



**ELMAS HECE MÜHENDİSLİK YÖNTEMİ
İLE METİN SESLENDİRME**

AHMET KÖROĞLU

**YÜKSEK LİSANS TEZİ
MEKATRONİK MÜHENDİSLİĞİ ANA BİLİM DALI**

DR. ÖĞR. ÜYESİ MÜRSEL ÖNDER

**Ağustos - 2019
Her hakkı saklıdır**

T.C.
TOKAT GAZİOSMANPAŞA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
MEKATRONİK MÜHENDİSLİĞİ ANA BİLİM DALI

YÜKSEK LİSANS TEZİ

ELMAS HECE MÜHENDİSLİK YÖNTEMİ İLE METİN
SESLENDİRME

AHMET KÖROĞLU

TOKAT
Ağustos - 2019

Her hakkı saklıdır

AHMET KÖROĞLU tarafından hazırlanan “**Elmas Hece Mühendislik Yöntemi ile Metin Seslendirme**” adlı tez çalışmasının savunma sınavı 5 AĞUSTOS 2019 tarihinde yapılmış olup aşağıda verilen Jüri tarafından Oy Birliği / ~~Oy Çokluğu~~ ile Gaziosmanpaşa Üniversitesi Fen Bilimleri Enstitüsü Mekatronik Mühendisliği Ana Bilim Dalı'nda YÜKSEK LİSANS TEZİ olarak kabul edilmiştir.

Jüri Üyeleri

İmza

Danışman

Dr. Öğr. Üyesi Mürsel ÖNDER
Gaziosmanpaşa Üniversitesi



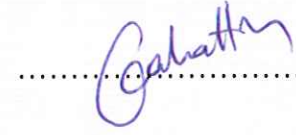
Üye

Dr. Öğr. Üyesi Ebubekir YAŞAR
Gaziosmanpaşa Üniversitesi



Üye

Dr. Öğr. Üyesi Bahattin KARAKAYA
İstanbul Üniversitesi-Cerrahpaşa



ONAY



Prof. Dr. Çetin ÇEKİÇ
Fen Bilimleri Enstitüsü Müdürü
21/08/2019

TEZ BEYANI

Tez yazım kurallarına uygun olarak hazırlanan bu tezin yazılmasında bilimsel ahlak kurallarına uyulduğunu, başkalarının eserlerinden yararlanılması durumunda bilimsel normlara uygun olarak atıfta bulunulduğunu, tezin içerdiği yenilik ve sonuçların başka bir yerden alınmadığını, kullanılan verilerde herhangi bir tahrifat yapılmadığını, tezin herhangi bir kısmının bu üniversite veya başka bir üniversitedeki başka bir tez çalışması olarak sunulmadığını beyan ederim.

AHMET KÖROĞLU

5 Ağustos 2019

ÖZET

YÜKSEK LİSANS TEZİ

ELMAS HECE MÜHENDİSLİK YÖNTEMİ İLE METİN SESLENDİRME

AHMET KÖROĞLU

TOKAT GAZİOSMANPAŞA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

MEKATRONİK MÜHENDİSLİĞİ ANA BİLİM DALI

(TEZ DANIŞMANI: DR. ÖĞR. ÜYESİ MÜRSEL ÖNDER)

Bilim adamları sesin yapay olarak üretilmesi konusunda iki yüzyılı aşkın bir süredir çalışmaktadır. Bilgisayarın icadından sonra, ses üretmek için bilgisayarlar kullanılmaya başlanmıştır. Bu yeni teknolojinin yardımıyla Metin Seslendirme sistemleri üretilmeye başlanmıştır. Bu sistemler girdi olarak bir metin alıp bu metnin sesli olarak okunmuş halini üretmektedir. İngilizce, Almanca ve Fransızca gibi bazı diller araştırmacıların ilgisini çekerken, Arapça gibi diller konusunda çok fazla çalışma yapılmamıştır.

Bu tez çalışmasında Arapça üzerine Kur'an-ı Kerim metni esas alınarak bir metin seslendirme sistemi geliştirilmiştir. Bu sistemi oluşturmak için yeni ve özgün bir yöntem olan Elmas Hece Mühendislik Yöntemi (EHMY) kullanılmıştır. Yapılan çalışmada bu yöntemi temel alan bir Metin Seslendirme sisteminin Visual C# yazılım dili kullanılarak oluşturulma aşamaları anlatılmıştır. Oluşturulan bu sistem girdi olarak bir ayet alarak çıktı olarak bu ayete karşılık gelen Arapça sesleri üretmektedir. Çalışmalar sonunda çok büyük miktarda veri kazancı sağlanmıştır. Ayrıca bu sistemin bazı ufak değişikliklerle benzer diller için de kullanılabilmesi görülmüştür. Çalışmada hedeflenen sonuçlara ulaşılmıştır.

2019, 78 Sayfa

ANAHTAR KELİMELELER: Metin Seslendirme, Elmas Hece Mühendislik Yöntemi, Arapça, Kur'an-ı Kerim

ABSTRACT

MASTER THESIS

TEXT TO SPEECH WITH DIAMOND SYLLABLE ENGINEERING METHOD

AHMET KÖROĞLU

**TOKAT GAZIOSMANPASA UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES**

DEPARTMENT OF MECHATRONICS ENGINEERING

SUPERVISOR: ASSIST. PROF. DR. MÜRSEL ÖNDER

Scientists have been working on the artificial production of sound for more than two centuries. Computers were used to produce sound after the invention of computers. Text to Speech (TTS) Systems have started to be produced with the help of this new technology. These systems take a text as input and produce an aloud reading of this text. While some popular world languages, such as English, German and French, have attracted researchers attention, much has not been done on languages such as Arabic.

In this thesis, a text to speech system was developed based on the text of the Holy Quran on Arabic. In order to create this system, a new and unique method, Diamond Syllable Engineering method was used. In this study, the steps of creating a text to speech system based on this method using Visual C # software language are explained. This generated system takes a verse text as input produces Arabic sounds corresponding to the verse as output. At the end of the studies, a large amount of data was obtained. It has also been found that the system can be used for similar languages with some minor changes. In the study, the targeted results were achieved.

2019, 78 PAGE

KEYWORDS: Text to Speech, Diamond Syllable Engineering Method, Arabic, Holy Quran

ÖNSÖZ

Bu tezin hazırlanmasında öncelikle lisans eğitimim sürecinde kendisinden çok şey öğrendiğim ve yüksek lisans çalışmam boyunca da manevi desteğini, bilgi ve tecrübesini esirgemeyen kıymetli danışman hocam Dr. Öğr. Üyesi Mürsel Önder' e sonsuz teşekkürlerimi sunarım.

Tez çalışmalarım boyunca yardım ve desteklerini her daim üzerimde hissettiğim çok değerli hocam Özkan İNİK' e en içten teşekkürlerimi sunarım.

Çalışmam boyunca yardımlarını esirgemeyen Gaziosmanpaşa Üniversitesi Mekatronik Mühendisliği öğretim üyelerine en içten teşekkürlerimi sunarım

Çocukluğumdan bugüne kadarki tüm eğitim hayatım boyunca bana her türlü desteği veren, hiçbir fedakarlıktan kaçınmayan başta annem Sn. Birgül KÖROĞLU ve babam Sn. Cafer KÖROĞLU olmak üzere tüm aileme sosuz teşekkür ve saygılarımı sunarım.

Bu süreçte her zaman pozitif enerjisi ile verdiği moral ve motivasyondan dolayı sevgili arkadaşım Sn. Şura DURMUŞ' a ayrıca teşekkür ederim.

AHMET KÖROĞLU

5 Ağustos 2019

İÇİNDEKİLER

Sayfa

ÖZET	i
ABSTRACT.....	ii
ÖNSÖZ	iii
İÇİNDEKİLER	iv
KISALTMALAR	vi
ŞEKİL LİSTESİ.....	vii
ÇİZELGE LİSTESİ.....	viii
1. GİRİŞ	1
2. KAYNAK ÖZETLERİ	3
2.1. Konuşma Sentezi Nedir?.....	3
2.2. Konuşma Sentezindeki Zorluklar.....	4
2.3. Metin Seslendirme Uygulamaları	5
2.4. Metin Seslendirmenin Kısa Geçmişi.....	6
2.5. Başlıca Metin Seslendirme Sistemleri	9
2.5.1. MITalk sistemi	9
2.5.2. Infovox sistemi.....	9
2.5.3. Bell Labs konuşma sentezi sistemi	9
2.5.4. CNET PSOLA sistemi	10
2.5.5. ETI Eloquence sistemi	10
2.5.6. Festival konuşma sentezi sistemi	10
2.5.7. MBROLA sistemi	11
2.5.8. Whistler sistemi	11
2.6. İnsan Konuşma Üretimi	11
2.7. Konuşma Sentezi Teknikleri.....	13
2.7.1. Artikülatör sentezi.....	13
2.7.2. Format sentez	14
2.7.3. Birleştirmeli sentez	15
3. MATERYAL VE YÖNTEM.....	18
3.1. Metin İşleme Bölümünün Hazırlanması	18
3.1.1. Elmas hece nedir?	18
3.1.2. Elmas hece algoritmasının geliştirilmesi	19
3.1.3. Elmas hece algoritmasının çalışma prensibi	22
3.1.4. Elmas hece tablosunun elde edilmesi	25

3.1.5. Elmas hece tablosunun tasnifi.....	28
3.2. Metin Seslendirme Bölümünün Hazırlanması	30
3.2.1. Elmas hece ses tablosunun oluşturulması	30
3.2.2. Seslendirme algoritması.....	31
4. BULGULAR.....	355
5. SONUÇ.....	366
6. KAYNAKLAR	377
7. EKLER.....	39
7.1. Elif “ ا ” Harfi için Elde Edilen Tasnif Tablosu	39
7.2. Be “ ب ” Harfi için Elde Edilen Tasnif Tablosu	41
7.3. Te “ ت ” Harfi için Elde Edilen Tasnif Tablosu.....	43
7.4. Se “ ث ” Harfi için Elde Edilen Tasnif Tablosu	45
7.5. Cim “ ج ” Harfi için Elde Edilen Tasnif Tablosu.....	46
7.6. Ha “ ح ” Harfi için Elde Edilen Tasnif Tablosu.....	47
7.7. Hı “ خ ” Harfi için Elde Edilen Tasnif Tablosu	48
7.8. Dal “ د ” Harfi için Elde Edilen Tasnif Tablosu.....	49
7.9. Zel “ ذ ” Harfi için Elde Edilen Tasnif Tablosu	51
7.10.Ra “ ر ” Harfi için Elde Edilen Tasnif Tablosu	52
7.11.Ze “ ز ” Harfi için Elde Edilen Tasnif Tablosu.....	54
7.12.Sin “ س ” Harfi için Elde Edilen Tasnif Tablosu	55
7.13.Şın “ ش ” Harfi için Elde Edilen Tasnif Tablosu	56
7.14.Sat “ ص ” Harfi için Elde Edilen Tasnif Tablosu	57
7.15.Dat “ ض ” Harfi için Elde Edilen Tasnif Tablosu.....	58
7.16.Tı “ ط ” Harfi için Elde Edilen Tasnif Tablosu.....	59
7.17.Zı “ ظ ” Harfi için Elde Edilen Tasnif Tablosu.....	60
7.18.Ayın “ ع ” Harfi için Elde Edilen Tasnif Tablosu.....	61
7.19.Gayın “ غ ” Harfi için Elde Edilen Tasnif Tablosu.....	62
7.20.Fe “ ف ” Harfi için Elde Edilen Tasnif Tablosu.....	63
7.21.Kaf “ ق ” Harfi için Elde Edilen Tasnif Tablosu	64
7.22.Kef “ ك ” Harfi için Elde Edilen Tasnif Tablosu	65
7.23.Lam “ ل ” Harfi için Elde Edilen Tasnif Tablosu	66
7.24.Mim “ م ” Harfi için Elde Edilen Tasnif Tablosu.....	68
7.25.Nun “ ن ” Harfi için Elde Edilen Tasnif Tablosu.....	70
7.26.Vav “ و ” Harfi için Elde Edilen Tasnif Tablosu	72
7.27.He “ ه ” Harfi için Elde Edilen Tasnif Tablosu	73
7.28.Ye “ ي ” Harfi için Elde Edilen Tasnif Tablosu.....	75
7.29.Elmas Hece Ses Tablosu	76
8. ÖZGEÇMİŞ.....	78

KISALTMALAR

Kısaltmalar

Açıklama

EHA

Elmas Hece Algoritması

EHMY

Elmas Hece Mühendislik Yöntemi

EHT

Elmas Hece Tablosu

EHST

Elmas Hece Ses Tablosu

ŞEKİL LİSTESİ

<u>Şekil</u>	<u>Sayfa</u>
Şekil 2.1 Kempelen'in makinesinin Wheatstone sürümü (Flanagan, 2009).....	7
Şekil 2.2 Elektronik mekanizmaya sahip ilk konuşma oluşturma sistemi.....	8
Şekil 2.3 İnsan konuşma sistemi (Flanagan ve ark., 2009).....	12
Şekil 3.1 Resmi Osmanî Kur'an-ı Kerim metni	20
Şekil 3.2 Ali el-Kârî Kur'an-ı Kerim metni.....	20
Şekil 3.3 Elmas Hece Algoritması.....	21
Şekil 3.4 Elmas Hece Mühendislik Yöntemi Program Arayüzü	22
Şekil 3.5 "Ayrıntılı Elmas Hece" sekmesi.....	23
Şekil 3.6 Seçilen ayet ve o ayetin elmas hecelere ayrılmış hali	24
Şekil 3.7 Bulunan hecelerin tabloya aktarılması	25
Şekil 3.8 "Elmas Hece Tablosu" sekmesi	26
Şekil 3.9 Kur'an-ı Kerim'in tamamı için elmas hecelerin bulunması	27
Şekil 3.10 Seslendirme algoritması	32
Şekil 3.11 Seslendirmenin yapılacağı sekmenin arayüzü.....	33
Şekil 3.12 "ب" sesine karşılık gelen ses çalınırken program çıktısı	34
Şekil 3.13 "س" sesine karşılık gelen ses çalınırken program çıktısı.....	34

ÇİZELGE LİSTESİ

<u>Çizelge</u>	<u>Sayfa</u>
Tablo 3.1 Elmas Hece Tablosu Kesit.....	28
Tablo 3.2 "Bi" Sesine Karşılık Gelen Elmas Heceler.....	28
Tablo 3.3 Tek Bir Harf İçin Örnek Tasnif Tablosu	29
Tablo 3.4 Elmas Hece Ses Tablosu Kesit	31



1. GİRİŞ

Bilim adamları sesin yapay olarak üretilmesi konusunda iki yüzyılı aşkın bir süredir çalışmaktadır. Bilgisayarın icadından sonra, ses üretmek için bilgisayarlar kullanılmaya başlanmıştır. Bu yeni teknolojinin yardımıyla girdi olarak bir metin alıp bu metnin sesli olarak okunmuş halini üreten “Metin Seslendirme” sistemleri üretilmeye başlanmıştır. Metin seslendirme konusu, özellikle son yirmi yıldır, üzerinde yoğun çalışmaların yapıldığı önemli konulardan biridir. Metin seslendirme sistemleri özellikle bazı fiziki engelleri bulunan insanların hayatlarını kolaylaştırma konusunda önemli bir rol oynamaktadır. Bu insanların kısıtlı iletişim olanaklarının artırılması bakımından çok büyük bir öneme sahiptir. Metin seslendirebilen bir sistem görme zorluğu yaşayan ya da görme engeli bulunan insanların, okuyamadıkları metinleri seslendirerek anlayabilmeleri; konuşma zorluğu yaşayan ya da konuşma engeli bulunan insanların da karşı tarafa iletmek istedikleri şeyleri yazıp bu uygulamalarla seslendirerek iletmeleri için olanak sağlamaktadır. Ayrıca eğitim öğretim, haberleşme ve kişisel gelişim gibi pek çok amaç için kullanılabilmesi gibi uzaktan iletişim kurma uygulamalarında ve insan yaşamının hemen her alanda kullanılabilir. Metin seslendirme sistemi geliştirilmiştir.

Bu tez çalışmasında Kur'an-ı Kerim metni esas alınarak geliştirilen bir metin seslendirme sistemi geliştirilmiştir. Kur'an-ı Kerim Arapçası günümüze kadar değişmediği için Standart Arapça olarak isimlendirilmektedir. Çalışmanın Arapça üzerine yapılma nedenleri ise kısaca şunlardır: Metin seslendirme sistemleri incelendiğinde: İngilizce, Almanca ve Fransızca gibi bazı popüler dünya dilleri araştırmacıların ilgisini çekerken, Arapça gibi diller konusunda çok fazla çalışma yapılmamıştır. Ayrıca Kur'an okuma öğretimi, Müslüman ülkelerde insanların karşısına çıkan ilk eğitim safhalarından biridir. Bu eğitimi alanlar çoğunlukla ilköğretim çağındaki çocuklardır. Günümüzde de Kur'an okuma öğretimine yönelik camilerde, Kur'an kurslarında hatta anaokulu ve kreşlerde dahi birçok eğitimler verilmektedir. Fakat çoğu insan bu eğitimlerin sonunda genellikle Kur'an okumayı öğrenememektedir. Kur'an okumayı öğrenenler ise kısa sürede öğrendiklerini unutmaktadır. Bu sorunları en aza indirmek için Kur'an-ı Kerim metni esas alınarak EHYMY geliştirilmiştir (Mürsel, 2017). Aynı zamanda Visual C# yazılım dili yardımıyla bu yöntemi temel alan bir metin seslendirme sistemi geliştirilmiştir. Bir MS sistemi temel olarak metin işleme bölümü

ve metin seslendirme bölümlerinden oluşmaktadır. Metin işleme bölümü, giriş metninin ayrıştırılması ve metin seslendirme bölümü için giriş metninin hazırlanmasında rol oynamaktadır. Bu doğrultuda metin işleme bölümünün düzgün çalışabilmesi için Elmas Hece Algoritması (EHA) geliştirilmiştir. EHA giriş metnini metin seslendirme bölümünün algılayacağı formata dönüştürmektedir.

Metin seslendirme bölümü için gereken altyapıyı oluşturmak amacıyla EHA tüm Kur'an-ı Kerim metnine uygulanarak Kur'an-ı Kerim' de geçen bütün Elmas Hece (EH) varyasyonları bulunmuştur. Bu heceler Elmas Hece Tablosu (EHT) denilen bir tabloda tutulmaktadır. Elde edilen EHT' nda heceler karmaşık halde olduğundan her bir harfin hareketlere yani hecelere karşılık gelecek seslere göre tasnifi yapılmıştır. Daha sonra tasnif tablosundan ses tekrarları atılarak Elmas Hece Ses Tablosu (EHST) elde edilmiştir. Daha sonra elde edilen EHST'nda ki seslere karşılık gelecek seslerin kayıtları alınarak EH mantığını esas alan bir ses veri tabanı oluşturulmuştur. Son olarak tez konusu sistemin ve genel olarak metin seslendirme sistemlerinin gelecek çalışmalarını yönlendirebilecek birtakım yaklaşımlar önerilmiştir. Tez için oluşturulan özgün metin seslendirme sisteminin Visual C# yazılımıyla oluşturulmuş program kodları ve elde edilen veriler de ek olarak sunulmuştur.

Bu tez çalışması, beş bölümden oluşmaktadır. 1. Bölüm "GİRİŞ" bölümüdür. Bu bölümde çalışma hakkında bilgilere yer verilmiştir. 2. Bölüm "KAYNAK ÖZETLERİ" bölümüdür. Bu bölümde ise çalışma ile ilgili literatür taraması yapılmıştır. 3. Bölüm "MATERİYAL ve YÖNTEM" bölümüdür. Bu bölümde Elmas Hece yöntemi, çalışma prensibi ve özellikleri hakkında bilgi verilmiştir. Bu bölümde ayrıca elmas hece yöntemini esas alan bir metin seslendirme uygulamasının geliştirilme aşamaları ve programın çalışma mantığından bahsedilmiştir. 4. Bölüm "BULGULAR ve TARTIŞMA" bölümüdür. Burada metin seslendirme uygulamasının tasarım aşamalarından sonra elde edilen bulgular anlatılmıştır. Son bölüm ise "SONUÇ" bölümüdür. Elde edilen sonuçların yorumlanması ve bu konu hakkında yapılacak olan yeni çalışmalar için önerilere yer verilmiştir.

2. KAYNAK ÖZETLERİ

Bu bölümde Metin Seslendirme hakkında genel bilgilere, Metin Seslendirme Yöntemleri ve bu yöntemlerin içinde yer aldığı çalışmalar kısaca anlatılmıştır. Yapılan çalışmaların özellikle popüler dünya dilleri üzerine olduğu ön plana çıkmaktadır.

2.1. Konuşma Sentezi Nedir?

İnsanla konuşma yoluyla etkileşime girebilen bilgisayarlar, bilgisayar çağının başından beri bilim adamları için bir hayal haline gelmiştir. Konuşabilen ve konuşmayı tanıyan bilgisayarlar bilim kurgu filmlerinin favorisi haline gelmiştir. Bilgisayar mühendisliği, elektronik mühendisliği gibi farklı alanlardan birçok bilim adamı bu hayale ulaşmak için bu konularda araştırmalar yapmıştır. Bu konu iki temel araştırma alanından oluşmaktadır: Metin Seslendirme ve Konuşma Tanıma. Bu iki problem farklı analiz edilmektedir. Bilim adamları metin seslendirme ve konuşma tanıma sistemleri üzerinde ayrı ayrı çalışmaktadır, ancak sistemler yeterince doğru olduğunda gelecekte konuşmayı anlayabilen ve konuşabilen bir bilgisayar oluşturmak için bu iki sistemin birleştirilmesi mümkün olacaktır.

Belirli bir metni konuşma sinyallerine dönüştürebilen sistemlere metin seslendirme sistemi denilmektedir. Bu sistemler için metnin kaynağı çok farklı olabilmektedir. Bir OCR çıktısı bu sistem için bir girdi olsa da aynı zamanda bir dil üretme sistemi tarafından üretilen metin de bir metin seslendirme sistemi için girdi olabilmektedir. İdeal bir metin seslendirme sisteminin amacı, bir insanın okuyabileceği herhangi bir metni işleyebilmektir. Örneğin, bir metin seslendirme sistemi sayıları okuyabilir, kısaltmaları kaldırabilir, bir kelime için farklı yazımları çözebilir. Bir metin seslendirme sistemi temel olarak metin işleme bölümü ve metin seslendirme bölümü olmak üzere iki bölümden oluşmaktadır. Metin işleme bölümü, giriş metninin ayrıştırılması ve metin seslendirme bölümü için giriş metninin hazırlanmasında rol oynamaktadır. Metin işleme bölümü ideal bir sistemde çok karmaşık olabilmektedir. Çünkü herhangi bir metni işleyebilmek ve doğru bir sonuç elde edebilmek için metnin çok doğru bir şekilde ayrıştırılması gerekmektedir. Metin işleme bölümünün ilk amacı, giriş metnini metin seslendirme bölümü tarafından işlenebilecek doğru alt bölümlere bölmektir. Doğru alt bölümler metin seslendirme bölümünde kullanılan sentez tekniğine göre

değişebilmektedir. Metin işleme bölümünün diğer amacı, tonlamayı bir sözcük ve cümle ile belirlemektir. Bu bilgi aynı zamanda metin seslendirme bölümüne anlayabileceği bir biçimde aktarılmalıdır. Bu amaçlara, metin seslendirme sistemlerinin metin işleme bölümlerinde sistemin kalitesiyle ilgili bir dereceye kadar ulaşılmaktadır. Metin seslendirme kısmı, son konuşmanın sentezlenmesinden sorumludur. Metin işleme bölümünden gelen girdiyi alarak çıktı olarak konuşma üretmektedir. Kural tabanlı ve birleşik sentez olmak üzere iki popüler konuşma sentezi tekniği bulunmaktadır. Kullanılan tekniğin türüne göre, sistem için bazı ön işlemler yapılmalıdır. Örneğin, sentezde kullanılacak olan ana ses birimleri için bir veritabanı, sistem çalışmaya başlamadan önce kaydedilmelidir.

2.2. Konuşma Sentezindeki Zorluklar

Her durumda doğru çalışması için, bir bilgisayar uygun şekilde programlanmalıdır. Programcı, programın karşılaşılabileceği durumları tahmin etmelidir. Bazı görevler için bu kolay olsa da doğal dil işleme gibi bazı görevler için yönetimi çok zor olabilmektedir. Sisteme olası her girişi belirlemek çok zor olduğundan, klasik programlama yaklaşımından farklı olan bazı teknikler kullanılmalıdır. Bununla birlikte, bu teknikler genellikle çoğu durumda doğru sonuç veren bazı bulguları sunsa da, her durumda bu geçerli olmamaktadır. Bir MS sistemi doğal dil metni ile ilgilenmektedir. Bu nedenle bir MS sistemi de bu tür problemlere çözüm sunmaktadır.

MS sistemlerindeki en büyük sorunlardan biride telaffuz sorunudur. Bir dil fonetik değilse, MS sistemi telaffuz ile ilgilenmelidir. Bu sorunun çözümlerinden biri, bir dilin tüm kelimelerinin telaffuzlarını kaydetmektir, fakat bu hafıza açısından pahalı bir çözümdür. Diğer bir çözüm ise telaffuz hakkında bazı genel kurallar oluşturarak bu kuralları giriş sözcüklerine uygulamaktır. Bu bellek açısından daha iyi bir çözümdür, ancak çok iyi bir dilbilimsel araştırma gerektirir ve bu istisnai durumlarda başarısız olabilir, çünkü bu kurallar her kelime için geçerli olmayabilir. Bir diğer sorun da telaffuzdaki belirsizlik; Bir kelime için birkaç farklı olası telaffuz olabilir ve hangisinin doğru olduğuna karar verilmesi gerekir. Metin işleme bölümü de kısaltmalar ile ilgilenmelidir. Her gün yeni kısaltmalar günlük hayata eklendiğinden, tüm kısaltmaların üstesinden gelebilecek bir sistem oluşturmak çok zordur. Bunun yanında kısaltmalarda belirsizlikte olabilir. Sayıları okumak metin seslendirme sistemlerinde ki başka bir

zorluktur. Sistem önce ne tür bir sayı olduğunu anlamalı ve ona göre davranmalıdır. Normal numaralar telefon numaralarından farklı okunur. Ayrıca, bir markanın seri numarasıysa, bir sayı farklı şekilde okunabilir. Örneğin, “Nokia 8850” de, “sekiz bin sekiz yüz elli” değil, “Nokia seksen sekiz, elli” olarak okunmalıdır. Tonlama hakkında karar vermek tonlama değişebileceğinden en zor görevlerden biridir.

Bu iki durumu göz önünde bulundurulmalıdır. Metin işleme bölümü bunu bağlamdan anlamalıdır. Metin işleme bölümü, sorunu en aza indirmelidir. Doğru telaffuz ve tonlamayı seçtikten sonra, metin seslendirme kısmı bunu anlamalıdır. Bu çok zor bir iştir, çünkü algısal olarak aynı sesler farklı bağlamlarda akustik olarak farklı olabilir. Bir sesin kesin süresi ve frekansları, ondan önce ve takip ettiği bölümler, kelime içindeki yeri, vurgulanıp vurgulanmaması gibi pek çok faktöre bağlıdır. Metin işleme bölümü tonlama ile uğraşırken, yalnızca tonlamanın doğal bir konuşma için nerede gerçekleşmesi gerektiğini belirler ancak bunu gerçekleştirmek metin seslendirme bölümünün rolüdür. Tonlama mekanizması henüz tam olarak anlaşılmamıştır; farklı tonlama modelleri var, ancak bunların hiçbiri doğru çalışması için yeterince başarılı değildir.

İdeal bir MS sistemi bu sorunlara iyi çözümler getirebilmelidir. Henüz bu problemleri tam olarak çözebilecek bir sistem yoktur, tüm sistemler elinden gelenin en iyisini yapmaya çalışır, bu problemleri çözmedeki başarıları ise sistemin kalitesini belirler.

2.3. Metin Seslendirme Uygulamaları

Gündelik hayatımızda konuşmayı çok kullanıyor olsak da aslında konuşarak çok şey öğrenmiyoruz. Genellikle bir şeyler öğrenmek için kitap okumak yerine dinlemeyi tercih ediyoruz. Konuşurken bir yöne bakmamıza gerek yok bir şey tutmamız gerekmiyor. Bu nedenle ellerimiz ve gözlerimiz serbesttir. Fazla konsantrasyon gerektirmeyen başka bir şey yapabiliriz. Dahası konuşma teknolojisi kör insanlar için bilgisayar dünyasının kapısını açar. Her metni okuyabilir, yazdıklarını kontrol edebilir ve internetin metin bölümlerine ulaşabilirler.

Metin seslendirme sistemlerinin başka bir kullanımı ise telefon hattı yardımı ile bilgisayara uzak bir yerden ulaşmaktır. Rezervasyon sistemleri (örneğin, havayolu,

otobüs, vb.) tarafından kullanılabilir. Ek olarak, bankacılık ve finans kuruluşları bu teknolojiyi kullanıcıya hesap bilgilerini vermek veya telefon hattıyla yeni işlemler yapmak için kullanabilirler. Bu nedenle insanların bankaya gitmeleri ve basit işlemlerin gerçekleşmesi için sıralarını beklemeleri gerekmez. Ayrıca özel bir donanım gerektirmez. Dünyanın herhangi bir yerinde bir telefon bulunursa bir banka sistemine ulaşılabilir.

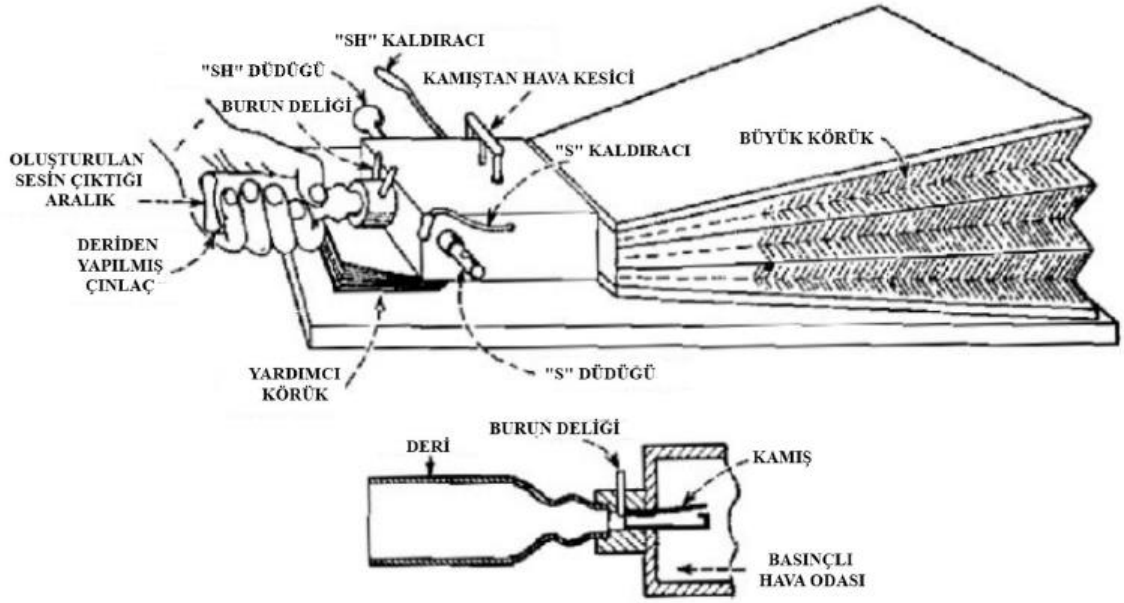
Sentezlenmiş konuşma aynı zamanda birçok eğitim durumunda da kullanılabilir. Konuşma sentezleyicili bir bilgisayar günde 24 saat ve yılda 365 gün ders verebilir. Birçok farklı dilin yazım ve telaffuz öğretimi gibi özel görevler için programlanabilir. Elektronik posta en çok kullanılan iletişim yollarından biri haline gelmiştir. Ancak, bazen e-postalara ulaşmak mümkün değildir. Bu sorunu çözmek için, e-postaları okuyabilen bazı sistemler geliştirilmiştir. Müşteri e-postalarını dinlemek için telefonunu kullanır. Bu sistemlerin yararlı olması için etkileşimli olması gerekir. İdeal olarak insanlar konuşma yoluyla sistemle etkileşime girmelidir, bunun için otomatik bir konuşma tanıma sistemi gerekir.

2.4. Metin Seslendirme'nin Kısa Geçmişi

MS'nin tarihi ilk bilgisayarın icadından sonra başlamaktadır. Çünkü metin seslendirme sisteminin bir metin ve bilgisayara ihtiyacı vardır. Belirli bir metni otomatik olarak konuşmaya dönüştürebilmelidir. Bununla birlikte metin seslendirme ile ilgili ilk çalışmalar iki yüzyıl önce yapılmıştır. Rus Profesör Kratzenstein 1779 yılında St. Petersburg'da beş uzun ünlü /a/, /e/, /i/, /o/ ve /u/) arasındaki fizyolojik farklılıkları anlattı ve bu sesleri yapay olarak üretmek için bir sistem geliştirmiştir (Flanagan, 1972; Schroeder, 1993).

Wolfgang von Kempelen, 1791 yılında Viyana'da tek sesler ve bazı ses kombinasyonları üretebilen "Akustik Mekanik Konuşma Makinesi" adı verilen bir makine yapmıştır (Klatt, 1987; Schroeder, 1993). Aslında, araştırmasına 1769'da Kratzenstein'dan önce başlamış ve insan konuşma üretimi üzerine yaptığı çalışmalar ve konuşma makinesiyle ilgili deneyleri hakkında bir kitap yayınlamıştır.

Charles Wheatstone, von Kempelen'in konuşma makinesinin ünlü versiyonunu 1800'lerin ortalarında inşa etmiştir (Flanagan, 1972). Bu biraz daha karmaşık bir makineydi ve ünlüler ve ünsüzlerin çoğunu üretebilmektedir. Ayrıca bazı ses kombinasyonları hatta bazı kelimeler üretebilmiştir.

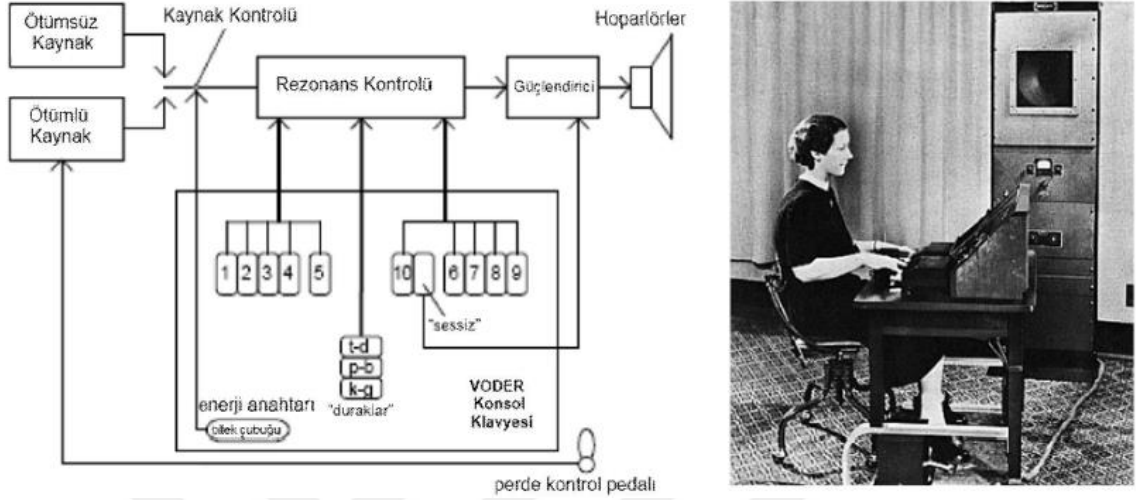


Şekil 2.1 Kempelen'in makinesinin Wheatstone sürümü (Flanagan, 2009)

1838'de Willis belirli bir sesli harf sesi ile ses kanalının geometrisi arasındaki bağlantıyı bulmuştur. Organ borularıyla aynı olan tüp rezonatörlerini kullanarak farklı ünlü sesler üretmiştir. Ünlü ses kalitesinin sadece tüpün uzunluğuna bağlı olduğunu ancak çapına bağlı olmadığını keşfetmiştir (Schroeder, 1993).

1922'de Stewart yalnızca statik sesli harfler üretebilen ilk tam elektrik konuşma sentezleme cihazını tanıtmıştır. Ünsüz sesler veya bağlı sesler bu sistemde üretilmesi mümkün değildi (Klatt, 1987). Wagner de benzer bir sistem geliştirmiştir. 1939'da ilk konuşma sentezleyicisi olarak kabul edilen cihaz olan VODER, New York'ta Homer Dudley tarafından tanıtılmıştır (Flanagan, 1972; Klatt, 1987). Konuşma kalitesi ve anlaşılabilirliği iyi olmaktan uzak olmasına rağmen, yapay konuşma üretme potansiyeli göstermiştir. İlk format sentezleyici PAT (Parametric Artificial Talker) 1953 yılında Walter Lawrence tarafından tanıtılmıştır (Klatt, 1987). 1972'de, John Holmes sentezleyicisine, "basit hayattan zevk alıyorum" adlı sentezlenmiş cümleyi elle

ayarladığını duyurmuştur. Kalite o kadar iyiydi ki, ortalama bir dinleyici sentezlenen ses ile doğal olan arasındaki farkı söyleyememiştir (Klatt, 1987).



Şekil 2.2 Elektronik mekanizmaya sahip ilk konuşma oluşturma sistemi

Noriko Umeda ve arkadaşları Japonya'da İngilizce için ilk tam metin konuşma sistemini geliştirmiştir. Mevcut sistemlerin kalitesinden çok uzak olmasına rağmen konuşma monotondur ama oldukça anlaşılırdı. Allen, Hunnicutt ve Klatt, 1979'da M.I.T.'de Telesensory Systems Inc. ticari metin seslendirme sisteminde kullanılan MITalk'u üretmiştir (Allen ve ark., 1987; Klatt, 1987).

1970'lerin sonunda ve 1980'lerin başında, ticari olarak önemli miktarda metin seslendirme ürünü üretilmiştir. Bilgisayarlarda çalışan yazılım ürünlerinin yanında donanım çözümleri sunan farklı metin seslendirme çipleri bulunmaktadır. Bu günlerden sonra, DecTalk, Whistler, Mbrola, vb. Gibi başarılı sayılabilecek birçok metin seslendirme sistemi farklı diller için geliştirilmiştir. Bununla birlikte hem anlaşılabilirlik hem de doğallık açısından kaliteli ses üreten bir sistem için daha fazla ilerlemeye ihtiyaç duyulmaktadır (Dutoit ve Leich, 1993; Dutoit ve ark., 1996; Huang ve ark., 1996; Anonim, 2018).

2.5. Başlıca Metin Seslendirme Sistemleri

Bu bölümde bazı metin seslendirme sistemleri, geliştirme amaçları ve devam eden projeler açıklanmaktadır. Tüm sistemleri anlatmak mümkün olmasa da iyi bilinen sistemler kısaca anlatılmıştır.

2.5.1. MITalk sistemi

Bu sistem 1979 yılında Allen, Hunnicutt ve Klatt tarafından gösterilmiştir. Bu sistem MIT laboratuvarlarında format sentezleyici bir sistem olarak geliştirilmiştir. Bu sistemde kullanılan teknoloji bugün için birçok sistem için temel oluşturmaktadır (Allen ve ark., 1987).

2.5.2. Infovox sistemi

Telia Promotor AB Infovox en ünlü çok dilli MS sistemlerinden biridir. İlk ticari versiyon 1982'de İsveç Kraliyet Teknoloji Enstitüsü'nde geliştirilmiştir. Bu sistemde sentez yöntemi olarak kademeli format sentezi kullanılmaktadır. (Ljungqvist ve ark., 1994). Şu anda bu sistemin hem yazılım hem de donanım uygulamaları bulunmaktadır.

En son ticari versiyon Infovox 230, Amerikan ve İngiliz İngilizcesi, Danimarka'ca, Fince, Fransızca, Almanca, İzlandaca, İtalyanca, Norveççe, İspanyolca, İsveççe ve Hollandaca için geçerlidir (Ljungqvist ve ark., 1994). Konuşma anlaşılır ve sistemin erkek, kadın ve çocuk da dahil olmak üzere 5 farklı ses seçeneği bulunmaktadır. Yeni sesler kullanıcı tarafından da eklenebilmektedir.

Son zamanlarda Infovox 330 tanıtılmıştır. Bu sürüm İngilizce, Almanca ve Hollandaca dillerini desteklemekte ve diğer diller geliştirilmektedir. Daha önceki Infovox sistemlerinden farklı olarak, bu sürüm diphone birleştirme yöntemine kullanılmaktadır. Bu yüzden bu sistem daha karmaşıktır ve daha fazla hesaplama yükü gerektirmektedir.

2.5.3. Bell Labs konuşma sentezi sistemi

İngilizce, Almanca, Fransızca, İspanyolca, İtalyanca, Rusça, Romence, Çince ve Japonca için kullanılabilir (Möbius ve ark., 1996). Diğer diller geliştirilme aşamasındadır. Yazılım, İngilizce dışındaki tüm diller için aynıdır. Bu nedenle çok dilli

bir sistem olarak görülmektedir. Dile özgü bilgiler ayrı tablolarda ve parametre dosyalarında saklanmaktadır.

Sistemin iyi metin analiz yetenekleri, kelime veya doğru isim telaffuzu, tonlama, vurgulama ve yazılı ifadeleri bölümlene süresi gibi özellikleri vardır. Sistemin en iyi özelliklerinden biri de farklı araştırma gruplarının farklı modüller üzerinde bağımsız olarak çalışabilmeleri için tamamen modüler olmasıdır. Geliştirilmiş modüller, modüller arasında iletilen bilgiler doğru bir şekilde tanımlandığı sürece herhangi bir zamanda entegre edilebilmektedir (Möbius ve ark., 1996).

2.5.4. CNET PSOLA sistemi

France Telecom CNET, 1980'lerin ortalarında birleştirmeli sentez için en umut verici yöntemlerden biri olan PSOLA'yı kullanan bir telefon tabanlı sentezleyici tanıtmıştır. En son ticari ürün Elan Informatique'den ProVerbe metin seslendirme sistemi olarak temin edilebilmektedir. Konuşma hızı sistemde ayarlanabilmektedir. Sistem şu anda Amerikan ve İngiliz İngilizcesi, Almanca, Fransızca ve İspanyolca dillerinde kullanılabilmektedir.

2.5.5. ETI Eloquence sistemi

Bu sistem ABD' de bulunan Eloquent Technology Inc. Şirketi tarafından geliştirilmiştir. Şu anda İngiliz ve Amerikan İngilizcesi, Almanca, Fransızca, İtalyanca, Meksika ve İspanyolca dilleri için kullanılabilmektedir. Çince gibi bazı diller ise geliştirilme aşamasındadır. Her dil için 7 farklı ses vardır ve bunlar kullanıcı tarafından kolayca özelleştirilebilmektedir.

2.5.6. Festival konuşma sentezi sistemi

Bu sistem Edinburgh Üniversitesi'ndeki CSTR'de geliştirilmiştir. İngiliz ve Amerikan İngilizcesi, İspanyolca ve Galce şu anda sistemde ki mevcut dillerdir. Eğitim, araştırma ve kişisel kullanım için sistem serbestçe kullanılabilmektedir (Black ve Taylor, 1997).

2.5.7. MBROLA sistemi

Belçika'daki Facul Polytechnique de Mons'daki TCTS Laboratuvarı tarafından başlatılan bir projedir. Projenin temel amacı, ticari olmayan amaçlar ve araştırma amaçlı kullanımlar için çok dilli bir metin seslendirme sistemi oluşturmaktır. Bu projede kullanılan yöntem PSOLA'ya çok benzerdir. Ancak, PSOLA, CNET'in ticari markası olduğundan, bu projenin adı MBROLA'dır.

MBROLA tam bir metin seslendirme sistemi değildir. Çünkü ham metni girdi olarak kabul etmez. Süre ve ses tonu gibi bazı bilgilerin yer aldığı bir fonem listesi alarak ve çıkış konuşmasını yapmaktadır. Amerikan, İngiliz, Breton İngilizcesi, Brezilya Portekizcesi, Fransızca, Felemenkçe, Almanca, Romence ve İspanyolca dillerinde veri tabanları mevcuttur. Diğer diller için veri tabanları üzerinde çalışılmaktadır (Dutoit ve Leich, 1993; Dutoit ve ark., 1996; Anonim, 2018).

2.5.8. Whistler sistemi

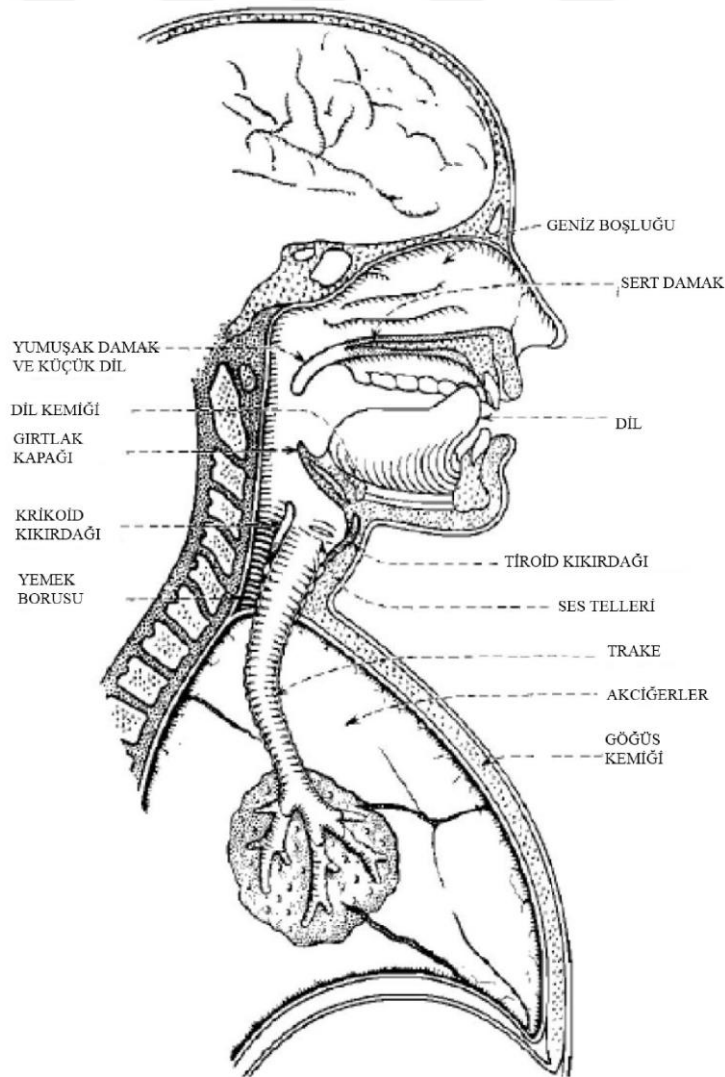
Microsoft'ta geliştirilmekte olan eğitilebilir bir metin seslendirme sistemidir. Sistemin amacı, doğal seslendirme konuşması üretmek ve orijinal konuşmacının akustik özelliklerini andıran bir çıktı üretmektir. Konuşma motoru, birleştirilmiş sentez ve eğitim prosedürüne dayanmaktadır (Huang ve ark., 1996).

2.6. İnsan Konuşma Üretimi

Konuşma, insanda ses organlarında üretilmektedir. Ses organları Şekil 2.3' te gösterilmiştir. Akciğerler diyafram ile birlikte konuşma üretimi için ana enerji kaynağıdır. Hava akımı, ses telleri ve gırtlak arasındaki glottis (nefes borusu ağzı) vasıtasıyla ses yolu, gırtlak ve burun boşluğu olmak üzere üç ana boşluğa zorlanır. Hava akımı sırasıyla burun ve ağız boşluklarından çıkmaktadır. Glottis, ses sistemindeki en önemli ses kaynağıdır. Ses telleri arasında V şeklinde bir açıklıktır. Ses telleri konuşma sırasında farklı sesler üretmek için farklı hareket etmektedir. Hızla kapanıp açılarak hava akımını düzenleyerek sesli harflerin oluşmasına yardımcı olmaktadır. Ses tellerinin temel titreşim frekansı erkeklerde yaklaşık 110 Hz, kadınlarda 200 Hz ve çocuklarda ise 300 Hz' dir. Dur ünsüz sesleri üretmek için ses telleri, hava akışını tamamen açık bir pozisyona engelleyen tamamen kapalı bir pozisyondan hareket edebilir. Ancak /s/ veya

/f/ gibi dengesiz ünsüzler için bunlar tamamen açık olabilir. /H/ gibi fonemler için ara konum oluşabilir.

Gırtlak boşluğu gırtlak ağız boşluğuna bağlıdır. Neredeyse sabit boyutlara sahiptir, ancak bir ucundaki gırtlak ve diğer ucundaki yumuşak damak yükselterek veya indirilerek uzunluğu hafifçe değiştirilebilmektedir. Yumuşak damak havayı izole ederek rotayı gırtlak boşluğundan burun boşluğuna bağlamaktadır. Gırtlak boşluğunun dibinde yiyeceğin gırtlak ağıza ulaşmasını önlemek ve yemek borusunu akustik olarak vokal kanaldan izole etmek için krikoid ve sahte ses telleri bulunmaktadır. Krikoid yutma sırasında kapanıp ve normal solunum sırasında açılarak sahte ses tellerini ve ses tellerini izole etmektedir.



Şekil 2.3 İnsan konuşma sistemi (Flanagan ve ark., 2009)

Ağız boşluğu, ses yolunun en önemli kısımlarından biridir. Büyüklüğü şekli ve akustiği damak, dil, dudak, yanak ve diş hareketleri ile değişebilmektedir. Özellikle dil çok esnek, uç ve kenarlar bağımsız olarak hareket ettirilebilmektedir. Aynı zamanda tüm dil ileri, geri, yukarı ve aşağı hareket edebilir. Dudaklar, sesin yayıldığı ağız açıklığının boyutunu ve şeklini kontrol etmektedir. Ağız boşluğunun aksine, burun boşluğu sabit boyutlara ve şekle sahiptir. Uzunluğu yaklaşık 12 cm ve hacmi 60 cm³ 'tür. Burun boşluğuna giden hava akımı yumuşak damakla kontrol edilmektedir. (Rabiner ve Juang, 1993; Deller ve ark., 1987).

2.7. Konuşma Sentezi Teknikleri

Sentezlenmiş konuşma üretmek için çeşitli yöntemler vardır. Bu yöntemlerin hepsinin bazı avantajları ve dezavantajları bulunmaktadır. Bu yöntemler genellikle 3 grupta kategorize edilir: artikülatör sentez, format sentez ve birleşik sentez. Bu bölümde, her yönteme genel bir bakış sunacağız.

2.7.1. Artikülatör sentezi

Artikülatör sentez yöntemi, insan ses organlarını mümkün olduğunca mükemmel bir şekilde modellemeye çalışmaktadır bu nedenle potansiyel olarak en umut verici yöntemdir. Bununla birlikte, uygulanması çok zor bir yöntemdir ve çok fazla hesaplama gerektirmektedir. Bu nedenlerden dolayı, bu yöntem diğer sentez yöntemlerine göre daha az dikkat çekmiştir ve diğer yöntemlerle aynı başarı seviyesinde uygulanamamıştır (Kröger, 1992; Rahim ve ark., 1993).

Vokal kanal kasları, vokal kanalın şeklini değiştirmesine ve artikülatörlerin değişmesine neden olmaktadır. Bu durum da farklı seslerin çıkmasına yol açmaktadır. Artikülatör modelinin analizi için gerekli veriler genellikle doğal konuşmanın X-ışını analizinden elde edilmektedir. Bununla birlikte elde edilen veriler, gerçek ses yolu doğal olarak 3 boyutlu olmasına rağmen genellikle 2 boyutlu olarak elde edilmektedir. Bu nedenle artikülatör sentezi, konuşma sırasında artikülatörlerin hareketleri yeterli verilerin bulunmaması nedeniyle modellenmesi çok zordur. Bu yöntemle, artikülatörlerin kütle ve serbestlik derecesi de göz önüne alınamadığı için bu da bazı eksikliklere neden

olmaktadır. Dilin hareketleri aşırı derece karmaşık olduğundan tam olarak modellenmesi de oldukça zordur.

Artikülator sentezi, mevcut sistemlerde çok nadiren kullanılmaktadır. Çünkü bu model için gerekli parametreleri elde etmek için yapılması gereken analiz işlemi çok zor bir iştir. Aynı zamanda bu model çalışma süresinde çok fazla hesaplama gerektirmektedir. Bununla birlikte, analiz yöntemlerinin gelişmesi ve hesaplama gücündeki artış, insan konuşma sistemi için en iyi model gibi görüldüğü için geleceğin yöntemi olarak görülmektedir.

2.7.2. Format sentez

Bu muhtemelen son yıllarda en yaygın kullanılan yöntemdir. Bu kaynak filtre modeline dayanmaktadır. Burada kaynak modeller akciğerler ve filtre modelleri ise vokal kanaldır. Bu teknikte kademeli ve paralel olmak üzere kullanılan iki temel yapı vardır. Bununla birlikte, daha iyi bir performans için, genellikle bunların bir kombinasyonu kullanılır. Bu teknik aynı zamanda sonsuz sayıda ses üretmeyi mümkün kılmaktadır, böylece birleşik sentezden daha esnektir. DECTalk, MITalk, Infovox'un önceki sürümleri format sentez yöntemini kullanan sistemlere örnektir (Allen ve ark., 1987; Ljungqvist ve ark., 1994; Hallahan, 1996).

Kademeli yapı, genizsel olmayan sesler için daha iyidir. Paralel yapıdan daha az kontrol bilgisi gerektirdiğinden uygulanması da daha kolaydır. Bununla birlikte, sürtünmeli ve patlayıcı seslerin üretilmesi bir problemdir. Paralel yapı genizsel, sürtünmeli ve sessiz ünsüzler için daha iyidir ancak bazı ünlüler bu yapı ile modellenememektedir. Bu tekniklerin hiçbiri tatmin edici ses üretmek için yeterli olmadığından sistemlerin çoğunda bu ikisinin kombinasyonu kullanılmaktadır. (Allen ve ark., 1987; Klatt, 1980; O'Saughnessy, 1987).

Format sentez sistemleri bir konuşma veritabanı gerektirmemektedir. Bu nedenle küçük bellek alanına ihtiyaç duymaktadır. Ayrıca farklı sesler üretebilirler yani bir konuşmacıya bağlı değildirler. Bunlar format sentez sisteminin avantajlarıdır. Ancak insan konuşmasını modellemek için bu modelde kullanılacak filtre için bazı parametreler çıkarılmalıdır. Bu parametrelerin elde edilmesi çok zor bir iştir. Ayrıca bu

parametreler bir formüle göre kullanılmalıdır, buda çalışma süresinde bir miktar hesaplama yükü gerektirmektedir. Her ne kadar bu sistemlerdeki hesaplama yükü artikülatör sentez sistemlerinden çok daha az olsa da birleşik sentez sistemlerinden daha fazladır.

2.7.3. Birleştirmeli sentez

Önceden kaydedilmiş ifadelerin birleştirilmesi ile yapılmaktadır. Anlaşılır ve doğal bir konuşma oluşturmanın en kolay yolu olarak düşünülebilir. Bununla birlikte birleştirmeli sentezleyiciler genellikle diğer tekniklerin gerektirdiğinden daha fazla bellek gerektirmektedir.

Birleştirmeli sentez yönteminde en önemli adımlardan biri doğru birim uzunluğuna karar vermektir. Genellikle daha uzun ve daha kısa birimler arasında bir alışveriş vardır. Daha uzun birimlerde daha doğal ses elde etmek daha kolaydır; ancak bu durumda gereken birim ve gereken bellek miktarı daha fazla olmaktadır. Daha kısa birimler kullanıldığında genellikle çıkış konuşması daha az doğaldır ve daha az bellek gerekir ancak veritabanı hazırlamak daha kesin bir işlem gerektirmektedir.

Birim uzunluk kararı uygulamanın ihtiyacına göre değişmektedir. Uygulamanın ihtiyacına göre birim uzunluğu harften başlayıp kelimeye kadar çok farklı aralıkta olabilir.

Kelime birleştirme

Havayolu rezervasyon sistemi gibi küçük kelime bilgisine sahip bir uygulamaya ihtiyaç duyulursa bu yöntem kolayca kullanılabilir. Sözcükler ayrı ayrı tonlamalı bir şekilde kaybedilir. Ayrıca sistem önceden kaydedilmiş kelimelerle sınırlandırılacaktır. Buna ek olarak kelimeleri birleştirmek, kelimeler sıfır seviyeden başlayıp ve sıfır seviyede bittiği için fazla işlem yükü gerektirmeyecektir. Her kelimedden sonra bir süre durulduğundan cümle kesik kesikmiş gibi okunacaktır. Eğer kelimeler bir cümleden alınırsa ve cümlede tonlama varsa bu sistemi etkileyecektir. Örneğin, “Ben okula gitmek istiyorum” cümlesinde “ben” üzerinde tonlama varsa ve bu cümleden “ben” kelimesi kaydedilmişse, sistem istenmediğinde bile daima “ben” kelimesini tonlamalı olarak okuyacaktır. Bu sistemler çalışma zamanında çok az hesaplama gerektirir, çünkü

önceden kaydedilmiş sözcükleri bir araya getirirler. Bu sistemlerin bir diğer avantajı ise çıkış kelimeleri içindeki konuşma kalitesinin önceden kaydedilmiş olduklarından çalışma zamanında oluşturulmadıkları için neredeyse mükemmel olmasıdır (Allen ve ark., 1987).

Ses birimi (harf) birleştirme

Bu yöntemde öncelikle dildeki harfler çıkarılmaktadır. Ardından harflere ait sesler kaydedilmektedir. Bunu yapmanın en kolay yolu bazı örnek kelimeler veya cümleler kaydedip daha sonra ses düzenleyicileri ile elle veya bir ses editörü yardımıyla bazı otomatikleştirilmiş teknikler kullanılarak bunlardan harflere ait sesleri çıkarmaktır. Daha sonra farklı harf seslerinin enerji değerleri normalleştirilir. Son olarak bu sesler kelimeleri oluşturmak için bir araya getirilir. Bu birleştirme işleminde sesler arasında geçiş düzgünlüğü sağlamak için bazı dijital sinyal işleme teknikleri kullanılmalıdır. Bu sistemlerdeki güçlüklerden biri konuşma sinyalindeki bir sesin başlangıcı ve bitişi kesin olarak tespit edilemediğinden, sesleri giriş konuşmasından doğru bir şekilde elde edememektir. Bu nedenle son harf ses setini almak için çok fazla denemeye ihtiyaç duyulabilir. Harf ses setini aldıktan sonra, bu seslerin düzgün bir şekilde birleştirilmesi gerekir; ancak bu aynı zamanda zor bir iştir. Ayrıca bir noktada birden fazla ses kullanmak mümkünse hangisinin kullanılacağına sistem tarafından doğru şekilde karar verilmelidir. Öte yandan bu sistemler çok fazla hafıza gerektirmez çünkü sadece birkaç konuşma parçası önceden kaydedilmiştir. Ayrıca sesler üzerindeki bazı işlemlerle farklı tipte sesler elde edilebilir (Allen ve ark., 1987).

Diphone (hece) birleştirme

Hece harften daha uzun ve kelimedenden daha küçük olan temel konuşma birimleridir. Bu birimleri birleştirmek için birkaç yöntem vardır. İyi bir sonuç elde etmek için sesin kararlı kısımlarından birleştirme yapılmalıdır. Bu nedenle konuşmanın durağan kısımları yani ses sürelerinin sona eren veya neredeyse sıfır olabilen seslendirilmemiş kısımları kullanılmalıdır. Konuşmanın seslenen bölümlerine birleştirme yapmak iyi sonuçlar verir, ancak farklı zamanlarda kaydedildikleri takdirde iki sesli bölüm arasında dengesiz geçişler olabilir. Bu sorunları aşmak için PSOLA gibi bazı algoritmalar kullanılır. PSOLA yöntemi iki konuşma sinyalini alır. Bu sinyalin biri sesli kısmı ile

sona erer ve diğeri ise sesli bir bölümle başlar. Bu seslendirilen parçalar aynı sese karşılık gelmelidir. PSOLA bu iki sinyalin perde değerlerini değiştirerek her iki taraftaki perde değerlerinin eşit olmasını sağlar. Böylece segmentler arasında daha yumuşak bir geçiş sağlanır. Bu tekniğin avantajı, diğer tekniklerle karşılaştırıldığında daha iyi bir çıktı konuşması elde etmektir. Bununla birlikte, bu sistemler çok fazla bellek gerektirir çünkü çok fazla konuşma biriminin önceden kaydedilmesi gerekir. Çıkış konuşması üzerindeki kontrol, bu teknikte diğerlerine kıyasla daha azdır, bu nedenle bu sistemler genellikle bir konuşmacı ile sınırlıdır. Yeni bir konuşmacı eklemek, genellikle yeni konuşmacı için veri tabanının tamamını kaydetmek anlamına gelir. Bell Labs metin seslendirme Sistemi, daha sonra Infovox sürümleri olan CNET PSOLA, hece birleştirme yöntemini kullanan sistemlere örnek olarak verilebilir (Möbius ve ark., 1996; Donovan, 1996).

3. MATERYAL VE YÖNTEM

Bu bölümde EHYMY kullanılarak Kur'ân-ı Kerîm metnini esas alan metin seslendirme uygulamasının yapım aşamaları anlatılmıştır. Geliştirilen metin seslendirme sistemi iki ana bölümden oluşmaktadır. Bu bölümler metin işleme bölümü ve metin seslendirme bölümüdür. Metin işleme bölümü seslendirilecek ayet metnini işleyerek, metnin seslendirme bölümünün anlayabileceği bir şekle dönüştürür. İşlenen ayet metni daha sonra metin seslendirme bölümüne girerek bu metne karşılık gelen konuşmanın üretilmesi sağlanır.

3.1. Metin İşleme Bölümünün Hazırlanması

EHYMY yeni ve özgün bir yöntem olduğu için öncelikle elmas hece mantığının anlaşılması gerekmektedir. Çünkü metin işleme bölümü seslendirilecek ayeti metin seslendirme bölümünün anlayabileceği formata dönüştürürken elmas hece mantığını kullanmaktadır.

3.1.1. Elmas hece nedir?

Önder (2017) kitabında Elmas heceyi şu şekilde tanımlamıştır. Kur'ân-ı Kerîm'i yeni öğrenenler ve orta düzeydeki okuyucular heceleyerek ve hatta kekeleyerek okumaktadır. Kekelerken: okunmayan harf, secavendler, uzatmalar vs. zihinden geçmektedir. Bu heceleme işlemi, Kur'ân-ı Kerîm ve Arapça yazı karakterlerine özel aşağıdaki adımları içermektedir.

1. temel karakterin belirlenmesi,
2. karakterin harekesine/harekelerine ve med (uzatma harfleri; elif, vav, ye) harflerine göre nasıl okunacağı, uzatılıp uzatılmayacağı
3. seslendirilecek bir sonraki temel karakterle arada okunmayan harf ve secavend harfleri olup olmadığı,
4. okunmayan harfler ve secavend harflerini de gözeterek ele alınan kısma bir ses atanması
5. Atanan sesin bir sonraki ses ile ilişkilendirilmesi

Bu ses atama işlemi yapıldığında bir sonraki ses ile ilişkilendirme yapılmadan vücuda gelen hem karakter grubu hem de ses karşılığı “Elmas Hece” olarak isimlendirilmiştir. Normal hece ve Elmas Hece arasındaki fark aşağıda gösterilmiştir.

بِسْمِ اللّٰهِ

Normal Hece - بِسْمِ - مِ اللّٰ - ل - هِ

Elmas Hece - بِ - س - مِ ال - ل - هِ

Bu işlemler sırasında ayet başından sonuna kadar duraksız ilerlendiği varsayılmaktadır. Zaten durulması lazım ve caiz olan noktalar ve bunu gösteren Secavend Harfleri elmas hece sonunu oluşturmaktadır.

Metin işleme bölümünde ayetler bu mantık esas alınarak metin seslendirme bölümüne hazırlanmaktadır. Ancak bunun elle tek tek tüm Kur’ân-ı Kerîm’ e uygulanması mümkündür fakat çok fazla zaman gerektirecektir. Bu nedenle bunun otomatikleştirilmesi gerekmektedir. Bu bağlamda bu işlem için Elmas Hece Algoritması (EHA) adı verilen bir algoritma geliştirilmiştir.

3.1.2. Elmas hece algoritmasının geliştirilmesi

EHA geliştirilmeye başlamadan önce doğru veriler içeren Kur’an-ı Kerim metninin elde edilmesi gerekmektedir. Bu doğrultuda iki farklı teknikle yazılmış Kur’an-ı Kerim metinlerine ulaşmak mümkündür.

- Resmi Osmani denilen yazım tekniği ile yazılmış olan Kur’an-ı Kerim metni. Hz. Osman R.A. zamanında hazırlanan mushaflarda bu teknik kullanılmıştır (Abay, 2015; Tanzil, 2007).
- Ali el-Kārî denilen yazım tekniği ile yazılmış olan Kur’an-ı Kerim metni. Türkiye’de mushaflar bu yazım tekniğine göre yazılmıştır (Abay, 2015; Biçer, 2010).

İki yazım tekniği arasındaki farklar aşağıdaki Şekil 3.1 ve Şekil 3.2’de gösterilmiştir.



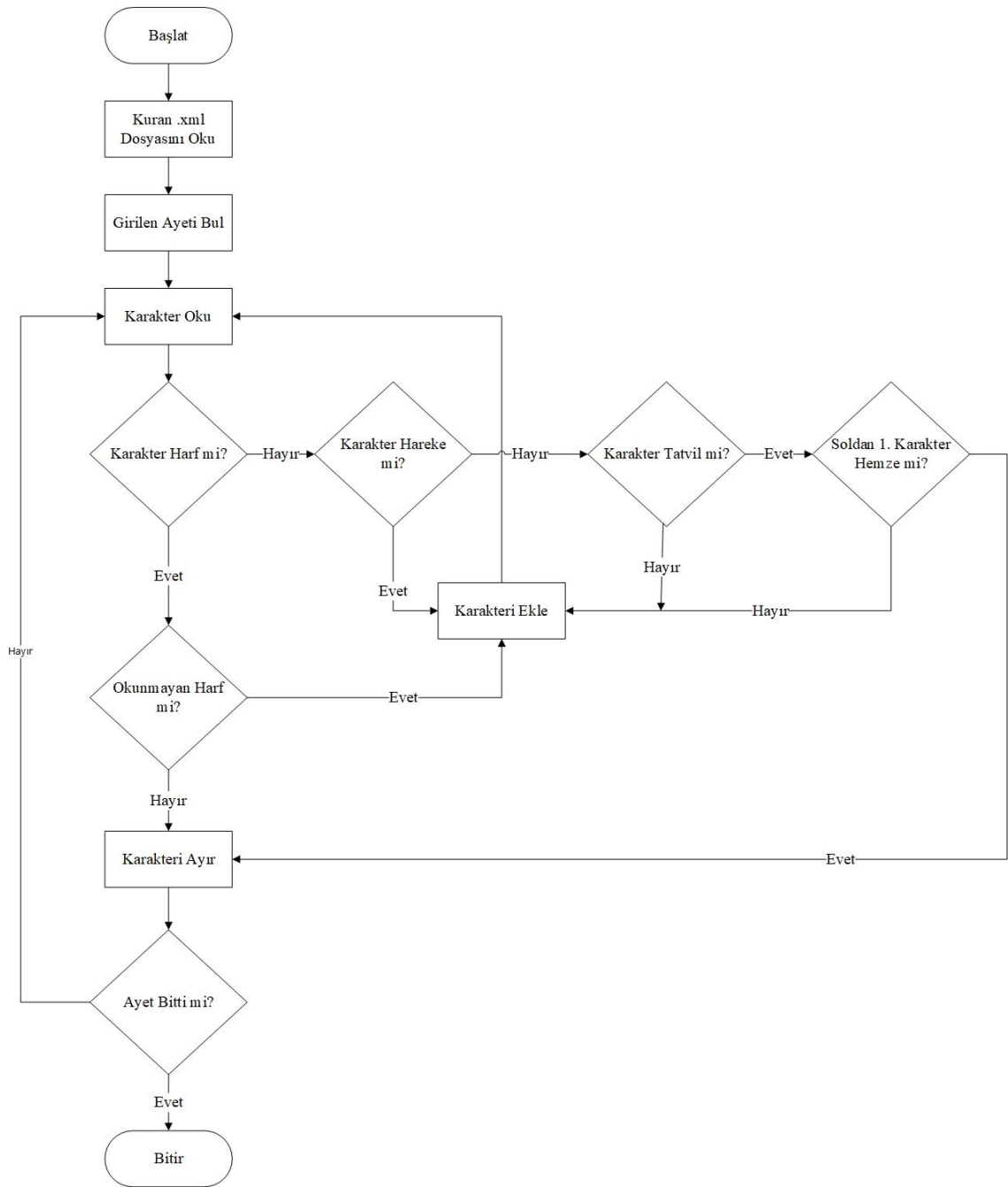
Şekil 3.1 Resmi Osmanî Kur'an-ı Kerim metni



Şekil 3.2 Ali el-Kârî Kur'an-ı Kerim metni

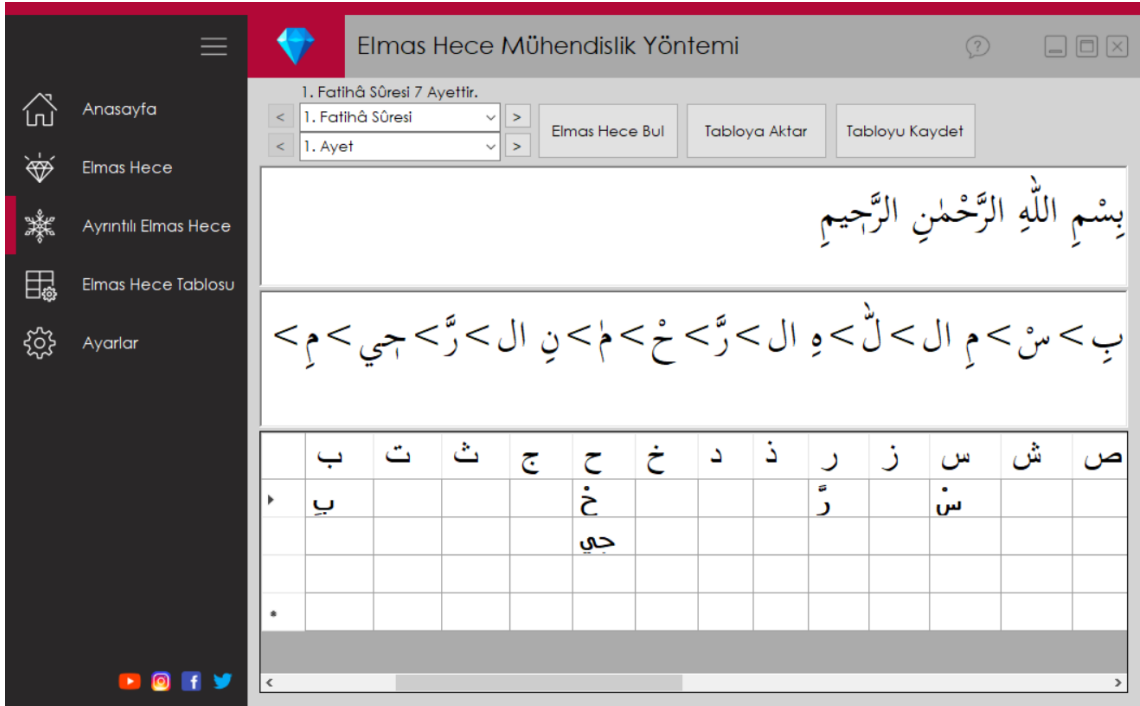
EHA Ali el-Kârî yazım tekniği ile yazılmış olan Kur'an-ı Kerim metni esas alınarak geliştirilmiştir. Bu yazım tekniğinin seçilme nedeni ise; el yazması Kur'an-ı Kerim'ler ve bilgisayar hatlı Kur'an-ı Kerim'ler hep bu tarzı kullanmaktadır. T.C. Diyanet İşleri Başkanlığı resmi internet sitesi başta olmak üzere pek çok internet sitesi bu tarzdaki metinler ile hizmet vermektedir. Bilgisayar programları, mobil uygulamalar, hatim videoları gibi pek çok alanda da Ali el-Kârî tarzı ile yazılan metinler kullanılmaktadır. Ali el-Kârî tarzında ve programlamaya uygun formatta çalışma dosyası <http://www.quraninfo.net/> sitesinden .xml dosya formatında elde edilmiştir.

Gerekli veriler elde edildikten sonra Kur'an-ı Kerim'in tamamı için, alfabedeki temel harflere ait elmas hecelerinin tespiti için EHA geliştirilmiştir. EHA Şekil 3.3'te görüldüğü gibidir.



Şekil 3.3 Elmas Hece Algoritması

EHA Kur'an-ı Kerim'in tamamını kuşatacak şekilde geliştirilmiştir. Bu algoritma Visual C# programlama dili kullanılarak hayata geçirilmiştir. Doğruluğu bu program vasıtasıyla kontrol edilmiştir. Visual Studio ortamında geliştirilen programın arayüzü Şekil 3.4'te gösterilmiştir.

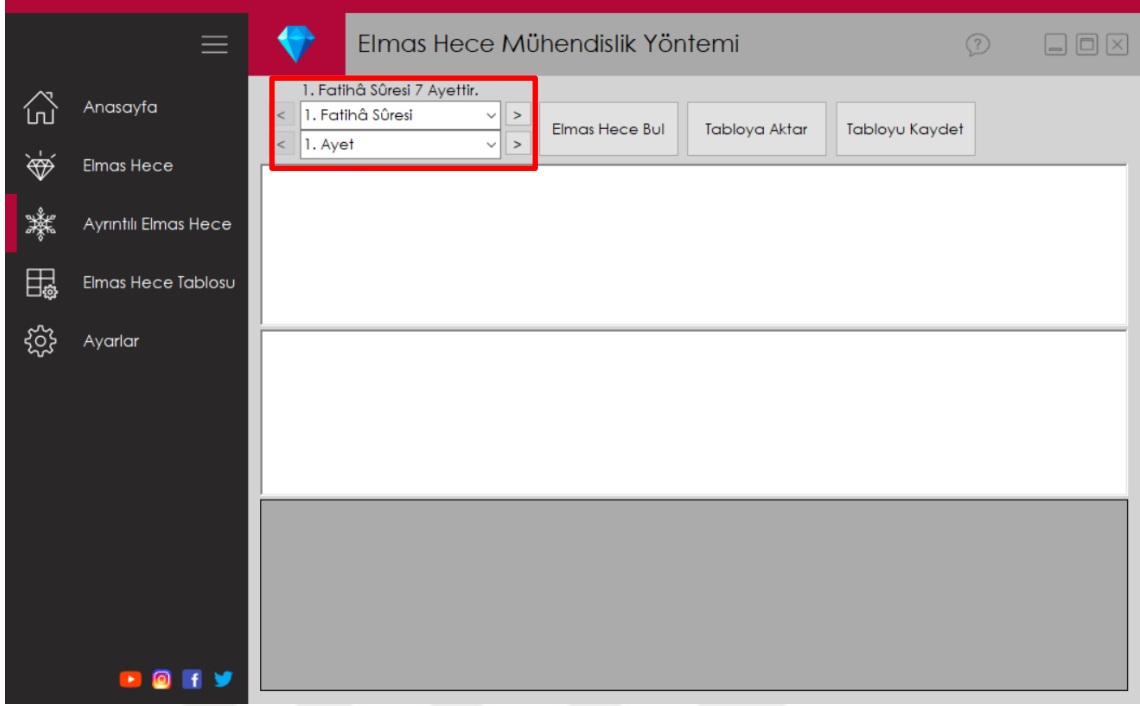


Şekil 3.4 Elmas Hece Mühendislik Yöntemi Program Arayüzü

3.1.3. Elmas hece algoritmasının çalışma prensibi

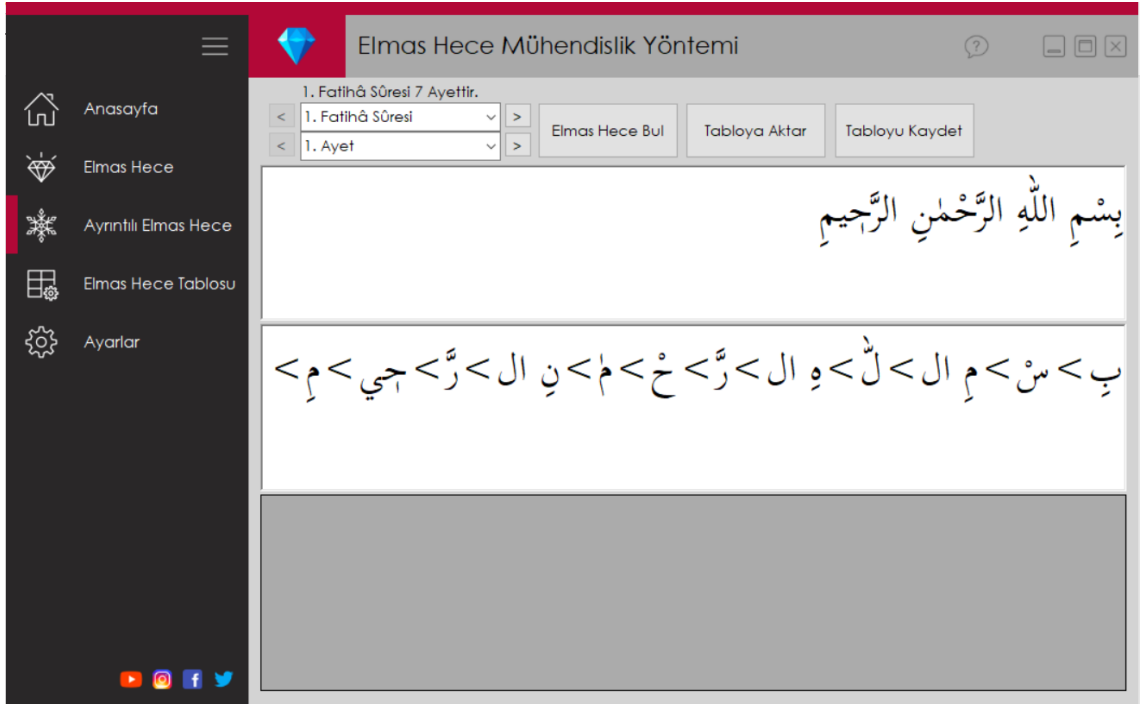
EHA öncelikle xml dosya türündeki metinde istenilen ayeti okur. Okuma işleminden sonra tek tek bu ayetteki karakterleri analiz ederek belirlenen kurallar doğrusunda ayeti elmas hecelere bölerek hafızada tutar. Daha sonra hafızada tuttuğu bu bilgileri metin seslendirme bölümüne göndererek işlemi tamamlar.

EHA' nın çalışma prensibini daha iyi anlamak için Şekil 3.5'te gösterilen geliştirilen programdaki "Ayrıntılı Elmas Hece" sekmesine tıklanır.



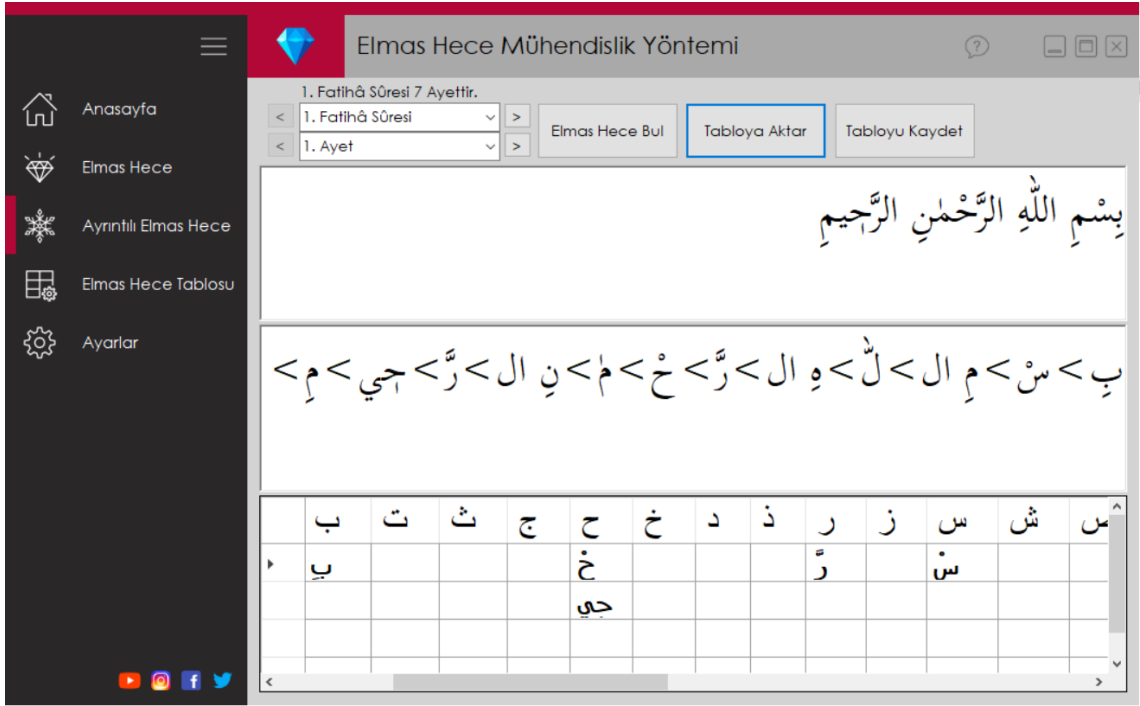
Şekil 3.5 “Ayrıntılı Elmas Hece” sekmesi

Daha sonra açılan ekranda elmas heceleri bulunmak istenen Sure ve Ayet Şekil 3.e’te ki işaretli alandan seçilir. Gerekli seçimler yapıldıktan “Elmas Hece Bul” butonuna tıklanır. Bu işlem sonucunda ilgili ayet ve o ayete ait elmas heceler bulunarak Şekil 3.6’da görüldüğü gibi ekrana yazdırılır.



Şekil 3.6 Seçilen ayet ve o ayetin elmas hecelere ayrılmış hali

Bu şekilde elmas heceler bulunduğundan sonra elde edilen elmas heceler “Tabloya Aktar” butonuna basılarak Şekil 3.7’de ki gibi tabloya aktarılabilir. Daha sonra bu tablo kaydedilmek istenildiği takdirde “Tabloyu Kaydet” butonuna tıklanarak Excel formatında kaydedilir.

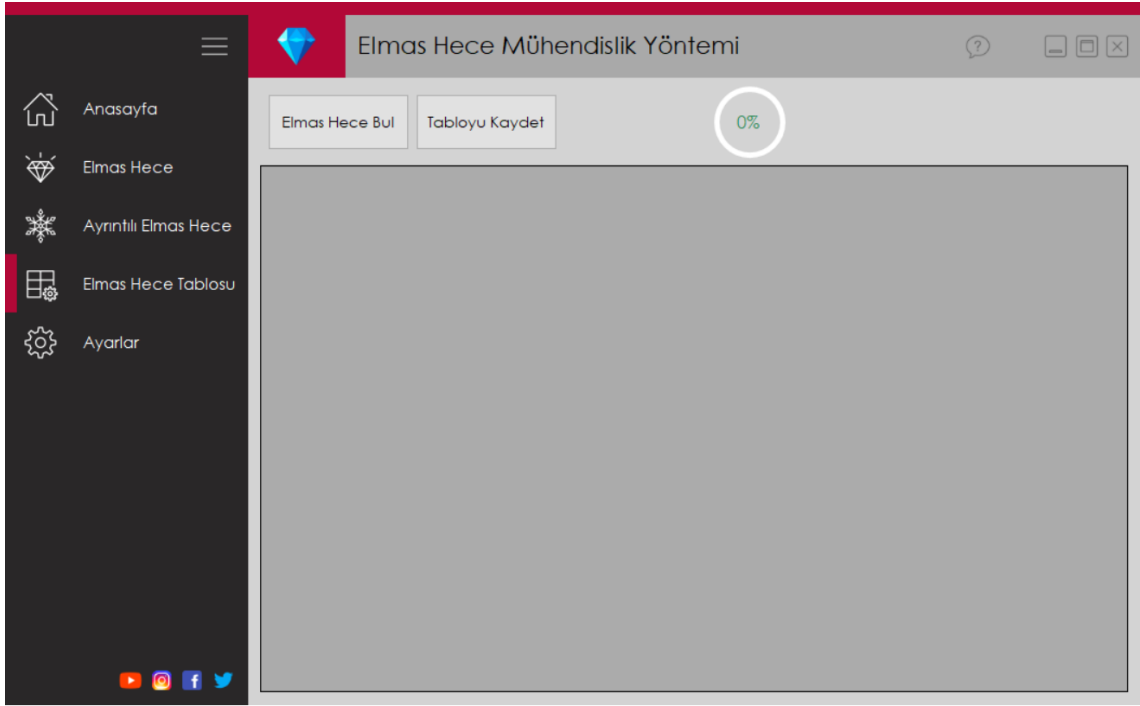


Şekil 3.7 Bulunan hecelerin tabloya aktarılması

Şekil 3.7’de görüldüğü gibi Kur’an-ı Kerim’in tamamı için hangi ayet seçilirse seçilsin EHA sorunsuz bir şekilde çalışmaktadır. EHA’ nın en önemli özelliği tüm Kur’an-ı Kerim’i kuşatacak şekilde geliştirilmesidir.

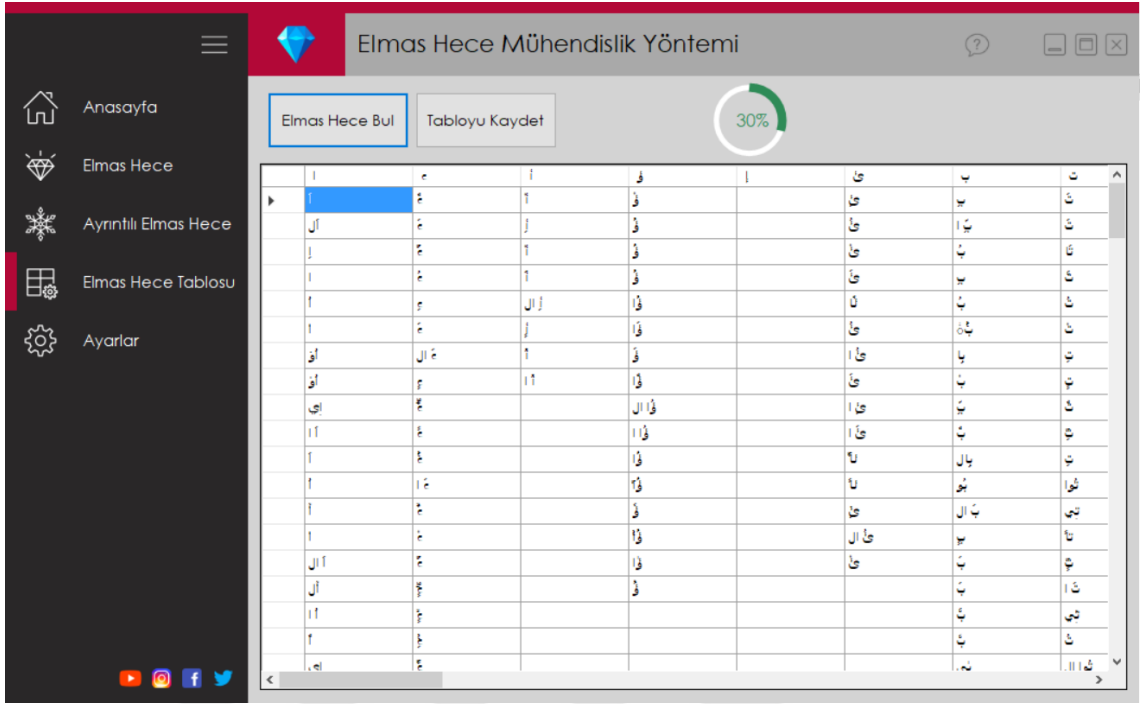
3.1.4. Elmas hece tablosunun elde edilmesi

Kur’an-ı Kerim’in tamamı için, bu program kullanılarak alfabedeki temel harflere ait elmas hecelerin tespiti otomatik yapılmaktadır. Bu işlem için öncelikle program arayüzünde bulunan Şekil 3.8’de gösterilen “Elmas Hece Tablosu” sekmesine girilir.



Şekil 3.8 "Elmas Hece Tablosu" sekmesi

İlgili sekme girildikten sonra Şekil 3.8’te görülen “Elmas Hece Bul” butonuna tıklanır. Butona tıklandığı anda program otomatik olarak tüm Kur’an-ı Kerim de geçen elmas heceleri bulmaya başlamakta ve bulduğu heceleri tabloya aktarmaktadır. Aynı zamanda arayüzde bulunan ProgressBar bu işlemin ne kadarlık kısmının tamamlandığını göstermektedir. Bu işlem Şekil 3.9’ta görülmektedir.



Şekil 3.9 Kur'an-ı Kerim'in tamamı için elmas hecelerinin bulunması

Kur'an-ı Kerim'in tamamı için tüm elmas hecelerinin bulunması yaklaşık 10 saniye sürmektedir. Bu işlem sonunda elde edilen heceler tablo halinde kaydedilmektedir. Bu tabloya Elmas Hece Tablosu (EHT) adı verilmiştir. EHT Kur'an-ı Kerim de geçen bütün elmas heceleri içermektedir. Bu tablo aynı zamanda elmas hece mantığının temel alan bir veri tabanı niteliği taşımaktadır. EHT tasnifi yapılarak metin seslendirme bölümü için gerekli ses veri tabanının oluşturulması gereken altyapının oluşturulacaktır. EHT'nden bir kesit Tablo 1'de gösterilmiştir.

Tablo 3.1 Elmas Hece Tablosu Kesit

ا	ب	ت	ث	ج	ح	خ	...	ك	ل	م	ن	و	ه	ى
اَ	بَ	تَ	ثَ	جَا	حَا	خَا	...	كُ	لُ	مِال	نِال	وُ	هَال	ىَ
آل	بِا	تَ	ثَ	جِ	حِ	خِ	...	كُال	لُ	مُ	نُ	وَا	هَ	يَا
اِ	بُب	تَا	ثَا	جِ	حِ	خَا	...	كِبِ	لِ	مِ	نِ	وَا	هَ	ىِ
ا	بِ	تَ	ثِ	جِ	حِ	خِ	...	كَا	لَا	مَ	نَ	وَا	هَ	ىِ
اُ	بُب	تُ	ثُ	جِ	حِ	خِ	...	كُ	لُ	مِ	نِ	وَا	هَ	ىِ
اِ	بِب	تَ	ثَ	جِ	حِ	خِ	...	كُ	لَاال	مِ	نِال	وَال	هَ	يَا
اُو	ءال	أُ	وَا	بِا	تِ	كَا	لِ	مِ	نِ	وَا	هَ	ىِ
اُو	ءِ	أُ	وَا	بِ	تِ	كُ	لِ	مِ	نِ	وَال	هَ	ىِ
اِى	ءِ	أِ	وَال	بِا	تِ	كُ	لَا	مِ	نِ	وَا	هَ	ىِ
...

3.1.5. Elmas hece tablosunun tasnifi

Her bir harfin elmas hece varyasyonlarının tamamının bulunduğu EHT’ de, her bir harf için elmas heceler karışık halde bulunmaktadır. Tablonun daha anlaşılır olması için Kur’an-ı Kerim’in tamamı için, alfabedeki temel harflere ait elmas hece varyasyonlarının her bir harf için ayrı ayrı tespitinin gerekli olduğu görülmüştür. Bu gereklilik şu şekilde izah edilebilir:

Örneğin herhangi bir ayetteki “ne” sesine karşılık “ن” karakterleri yer almaktadır. Hâlbuki insan zihni Kur’an-ı Kerim okurken aynı “ne” sesini Tablo 3.2’deki varyasyonlara da atamaktadır.

Tablo 3.2 "Ne" Sesine Karşılık Gelen Elmas Heceler

نُ	نَ	نَاال	نَ	نَال	نُ	نُ	نَا	نِ	نِ	نِ	نِ	نِ	نِ	نِ
----	----	-------	----	------	----	----	-----	----	----	----	----	----	----	----

Bu doğrultuda EHT' nin tasnifi hareketler esas alınarak her bir harf için ayrı ayrı yapılmıştır. Bu işlem ile tek bir ses dosyası ile birden çok varyasyonun seslendirilmesi sağlanmaktadır. Bu da büyük miktarda veri kazancı demektir. Tasnifler yapıldıktan sonra seslendirme açısından istisnai durumlar ayrıca bu tablolar üzerinden tespit edilmiştir.

Örneğin: “و” elmas hecesi muttasıl zamir (bitişik zamir) olarak kelime sonuna gelirse uzatılmakta, harici durumlarda uzatılmamaktadır. Başka bir örnek verecek olursak yapısı gereği normalde uzun okunması gereken “ا” elmas hecesi kendinden sonra cezmlı veya şeddeli elmas hece geldiğinde uzun okunmamaktadır. Uzun okunması gereken yerlerde/hecelerde uzatma alameti olacak hareketler konmaktadır.

Örneklere verildiği gibi her bir harf için ayrı ayrı istisnaların da yönetilmesi ile yapılan tasnifler neticesinde ses tablosunun alt yapısı oluşturulmuştur. Tablo 3.3'te nun (ن) harfi için yapılan tasniften örnek bir kesit verilmiştir. Tüm harfler için yapılan tasniflerin tamamı Ek-1 – Ek-28' te verilmiştir.

Tablo 3.3 Tek Bir Harf İçin Örnek Tasnif Tablosu

َ	ِ	ُ	ْ	ٍ	ٌ	ِ	ِ	ِ	ِ	ِ	ِ	ِ	ِ	...
نَ	نِ	نُ	نْ	نٍ	نٌ	نِ	نِ	نِ	نِ	نِ	نِ	نِ	نِ	...
نَ	نِ	نُ	نْ	نٍ	نٌ	نِ	نِ	نِ	نِ	نِ	نِ	نِ	نِ	...
نَا	نِ	نُ	نْ	نٍ	نٌ	نِ	نِ	نِ	نِ	نِ	نِ	نِ	نِ	...
نَ	نِ	نُ	نْ	نٍ	نٌ	نِ	نِ	نِ	نِ	نِ	نِ	نِ	نِ	...
نَا	نِ	نُ	نْ	نٍ	نٌ	نِ	نِ	نِ	نِ	نِ	نِ	نِ	نِ	...
نَا	نِ	نُ	نْ	نٍ	نٌ	نِ	نِ	نِ	نِ	نِ	نِ	نِ	نِ	...
...

3.2. Metin Seslendirme Bölümünün Hazırlanması

Bu bölümde önceki bölümde hazırlanan metnin seslendirilmesi yapılacaktır. Seslendirmenin yapılabilmesi için öncelikle Elmas Hece Ses Tablosu (EHST) oluşturulmuştur. Daha sonra elde edilen bu tablodaki ifadeler karşılık gelecek seslerin kayıtları alınmıştır. Alınan ses kayıtları işlenerek bir ses veri tabanının oluşturulmuştur. Ses veri tabanı oluşturulduktan sonra seslerle hecelerin senkron bir şekilde seslendirilmesi için bir algoritma geliştirilmiştir.

3.2.1. Elmas hece ses tablosunun oluşturulması

Tasnif işlemleri yapıp elmas hece tasnif tablosu elde edildikten sonra EHST oluşturmak üzere harflerin her bir hareke için tasnifi elde edilen elmas heceler ses tekrarları atılarak sadeleştirilmiştir. Bu sınıflama ile EHST oluşturulmuştur. Tüm bu işlemler sonucunda elde edilen ses alfabetinde bulunan ifadeler karşılık gelecek ses kayıtları alınarak işlenmiştir. Daha sonra geliştirilen programda elmas hece gösterim üzerine elmas hece seslendirmesi yapılmıştır.

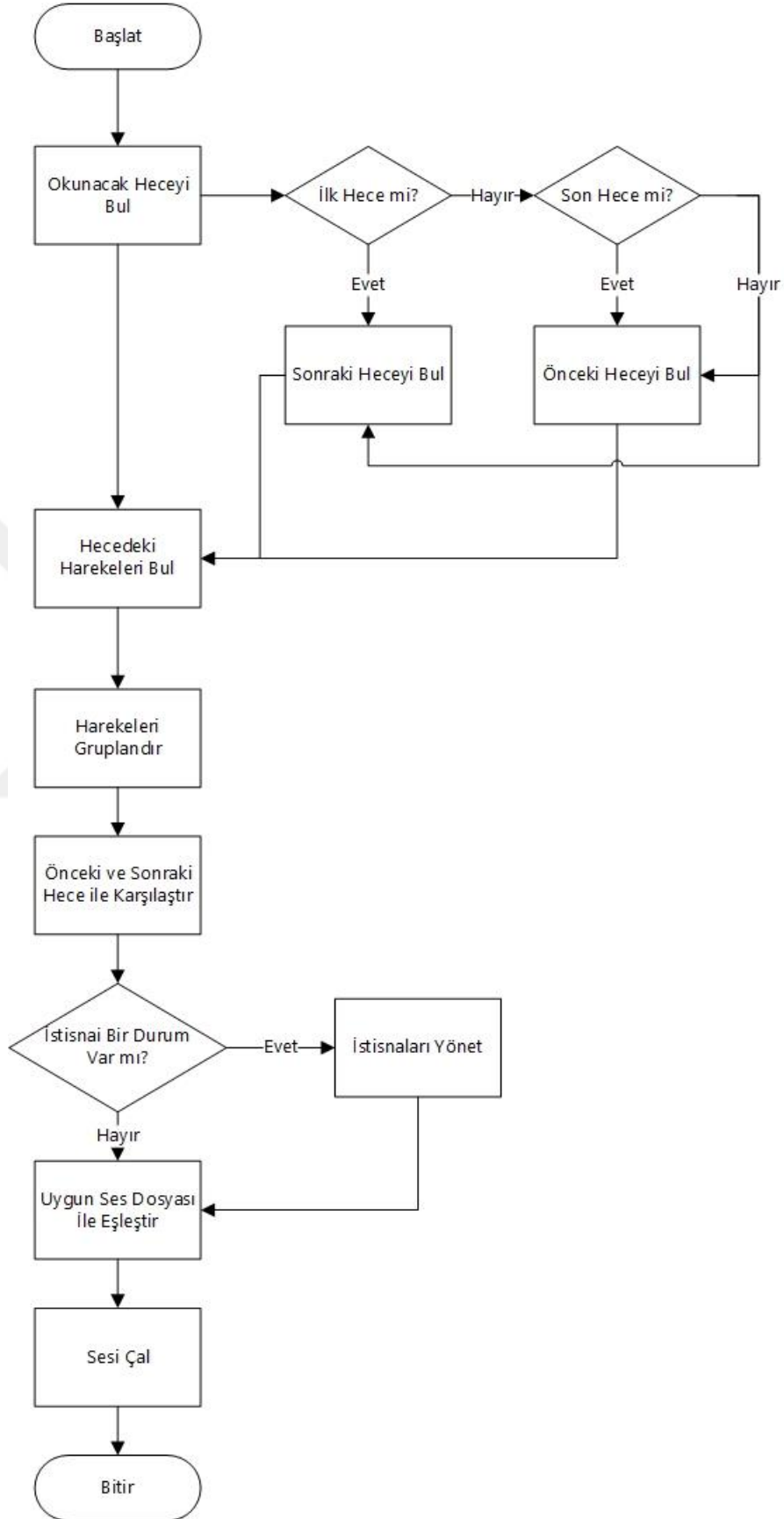
Tablo 3.4'te EHST'ndan bir kesit verilmiştir. Tablo 3.4'te görüldüğü gibi üstteki hareketler esas alınarak yapılan tasnifte satırlarda görülen boşluklar harfin o varyasyonunun olmadığını göstermektedir. EHST'nun tamamı Ek-29'de verilmiştir.

Tablo 3.4 Elmas Hece Ses Tablosu Kesit

	اَ	اِ	اُ	آَ	آِ	آُ	أَ	أِ	أُ	إَ	إِ	إُ	...
ا	اَ	اِ	اُ	آَ	آِ	آُ	أَ	أِ	أُ	إَ	إِ	إُ	...
ب	بَ	بِ	بُ	بَٓ	بِٓ	بُٓ	بَٔ	بِٔ	بُٔ	بَٕ	بِٕ	بُٕ	...
ت	تَ	تِ	تُ	تَٓ	تِٓ	تُٓ	تَٔ	تِٔ	تُٔ	تَٕ	تِٕ	تُٕ	...
ث	ثَ	ثِ	ثُ	ثَٓ	ثِٓ	ثُٓ	ثَٔ	ثِٔ	ثُٔ	ثَٕ	ثِٕ	ثُٕ	...
ج	جَ	جِ	جُ	جَٓ	جِٓ	جُٓ	جَٔ	جِٔ	جُٔ	جَٕ	جِٕ	جُٕ	...
ح	حَ	حِ	حُ	حَٓ	حِٓ	حُٓ	حَٔ	حِٔ	حُٔ	حَٕ	حِٕ	حُٕ	...
...
ل	لَ	لِ	لُ	لَٓ	لِٓ	لُٓ	لَٔ	لِٔ	لُٔ	لَٕ	لِٕ	لُٕ	...
م	مَ	مِ	مُ	مَٓ	مِٓ	مُٓ	مَٔ	مِٔ	مُٔ	مَٕ	مِٕ	مُٕ	...
ن	نَ	نِ	نُ	نَٓ	نِٓ	نُٓ	نَٔ	نِٔ	نُٔ	نَٕ	نِٕ	نُٕ	...
و	وَ	وِ	وُ	وَٓ	وِٓ	وُٓ	ؤَ	ؤِ	ؤُ	وَٕ	وِٕ	وُٕ	...
ه	هَ	هِ	هُ	هَٓ	هِٓ	هُٓ	هَٔ	هِٔ	هُٔ	هَٕ	هِٕ	هُٕ	...
ی	یَ	یِ	یُ	یَٓ	یِٓ	یُٓ	یَٔ	یِٔ	یُٔ	یَٕ	یِٕ	یُٕ	...

3.2.2. Seslendirme algoritması

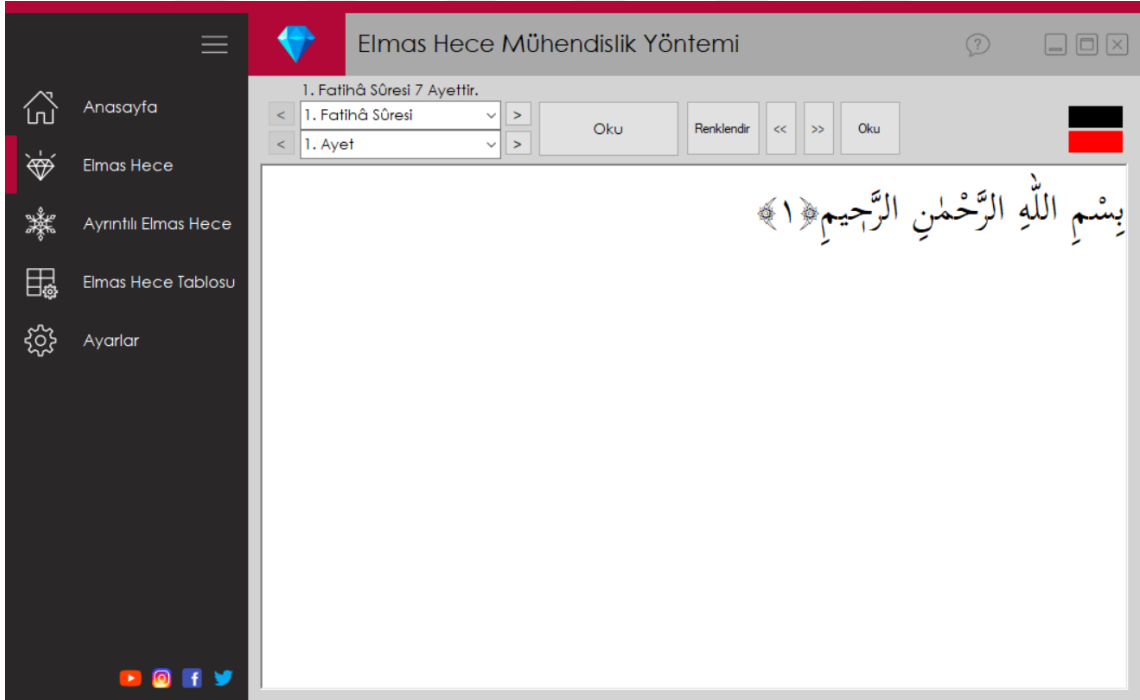
Kur'an-ı Kerim de geçen tüm heceler ve bu hecelere karşılık gelen sesler eşleştirilerek seslendire bilmek için seslendirme algoritması geliştirilmiştir. Bu algoritmanın da EHA'nda olduğu gibi tüm Kur'an-ı Kerim'i kapsayacak şekilde hatasız ve stabil olması gerekmektedir. Tüm bu hususlar dikkate alınarak seslendirmenin hatasız ve stabil olabilmesi için geliştirilen algoritma Şekil 3.10'da verilmiştir.



Şekil 3.10 Seslendirme algoritması

Algoritmanın çalışma mantığını kısaca anlatacak olursak öncelikle seslendirilecek olan hece algoritma tarafından tespit edilerek ayetin, ilk hecesi mi yoksa son hecesi mi olduğuna bakılmıştır. Buna bakılmaktaki amaç seslendirmesi yapılacak olan hecenin kendinden önce ya da kendinden sonra gelen heceye ve bu hecelerdeki harekelere göre farklı okunma ihtimalinin olmasıdır. Eğer ayetin ilk hecesiyse kendinden önce başka hece olmadığı için sadece sonraki heceye bakarak uygun ses dosyasının kararı verilecektir. Fakat hece ayetin son hecesiyse de kendinden sonra başka bir hece olmadığı için önceki heceye bakarak uygun ses dosyasının kararı verilecektir. Hece ayetin ilk veya son hecesi değilse hem seslendirilecek heceden önceki heceye hem de sonraki heceye bakılarak uygun ses dosyasına karar verilecektir. Tüm bu işlemler gerçekleştirildikten sonra Kur'an-ı Kerim'in tamamı "Seslendirme Algoritması" kullanılarak sorunsuz bir şekilde seslendirilmiştir.

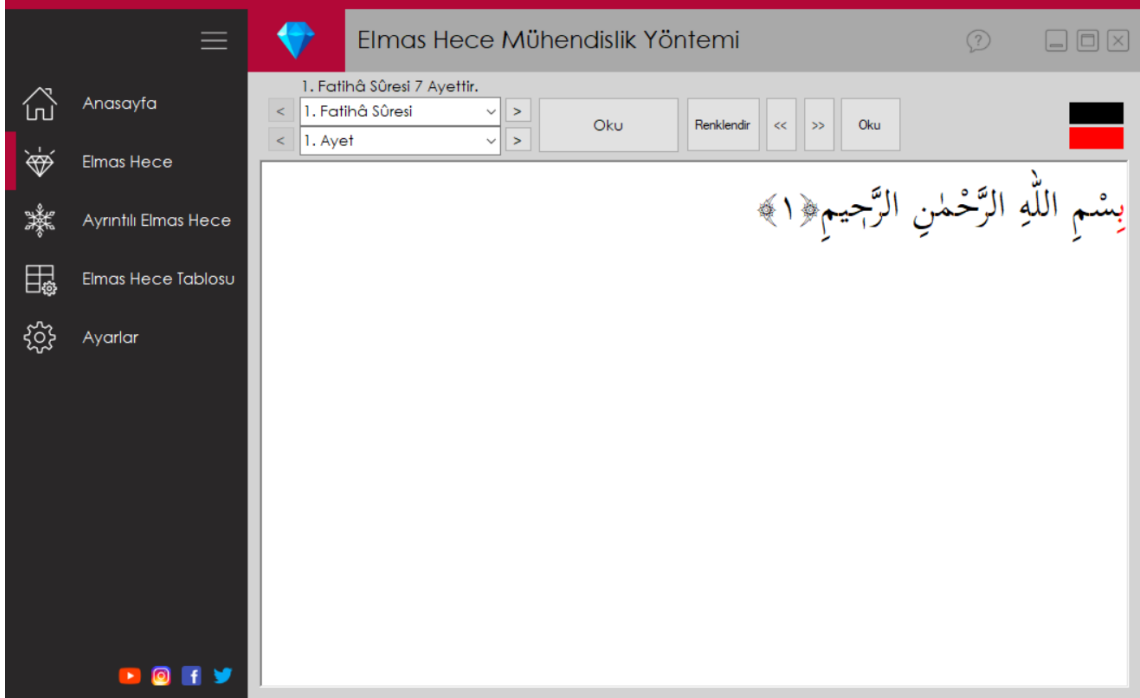
Bu yöntemle seslendirme yapılırken seslendirilen hece senkron olarak program tarafından edilmektedir. Bu daha iyi anlaşılması için görsellerle anlatılmıştır.



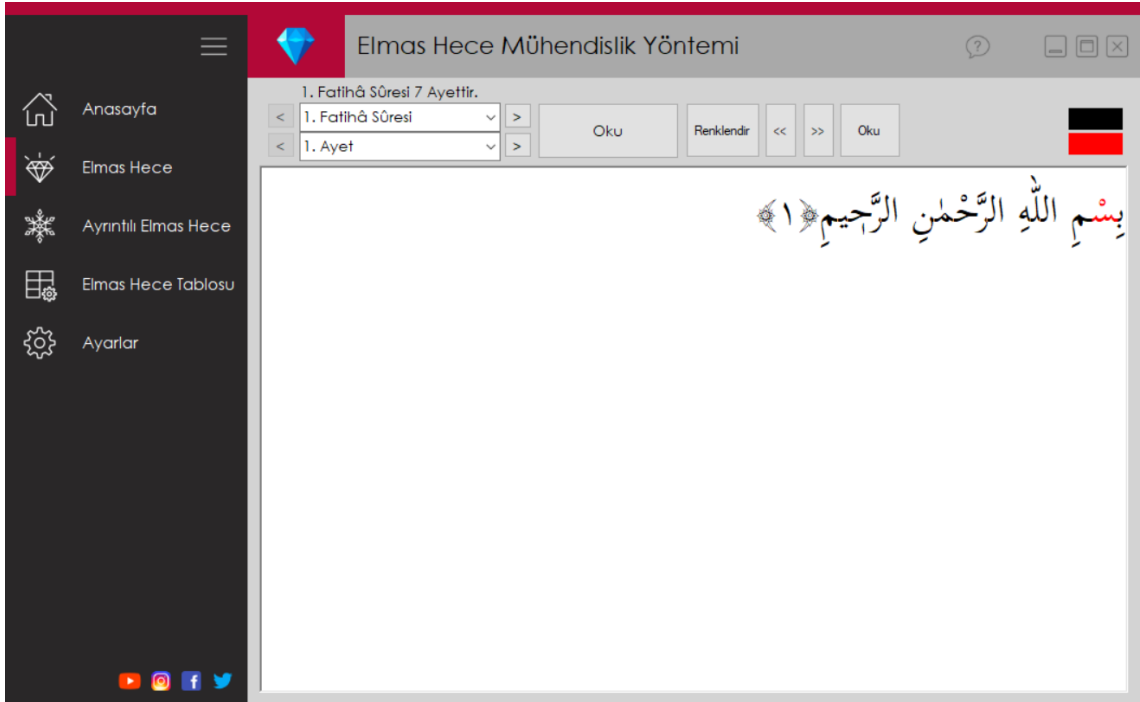
Şekil 3.11 Seslendirmenin yapılacağı sekmenin arayüzü

Şekil 3.11' de görülen "Elmas Hece" sekmesine tıklanarak seslendirmenin yapılacağı sekme açılır. Daha sonra seslendirilmek istenen ayet seçilir. Daha sonra "Oku" butonuna basılarak seslendirme sağlanmaktadır.

Seslendirme aynı zamanda program arayüzünde senkron olarak takip edilmektedir. Bu işlem Şekil 3.12 ve Şekil 3.13 ile gösterilmiştir.



Şekil 3.12 “ب” sesine karşılık gelen ses çalınırken program çıktısı



Şekil 3.13 “س” sesine karşılık gelen ses çalınırken program çıktısı

4. BULGULAR

Her bir harfe ait elmas hece varyasyonları EHA kullanılarak tespit edilerek EHT aktarılmıştır. Kur'an-ı Kerim'in tamamında 3141 farklı elmas hece olduğu görülmüştür. Daha sonra bu EHT üzerinden ses tekrarları çıkarılıp EHST elde edilmiştir. Tüm bu çalışmalar sonucunda elde edilen EHST'nda toplam ses sayısının 569 olduğu tespit edilmiştir. Bu da 569 Elmas Hecenin ses kaydı alınarak Kur'an-ı Kerim'in tamamının seslendirilmesinin mümkün olduğu anlamına gelmektedir.

Seslendirme işlemi yapılırken program incelendiğinde bazı istisnai durumların olduğu tespit edilmiştir. Bu istisnaların nedeni ise Kur'an-ı Kerim'in kendi has bir okunma şeklinin olmasıdır. Eğer bu istisnalara dikkat edilmezse okunuştaki bozukluklar nedeni ile anlamda da bozukluk ortaya çıkabilmektedir. Bu istisnai durumlar şunlardır;

- Lafzatullah' daki (اللّٰه) Lam eğer kendinden önceki hecenin harekesi üstün veya ötre ise kalın esre ise ince okunmalıdır.
- Sat (ص) harfi bazı ayetlerde sin (س) gibi okunmalıdır.
- Seslendirilecek hecede kasır varsa normalde uzun okunması gereken hece kısa okunmalıdır.
- İki üstünlü bir hecede küçük nun varsa bu hece okunurken harfin üstünlü hali ve ni sesi eklenerek okunur. Örneğin;

اَللّٰه hecesi okunurken ن - ل diye okunmalıdır.

bu durum iki esre iki ötreli heceler içinde geçerlidir

- Ra harfi (مَجْرِيهَا) kelimesinde ince okunmalıdır.

Bu istisnai durumlar için özel çözümler üretilerek programın daha stabil hale gelmesi sağlanmıştır.

5. SONUÇ

Mevcut metin seslendirme sistemleri popüler dünya dilleri üzerine odaklanmışken Arapça gibi diller geri planda kalmıştır. Arapçayı destekleyen metin seslendirme sistemlerinin en büyük eksiği ise Arapçanın kendine has okunuş şeklini göz ardı etmeleridir. Bu durumda özellikle İslam aleminin kutsal kitabı Kur'ân-ı Kerim' in yanlış okunmasına ve anlamının bozulmasına neden olmaktadır. Yaptığımız çalışmada EHYMY kullanılarak geliştirilen metin seslendirme sistemi bu sorunları çözüm sunan özgün bir sistemdir. Kur'ân-ı Kerim' in doğru okunup doğru anlaşılabilmesi için bu sistem kullanılarak geliştirilecek yazılı, görsel ve işitsel ürünlerle başlangıç düzeyindeki okuyucuya kılavuzluk edilebileceği düşünülmektedir. Aynı zamanda EHYMY etkili kişisel gelişim araçları, hatim setleri gibi birçok ürüne de altyapı sağlayacaktır.

Günümüzdeki hatim programlarında ki en büyük sorun yüksek veri miktarlarına ihtiyaç duymalarıdır. EHYMY ile geliştirilen metin seslendirme sistemi 569 ses dosyası ile Kur'ân-ı Kerim' in tamamının seslendirilmesi sağlanarak çok büyük miktarda veri kazancı sağlanmıştır. Bu sistemin en büyük avantajlarından biride çok büyük veri miktarına ihtiyaç duyulmadan yazılı ve görsel ürünlerin tasarlanmasına alt yapı sağlar nitelikte olmasıdır.

Aynı zamanda geliştirilen bu metin seslendirme sistemi küçük değişiklikler ile benzer dillere de uygulanabileceği görülmüştür.

Metin seslendirme sistemindeki okuyuş doğal dil işleme yöntemleri iyileştirilebilir. Ayrıca sistem konuşma tanıma gibi yöntemler ilave edilerek daha dinamik ve kullanıcı ile daha etkileşimli bir sistem haline getirilmesi mümkündür.

6. KAYNAKLAR

- Abay, M., 2015. Mushaf İmlasında Ali el-Kārî Tarzı Meselesi. *Usul*, 23(1), pp. 7-44.
- Allen, J., Hunnicutt, S. ve Klatt, D., 1987. *From Text To Speech: The MITalk System*. Cambridge: Cambridge University.
- Anonim, 2018. MBROLA Project Homepage. <http://tcts.fpms.ac.be/synthesis/mbrola.html>, (29.05.2019)
- Biçer, S., 2010. Quraninfo Homepage. <http://www.quraninfo.net>, (11.12.2018)
- Black, A. ve Taylor, P., 1997. *Festival Speech Synthesis System: System Documentation*, basım yeri bilinmiyor: Human Communication Research Centre Technical Report HCRC/TR-83.
- Deller Jr, J. R., Proakis, J. G. ve Hansen, J. L., 1987. *Discrete-Time Processing of Speech Signals*. New Jersey: Prentice Hall.
- Donovan, R., 1996. *Trainable Speech Synthesis*, PhD Thesis. England: Cambridge University Engineering Department.
- Dutoit, T. ve Leich, H., 1993. MBR-PSOLA: Text-to-Speech Synthesis Based on an MBE Re-Synthesis of the Segments Database. *Speech Communication*, Issue 13, 435-440.
- Dutoit, T., Pagel, V., Pierret, N., Bataille, F. ve Vrecken, O., 1996. The MBROLA Project: Towards a Set of High Quality Speech Synthesizers Free of Use for Non Commercial Purposes. *Proceedings of ICSLP*, 96 (3).
- Flanagan, J., 1972. *Speech Analysis, Synthesis and Perception*. Berlin-Heidelberg-New York: Springer-Verlag.
- Flanagan, J., Allen, J. B. ve Hasegawa-Johnson, M. A., 2009. *Speech Analysis Synthesis & Perception*. Berlin, Springer.
- Hallahan, W., 1996. *DECTalk Software: Text-to-Speech Technology and Implementation*. Dijital Technical Journal.
- Huang, X., Acero, A., Adcock, J., Hon, H., Goldsmith, J., Liu, J. ve Plumpe, M., 1996. Whistler: A Trainable Text-to-Speech System. *Proceeding of ICSLP*, 96 (4).
- Klatt, D., 1980. Software for a Cascade/Parallel Format Synthesizer. *Journal of the Acoustical Society of America*, Issue 67, 971-995.
- Klatt, D., 1987. Review of Text-to-Speech Conversion For English. *Journal of the Acoustical Society of America*, 82 (3), 737-793.
- Kröger, B., 1992. Minimal Rules for Articulatory Speech Synthesis. *Proceedings of EUSIP*, 92 (1), 331-334.
- Ljungqvist, M., Lindström, A. ve Guastafson, K., 1994. A New System for Text-to-Speech and Its Applications to Swedish. *ICSLP*, 94 (4), 1779-1782.
- Möbius, B., Schroeter, J., Santen, J., Sproat, R. ve Olive, J., 1996. *Recent Advances in Multilingual Text-to-Speech Synthesis*. Berlin, DAGA.
- Mürsel, Ö., 2017. *Kuran Eğitiminde Elmas Hece Mühendislik Yöntemi*. (Yayın Aşamasında)
- O'Saughnessy, D., 1987. *Speech Communication – Human and Machine*. Boston: Addison-Wesley.
- Rabiner, L. ve Juang, B. H., 1993. *Fundamentals of Speech Recognition*. New Jersey: Prentice Hall.
- Rahim, M., Goodyear, C., Klejin, B., Schroeter, J. ve Sondhi, M., 1993. On the Use of Neural Networks in Articulatory Speech Synthesis. *Journal of the Acoustical Society of America*, 93 (2), 1109-1221.

Schroeder, M., 1993. A Brief History of Synthetic Speech. *Speech Communication*, Issue 13, 231-237.

Tanzil, 2007. Tanzil Homepage. <http://tanzil.net>, (11.12.2018)



8. ÖZGEÇMİŞ

Kişisel Bilgiler

Adı Soyadı : Ahmet Köroğlu
Uyruğu : Türkiye Cumhuriyeti
Doğum Tarihi ve Yeri : 02.03.1992 - Tonya
Yabancı Dili : İngilizce
Telefon : 543 214 1035
E-mail : 02.03.1992@hotmail.com.tr

Eğitim Bilgileri

Derece	Alan	Eğitim Birimi	Mezuniyet
Yüksek Lisans	Mekatronik Mühendisliği Ana Bilim Dalı	Gaziosmanpaşa Üniversitesi, Fen bilimleri Enstitüsü	2019
Lisans	Mekatronik Mühendisliği	Gaziosmanpaşa Üniversitesi	2015
Lise	Fen Bilimleri	Tonya Anadolu Lisesi	2010

İş Deneyimi

Yer	Görev	Yıl
Almix Profesyonel Çiftlik Ekipmanları	Tasarım Mühendisi	2015-2016
İsemmak Hayvancılık ve Çiftlik Ekipmanları	İmalat Mühendisi	2016-2017