğa	ağalak
ğb	ağba
ğc	ağca
ğç	ağça
ğd	ağda
ğe	ağek
ğf	ağfa
ğg	ağga
ğğ	ağğa
ğh	ağhe
ğı	ağır
ği	ağir
ğj	ağje
ğk	ağka
ğl	ağla
ğm	ağma
ğn	ağna
ğo	ağort
ğö	ağört
ğp	ağpa
ğr	ağra
ğs	ağsa
ğş	ağşa
ğt	ağta
ğu	ağur
ğü	ağüz
ğv	eğve

H Grubu

Difon	Taşıyıcı
h_	ah
ha	hamal
hb	ahbe
hc	ahce
hç	ahçe
hd	ahde
he	hedef
hf	ahfe
hg	ahge
hğ	ahğe
hh	ahhat
hı	ahıra
hi	hikaye
hj	ahje
hk	ahke
hl	ahle
hm	ahme
hn	ahne
ho	horoz
hö	hötöröf
hp	ahpe
hr	ahre
hs	ahse
hş	ahşe
ht	ahte
hu	hudut
hü	ahüde
hv	ahve

I Grubu

Difon	Taşıyıcı
ı_	kızı
ia	kıar
ib	ıbrık
ic	ıcmak
iç	ıçmak
id	ıdmak
ie	kıer
if	ıflak
ig	ıglak
iğ	ığmak
ih	ıhlamak
ıı	kıık
ii	kıik
ij	ıjmak
ık	ıklamak
ıl	ıldamak
ım	ımramak
ın	ınlamak
ıo	kıor
ıö	kıör
ıp	ıplamak
ır	ırmak
ıs	ıska
iş	ışlamak
ıt	ıtlamak
ıu	kıur
ü	kür
ıv	ıvlamak

ğy	ağye	hy	ahye	ıy	ıylamak
ğz	ağzı	hz	ahze	ız	ızgara

İ Grubu

Difon	Taşıyıcı
i_	kirpi
ia	ziar
ib	ibrik
ic	icmal
iç	içmek
id	idman
ie	sier
if	iflemek
ig	iglemek
iğ	iğne
ih	ihlemek
ii	siir
ii	siir
ij	ijmek
ik	ikmal
il	ilmek
im	imza
in	insan
io	sior
iö	diör
ip	iplemek
ir	irdelemek
is	istek
iş	işmek
it	itmek
iu	diur
iü	diür
iv	ivmek
iy	iymek
iz	izlemek

J Grubu

Difon	Taşıyıcı
j_	nikelaj
ja	jamalak
jb	ejbe
jc	ejce
jç	ejçe
jd	ejde
je	jelibon
jf	ejfe
jg	ejge
jğ	ejğe
jh	ejhe
ji	ajımak
ji	jilet
jj	ajja
jk	ajka
jl	ajla
jm	ajma
jn	ajna
jo	jomak
jö	jömek
jp	ajpa
jr	ajra
js	ajsa
jş	ajşa
jt	ajta
ju	julak
jü	jürek
jv	ejve
jy	ejye
jz	ejze

K Grubu

Difon	Taşıyıcı
k_	kabak
ka	kabak
kb	ekbe
kc	ekce
kç	akçe
kd	akde
ke	kefen
kf	ekfe
kg	ekge
kğ	ekğe
kh	ekhe
kı	akıra
ki	kitap
kj	akja
kk	akka
kl	akla
km	akma
kn	akne
ko	akoma
kö	köpek
kp	akpa
kr	akra
ks	aksa
kş	akşa
kt	akta
ku	kulaç
kü	küvet
akva	
ky	akya
kz	akza

L Grubu

Difon	Taşıyıcı
l_	mal
la	lama
lb	alba
lc	alca
lç	alçı

M Grubu

Difon	Taşıyıcı
m_	dem
ma	masa
mb	embiya
mc	amca
mç	amça

N Grubu

Difon	Taşıyıcı
n_	keten
na	nazik
nb	enbiya
nc	enci
nç	ençi

ld	alda	md	emde	nd	endi
le	alema	me	melek	ne	nefer
lf	elfi	mf	emfe	nf	enfi
lg	elgime	mg	emge	ng	engi
lğ	elğime	mğ	emğe	nğ	enği
lh	elhime	mh	emhe	nh	enhi
lı	alımak	mı	mızık	nı	anılar
li	lira	mi	miço	ni	niyet
lj	elji	mj	emji	nj	anje
lk	elki	mk	emke	nk	anke
ll	elli	ml	emle	nl	anla
lm	elmi	mm	emmi	nm	anma
ln	elni	mn	emni	nn	anna
lo	aloma	mo	emola	no	noyan
lö	alö mü	mö	emöfe	nö	nöbet
lp	elpi	mp	empi	np	enpe
lr	elre	mr	emre	nr	enre
ls	else	ms	emse	ns	ense
lş	elşe	mş	emşe	nş	enşe
lt	elti	mt	emte	nt	ente
lu	aluma	mu	umuma	nu	nuray
lü	alüme	mü	müze	nü	nülay
lv	alvi	mv	emvi	nv	enve
ly	alyi	my	emyi	ny	enye
lz	elzi	mz	emzik	nz	enze

O Grubu

Difon	Taşıyıcı
o_	yoyo
oa	moar
ob	obruk
oc	ocmak
oç	oçmak
od	odmak
oe	moer
of	oflu
og	ogmak
oğ	oğlan
oh	ohlamak
oı	koır
oi	koir
oj	ojmak
ok	okmak

Ö Grubu

Difon	Taşıyıcı
ö_	mötö
öa	töar
öb	öblemek
öc	öcmak
öç	öçmek
öd	ödlek
öe	döer
öf	öfmek
ög	ögmek
oğ	öğren
öh	öhmek
öı	köır
öi	böir
öj	öjmek
ök	ökmek

P Grubu

Difon	Taşıyıcı
p_	kasap
pa	para
pb	apba
pc	apca
pç	apçe
pd	apda
pe	peri
pf	apfa
pg	apga
pğ	apğa
ph	apha
pı	apılar
pi	apiler
pj	apje
pk	apke

ol	olmuş	öl	ölmek	pl	apla
om	omlak	öm	ömlmek	pm	apma
on	ondan	ön	önden	pn	apna
oo	koop	öo	toör	po	polat
oö	koör	öö	föör	pö	apölek
op	oplamak	öp	öpmek	pp	appamak
or	ortak	ör	örmek	pr	apra
os	osman	ös	ösmek	ps	apse
oş	oşmak	öş	öşmek	pş	apşe
ot	otlak	öt	ötmek	pt	apte
ou	kour	öu	möür	pu	puma
ou	toür	öü	vöür	pü	pürüz
ov	ovmak	öv	övmek	pv	apve
oy	oylamak	öy	öymek	py	apye
oz	ozmak	öz	özmek	pz	apze

R Grubu

Difon	Taşıyıcı
r_	kader
ra	rahip
rb	arbede
rc	arce
rç	arçe
rd	arda
re	resim
rf	arfa
rg	arge
rğ	arğe
rh	arha
rı	rıza
ri	riya
rj	arja
rk	arka
rl	arla
rm	arma
rn	arna
ro	roma
rö	römork
rp	arpa
rr	arramak
rs	arsa
rş	arşa
rt	arta

S Grubu

Difon	Taşıyıcı
s_	makas
sa	salim
sb	asba
sc	asca
sç	asça
sd	asda
se	sema
sf	asfa
sg	asga
sğ	asğa
sh	asha
sı	sıva
si	sinek
sj	aşja
sk	aska
sl	asla
sm	asma
sn	asna
so	soba
sö	asömek
sp	aspa
sr	asra
ss	assa
sş	asşa
st	asta

Ş Grubu

Difon	Taşıyıcı
ş_	kaş
şa	şair
şb	aşba
şc	aşca
şç	aşça
şd	aşda
şe	şebek
şf	aşfa
şg	aşga
şğ	aşğa
şh	aşha
şı	şıra
şi	şimal
şj	aşja
şk	aşka
şl	aşla
şm	aşma
şn	aşna
şo	şoför
şö	şöbiyet
şp	aşpa
şr	aşra
şs	aşsa
şş	aşşa
şt	aştı

ru	rulet	su	sulak	şu	şurup
rü	rüya	sü	sümük	şü	aşüre
rv	arva	sv	asva	şv	eşve
ry	arya	sy	asya	şy	eşye
rz	arza	sz	asza	şz	eşze

T Grubu

Difon	Taşıyıcı
t_	baret
ta	tabak
tb	atba
tc	atca
tç	atça
td	atda
te	temel
tf	atfa
tğ	atga
tğ	atğa
th	atha
tı	tımar
ti	tilavet
tj	atja
tk	atkı
tl	atla
tm	atma
tn	atna
to	torun
tö	töre
tp	atpa
tr	atra
ts	atsa
tş	atşa
tt	attama
tu	tutum
tü	tümen
tv	atva
ty	atya
tz	atza

U Grubu

Difon	Taşıyıcı
u_	kuyu
ua	suar
ub	ublamak
uc	uclamak
uç	uçmak
ud	udlamak
ue	suer
uf	uflamak
ug	uglamak
uğ	uğramak
uh	uhrevi
ui	suir
ui	suir
uj	ujmak
uk	ukmak
ul	ulmak
um	umarsız
un	uncu
uo	tuor
uö	tuör
up	uplamak
ur	urgan
us	uslu
uş	uşlamak
ut	utku
uu	tuur
üü	kuür
uv	uvmak
uy	uyku
uz	uzman

Ü Grubu

Difon	Taşıyıcı
ü_	büdü
üa	küar
üb	übmek
üc	ücmek
üç	üçgen
üd	üdmek
üe	vüer
üf	üflemek
üg	ügmek
üğ	üğmek
üh	ühmek
üi	vüir
üi	vüir
üj	üjmek
ük	ükmek
ül	ülkü
üm	ümlmek
ün	ünlü
üo	tüor
üö	tüör
üp	üplemek
ür	ürkek
üs	üstün
üş	üşmek
üt	ütmek
üu	tüur
üü	tüüt
üv	üvmek
üy	üymek
üz	üzgün

V Grubu

Difon	Taşıyıcı
v_	tav
va	vatan

Y Grubu

Difon	Taşıyıcı
y_	kay
ya	yara

Z Grubu

Difon	Taşıyıcı
z_	baz
za	zaman

vb	evbe	yb	eybe	zb	azba
vc	evce	yc	eyce	zc	azca
vç	evçe	yç	eyçe	zç	azça
vd	evde	yd	eyde	zd	azda
ve	veda	ye	yemek	ze	zehir
vf	evfe	yf	eyfe	zf	azfa
vg	evge	yg	eyge	zg	azga
vğ	evğe	yğ	eyğe	zğ	azğa
vh	evhe	yh	eyhe	zh	azha
vı	avıza	yı	yılan	zı	zımak
vi	virane	yi	yitik	zi	ziyan
vj	evje	yj	eyje	zj	ezje
vk	evke	yk	eyke	zk	ezke
vl	evle	yl	eyle	zl	ezle
vm	evme	ym	eyme	zm	ezme
vn	evne	yn	eyne	zn	ezne
vo	vole	yo	yoğurt	zo	azoma
vö	avöler	yö	yöre	zö	azöme
vp	evpe	yp	eype	zp	ezpe
vr	evre	yr	eyre	zr	ezre
vs	evse	ys	eyse	zs	ezse
vş	evşe	yş	eyşe	zş	ezşe
vt	evte	yt	eyte	zt	ezte
vu	avular	yu	yumak	zu	zuhal
vü	vücut	yü	yürek	zü	zürafa
vv	evvel	yv	eyve	zv	ezve
vy	evye	yy	eyyeme	zy	ezye
vz	evze	yz	eyze	zz	ezze

ÖZGEÇMİŞ

Erdem ERKAN 1977 yılında Bartın'da doğdu; ilk ve orta öğrenimini aynı şehirde tamamladı. 1995-1997 yılları arasında Dokuz Eylül Üniversitesinde Bilgisayar Programcılığı, 1999-2003 yılları arasında da Niğde Üniversitesi Mühendislik Fakültesinde Elektrik-Elektronik Mühendisliği eğitimini tamamladı. 2003-2004 yılları arasında Amerika'da Boston School Of Modern Languages'de dil eğitimi aldı. 2005-2011 Yılları arasında Birko A.Ş. Koyunlu Halı Fabrikasında Bilgi İşlem Şefi olarak çalıştı. Otomasyon ve ERP projelerinde görev aldı. 2010-2011 Yılları arasında Niğde Üniversitesi Bilgi İşlem Daire Başkanlığında Uzman olarak görev yaptı. Halen 2011 yılında başladığı Bartın Üniversitesi Uzaktan Eğitim Uygulama ve Araştırma Merkezinde Sistem Yönetim Uzmanı olarak görev yapmaktadır. Evli ve 2 çocuk babasıdır. Karabük Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği Bölümünde Yüksek Lisans eğitimine devam etmektedir.

ADRES BİLGİLERİ

Adres : Bartın Üniversitesi
Uzaktan Eğitim Uygulama ve Araştırma Merkezi
Ağdacı Yerleşkesi / BARTIN

Tel : (542) 412 9967

E-posta : erdem__erkan@bartin.edu.tr