T.C.
KIRŞEHİR AHİ EVRAN ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
İLERİ TEKNOLOJİLER ANABİLİM DALI

TÜRKÇE METİNDEN KONUŞMA SENTEZLEMEYE YÖNELİK YAPILAN ÇALIŞMALARIN İNCELENMESİ

GAMZE YILMAZ

YÜKSEK LİSANS TEZİ

KIRŞEHİR / 2019

T.C.
KIRŞEHİR AHİ EVRAN ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
İLERİ TEKNOLOJİLER ANABİLİM DALI

TÜRKÇE METİNDEN KONUŞMA SENTEZLEMEYE YÖNELİK YAPILAN ÇALIŞMALARIN İNCELENMESİ

Gamze YILMAZ

YÜKSEK LİSANS TEZİ

DANIŞMAN
Doç. Dr. Osman ÖRNEK

II. DANIŞMAN
Doç. Dr. Mustafa YAĞCI

KIRŞEHİR / 2019

Bu çalışma 23.10.19 tarihinde aşağıdaki jüri tarafından İleri Teknolojiler Anabilim Dalında Yüksek Lisans tezi olarak kabul edilmiştir.

Tez Jürisi

Doç. Dr. Osman ÖRNEK
Kırşehir Ahi Evran Üniversitesi
Mühendislik-Mimarlık Fakültesi

Doç. Dr. Mustafa YAĞCI
Kırşehir Ahi Evran Üniversitesi
Mühendislik-Mimarlık Fakültesi

Prof. Dr. Raşit Köker
Sakarya Uygulamalı Bilimler Üniversitesi
Teknoloji Fakültesi

Doç. Dr. Levent URTEKİN
Kırşehir Ahi Evran Üniversitesi
Mühendislik-Mimarlık Fakültesi

Doç. Dr. Hakan SEPET
Kırşehir Ahi Evran Üniversitesi
Mühendislik-Mimarlık Fakültesi

TEZ BİLDİRİMİ

Tez içindeki bütün bilgilerin etik davranış ve akademik kurallar çerçevesinde elde edilerek sunulduğunu, ayrıca tez yazım kurallarına uygun olarak hazırlanan bu çalışmada bana ait olmayan her türlü ifade bilginin kaynağına eksiksiz atıf yapıldığını bildiririm.

Gamze YILMAZ

20.04.2016 tarihli Resmi Gazete'de yayımlanan Lisansüstü Eğitim ve Öğretim Yönetmeliğinin 9/2 ve 22/2 maddeleri gereğince; Bu Lisansüstü teze, Kırşehir Ahi Evran Üniversitesi'nin abonesi olduğu intihal yazılım programı kullanılarak Fen Bilimleri Enstitüsü'nün belirlemiş olduğu ölçütlere uygun rapor alınmıştır.

ÖNSÖZ

Yüksek Lisans ders ve tez sürecinde gösterdiği yardımlarla bana her zaman destek olan değerli danışmanım Doç. Dr. Osman ÖRNEK'e, tez konumu seçmemde yardımcı olan ve ileriki süreçte de bana sık sık yol gösteren değerli hocam ve danışmanım Doç. Dr. Mustafa YAĞCI'ya, tezim bitme aşamasındayken sorduğum sorulara içtenlikle cevap verip bana destek olan Prof. Dr. Raşit KÖKER hocama yürekten teşekkür ederim.

Bu süreçte her zaman yanımda hissettiğim aileme ve dostlarıma minnettarım. Tezimi, bundan çok daha iyisini hak eden anneme ve babama ithaf ederim.

Ekim, 2019 Gamze YILMAZ

İÇİNDEKİLER

Sayfa No

ŞEKİL LİSTESİ

Sayfa No

TABLO LİSTESİ

Sayfa No

KISALTMA LİSTESİ

Kısaltmalar	Açıklama
ATN	**:** Augmented Transition Network
DDİ	**:** Doğal Dil İşleme
DMA	**:** Direct Memory Access
DRT	**:** Diagnostic Rhyme Test
HPGS	**:** Head-Driven Phrase Structure Grammar
LFG	**:** Lexical Functional Grammar
MKS	**:** Metinden Konuşma Sentezleme
OLA	**:** OverLap Add
PSOLA	**:** Pitch Synchronous Overlap and Add
SAT	**:** Kompozisyon Notlama
SOLA	**:** Synchronous Overlap-Add
TD-PSOLA	**:** Time Domain- Pitch Synchronous Overlap and Add
TTS	**:** Text to Speech
URL	**:** Uniform Resource Locator
WSOLA	**:** Waveform Similarity OverLap Add
XML	**:** Extensible Markup Language

ÖZET

YÜKSEK LİSANS TEZİ

TÜRKÇE METİNDEN KONUŞMA SENTEZLEMEYE YÖNELİK YAPILAN ÇALIŞMALARIN İNCELENMESİ

Gamze YILMAZ

Kırşehir Ahi Evran Üniversitesi
Fen Bilimleri Enstitüsü
İleri Teknolojiler Anabilim Dalı

Danışman: Doç. Dr. Osman ÖRNEK
II. Danışman: Doç. Dr. Mustafa YAĞCI

Metinden konuşma sentezleme çalışmaları dijital ortamdaki her türlü bilginin sesli olarak iletilmesini sağlamaktadır. Doğal dil işleme çalışmaları ise doğal dilin yapısını daha iyi anlamak ve bilgisayar-insan etkileşimini kolaylaştırmak amacıyla yapılmaktadır. Bu tez çalışmasında, öncelikle Türkçenin biçimbilimsel analizine yönelik bilgiler örneklerle birlikte verilmiştir. Ayrıca doğal dil işleme ile metinden konuşma sentezlemeye yönelik bir çalışmanın aşamaları incelenmiş ve literatürde kullanılan farklı yaklaşımlar sonuçları ile birlikte değerlendirilmiştir. Yapılan literatür taramasına göre belirli kriterlerle araştırmaya dahil edilen 19 çalışmanın incelemesi yapılmıştır. Çalışmacıların 12'si; yani yaklaşık %63'ü, vurgu ve tonlamalarda eksiklikler olduğunu, bu durumun iyileştirilmesinin; doğal dil işleme bilimi ile mümkün olduğunu savunmaktadır.

Ekim 2019, 85 Sayfa.

Anahtar Kelimeler: Metinden Konuşma Sentezleme, Doğal Dil İşleme, Literatür Analizi

ABSTRACT

M.Sc. THESIS

ANALYSIS OF STUDIES IN TURKISH TEXT TO SPEECH SYNTHESIS

Gamze YILMAZ

Kirsehir Ahi Evran University
Graduate School of Sciences and Engineering
Advanced Technologies Department

Supervisor: Doç. Dr. Osman ÖRNEK
II. Supervisor: Doç. Dr. Mustafa YAĞCI

Text-to-speech synthesis studies enables all kinds of information to be transmitted by voice. Natural language processing studies, on the other hand, performed in order to better understand the structure of natural language and to facilitate computer-human interaction. In this thesis study, first of all, information about morphological analysis of Turkish is given with examples. In addition, the stages of a study on natural language processing and text to speech synthesis were examined and the different approaches used in the literature were evaluated together with the results. According to the literature review, 19 studies included in the study with certain criteria were examined. 12 of the researchers; that is, about 63% stated that there were deficiencies in the emphasis and intonation; argues that it is possible with natural language processing.

October 2019, 85 Pages.

Keywords: Synthesizing Text-To-Speech, Natural Language Processing, Literature Analysis

1. GİRİŞ

Metinden konuşma sentezleme sistemleri, dijital ortamdaki metinlerin sesli olarak kullanıcıya iletilmesini sağlamaktadır. Doğal dil işleme çalışmaları ise, doğal dilin yapısını anlayarak insan-bilgisayar etkileşimini kolaylaştırmaya yarayan bir bilim dalıdır. Metin sentezleme işleminde seslendirilen metnin doğru ve anlaşılır bir biçimde elde edilmesi, çalışmadan elde edilen verim açısından önemlidir. Bunun için ilk olarak dilin biçimbilimsel analizinin iyi yapılması gerekmektedir.

1.1. Amaç

Doğal dil işleme (DDİ) çalışmaları, doğal dilin yapısını daha iyi anlamak ve bilgisayar-insan etkileşimini kolaylaştırmak amacıyla yapılmaktadır. Metinden konuşma sentezleme (MKS) çalışmaları ise bireylerin erişmek istedikleri her türlü bilginin onlara sesli olarak iletilmesini sağlamaktadır.

MKS uygulamalarının en verimli şekli, insanın doğal konuşmasındaki vurgu ve tonlamalara en yakın konuşmanın üretilebilmesidir. Doğal dillerin analizinin yapılması ve bunun yardımcı bir teknolojiye öğretilebilmesi ise DDİ teknikleriyle mümkündür.

Bu tez çalışmasında, Türkçenin biçimbilimsel analizine yönelik bilgiler sunulmuş ve daha sonra DDİ biliminin amacı, karşılaşılan zorlukları ve aşamaları verilmiştir. Daha sonra MKS'ye yönelik bir uygulamanın aşamaları incelenmiş ve literatürde kullanılan başka yaklaşımlar incelenip sonuçları değerlendirilmiştir. Bu tezin amacı MKS sistemlerinde insan konuşmasına yakın bir konuşma şeklinin nasıl sağlanabileceğine dair bir inceleme sunmaktır.

1.2. Önem

Günümüzde bilişim ve iletişim teknolojilerinin artmasıyla birlikte bilgiye erişim ve bilginin üretimi konularında büyük kolaylıklar sağlanmaktadır. Toplumdaki bireylerin bu bilgilere erişim sağlayabilmesi ülkenin gelişiminin devam etmesi açısından da son derece önemlidir.

Bilgiye erişim, etkili ve doğru iletişimle mümkündür. Toplumdaki her bireyin yazılı ve dijital olan her ortamdaki bilgiye erişiminin sağlanabilmesi bilgi çağının gereklerindendir.

Sesli okunan metnin, insan sesine ve günlük konuşma dilindeki vurgu ve tonlamalara uygun olması uygulamanın kalitesi açısından oldukça önemlidir. Bu ise yapay zekâ ve DDİ

algoritmaları ile mümkündür. Bu çalışma, MKS sistemlerine DDİ tekniklerinin entegre edilmesi fikrini içermesi bakımından önemlidir.

2. GENEL KISIMLAR

2.1. Türkçenin Özellikleri

2.1.1. Türkçe Dil Ailesi ve Özellikleri

Türkçe, ek ve köklerden oluşan sözcüklerden meydana gelmektedir. Köklere eklenen eklerden dolayı, sondan eklemeli diller grubuna girmektedir. Bu doğrultuda, MKS çalışmalarında kök ve ek ayırt etme basamağı bulunmaktadır. Aynı şekilde Türkçe için yapılan DDİ çalışmalarında ise sözcüğü ek ve köklerine ayıran çeşitli algoritmalar kullanılmaktadır.

"Türkçe, Altay dil ailesine mensup bir dildir. Bu ailenin diğer üyeleri Moğolca, Mançu-Tunguzca ve Korece'dir" (Sel, 2013). Türkçe, sözcüklerin ekler vasıtasıyla üretildiği, sondan eklemeli bir dil olduğu için bir sözcük kökünden birden fazla sözcük türetilerek sözcük dağarcığı genişletilebilmektedir.

Türkçe birçok lehçeye sahip bir dildir ve dünya dilleri içinde en eski tarihe sahip dillerden biridir. Türkçeyi ana dili olarak konuşan yaklaşık 60 milyon insan bulunmaktadır. Bu insanlar Türkiye'de, Ortadoğu'da ve bazı Avrupa ülkelerinde yaşamaktadır. Türk dilleri ailesine mensup yaklaşık 40 dil vardır fakat bu dillerin bazıları günümüzde kullanılmamaktadır. Konuşulan Türk dilleri ise yaklaşık 165-200 milyon kişi tarafında anadili olarak konuşulmaktadır (Oflazer, 2016).

En fazla kullanılan Türk dili; %43 ile Türkçedir. Türkçeyi, Azerice ve Özbekçe takip etmektedir. Tablo 2.1'de bazı Türk dillerinin konuşulma oranları verilmiştir.

Tablo 2.1: Bazı Türk Dillerinin Konuşulma Oranları

Dil	**%**
Türkiye Türkçesi	43
Azerice	15
Özbekçe	14
Kazakça	7
Uygurca	6
Türkmence	2
Diğerleri	13

2.1.2. Türkçe Dilinin Özellikleri ve Biçimbilimsel Yapısı

Tanzimat döneminde dönemin aydınlarının Türkçe sözcükleri kullanma gayreti ve Arap alfabesinde yaptıkları yenileşme çabalarıyla başlayan sadeleşme dönemi, cumhuriyetin ilanı

ile birlikte hız kazanmıştır. Cumhuriyetle birlikte çağdaş Türkçenin temelleri atılmıştır (Sel, 2013). 1 Kasım 1928'de Latin alfabesinin kabulüyle birlikte tarama, türetme, derleme yollarıyla Türkçe sözcük sayısı büyük oranda artmıştır.

Türkçedeki sözcükler 30 bin kadar kök halinde bulunan sözcüğe eklerin eklenmesiyle oluşturulur. Sözcük dağarcığı, coğrafi ve tarihsel nedenlerden dolayı zamanla Arapça, Farsça, Ermenice gibi dillerden etkilenmiştir. Son 50-60 sene içinde ise İngilizceden etkilenmiştir (Oflazer, 2016).

Türkçe, biçimbilimsel açıdan bitişken bir dildir. Birleşik isimler, bitişik yazılan farklı sözcüklerden oluşur fakat birleşik ismin kendi anlamı, içerdiği sözcüklerin anlamlarından bağımsızdır (örneğin; "kasımpatı" sözcüğü kendini oluşturan sözcüklerin anlamından tamamen farklıdır ve bir çiçek ismidir).

Türkçe sözcüklerin bir diğer özelliği de yapım eklerinin çok fazla kullanılmasıdır. Tek bir fiil kökünden neredeyse 1,5 milyon farklı sözcük üretilebilir. Bu da Türkçenin biçimbilim yapısının üretim gücünü gösterir. Fakat sözcük çeşitliliğinin bu denli fazla olması DDİ çalışmalarında bazı sorunlara da yol açmaktadır.

2.1.2.1. Ses, Hece ve Ek Kavramları

Bir ses, ünlü (sesli) harfler ve ünsüz (sessiz) harfler olmak üzere ikiye ayrılır. Ünlüler tek başına seslendirilebilir fakat ünsüzler seslendirilirken yanında bir ünlü ifadeye ihtiyaç duyarlar.

Tablo 2.2: Ünlü ve Ünsüz Harfler

Ünlüler	**Ünsüzler**
a, e, ı, i, o, ö, u, ü	b, c, ç, d, f, g, ğ, h, j, k, l, m, n, p, r, s, ş, t, v, y, z

Tablo 2.2'de ünlü ve ünsüz harfler gösterilmiştir. Gösterilen her bir ifade bir sestir ve yazılı ifadelerde her sesin karşılığı olarak bir alfabetik simge bulunmaktadır. Alfabedeki ses sayısı ne kadar fazla olursa yazılı ifadede bunların gösterilişi o kadar karmaşık olmaktadır. Bu yüzden kümelendirme yolu kullanılmaktadır. "Fonemler, anlam ayırıcı özelliği bulunan ses kümeleridir" (Sel, 2016). Anlam ayırıcı özelliğinin bulunması sebebiyle fonemler belli bir dile özgüdür.

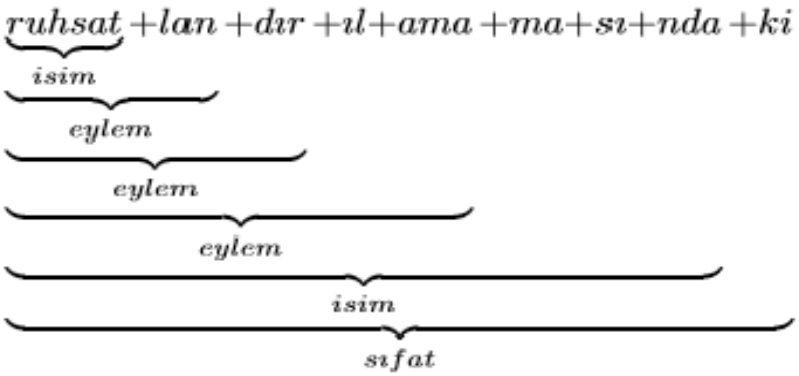
“Ağzın bir hareketiyle bir defada söylenebilen sözcük parçalarına hece denir” (Arık, 2011). Türkçede ünlü harfler tek başına seslendirilebildikleri için hece özelliği gösterebilirler fakat ünsüz harfler tek başına seslendirilemedikleri için yanına bir ünlü harf almadıkları sürece hece olarak kullanılamazlar.

Türkçede her hece içinde mutlaka bir sesli harf bulunmak zorundadır. Tablo 2.3’te Türkçe hecelerin genel yapısı verilmiştir. (A → sessiz harf, B → sesli harf)

Tablo 2.3: Türkçede Hecelerin Genel Yapısı

Hece	**Örnek**
B	a, e, o, ü, …
BA	ab, az, ek, öç, …
AB	ba, za, ce, tı, …
ABA	gel, yak, tır, …
BAA	ırk, alt, …
AAB	Bre
ABAA	yurt, Türk, sert, …

Sözcükler cümle içinde kullanıldıklarında bazı yapım ve çekim ekleri alırlar. Kök sözcüğe eklendiğinde sözcüğün anlamını değiştiren eklere yapım eki, değiştirmeyenlere ise çekim eki denir. Yapım ekleri sözcüğün anlamını değiştirebildiği için sözcük türetebilme açısından sıklıkla kullanılmaktadır. Çok fazla yapım ekine sahip bir sözcük görüntü açısından bazen çok karmaşık olabilir. Örneğin; “ruhsatlandırılamamasındaki” sözcüğü yapı ve görüntü olarak çok karmaşıktır. Bu sözcükte 5 tane yapım eki vardır ve isim kökü ile başlayıp 5 türetme sonrasında yeni bir sözcüğe dönüşmüştür.



Şekil 2.1: Karmaşık Bir Türkçe Sözcükteki Türetmeler [4]

2.1.2.2. Türkçe Sözcüklerde Vurgu ve Cümlelerde Öğe Sırası

"Sözcüklerde, kuvvetli söylenen hece üzerindeki baskıya 'vurgu' denir" (Şentürk ve Adalı, 2016). Konuşma sırasında sözcüklerin tekdüze heceler şeklinde çıkmamasının nedeni, her hecenin üzerine aynı kuvvetle basılmaması; yani vurgudur. Hecelere farklı kuvvetlerle basılması anlam farklılıklarını beraberinde getirir.

Türkçede vurgulu söylenen kısım genellikle ilk veya son hecede bulunur (Büyük bir oranda son hecededir). Ortadaki hecelerde vurgu olmaz. Hala tam olarak benimsenemeyen yabancı kökenli sözcüklerde ise genellikle vurgu ilk hecede olur (posta, banka, vb.). Hitap sözcüklerinde de sözcüğün kökeninin ne olduğu fark etmeksizin vurgu ilk hecededir.

Türkçe cümlelerde doğal öğe sırası Özne – Nesne - Yüklem şeklindedir. Diğer belirteç öğeler cümlenin herhangi bir yerinde bulunabilirler. Özne – Nesne – Yüklem sıralaması da kendi içinde yer değiştirerek 6 farklı şekilde kullanılabilirler. Aşağıdaki örneklerde ana eylem Ali'nin eve gitmesidir. Sözcüklerdeki sıra değişiklikleri cümlelerde anlam farklılıkları olmasına yol açmaktadır.

- Ali eve gitti.
- Eve Ali gitti. (Giden Ali idi, başkası değil.)
- Gitti Ali eve. (Gitmemesi gerekiyordu.)
- Gitti eve Ali. (Zaten gitmesi bekleniyordu.)
- Ali gitti eve. (Başkası da gidebilirdi.)
- Eve gitti Ali. (Başka bir yere de gidebilirdi.)

Yukarıdaki örneklerde de görüldüğü gibi, Türkçede sözcüklerin cümle içindeki yerleri cümlenin anlamını tamamen değiştirmektedir. Bu konu, DDİ biliminin çalışma alanlarından biridir.

2.2. Yapay Zekâ ve Doğal Dil İşleme

2.2.1. Zekâ ve Yapay Zekâ

Akıl ve zekâ kavramları birbirinden farklı kavramlardır. Akıl; bir olguyu düşünüp idrak etmek ve belli kavramlar arasında yapılabilen muhakeme gücüdür. Sürekli olarak gelişebilir ve artabilir. "Akıl, makine, bilgisayar, yazılım veya başka bir yolla taklit edilemez" (Elmas,2011). Zekâ ise; gerçeklerden sonuç çıkarma yeteneğidir. İnsanlar belli bir zekâya sahip olarak doğarlar. Zekâ; çalışarak, öğrenerek, edinilen bilgi birikimleriyle yükseltilebilir.

Dünyanın en karmaşık makinesi insan beynidir. Sayısal bir işlemi sonuca ulaştırma, bir olayı idrak etme gibi görevleri çok kısa bir sürede yapabilir. Bilgisayarlar ise sayısal bir işlemi sonuca ulaştırma konusunda insanlardan daha hızlı olsalar da, bir olayı idrak etme veya tecrübelerle edinilmiş bilgileri kullanabilme konusunda insanların çok gerisindedir.

Yapay zekâ ise, en basit tabiriyle, zekânın taklit edilmiş halidir. "Bir bilgisayarın ya da makinenin, insana özgü olan akıl yürütme, anlam çıkartma, genelleme ve geçmiş deneyimlerden öğrenme gibi yüksek zihinsel süreçlere ilişkin görevleri yerine getirme yeteneği olarak tanımlanmaktadır" (Nabiyev, 2012).

2.2.1.1. Yapay Zekâ ve Doğal Zekânın Karşılaştırılması

- Doğal zekâya sahip olan insanlar mevcut bilgilerini zamanla unutabilir veya doğal zekâ zamanla değişebilir. Yapay zekâ daha kalıcıdır.
- Doğal zekâ söz konusuyken, uzmanlığın başka bir insana aktarılması uzun bir zaman alabilir. Yapay zekâ kolayca geniş kitlelere aktarılabilir ve kopyalanabilir.
- Yapay zekânın maliyeti, kaliteli bir personelin eğitilip yetiştirilmesinden çok daha düşüktür.
- Doğal zekâya sahip olan insanoğlu kararsız ve değişken olabilir. Yapay zekâ bir bilgisayar sistemi olduğundan tutarlıdır.
- Doğal zekânın tekrar üretimi güçtür. Yapay zekâ belgelenebilir.
- Yapay zekâ insana bağlı olduğundan, günümüz teknolojisi itibari ile yaratıcı ve yenilikçi değildir. Doğal zekâ yaratıcı ve doğurgandır.

- Yapay zekâ sistemleri sembolik girdilerle çalışır. Doğal zekâ, insana öğrendiği deneyimleri kullanma ve bunlardan faydalanma yeteneği sağlar.
- Doğal zekâya sahip bir insan muhakeme gücünü ve tecrübelerini karşılaştığı problemlerde kullanabilir. Bu da doğal zekâ avantajlarının en önemlisidir.

2.2.2. Türkçe Doğal Dil İşleme

2.2.2.1. Doğal Dilin Temel Özellikleri

İnsanlık tarihindeki bazı gelişmeler, insanlığın gelişiminde önemli atılımlar sağlamıştır. Dil ise bu atılımların en önemlilerinden biridir. Tarih boyunca, insanlar için dil önemli bir kavram olmuştur. İnsanlar diller sayesinde birbirleri ile iletişim kurabilmişlerdir. Diller zaman içerisinde gelişmiş ve gelişmeye devam etmektedir. Dillere yeni sözcükler eklenebilir veya dillerdeki var olan sözcükler zamanla unutulabilir. Hatta zaman ilerledikçe dillerin cümle kurma biçimi de değişebilir. Fakat dillerin kuralları ve yapıları genel anlamda değişmemektedir.

Türkçede sözcükler zaman içerisinde değişim göstermiştir. "Hatta Osmanlı Türkçesinde, Türkçe sözcüklerin, dilin toplam söz varlığının yarısına kadar düştüğü söylenebilir" (Adalı, 2016). 1932 yılında başlatılan Dil Devrimi ile Türkçeden yabancı sözcükler atılıp sadeleştirme çalışmaları hız kazanmıştır.

Dilin yapısı, insanların düşünebilme kabiliyetini de etkiler. Bir dilin sözcük çeşitliliği o dilin gelişmişliğinin göstergesidir. Burada çeşitlilikten kasıt; sözcük sayısı değildir. Dilin yapısına bakarak o dilin gelişmişlik düzeyine karar vermek daha mümkündür. Türkçe, sondan eklemeli ve bitişken bir dil olduğu için bir kökten çok fazla sözcük üretilebilmektedir. Bu açıdan bakıldığında Türkçenin ve diğer bitişken dillerin gelişmişlik bakımından diğer dillerden daha üstün olduğu söylenebilir.

Dil bilimciler, dilin kurallarını ortaya koyabilmek için aşağıdaki dört konu üzerinde çalışırlar:

- Ses Bilimi: Dünya üzerinde kullanılmış ve kullanılmakta olan bütün konuşma seslerini bütün özellikleriyle inceleyip ele alan bilim dalıdır (Topbaş ve Kopkallı, 1994).
- Biçim Bilimi: Bir dilin en küçük anlam ve biçim birimleri cümlenin yapı taşlarını oluşturur. Bu yapıtaşlarını inceleyen bilim dalına biçim bilimi denir.
- Söz Dizimi: Cümle yapısı araştırmasıdır.
- Anlam Bilimi: Bir dili anlam bakımından inceler.

Dil bilimciler, bu konular sayesinde dilin zaman içerisindeki değişimini incelerler. Dil bilimcilerin elde ettikleri sonuçlara göre ise doğal dil alanında çalışan bilişimciler, bilgisayar ile dili işlerler.

2.2.2.2. Doğal Dil İşleme Nedir?

Dillerin gelişmişlik düzeylerine göre, o dili kullanan toplumların sanat, bilim, kültür alanında ürettikleri eserler arasında bir bağ vardır. Özellikle bilişimdeki gelişmelerden sonra, bilgisayarla konuşma amacında olan bilim insanları, dil bilimi çalışmalarını hızlandırmışlardır. DDİ, dillerin bilgisayar yardımıyla işlenmesi üzerinde çalışmaktadır (Adalı, 2016). Daha önceleri insanlarla bilgisayarların etkileşimi için doğal dillerin kullanılması gayesi ile başlatılmış olsa da daha sonra bilgisayarlı dil bilimi adı altında toplanmıştır. Bilgisayarların konuşarak insanlara bir veriyi aktarmasına konuşma, insanların konuşarak bilgisayara bir veri girişi yapmasına ise konuşmayı anlama adı verilmektedir.

"DDİ; ana işlevi bir dili çözümleme, anlama, yorumlama ve üretme olan bilgisayar sistemlerinin tasarımını ve gerçekleştirilmesini konu alan bir mühendislik alanıdır" (Nabiyev, 2012). DDİ, birçok farklı alanda geliştirilen teknolojileri ve yöntemleri bir araya getirir. Bilişsel psikoloji, bilgisayar destekli veya kuramsal dilbilim, yapay zekâ, dil çözümleme bunlardan bazılarıdır.

Daha önceki yıllarda DDİ, yapay zekânın bir alt dalı olarak görülmekteydi; fakat kazanılan başarılar neticesinde günümüzde bilgisayar bilimlerinin konularından biri olarak kabul edilmektedir.

DDİ, dilbilim kuramlarına deney ortamı hazırlayarak daha kapsamlı ve çabuk sınanmalarını sağlamaktadır (Nabiyev, 2012). Bu açıdan, dilbilimciler ve bilgisayar bilimcileri DDİ çalışmaları için ortak hareket etmeye başlamışlardır.
DDİ iki temel problemi çözmeye çalışmaktadır:

- İnsan-bilgisayar etkileşimini doğal dil ile yapabilme
- İnsan-insan etkileşimini kolaylaştırma ve zenginleştirme

DDİ, bu iki problemi çözmek için bilgisayar sistemleri geliştiren bilgisayar mühendisliği alt alanıdır. Yapay zekâ, yapay öğrenme, biçimsel diller ve otomatlar kuramı, istatistik, dilbilim, yazılım mühendisliği alanlarından kavram ve teknikler kullanmaktadır.

2000 yılından önce bütün dünyada daha çok kural tabanlı DDİ çalışmaları yapılmıştır. Sözcük çözümlemesi, doğru sözcük çözümlerinin seçilmesi, Lexical Functional Grammar (LFG) / Head-Driven Phrase Structure Grammar (HPSG) / Augmented Transition Network (ATN) formalizmaları ile gramer modelleri geliştirilmesi, kural tabanlı kısıtlı Türkçe-İngilizce çeviri sistemleri geliştirilmesi, ilk istatistik temelli çalışmalar ve ilk konuşma işleme bu çalışmalardan bazılardır.

2000 yılından sonra ise daha çok kaynak oluşturma ve konuşma işleme çalışmalarına ağırlık verilmiştir. Bu çalışmalarla yapay öğrenme yöntemleri kullanılmaya başlanmıştır.

DDİ temel olarak dil çözümlemesini, dilin üretimini ve edinimini otomatik olarak yapabilmeyi amaçlar. Dil çözümlemesi anlama veya işleme, dil üretimi ise bunun tam tersi olarak adlandırılabilir.

Dil çözümlemesinde girdi dil (konuşma veya yazı), çıktı ise gösterim (söyleşi)'dir. Dil üretiminde ise girdi gösterim, çıktı dildir (Eryiğit ve Torunoğlu Selamet, 2017).

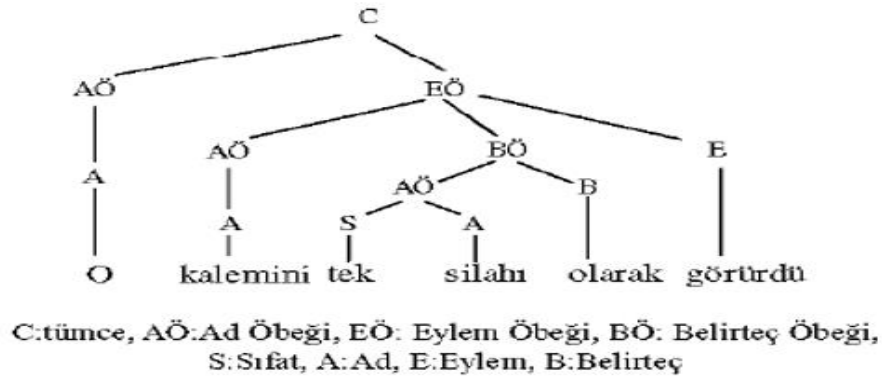
Şekil 2.2: Dilbilimsel Gösterim Düzeyleri (Oflazer, 2018)

2.2.2.3. Doğal Dil İşleme Çalışmalarının Başlıca Elemanları

DDİ biliminin temel amacı, doğal dil kurallarını bilgisayara öğretmektir. Bunun için kullanılacak bir sözlüğe ve bu sözlüğü bilgisayarın kullanabilmesi için çeşitli algoritmalara ihtiyaç vardır. Bu amaç doğrultusunda; sistemde genel olarak beş temel eleman bulunur. Bu elemanlar şu şekilde sıralanabilir:

- Ayrıştırıcı: DDİ çalışmalarının en temel elemanıdır. Verilen cümlenin sözdizimsel analizini yapıp ayrıştırıcı ağacını oluşturur. Bu alanda en çok kullanılan yaklaşım,

cümlenin öbeklere bölünmesidir. Cümle ilk olarak ad öbeği ve eylem öbeğine bölünür. Daha sonra bunlar kendi içlerinde daha küçük öbeklere bölünürler. Ayrıştırma işlemi sonucunda, görevleri belli olan bu öbekler anlamsal analiz işleminden geçirilerek bir çıkış cümlesi oluşturulur.



C:tümce, AÖ:Ad Öbeği, EÖ: Eylem Öbeği, BÖ: Belirteç Öbeği,
S:Sıfat, A:Ad, E:Eylem, B:Belirteç

Şekil 2.3: Ayrıştırma Ağacı Örneği (Delibaş, 2008)

- Sözlük: Kullanılan program tarafından bilinmesi gereken tüm sözcükleri içerisinde barındırır. Ayrıştırıcı elemanı ile birlikte sözdizimsel analiz yapar. DDİ çalışmalarında kullanılan sözlükler, çalışmanın içerdiği sözcüklerin köklerini ve bunların anlamlarını içerir.

  Sözlük üzerinde dört bölümden oluşan işlemler gerçekleşir. *Jeton seçme* bölümünde sözcükler ve noktalama işaretlerine jetonlar konularak cümle bölümlere ayrılır. *Köksel analiz* bölümünde sözcük hecelerine ayrılarak sözcüğün kökü bulunur. Türkçe sondan eklemeli bir dil olduğu için sözcükte kök bulma aşaması çok önemlidir. *Sözlüğe bakma* bölümünde bulunan sözcük kökünün sözlükteki anlamına bakılır. *Hata dönüşümü* bölümü ise sözcük kökünün anlamının sözlükte bulunmaması durumudur. Eğer kök anlamı sözlükte yoksa bir hata durumu söz konusudur. Bu hatanın kaynakları; sözcüğün yanlış hecelenmesi, özel isimlerin tespitinin hatası veya yanlış yazılan sözcüklerdir.

  DDİ çalışmalarında kullanılan sözlükler çok fazla sözcük içerdiği için büyük boyuttadır ve karmaşık bir yapı içerir. Bu yüzden sözlük oluşturmak büyük bir zaman ve yatırım gerektirmektedir.
- Bilgi Tabanı: Genel bilgi tabanı ve görev bağımlı bilgi tabanı olmak üzere ikiye ayrılır (Delibaş, 2008). Anlayıcı ile birlikte cümlenin anlamını tespit eder.

- Anlayıcı: Ayrıştırıcı ağacının bilgi tabanındaki anlamsal karşılığını bulan elemandır. Cümlenin anlamını bilgi tabanı ile birlikte tespit eder. Girdi cümlesine uygun cevabı hazırlar.
- Üretici: Belli cümleler ve sözcükler için depolanmış bazı kayıtların kullanıcıya gösterilmesini sağlar.

2.2.2.4. Doğal Dil İşleme Uygulamaları ve İlgi Alanları

DDİ günümüzde birçok bilim insanı tarafından çalışılan bir mühendislik alanıdır. DDİ'nin uygulama alanlarından bazıları aşağıda sıralanmıştır:

- Soru cevaplama
- Yazı ve konuşma çevirisi
- Bir dizi yazıyı özetleme
- Bir toplantıda konuşulanları takip edip özetleme
- E-postalara otomatik yanıtlar gönderme
- Verilen komutları yerine getirme
- Yazım ve söz dizim hatalarını düzeltme
- Kompozisyon notlama (SAT)
- Okuduğunu anlama sınavları için soru üretme
- Devamlı olarak yazı okuyup bilgi tabanı güncelleme
- Etkileşimli olarak dil öğretme
- Konuşma, duyma, görme kısıtlı kişilere yardımcı olma
- Bir metnin içinde geçen bir sözcüğü başka bir sözcükle değiştirme
- Basılı bir metni bilgisayar ortamına aktarıp okuma ve metni seslendirme
- Bir metni anlama ve metnin ana fikrini çıkarma
- Konuşmayı yazıya dökme

2.2.2.5. Türkçe Doğal Dil İşleme Yazılım Zinciri Evreleri

Türkçe DDİ yazılım zinciri; normalizasyon, sözcük analizi, varlık ismi tanıma ve cümle analizinden oluşur. Ayrıca kullanılacak modellerin eğitilmesi için elimizde verileri barındıracak olan dataset (veri kümeleri)'ler gereklidir.

Normalizasyon: DDİ'de girdi bazen bozuk formatta olmakta ve kendi hata paylarıyla kullanıcıya gelmektedir. Fax gönderileri, sosyal medya yazıları, gürültülü ses bu duruma bazı örneklerdir. Bu bozuk formattaki yazılar, normalizasyon ile doğru formata çevrilir.

Normalizasyonda girdimiz; harf yazımı dönüşümü, dönüştürme kuralları, özel isim tespiti, sesli harf üreteci, Türkçe karakter düzeltici, şive düzeltici, yazım hatası düzeltici gibi aşamalardan geçerek Türkçe yazım kurallarına uygun bir hale getirilir.

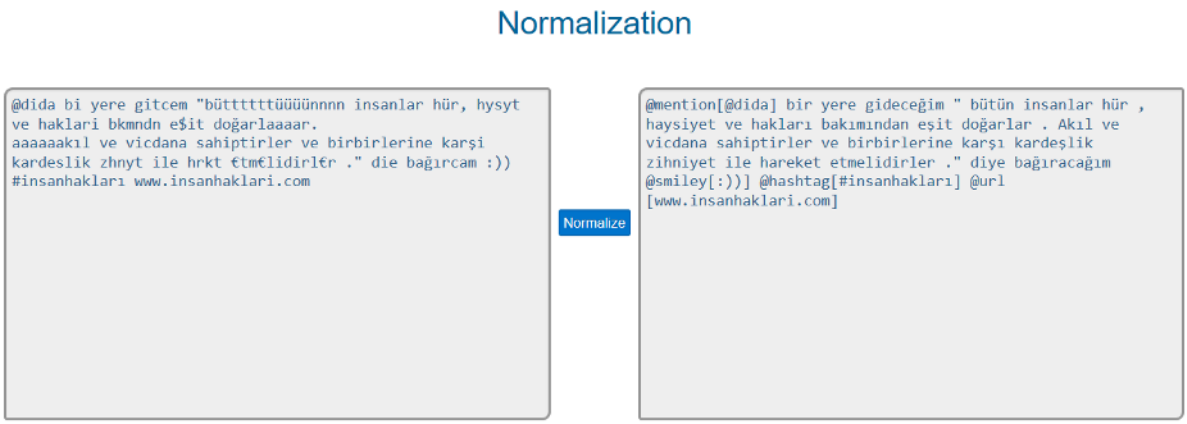
Normalizasyon basamağının içine varlık tanıma, özel isim tanıma gibi problemler de girmektedir. Örneğin; umut kelimesi, hem özel isim olarak hem de ümit etmek anlamındaki isim olarak kullanılabilir. Bu durumda cümlenin anlamından yola çıkarak kelimenin özel isim mi cins isim mi olduğuna karar verilir. Harf yazımı dönüşümünde özel/cins isim olma durumlarına göre doğru formata dönüştürme işlemi yapılır.

umuttan → umuttan (?), Umut'tan (?)

ayşenden → Ayşe'nden (?), Ayşen'den (?)

dünyada → dünyada (?), Dünya'da (?)

Özellikle bazı sosyal medya yazılarında kuralsız yazımlardan dolayı anlam konusunda güçlük çekilebilir. Yazıların içinde bir defa kullanılması gereken karakterler birden fazla kullanılabilir. Bazı harflerin yerine görsel anlamda benzeyen fakat Türkçe kurallarını ihlal eden karakterler kullanılabilir. Böyle durumlarda dönüştürme kuralları devreye girmektedir. Bu kurallar özellikle sosyal medya yazılarında e-posta isimlerini, hashtag, mention ve Uniform Resource Locator (URL)'leri yakalamaya yöneliktir.



Şekil 2.4: Örnek Bir Normalizasyon İşlemi (Eryiğit, 2014)

Bazı kelimelerde yazım yanlışlarından veya tercihlerden kaynaklı olarak sesli harfler kullanılmamaktadır. Türkçede sınıflandırıcının karar vermesi gereken çıktı olan 8 tane sesli harf bulunur. Hangi harfin doğru olduğuna karar veren kural kodlayıcılarıdır. Kural kodlayıcılarını eğitebilmek için elimizde veri olması gerekir (Örneğin; gazete haberlerinden toplanan veriler). Elimizde olan karakterler ve komşu harfler göz önünde bulundurularak doğru harfe karar verilir. Tablo 2.4'te "okuldan" kelimesinin sesli harfleri çıkarıldığında kelimeyi doğru formata çevirme işleminde uygulanan basamaklar gösterilmiştir.

Tablo 2.4: Okuldan Kelimesinin Örnek Temsili (Eryiğit, 2014)

Curr. Word	**Neigh. Ch(-3)**	**Neigh. Ch(-2)**	**Neigh. Ch(-1)**	**Neigh. Ch(+1)**	**Neigh. Ch(+2)**	**Neigh. Ch(+3)**	**Class Label**
kldn	-	-	-	k	l	d	o
kldn	-	-	k	l	d	-	u
kldn	-	k	l	d	n	-	null
kldn	k	l	d	n	-	-	A
kldn	l	d	n	-	-	-	null

Bazı yazılarda Türkçe karakter kullanılmaması veya yanlış kullanılması durumunda anlam kargaşası olabilmektedir. Normalizasyon işleminde bu şekilde yazılan kelimenin sağındaki ve solundaki sözcüklere; yani o andaki bağlama göre kelimenin doğru yazımına karar verilir. Tablo 2.5'te, Türkçe karakter kullanılmasına ve kullanılmamasına bağlı olarak anlamı değişen bazı kelimeler gösterilmiştir.

Tablo 2.5: Türkçe Karakter Kullanımına Göre Anlamı Değişen Bazı Kelimeler (Eryiğit, 2014)

Girdi	**Çıktı 1**	**Çıktı 2**	**Çıktı 3**
cin	cin (genie)	çin (China)	çın (ding)
kus	kus (vomit)	kuş (bird)	küs (sulk at)

Normalizasyonda şiveli yazılan bir sözcüğün Türkçe yazım kurallarına göre düzeltilmesi de sağlanır. Burada kural tabanlı sistemler kullanılır. Bu sistemlerde belli ek kalıpları mevcuttur. Sözcükte problemli yerler atılır. Sözcük köküne morfolojik açıdan, biçimbilimsel analiz örüntüsü ve eklerin analizlerini göz önünde bulundurarak doğru ekler üretilip eklenir. Aynı zamanda normalizasyon aşamasında klavyeye yanlışlıkla basılması durumunda ortaya çıkan

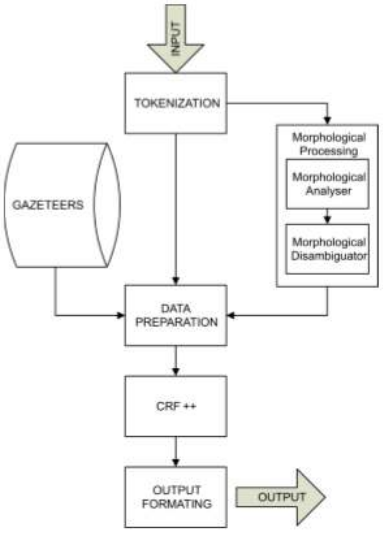
yazım yanlışları da düzeltilir. Temel olarak bir hata modeli ağacı ve bir dil ağacı bulunur. Mevcut bulunan hata modeline göre aday sözcük üretilip dil modeli ağacıyla karşılaştırılarak doğru sözcük üretilir.

yapacağım-yapacaksın-yapacak-yapacağız-yapacaksınız-yapacaklar (First sg) (Second sg) (Third sg) (First pl) (Second pl) (Third pl)
yapcam – yapcan – yapcak – yapcaz – yapcanız – yapcaklar yapçam – yapçan – yapçak – yapçaz – yapçaksınız – yapçaklar yapıcam – yapıcan – yapıcak – yapıcaz – yapıcaksınız – yapıcaklar yapacam – yapacan – yapacak – yapacaz – yapacanız – yapacaklar

Şekil 2.5: Yapmak Fiilinin Farklı Kişi Kiplerinde Gelecek Zamanda Bazı İdiyolektik Kullanımları (Eryiğit ve Torunoğlu Selamet, 2017)

Sözcük Analizi: Bu aşamada; özel isim, cins isim, sıfat, fiil edat, zarf gibi ayrımlar yapılmaktadır. Sonlu durum otomatları üzerinde kodlanmış kurallardır.

Varlık İsmi Tanıma: Bir metinde geçen kelimeleri kişi, lokasyon, tarih gibi belirli türlere göre işaretleme adımıdır. Akademi, özel sektör firmaları, bankalar vb. kısacası her ihtiyaç sahibi kendine özel türler belirler ve verisini ona göre işaretler. Buna varlık tanıma (entitiy recognition) denir. Bu işlemi yaparken elimizde veri kümeleri bulunur. Şimdiki duruma göre türler; kişi, isim, lokasyon, zaman ifadeleri, parasal ifadeler gibi başlıklar altında incelenir. Çalışan model aşağıdaki gibidir.



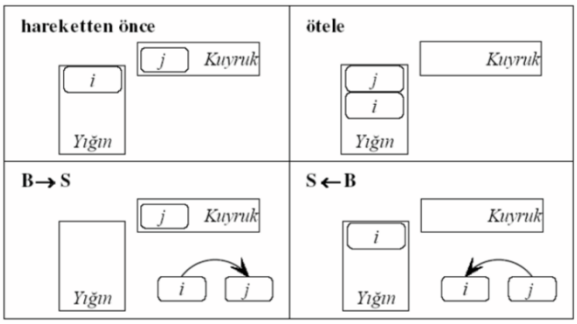
Şekil 2.6: Varlık İsmi Tanıma Modeli (Şeker ve Eryiğit, 2016)

Token	IOB2 Tags	RAW Tags
Mustafa	B-PERSON	PERSON
Kemal	I-PERSON	PERSON
Atatürk	I-PERSON	PERSON
1919	O	O
yılında	O	O
Samsun	B-LOCATION	LOCATION
'a	O	O
çıktı	O	O
.	O	O

Şekil 2.7: Varlık İsmi Tanıma Örneği (Şeker ve Eryiğit, 2016)

Cümle Analizi: Cümle analizinde bağlılık analizi denilen daha esnek sistemler geliştirilir, kural yazılmaz. Cümlenin öğeleri ikili ilişkiler şeklinde kurulur. Bağlılık analizi, "Doğal Dil Ayrıştırması (DDA)" alanında son yıllarda popüler hale gelmiş bir yöntemdir (Eryiğit ve diğ., 2006).

Bunun yanında cümle analizi geçiş tabanlı ayrıştırıcı yöntemle de çözülür. Bu yönteme göre bir yığın yapısı ve bir kuyruk yapısı bulunmaktadır. Yığında işlenmekte olan sözcükler, kuyrukta ise işlenmek için bekleyen sözcükler tutulur. Aslında yığında ve kuyrukta tutulanlar sözcükler değil çekim gruplarıdır. Ayrıştırma işlemi esnasında bu çekim grupları arasında ilişkiler oluşturulmaya çalışılır. Yığındaki çekim grupları ötelendikçe, kuyrukta bekleyen çekim gruplarına doğru ilerlenir. Son durumdaki yığın ve kuyruğun durumu aşağıdaki şekilde gösterilmiştir.



Şekil 2.8: Ayrıştırma Algoritması (Eryiğit ve diğ., 2006)

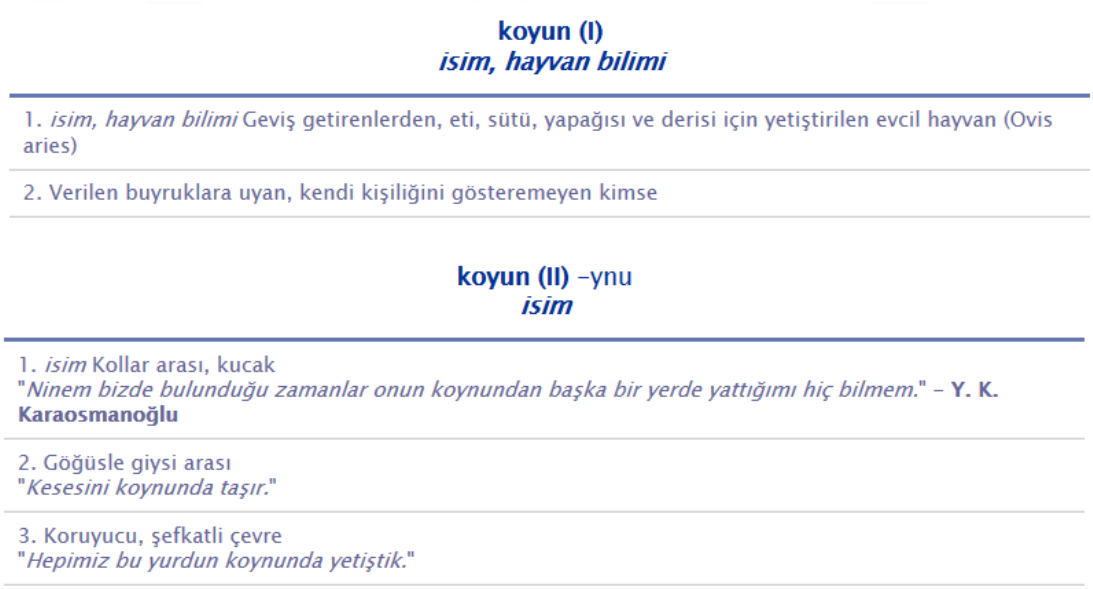
2.2.2.6. Doğal Dil İşlemede Karşılaşılan Zorluklar

Özellikle konuşma tanıma çalışmalarında kuralsız, anlaşılmaz ve gürültülü konuşmalar, bu konuşmaların algılanmasını zorlaştırmakta ve konuşma üzerinde çalışma yapılamamasına sebep olabilmektedir. Gürültü sorunu, kişiye ve konuya bağlılık açısından konuşmayı anlama çalışmalarında sıkıntı yaratabilmektedir. Bazı konuşmacıların bir cümleyi tamamlamadan

diğer bir cümleye başlaması, cümlenin nerde başlayıp nerde bittiğini anlamak konusunda zorluk yaratmaktadır.

Metin veya konuşma sentezlemede kök sözcükler özgün dillerindeki yazımları ile ifade edilirken eklenen Türkçe ekler sözcüğün özgün dilindeki söylenişine göre eklenmekte ve bu da anlam ve yazım bakımından bazı zorlukları beraberinde getirmektedir (örneğin; Godot'yu, Bordeaux'a, serverlar vb.).

Sesteş sözcükler; söylenişleri aynı, anlam ve kökleri farklı olan sözcüklerdir. Bu sözcükler bazen, özellikle yazılı ifadelerde anlam belirsizliklerine sebep olmaktadır. Aynı zamanda sözcükler farklı şekilde eklere bölünebilir veya ekler farklı işlevlere sahip olabilir. Aşağıda "koyun" sözcüğünün Türk Dil Kurumu'ndaki faklı anlamları gösterilmiştir.



Şekil 2.9: Örnek Sesteş Sözcükler (Sesteş Sözcükler, 2016)

Yazılı dilde kuralların denetlenmesine engel teşkil eden yazımlar olabilmektedir. Bu gibi durumlarda kök sözcüğün okunuşunun nasıl bittiğini bilmek durumunda kalınır. Kısaltmalar, saat ve tarih yazımları bu duruma örnektir.

Bazı diller söz dizim kuralları açısından oldukça katıyken, bazı diller bu konuda son derece esnektir. Doğal diller programlama dilleri gibi kesin değildir ve bu yüzden DDİ üzerine çalışmalar yapanlar bu konuda zorluklarla karşılaşabilirler.

DDİ'de girdi veya çıktı olan gösterimlerin detayları ve yeterliliği uygulamanın ne olduğuna bağlıdır. Gösterimler kuramsal ve idealdir. Bu yüzden gösterimleri hiçbir zaman doğrudan gözlemlenemez.

Bir dilin çok yapılı ve çok anlamlı olması, her düzeydeki girdilerin genelde çok sayıda yorumu olabilmesine sebep olmaktadır. Örneğin;

- Bir konuşma ses dalgası çok sayıda olası tümceye dönüştürülebilir.
- Bir tümcedeki her sözcüğün çok sayıda biçimbilimsel yorumu olabilir.
- Her kök sözcüğün birden fazla anlamı olabilir.
- Bir tümcenin çok sayıda yapısal çözümlemesi olabilir.
- Bir yapısal çözümlemenin birden fazla mantıksal çözümlemesi olabilir.

Yukarıdaki olasılıklardan esas zor olan ise, olası yorumların hangisinin doğru olduğunu çevrimden kestirmektir. İnsanlar bunu çok kolay bir şekilde yapabiliyorken, bilgisayarlar henüz yapamamaktadır.

2.3. Metinden Konuşma Sentezleme Sistemleri

Metnin sesli olarak kullanıcıya iletimini sağlayan MKS sistemleri, bilgiye erişim ve bilginin üretimi konularında günümüzde büyük kolaylıklar sağlayan teknolojilerdendir (Aydın, 2011).

Gerçeğe yakın seslerin üretilmeye çalışılması 18. yüzyılda başlamıştır. Tarihte bu alanda bilinen ilk önemli örnek, von Kempelen'in 1791 tarihli "Konuşma Makinesi"dir. 1950'li yıllarda daha önce kaydedilmiş sesleri art arda sıralayarak seslendiren çeşitli cihazlar geliştirilmiştir. 1960 ve sonrasında gelişen teknolojinin yardımıyla MKS çalışmaları hız kazanmıştır. Türkçe MKS çalışmaları ise 1990'lı yıllarda başlamış ve günümüzde farklı araştırmacılar tarafından yapılan çalışmalar hala devam etmektedir.

Genel olarak MKS sistemlerinde 3 çeşit yaklaşım vardır. Bu yaklaşımlardan birincisi kural tabanlı formant sentezleyiciler, ikincisi söyleyiş sentezleyicileri, üçüncüsü ise eklemeli sentezleyicilerdir (Sel ve diğ., 2011).

Kural tabanlı formant sentezleyiciler konuşma sinyalinin doğrusal öngürümlü kodlaması temeline dayanır. Söyleyiş sentezleyicileri sadece formant frekansların denetimine ihtiyaç duyar. Eklemeli sentezleyiciler ise önceden kaydedilmiş olan ses birimlerini bir araya getirerek konuşmayı oluşturur. Seslerin bir araya getirilmesi sinyal işleme ile gerçekleşir.

Türkçe MKS sistemlerinde genel olarak izlenen yol; Türkçe dilinin yapısından dolayı, yani eklemeli bir dil olduğundan dolayı, eklemeli yöntemlerdir. Benzer çalışmalar incelendiğinde genelde bu yöntemin kullanıldığı görülmüştür. Yine bu çalışmalardan bazıları eklenecek

parçalar olarak ikili, üçlü fonemleri kullanmıştır. Genel olarak literatür incelendiğinde Türkçenin en küçük yapı taşının hece olduğu ve heceler yardımıyla bir kelimeden çokça kelimeler üretildiği de bilinmektedir. O yüzden birleştirilecek seslerin hecelerden seçilmesi daha uygun olmaktadır (Sel, 2013).

Eklemeli sentezleyicilerde ilk önce ses veri tabanı hazırlanır. Daha sonra metin metnin ön incelemesine alınır. Burada metnin içerdiği sayılar, noktalama işaretleri vs. incelenir. Daha sonra metin parçalara ayrılır ve konuşma veri tabanındaki seslerle sayısal sinyal işlemeye alınır. Seslendirme kuralları da işlenerek konuşmaya çevrilir.

2.3.1. Metinden Konuşma Sentezleme Sisteminin Yapısı

MKS sistemleri genel olarak beş basamaktan oluşmaktadır. Bu basamakların detayları aşağıda verilmiştir.

2.3.1.1. Metin Ön İşlemleri

Bu aşamada sayılar, tarihler, kısaltmalar, kesirler gibi ifadelerin konuşma dilindeki sözcüklere dönüştürülmesi gerekir. Bu işleme normalizasyon da denilmektedir.

İlk olarak bir kısaltmalar sözlüğü oluşturulmalıdır. Kelimeleşmiş ve büyük harfle yazılanlar (örneğin; TUBİTAK), sonuna nokta koyulanlar (örneğin; Prof.), sonuna nokta koyulmayan ve küçük harfle yazılanlar (örneğin; kg, mm) gibi birçok kısaltma türü bulunmaktadır. Bu türlerin yazım kuralları çerçevesine göre okunuşları ses veri tabanına kaydedilmelidir. Konuşma sentezlemenin son aşamasında metinde tespit edilen kısaltmalar veri tabanındaki okunuşlarına göre seslendirilmelidir.

Sayıların seslendirilmesi de MKS uygulamalarında oldukça önemlidir. Tam sayı, kesirli sayı, TC kimlik numarası, saat, tarih gibi ifadelerin her biri farklı şekilde okunmaktadır. Örneğin; "2/5" ifadesi hem "iki bölü beş" hem de "beşte iki" şeklinde seslendirilebilmektedir. Aynı şekilde "05/05/2005" tarih ifadesi hem "beş beş ikibinbeş" hem de "beş mayıs ikibinbeş" şeklinde okunabilmektedir. Bu yüzden öncelikle sayılar yazılı hale getirilip ondan sonra seslendirmeleri yapılmalıdır.

Metin içerisindeki noktalama işaretleri de tespit edilmelidir. Bunun nedeni ise metin içerisinde nerelerde tam ölçü, nerelerde yarım ölçü sessizlik ekleneceğinin belirlenebilmesidir. Nokta (.) olan yerlerde tam ölçü, virgül (,) olan yerlerde yarım ölçü

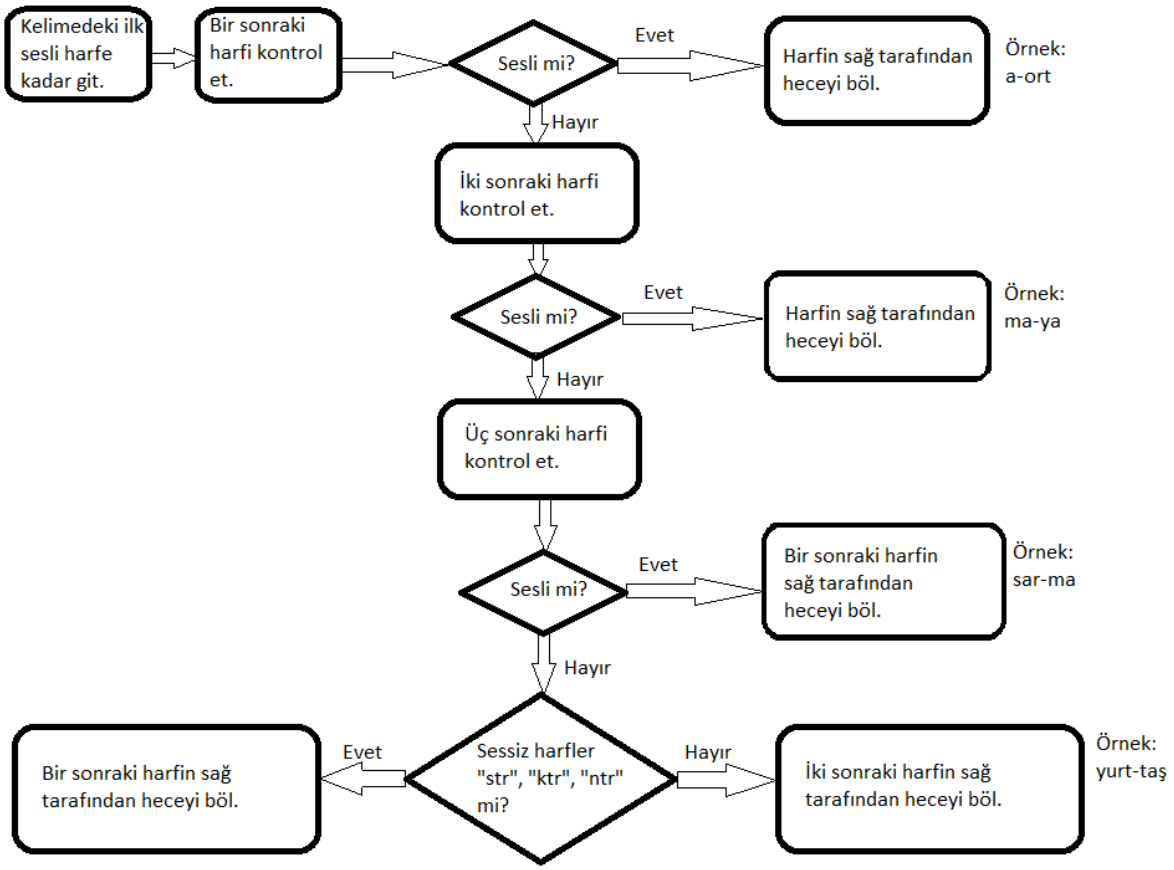
sessizlik olması metnin anlamının tam olarak verilebilmesi açısından önemlidir. Aynı şekilde metinde geçen "+, &, /, %, *, -" gibi ifadelerin doğru bir şekilde seslendirilmesi gerekmektedir.

Türkçede konuşma dili ile yazı dili arasında da bazı farklılıklar bulunabilir. Sessiz yumuşaması, sesli düşmesi, sessiz benzeşmesi bu durumlardan bazılarıdır. Örneğin, son harfi süreksiz sert sessiz (p,ç,t,k) olan bazı özel isimler ünlü ile başlayan bir ek aldığında okunuşu değişir. "Irak'a" şeklinde yazılan sözcük "Irağa" şeklinde okunur. Bu gibi durumlarda hangi sert sessiz harfin yumuşamaya uğrayacağı sorununa DDİ teknikleri ile çözüm getirilebilir.

Bazı durumlarda ise sözcük cümle içindeki anlamına göre seslendirilir. Bu tarz sözcüklere sesteş sözcükler denir. Örneğin "hala" sözcüğü sesteş bir sözcüktür. Birinci anlamı "babanın kız kardeşi" iken diğer anlamı "henüz"dür. Konuşma dilinde bu iki anlamı ayrı ayrı kullanabilmemizi sağlayan şey ise vurgudur. MKS uygulamalarında sesteş sözcüklerin hangi anlamda kullanıldığını, cümlenin anlamına ve kendinden önceki/sonraki sözcüklere bakarak ayırt etmeyi sağlayan bilim DDİ'dir. Bu konudaki çalışmalar hala devam etmektedir.

2.3.1.2. Metnin Hecelere Ayrılması

Türkçe MKS çalışmalarında okunacak metnin hecelere ayrılması için çok fazla algoritma geliştirilmiştir. Bu algoritmaların hepsi içerdiği çalışma alanı doğrultusunda başarılıdır. Metin ön işlemleri veya diğer adıyla normalizasyon aşamasında belirtildiği gibi Türkçenin bazı ses olaylarından dolayı farklı okunan veya bazı kurallar doğrultusunda okunması gereken sözcükler için geliştirilen algoritmalara yeni özellikler eklenmiştir. Örnek bir heceleme algoritmasının akış şeması aşağıda gösterilmiştir.



Şekil 2.10: Heceleme Algoritması Akış Şeması

2.3.1.3. Sesler

Ön işlemden geçerek hecelere ayrılan metnin bir sonraki aşaması eklemeli yöntem kullanılarak hecelerin birleştirilmesidir.

"Ses, gırtlaktaki ses tellerinin hava moleküllerini titreştirmesi sonucunda, bu hava moleküllerinin sıkışma ve dağılması ile enerjinin uygulandığı yöne paralel olarak boyuna basınç dalgaları oluşturması sonucunda meydana gelir" (Sel, 2013).

Bu tür çalışmalardaki kayıtlar genelde "wav" dosya formatında tutulur. Bu dosya formatında hiçbir sıkıştırma yöntemi kullanılmadığı için sesin kalitesi düşmez ancak çok yer kaplar.

Waw dosyası; riff, format ve data veri bölgesi olmak üzere iç veri bölgesinden oluşur (Sel, 2013). Tablo 2.6'da bu veri bölgeleri açıklanmıştır.

Riff bölgesinde (12 byte), dosyanın bir "wav" dosyası olduğu belirtilir. Format bölgesinde (24 byte), formatın parametreleri tutulur. Data bölgesinde ise gerçek örnekleme verileri tutulur.

Tablo 2.6: Riff Veri Bölgesi (Sel, 2013)

Byte Sırası	Açıklama
0-3	"RIFF" (ASCII karakterleri şeklinde)
4-7	LittleEndian şekilde paketin geri kalan boyutu
8-11	"WAVE" (ASCII karakterleri şeklinde)

Tablo 2.7: Format Veri Bölgesi (Sel, 2013)

Byte Sırası	Açıklama
0-3	"RIFF" (ASCII karakterleri şeklinde)
4-7	LittleEndian şekilde paketin geri kalan boyutu
8-9	"WAVE" (ASCII karakterleri şeklinde)
10-11	Kanal sayısı (0x01=Mono, 0x02=Stereo)
12-15	Hz olarak örnekleme oranı (binary)
16-19	Saniyedeki sekizli miktarı
20-21	Örnekteki sekizli miktarı 1=8 bit Mono, 2=8 bit Stereo yada 3=16 bit Mono, 4=16 bit Stereo
22-23	Örnekteki bit sayısı

Tablo 2.8: Data Veri Bölgesi (Sel, 2013)

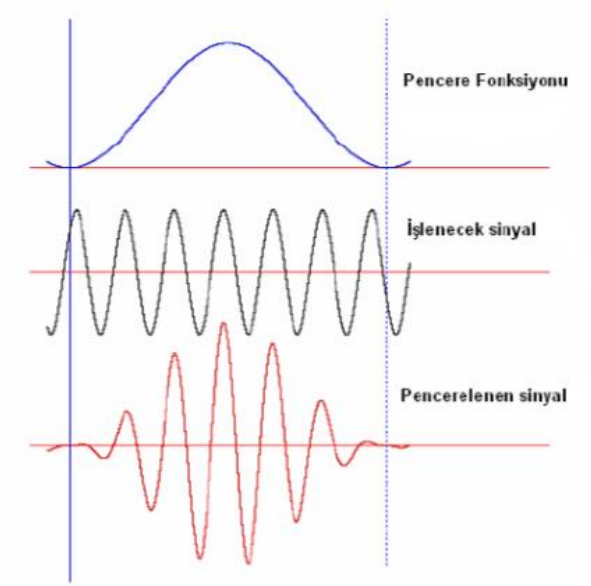
Byte Sırası	Açıklama
0-3	"data" (ASCII karakterleri şeklinde)
4-7	Verinin uzunluğu
8-son	Veri (Örnekler)

Ses parçaları birbirlerine fazları, frekansları ve enerjileri uyumlu olacak şekilde birleştirilmelidir. Bu yüzden sesler veri tabanına eklenmeden önce bu özelliklerinin birbirine uyumlu hale getirilmesi gerekmektedir. Bu işlem genellikle belli bir pencere fonksiyonu kullanılarak yapılır (Sel, 2013).

"Pencere fonksiyonları (veya kısaca pencere), sonlu impuls cevaplı (FIR, Finite Impulse Response) sayısal filtre tasarımında istenmeyen salınımları ortadan kaldırmak için kullanılan yapılardır" (Kaya ve İnce, 2010). İşlenmeden önce, sinyaller belli parçalara ayrılır. Bu parçalar örnek içeren parçalardır ve her birisine pencere ismi verilmektedir. Pencere fonksiyonları sayesinde parçalara ayrılan sinyallerin başlangıç ve bitiş kısımlarının söndürülmesi, orta kısımların ise vurgulanması sağlanmaktadır.

Eklemeli sentezleme kullanan sistemlerde genel olarak Hanning Pencere fonksiyonu kullanıldığı görülmüştür. Hanning penceresinden geçirilmiş ses verileri konuşma sentezleme

çalışması esnasında birleştirilen seslerin fazlarının ve frekanslarının eşit olması sağlayabilmektedir (Sel, 2013).



Şekil 2.11: Sinyalin Hanning Pencere Fonksiyonu Uygulandıktan Sonraki Görüntüsü (Sel, 2013)

2.3.1.4. Seslerin Veri Tabanına Kaydedilmesi

Sesler veri tabanına heceler şeklinde kaydedilir. Türkçede altı tip hece bulunmaktadır. Ancak günümüzde diğer dillerden dilimize geçmiş olan birçok sözcük bulunmaktadır. Yabancı dillerden gelen ve dilimizce benimsenmiş sözcüklerin hecelemelerinin doğru yapılmaması metin seslendirme çalışmalarında sorunlara yol açmaktadır. Bu yüzden bu tip hecelerin de eklenmesiyle Türkçede sekiz faklı hece tipi bulunmaktadır.

Hece tiplerinde sesli harf V, sessiz harf C olarak sembolize edilmiştir. Aşağıda, bu hece tipleri ve sayıları belirtilmiştir.

- V modeli (sesli) Örnek: a, e, ö, u,
  (8 adet hece/ses dosyası)
- VC modeli (sesli + sessiz) Örnek: et, üç, ön
  (21 x 8 = 168 adet hece/ses dosyası)
- CV modeli (sessiz + sesli) Örnek: ba, zo, ğü
  (21 x 8 = 168 adet hece/ses dosyası)
- CVC modeli (sessiz + sesli + sessiz) Örnek: gök, kal, bit

(21 x 8 x 21 = 3528 adet hece/ses dosyası)

- VCC modeli (sesli + sessiz + sessiz) Örnek: ilk, üst
  (8 x 21 x 21 = 3528 adet hece/ses dosyası)
- CVCC modeli (sessiz + sesli + sessiz + sessiz) Örnek: Sırp, dört
  (21 x 8 x 21 x 21 = 74088 adet hece/ses dosyası)
- CCV modeli (sessiz + sessiz + sesli) Örnek: tra,gri
  (21 x 21 x 8 = 3528 adet hece/ses dosyası)
- CCVC modeli (sessiz + sessiz + sesli + sessiz) Örnek: klor, drop
  (21 x 8 x 21 x 21 = 74088 adet hece/ses dosyası)

Yukarıdaki hece türleri için toplam hece sayısı:

8 + 168 + 168 + 3528 + 3528 + 74088 + 3528 + 74088 = 159104

Ancak bu hecelerin tamamı Türkçenin imla kuralları gereğince doğru olan heceler değillerdir. Türkçedeki imla kuralları göz önünde bulundurulduğunda ve hecelerin oluşmasını sağlayan kurallar dikkate alınarak bu sayıyı 7822 adede düşürmek mümkündür (Sel, 2013).

2.3.1.5. Seslerin Birleştirilmesi

Eklemeli MKS çalışmalarında daha önceden kaydedilmiş farklı tonlamadaki sesleri birbirlerine istenilen ton ve süre ile bağlamak ve bir süreklilik kazandırmak için zaman ölçeği modifikasyonları kullanılmaktadır. Genel olarak bilinen 4 tür modifikasyon teknikleri bulunmaktadır (Sel, 2013).

- OverLap Add (OLA / Örtüşme-Ekleme Algoritması): Birbirlerine bağlı olan girdi sinyali parçacıklarının orijinal faz ilişkilerini yok eder ve ardından hizalanmamış sinyal parçacıkları arasında ara değerleme yaparak yeni çıktı sinyalini oluşturur (Sel, 2013).
- Synchronous Overlap-Add (SOLA / Eşzamanlı Örtüşme-Ekleme Algoritması): Korelasyon tekniklerini temel alan zaman ölçeği sıkıştırma ve genişlemeye dayalı bir algoritmadır (Sel, 2013).
- Time Domain- Pitch Synchronous Overlap and Add (TD-PSOLA / Zaman-Alan Senkronize Ton Örtüşme ve Ekleme Algoritması): Sinyalin bir ton (örneğin, insan sesi ve monofonik müzik enstrümanları) aracılığıyla karakterize edildiği hipotezine dayanan TD-PSOLA, SOLA algoritmasının bir varyasyonu olarak üretilmiştir. Bu

yöntemin ana problemlerinden biri temel frekansın kayıp olduğu durumlarda sinyalin temel ton periyodunu tahmin etmede yaşanan sıkıntıdır (Sel, 2013).

- Waveform Similarity OverLap Add (WSOLA / Dalga Şekilli- Benzerlik Tabanlı Senkronize Örtüştürme ve Ekleme Algoritması): Bu işleme tekniğinde kullanılan temel prensip, konuşma sinyali içerisinde belirli zaman periyotları boyunca dalga şekli benzerliklerinin ortaya çıkarılarak ses örneklerinin azaltılmasıdır (Sel, 2013).

2.3.2. Doğal Dil İşleme İle Metinden Konuşma Sentezleme

Metin sentezleyicilerde bilgisayarda mevcut olan sayısal kod ardışıklığının ses sinyaline dönüştürülmesi amaçlanır. Bu bağlamda sentezleyiciler iki kısma ayrılır:

- Kısıtlı sözlüklü sentezleyiciler
- Genel amaçlı sentezleyiciler

Kısıtlı sözlüklü sentezleyicilerde sesli konuşma, ses elemanlarının birleşmesi yardımı ile elde edilmektedir. Ses elemanları, spiker tarafından girildikten sonra kodlanarak bellekte tutulur. Sesli veri çıkışı istendiğinde, metin, birimlere parçalanarak uygun birimlerle eşleşmeler yapılır ve sonunda onlar birleştirilerek sesli onarım gerçekleştirilir. Bu sentezleyiciler basit bir biçimde çalışmasına rağmen, doğal konuşmanın oluşturulmasında, duygusal durumları ifade etmekte zorlanmaktadırlar (Nabiyev, 2012).

Kelimeler bir bütündür. Onları parçalara ayrıştırarak tekrar onarmak, bazı kayıplara ve gecikmelere yol açar. Bu yüzden sentetik konuşma, doğal konuşmadan farklılık gösterir ve her dile özgü bir sözlüğün oluşturulması zor bir problemdir.

Genel amaçlı sentezleyicilerde ise doğal konuşmanın ritim, aksan, melodi dahil bütün özelliklerinin ifade edilmesi amaçlanmaktadır. Bu durumda bilgi tabanında, yalnız fonemler ve heceler değil, onların duruma bağımlı olarak değişebilen özellikleri de tutulmaktadır (Nabiyev, 2012).

Metne göre sentetik konuşmayı gerçekleştiren sentezleyiciler metne ilişkin sonsuz sözlüklerle çalışmaktadırlar. Bilgisayardaki giriş imla metni kelimelere ayrıştırılarak farklı aksan işaretleyicileri koyulur. Uygun eklemeler ve silmeler yapılır (harf düşmesi, harf yumuşaması vb). Sonra ise işaretlenmiş bu bilgiler metin-fonem dönüştürücülerine verilir. Bu birimde harf birleşmesine ve izlemesine uygun olarak fonem uyumluluğu dönüşümü yapılır. Sonuçta konuşmalı fonem metni elde edilir (Nabiyev, 2012).

Gerçekçi bir sentetik konuşmanın elde edilebilmesi için formantlar ve melodi-ritmik özellikler üzerinde çalışılmalıdır. Formant olarak frekansla, genlikle ve bant genişliği ile karakterize edilen ses yolunun taşıma fonksiyonunun bantları düşünülmektedir.

Uzun ya da kısa ünlüler sesyazar (sonagraph) denen bir aygıtla saptanmaktadır. Tek tek sesler, ya da onların oluşturdukları sözcükler banda okunur; sonra sonagrafta elde edilen görüntüler ölçüm çizelgesi göstergelerine göre değerlendirilir. Ses dalgası silindir biçiminde olan sesyazarın dış yüzüne seviyelerle kaydedilir. Oluşan bu halkacıklar için formant terimi de kullanılmaktadır. Ses dalgaları, formant 1, 2, 3 vb gibi yükselme gösterir. Çene açısı ne kadar darsa, söz konusu ünlünün formantı o kadar alçaktır. Çene açısının daralması dilin ağız boşluğunda aldığı biçimle bağlantılıdır, dilin yüksekliği arttıkça söz konusu ünlünün birinci formantı, yani ses dalgasının yayılma alanındaki birinci basamağı uygun olarak alçalır. Ünlülerin tanınması ya da birbirinden ayırt edilmesinde 1. Ve 2. Formantlar önem taşır. 3 ve 4. Formantlar kişinin ses tınısının özelliklerini yansıtmaktadır (Nabiyev, 2012).

Sonagrafta belirli alanlar için belirli sürelerle kayıtlar yapılmaktadır. Ayrıca her uygun alan-süre değerlerine göre bir sesin yalın halde çıkma süresi saptanmaktadır. Ünsüzler ünlülere karşın sesyazarda dikeyine görüntü vermektedir. Tablo 2.9'da Türkçe ünlüler için Hertz olarak 1. ve 2. formant değerleri verilmiştir (Nabiyev, 2012). (Formant değerleri değişkenlik gösterebilir.)

Tablo 2.9: Türkçedeki Ünlülerin Formant Değerleri (Nabiyev, 2012)

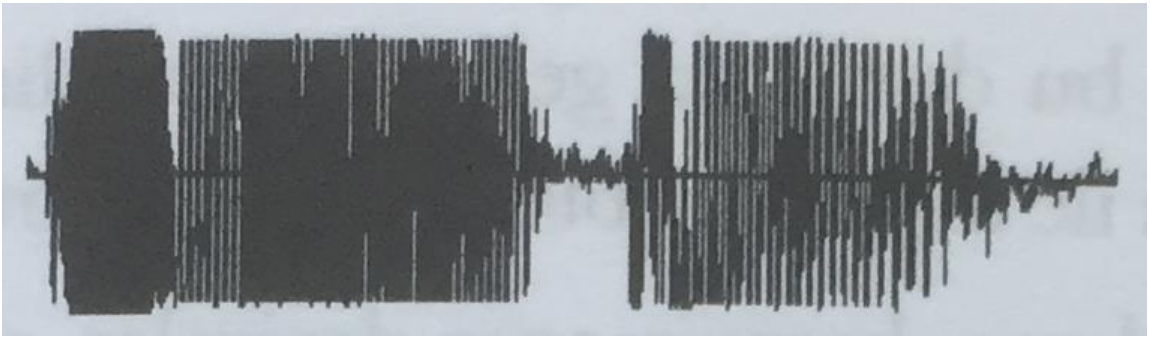
Ünlüler	**1. formant, Hz**	**2. formant, Hz**
A	800	1400
O	1200	2800
U	400	800
İ	400	1200
E	400	1800
Ö	400	1400
Ü	280	1700
I	320	2000

Ses üretmeyi amaçlayan sistemler yazılı metinden anlamlı konuşma gerçekleştirilmesi olarak tanımlanır. Bunun en kolay yolu ise Doğrudan Söz Sentezleme yöntemini kullanmaktır. Bu şekilde yapay bir konuşma elde edilir. Bunun için ise konuşması gerçekleştirilecek metnin ekleri de dahil olmak üzere bir konuşmacı tarafından doğal olarak okunması, anlamlı ses bilgilerinin örneklendirilerek kodlanması ve sistemde saklanması gerekir. Bu şekilde çalışan bütün sistemlerde analog-sayısal ve sayısal-analog dönüştürücüler kullanılır. Burada gerçekleştirilen işlemler şu şekildedir:

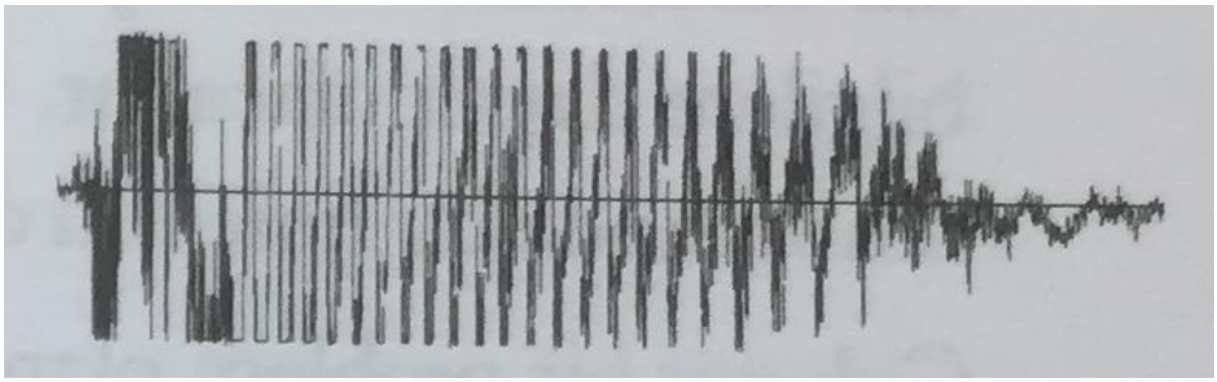
- Kullanıcı programı, sayısal ses bilgisinin bulunduğu bellek adresini Audio Controller'a yollar.
- Controller Direct Memory Access (DMA) kanalı ile ana bellekten bilgiyi alarak D/A dönüştürücüye iletir.
- Dönüştürücü, ayrık şekilde olan bu sayısal bilgiyi sürekli analog biçimine çevirerek hoparlöre uygular (Nabiyev, 2012).

Bu yöntemin en kötü yönü; metnin içerebileceği her kelimenin önceden bir kullanıcı tarafından okunmuş olmasının çok büyük bir bellek kullanımı gerektirmesidir. Kısıtlı sözlükler ise sistemin esnekliğini engelleyip başarısını düşürmektedir. Türkçe sondan eklemeli bir dil olduğu için kısıtlı sözlükler yaklaşımı çok basit kalmaktadır.

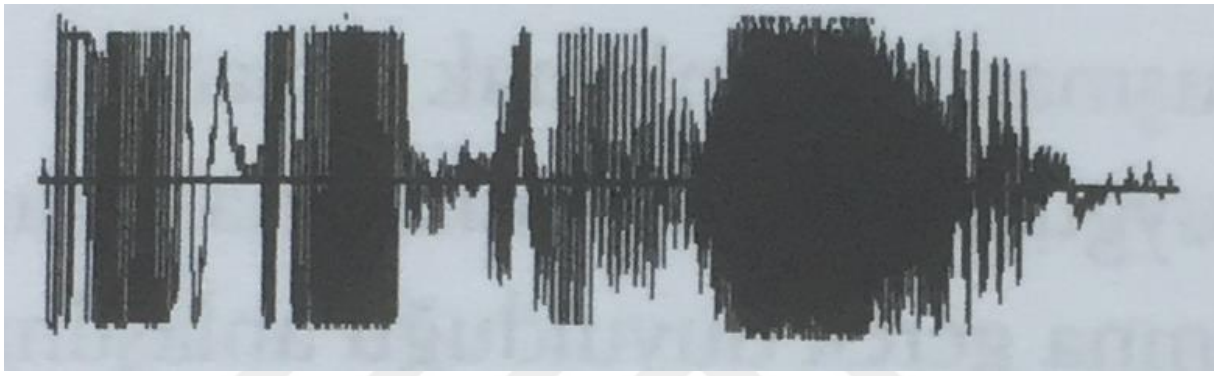
Hecelerin seslendirilmesi, hecelerden oluşan harflerin yerlerine göre harf seslerinin uygun tonda birleştirilmesi ile meydana gelir. Burada fonemler büyük önem taşımaktadır. Aynı durum, hecelerin bir araya gelerek kelimeleri oluşturmasında da vardır. Bir hecenin kelimenin ilk hecesi olarak söylenişi ile orta ve son hecesi olarak söylenişi farklıdır. Bu farklılıklar "çanta", "patates", "takım" kelimelerinin aşağıdaki şekilde gösterilen ses sinyallerinden de görülmektedir. Hatta aynı kelime içerinde aynı hecenin "görüntüsü" farklı olmaktadır (Nabiyev, 2012).



Şekil 2.12: "Çanta" Sözcüğünün Tamamının Ses Sinyali (Nabiyev, 2012)



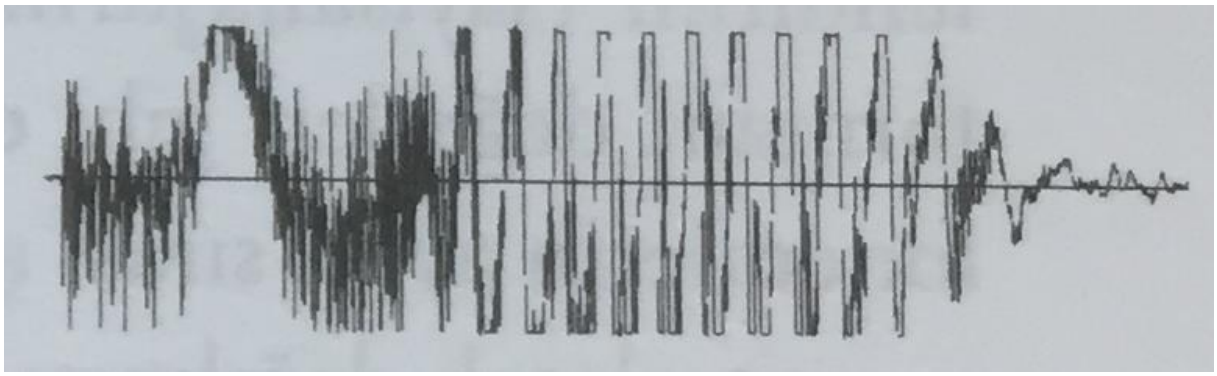
Şekil 2.13: "Çanta" Sözcüğündeki *–ta* Hecesinin Ses Sinyali (Hece Sonda) (Nabiyev, 2012)



Şekil 2.14: "Patates" Sözcüğünün Tamamının Ses Sinyali (Nabiyev, 2012)

Şekil 2.15: "Patates" Sözcüğündeki *–ta* Hecesinin Ses Sinyali (Hece Ortada) (Nabiyev, 2012)

Şekil 2.16: "Takım" Sözcüğünün Tamamının Ses Sinyali (Nabiyev, 2012)



Şekil 2.17: "Takım" Sözcüğündeki *–ta* Hecesinin Ses Sinyali (Hece Önde) (Nabiyev, 2012)

Şekil 2.18: "Tatatatatam" Sözcüğünün Ses Sinyali (Nabiyev, 2012)

Üretilen ses sinyallerinin anlaşılırlık testleri de yapılmalıdır. Bu testler özellikle görme engelli bireylerin ihtiyaçlarının karşılanmasına yönelik eksiklerin belirlenmesinde önemli rol oynamaktadır. Anlaşılırlık testleri birbirine yakın olan sesbirimlerin bireyler tarafından ayırt edilip edilmemesini de ölçer (örneğin; mil-nil, çark-şark kelimeleri). En yaygın kullanılan anlaşılırlık testi bir ANSI standardı olan, Kafiye Teşhis Deneyidir (DRT – Diagnostic Rhyme Test) (Nabiyev, 2012). Bu testte bireylere birbirine yakın veya tek harfle birbirinden ayrılan kelime çiftleri sunulur ve hangi kelimeyi duydukları sorulur. Kelime çiftleri birbirinden sadece bir harfle ayrılmakta ve seslilik, burunsal, duraklamalı, ıslıklamalı, sertlik ve bitişiklik özelliklerine göre sınıflandırılmaktadır.

Diğer dillerde olduğu gibi Türkçede de bazı sesteş kelimeler aynı yazılmasına rağmen farklı okunabilmektedir. Farklı okunması gereken kelimelerin aynı şekilde okunması ise MKS sistemlerinde bazı anlam kargaşalarına sebep olmaktadır.

İnsanlar metinde geçen olayları anlayabilmek için doğuştan bilişsel bir sisteme sahiptir. Anlam belirsizliğine sahip olan bir cümlenin anlaşılması öğrenilmiş olan anlam kümeleri içerisinden en uygun anlamın seçilmesiyle gerçekleşmektedir. Bilgisayar sistemlerinde ise bu durum uygun algoritmaların kullanılması ile çözüme kavuşur. Bu uygulamalar DDİ ile başlamış, hesaplamalı dilbilim çalışmaları ile devam etmektedir.

Sözcük anlamının belirginleştirilmesi konusu DDİ'nin çalışma alanlarından birisidir. Örneğin; "Yüzümde bir yara çıktı." Ve "Havuzda yüzdüm." cümlelerindeki "yüz" kelimesi isim veya fiil olma durumlarına göre adlandırılır. Bilgi çıkarımı yapılması amaçlanan bir uygulamada, söz konusu anahtar sözcük taranırken sözcüğün diğer anlamlarını elemek uygulamanın çalışma kalitesini arttıracaktır.

Sözcük anlamı belirginleştirmenin bir başka yaklaşımı ise metinde incelenen sözcük ile birlikte bir önceki ve bir sonraki sözcüğün de incelenip anlamsal olarak sınıflandırılmasıdır. "Sözcük kategorilerini oluşturan bu ontolojik sıradüzen aynı zamanda bir anlamsal ağ yapısı oluşturur" (Altan ve Orhan, 2005). Bu yaklaşım, günümüzde birçok DDİ uygulamasında kullanılmaktadır.

Sözcük anlamı belirginleştirmede kullanılan ve gerekli olan bilgi tipleri; öğeler, biçimbilim, yardımcı sözcükler, anlamsal sözcük birliktelikleri, seçimsel öncelikler, kullanım alanı ve şekli ve anlamların frekansıdır. SAB algoritmalarında kullanılan bilgi kaynakları arasında elektronik sözlükler (ES), ontolojiler ve derleme metinler yer alır, ayrıca bunların birkaçının birlikte kullanıldığı uygulamalar da vardır (Orhan, 2006).

SAB alanında yapılan ilk çalışmalar İngilizce dilinde olduğu için bu dilde birçok kaynağa ulaşmak mümkündür. "Özellikle Princeton Üniversitesi Bilişsel Bilimler Laboratuarı'nda 1985 yılında Prof. A.G. Miller tarafından başlatılan WordNet projesi anlamsal bir sözlük olarak İngilizce sözcükleri eşanlamlılar kümelerinde sınıflandırır ve sözcüklerin kısa, genel tanımlamalarını yaparak bu eşanlamlılar kümeleri arasındaki çeşitli anlamsal ilişkileri oluşturur" (Altan ve Orhan, 2005). Senseval projeleri ise bu alandaki çalışmaların yaygınlaştırılmasını sağlamıştır. "Farklı diller için geliştirilen yöntemlerin değerlendirmesinin yapıldığı uluslararası bir çalıştay olan Senseval'in dördüncüsü Semeval adıyla anılmaktadır ve 2007 yılında yapılmıştır. İlk kez bu çalıştayda ve sözcüksel örnekler alanında Türkçe için bir çalışma yer almıştır" (Tüysüz ve Güvenoğlu, 2014).

Bu alanda yapılan çalışmaların bir diğeri de ODTÜ-Sabancı derleme metnidir. Haziran 2003'te derleme metnin bir bölümü, Kasım 2003'te ise derleme metnin tamamı akademik araştırmalar için kullanıma açılmıştır.

Çalışmada bir ana derleme metin, bir de farklı kullanımlar için bu ana derleme metinden farklı bazı özellikleri olan ağaç bankası derleme metni geliştirilmiştir. Derleme metinde kullanılan metinler 1990 yılı sonrası basılan eserlerden seçilmiştir. Derleme metinde yaklaşık olarak

2.000.000 sözcük bulunmaktadır. 201 kitap, 87 makale ve 3 tane günlük gazeteden seçilmiş haberlerden oluşan 999 farklı yazılı metin kullanılmıştır (Orhan, 2006). Bu derlemede asıl görev işaretleyicilerindir. İşaretleyiciler düzeltme ve işaretlemeyi yaparlar. Bilgisayar programları ise işaretleyicilere analizde, çözümlemede ve belirsizlikleri gidermede yardımcı olmuştur.

Tablo 2.10'da, bu işlemler sonrasında ortaya çıkan Türkçe ağaç bankasının Extensible Markup Language (XML) biçimi gösterilmiştir.

Tablo 2.10: ODTÜ Ağaç Bankası Dosyalarının XML Biçimi (Orhan, 2006)

```
<?xml version="1.0" encoding="windows-1254" ?>
- <Set sentences="2">
- <S No="1">
  <W IX="1" LEM="" MORPH="" IG="[(1,"ben+Pron+PersP+A1sg+Pnon+Nom")]"
     REL="[3,1,(SUBJECT)]">Ben</W>
 <W IX="2" LEM="" MORPH="" IG="[(1,"bir+Det")]" REL="[3,1,(DETERMINER)]">bir</W>
 <W IX="3" LEM="" MORPH="" IG="[(1,"tutsak+Noun+A3sg+P1sg+Nom")]"
     REL="[4,1,(SENTENCE)]">tutsağım</W>
 <W IX="4" LEM="" MORPH="" IG="[(1,",+Punc")]" REL="[5,1,(OBJECT)]">,</W>
 <W IX="5" LEM="" MORPH="" IG="[(1,"de+Verb+Pos+Past+A3sg")]"
     REL="[6,1,(SENTENCE)]">dedi</W>
 <W IX="6" LEM="" MORPH="" IG="[(1,".+Punc")]" REL="[,( )]">.</W>
 </S>
 +<S No="2">
 </Set>
```

Sistemde cümlenin numarası, incelenen sözcük ve sözcüğün tipi, bu sözcükle ilişkili olan diğer sözcükler, ilişki biçimleri gibi veriler bulunmaktadır.

Tablo 2.11: ODTÜ Ağaç Bankasından Seçilen Sözcüklerin Cümle ve Anlam Sayıları (Orhan, 2006)

Sözcük	**Metinlerdeki Tümce Sayısı**	**Anlam Sayısı**
Yan	104	9
Git	189	10
Gör	133	9
Çık	231	15
Al	250	10
Gel	281	12
Yap	328	6
Ol	941	4

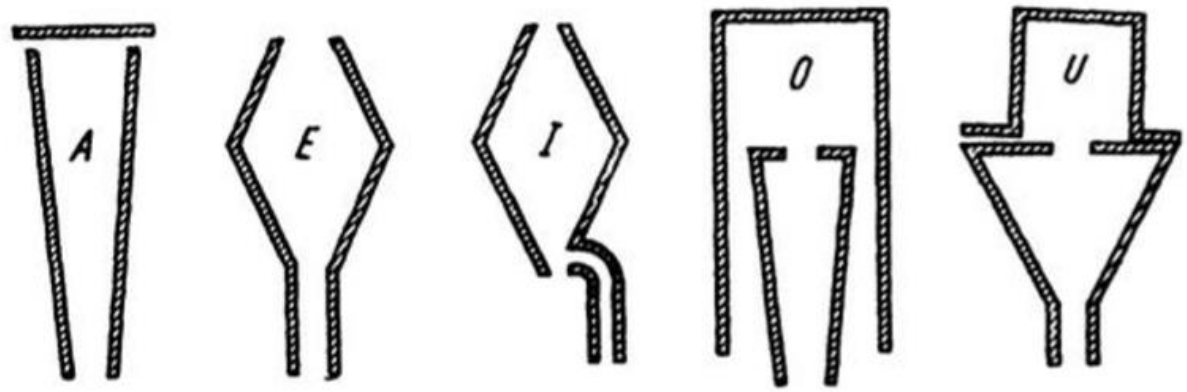
Tablo 2.11'de görüldüğü gibi sistemde bulunan 2.000.000 sözcükten seçilen bazı sözcüklerin anlam sayıları fazladır. Özellikle kelimenin anlamının vurgu ve tonlama ile verilebildiği durumlarda, her sözcüğün aynı şekilde seslendirilmesi MKS sistemlerinde zorluklara ve verimin düşmesine sebep olmaktadır.

3. MATERYAL VE YÖNTEM

Bu bölümde, ilk olarak geçmişten bugüne dünyada yapılan belli başlı MKS sistemlerinden bahsedilmiş, ardından araştırmanın sınırlılıkları doğrultusunda Türkçe MKS sistemleri hakkındaki literatür incelenmiştir. Türkçe MKS sistemlerinde genel olarak karşılaşılan sorunlar ve eksiklikler saptanmıştır. Araştırmaya dahil edilen çalışmaların sonuçları ve önerileri özet niteliğinde açıklanmıştır.

3.1. Metinden Konuşma Sentezleme Sistemlerinin Tarihsel Gelişimi ve Yapılan Çalışmalar

MKS sistemlerinin tarihi bilgisayarın icadından sonra başlamıştır, çünkü MKS sistemleri bir bilgisayara ihtiyaç duymaktadır. Bu sistemler metni otomatik olarak konuşmaya çevirebilmelidir. Konuşma sentezi ile ilgili ilk çabalar iki yüzyıl öncesine dayanmaktadır; Rus Profesör Christian Kratzenstein, beş uzun ünlü (a, e, i, o, u) arasındaki fizyolojik farklılıkları anlatıp bu sesleri 1779'da St. Petersburg'da yapay olarak üreten bir sistem oluşturmuştur (Schroeder, 1972).



Şekil 3.1: Kratzenstein'in çalışmasının gösterimi (Schroeder, 1972)

Wolfgang von Kempelen, 1791'de Viyana'da, tek sesler ve bazı ses kombinasyonları üretebilen "Akustik Mekanik Konuşma Makinesi" adında bir makine yapmıştır (Klatt, 1987). Charles Wheatstone ise, von Kempelen'in konuşma makinesinin ünlü versiyonunu 1800'lerin ortalarında inşa etmiştir. Bu makine öncekinden daha karmaşıktır çünkü ünlüler ve ünsüzlerin çoğunu üretebilmiştir. Makinenin biraz daha geliştirilmesinden sonra bazı kelimelerin üretimi sağlanabilmiştir.

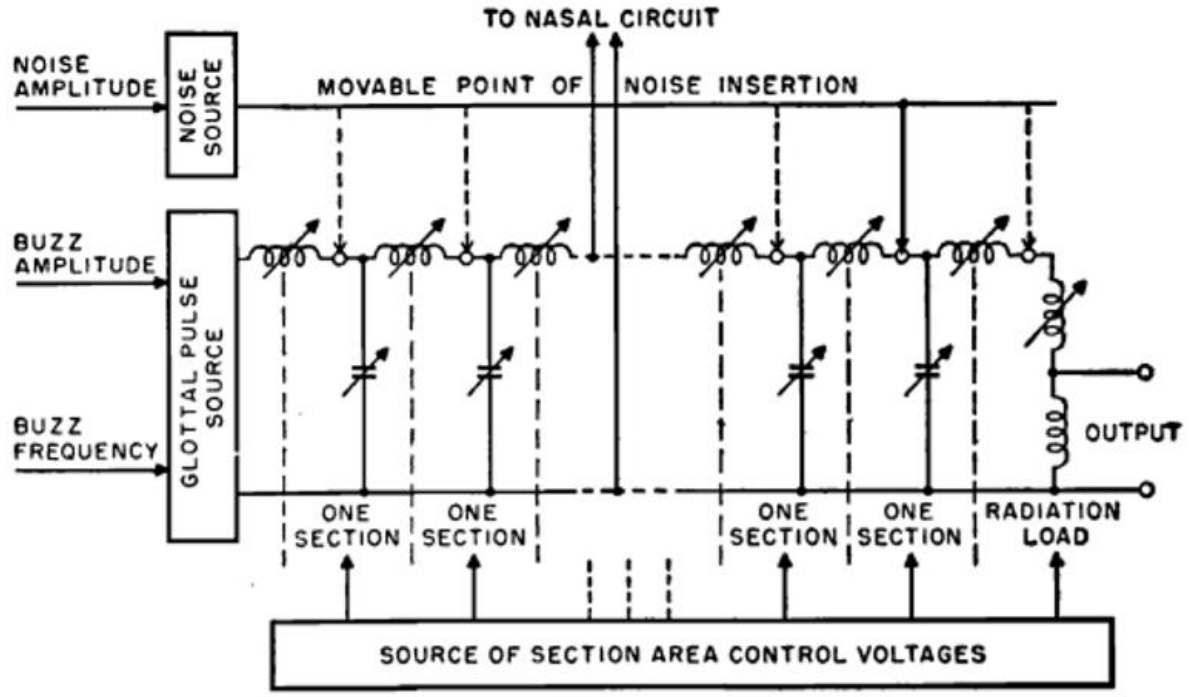
Willis, 1838'de bir sesli harf ile ses kanalının geometrisi arasındaki bağlantıyı bulmuştur (Klatt, 1987). Ses kanalları ile aynı yapıdaki tüp rezonatörleri kullanarak farklı ünlü sesler üretmiştir.

Stewart, 1922'de ilk tam elektrik konuşma sentezleme cihazını tanıtmıştır; bu cihazla sadece statik sesli olan sesleri üretebilmiştir, ünsüz seslerin veya bağlı seslerin üretimi bu sistemde mümkün olmamıştır (Klatt, 1987).

İlk konuşma sentezleyicisi olarak kabul edilen cihaz VODER, 1939'da New York'ta Homer Dudley tarafından tanıtılmıştır (Schroeder, 1972). Konuşma kalitesi ve anlaşılabilirlik açısından iyi olmamasına rağmen, yapay konuşma üretimi potansiyeli göstermesi açısından umut vadetmiştir.

Walter Lawrence tarafından ilk formant sentezleyici PAT (Parametric Artificial Talker) 1953'te tanıtılmıştır.

İlk artikülatör model konuşma sentezleyici, George Rosen tarafından 1958'de MIT (Massachusetts Institute of Technology)'de geliştirilmiştir (Klatt, 1987). Aşağıdaki şekilde gösterildiği gibi, artikülatör sentezleyici cihazı DAVO her bölüm için elle ayarlanmış değişken indüktörlere ve kapasitelere sahiptir.



Şekil 3.2: Konuşma sentezlemenin ilk artikülatör modeli (Klatt, 1987)

John Holmes, 1972'de "I enjoy the simple life." cümlesinin sentezleme işlemini elle ayarlamasına rağmen kalite olarak seslendirmede çok iyi sonuçlar almıştır. Ortalama bir dinleyici sentezlenen ses ile doğal ses arasında bir fark olmadığını söylemiştir (Klatt, 1987).

Noriko Umeda ve arkadaşları, Japonya'da İngilizce için ilk tam metin konuşma sistemini geliştirmiştir. Konuşma anlaşılırlık olarak iyi sonuçlar vermiştir fakat monoton bir seslendirme elde edilmiş ve mevcut sistemlerin kalitesinden oldukça uzak kalmıştır.

Allen, Hunnicutt ve Klatt (1987), MITalk'u üretmiştir.

1970'lerin sonunda ve 1980'lerin başında, ticari olarak çok sayıda MKS sistemi üretilmiştir. Bilgisayarda çalışan yazılım ürünlerinin yanında donanım çözümleri de sunan farklı MKS çipleri üretilmiştir. Bugünlerden sonra, aşağıda açıklamaları verilen, başarılı sayılabilecek birçok MKS sistemi farklı diller için oluşturulmuştur.

MKS konusunda dünyada yapılan birçok uygulama mevcuttur. Bu uygulamalar ticari ve ticari olmayan amaçlarla kullanılmıştır. Bu bölümde en bilinen uygulamalar açıklanmıştır. Tablo 3.1'de bu uygulamalar ve bunların temel amaçları verilmiştir.

Tablo 3.1: Dünyadaki temel MKS sistemleri

Çalışma	Tarih	Yöntem
ASY	1960'lar	Söyleyiş sentezleyici
MITalk	1979	Formant sentezleyici
Infovox	1982	Formant sentezleyici
Bell Labs TTS	1973	Çift / üçlü ses ekleme
ETI Eloquence	1988	Eklemeli sentezleme
CNET PSOLA	1980'lerin ortaları	Çift ses ekleme
Festival TTS	1990'ların sonları	Çift ses ekleme
MBROLA	1990'ların sonları	Çift ses ekleme
Whistler	2000'ler	Hece ekleme

ASY, 1960'larda ve 1970'lerde Paul Mermelstein, Cecil Coker ve arkadaşları tarafından vokal kanal modellerine dayanan hesaplamalı bir konuşma üretim modelidir. [30]. En bilinen söyleyiş sentezleyici olması bakımından önemlidir.

MITalk da günümüzde kullanılan ve geliştirilen birçok çalışmada temel teşkil etmesi bakımından önemlidir. J. Allen, S. Hunnicutt, D. Klatt (1987) tarafından yılında geliştirilmiştir. Telaffuz sözlüğü yerine harf sesine dayalı kurallar kullanmıştır.

Infovox, 1982 yılında ticari olarak kullanılmak amacıyla geliştirilmiştir. İngilizce, Almanca, Fransızca, İsveççe, Fince, Danimarkaca, İzlandaca, İspanyolca, İtalyanca, ve Türkçe dahil olmak üzere birçok dil desteği bulunmaktadır (Dutoit, 1999).

Bell Labs Text to Speech (TTS), çift ses veya üçlü ses ekleme yöntemine dayanmaktadır. Çok fazla modülü bir arada bulunduran bir yapısı vardır (metin işleme, vurgulama, telaffuz, kelimeyi bölme, duraksama vs.) (Dutoit, 1999). Çince, Japonca, Rusça, Romence, İtalyanca, İspanyolca desteği bulunmaktadır ve bu dillerin seslendirilmesi konusunda olumlu sonuçlar elde edilmiştir.

ETI Eloquence, eklemeli sentezleme yöntemine dayanan ve birçok dil desteği sağlayan bir uygulamadır. Uygulama günümüzde Realspeak adıyla kullanılmaktadır.

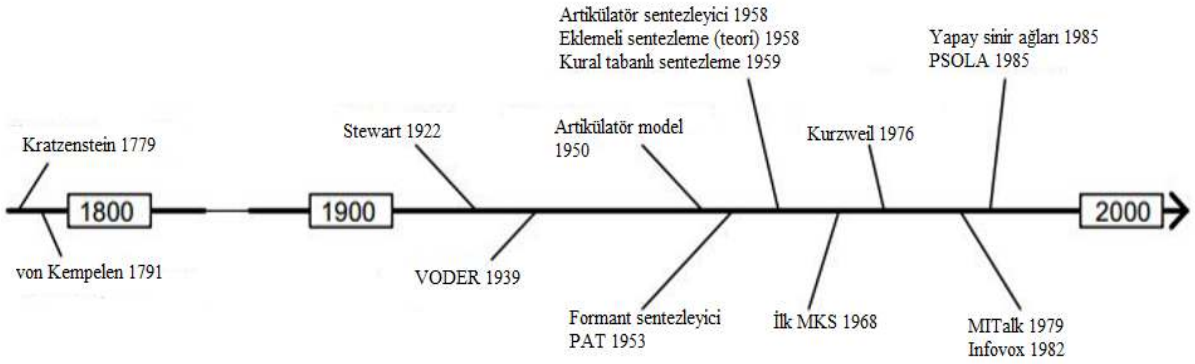
CNET PSOLA, 1980'lerin ortalarında France Telecom tarafından tanıtılmıştır. Pitch Synchronous Overlap and Add (PSOLA) algoritması kullanarak çift ses ekleme yöntemine dayanmaktadır. Uygulama geliştirilirken doğallık hedef alınmıştır. İngilizce, Almanca, Rusça, İspanyolca ve Fransızca dil desteği bulunmaktadır.

Festival TTS, çoklu dil desteği sağlayan C++ dilinde yazılmış ücretsiz bir yazılımdır. Açık kaynak kodlu olması bakımından önemlidir. Birçok Linux dağıtımının içinde kişisel bilgisayarlarda kullanılabilmektedir (Klatt, 1987). İspanyolca, Çekçe, Fince, Hintçe, İtalyanca, Rusça gibi dilleri desteklemektedir.

MBROLA, birçok dilde konuşma sentezleme sağlayan ve ticari olmayan ücretsiz bir yazılım olma amacıyla geliştirilmiştir. "Projede PSOLA benzeri algoritma kullanılmıştır ancak CNET patenti dolayısıyla bu isim yerine MBROLA kullanılmıştır" (Dutoit, 1999). Girdi olarak ses (metin yerine), süre ve frekans bilgilerini alıp çıktı olarak 16 kHz frekansında 16 bitlik örnekleri oluşturmaktadır. Türkçe dâhil olmak üzere birçok dil için destek sağlamaktadır.

Whistler, Microsoft'ta geliştirilen bir MKS sistemidir. Ses üretimi için eklemeli sentezleme yöntemi kullanılmış, eğitim modülünde ise saklı markov modeli esas alınmıştır.

Aşağıdaki şekilde MKS sistemlerinin tarihsel zaman çizelgesini gösteren bir şekil verilmiştir.



Şekil 3.3: Tarihsel Zaman Çizelgesi

İngilizce için MKS sistemlerinin geliştirilmesinden sonra dünyada konuşulan diğer diller için de MKS sistemleri ortaya çıkmaya başlamıştır. Aynı sentez teknikleri kullanılmasına rağmen, uygulamalar, ilgili dilin özelliklerine göre değişiklik göstermiştir. Bugün, modern teknolojinin avantajlarından yararlanılarak, yeni konuşma sentezleme sistemleri geliştirilmeye devam edilmektedir.

3.2. Literatür Taraması

Türkçe MKS konusunda yapılan literatür taraması sonucunda Türkçe MKS ve bu konuyla alakalı olarak yapılan benzer çalışma sayısı (materyal) 28'dir. Bunların çalışma türleri; doktora tezi, yüksek lisans tezi, konferans bildirileri ve makalelerdir. Bu çalışmalar genel olarak Türkçe'nin biçim bilimsel yapısı gereği eklemeli sentezleme yöntemini kullanmışlardır.

Yapılan araştırmalar incelendiğinde Türkçe MKS sistemlerinin birçoğunda eklemeli sentezleme yöntemi kullanıldığı görülmektedir [1]. Son yıllarda yapılan çalışmaların büyük çoğunluğu doğal bir konuşma elde edebilmek adına süre ve ezgisel modelleme üzerine odaklanmıştır [5].

3.3. Araştırmanın Sınırlılıkları ve Dâhil Edilme Kriterleri

Araştırmanın sınırlılıkları doğrultusunda çalışma kapsamına dâhil edilen tez, makale ve bildiri sayısı 28'den 19'a düşmüştür. Bu sınırlılıklar aşağıda sıralanmıştır:

- 2002 yılı ve sonrasında yapılan çalışmalar araştırmaya dâhil edilmiştir.
- Özellikle bir uygulama geliştirilmiş ve kullanıcı üzerinde test edilmiş çalışmalarla, mevcut yazılımların iyileştirilmesi için fikir öne süren çalışmalar araştırmamızın kapsamındadır.
- Araştırmaya dahil edilecek çalışmalar; İngilizce veya Türkçe olarak yayınlanmış makale, konferans/sempozyum bildirileri ve yüksek lisans/doktora tezleri ile sınırlandırılmıştır.
- Araştırmaya dâhil edilen tezler, Yükseköğretim Kurulu Başkanlığı Tez Merkezi'nden alınmıştır ve erişime açık olanlar araştırma kapsamına dâhil edilmiştir.

3.4. Verilerin Toplanması

Bu bölümde, yukarıda sıralanan araştırma sınırlılıkları dâhilinde incelenen çalışmalar kısaca açıklanmıştır. Türkçe metin sentezlemeye yönelik çalışmalar iki grupta incelenebilir. İlk grupta yapılan çalışmalar DDİ modüllerinden faydalanan veya elde edilen sonuçla DDİ'ye ihtiyaç duyduğunu ifade eden çalışmalardır. İkinci grupta açıklanan çalışmalar ise klasik MKS yöntemlerini kullanan çalışmalardır. Bu doğrultuda, araştırmamız kapsamına giren 19 çalışma iki başlık altında incelenecektir. Tablo 3.2'de bu çalışmaların yılı, adı, yazarı ve türü verilmiştir.

Tablo 3.2: Türkçe MKS çalışmaları

Yıl	Çalışmanın Adı	Yazar	Çalışma Türü
2002	Türkçe Metin Seslendirme	Barış Eker	Yüksek Lisans Tezi
2007	Taşınabilir Cihazlar İçin Türkçe Metinden Konuşma Sentezleme Sistemi	İlker Ünaldı	Yüksek Lisans Tezi
2007	Görme Engelliler İçin Basılı Doküman Yorumlama ve Seslendirme Sisteminin Gerçekleştirilmesi	Emre Uzun	Yüksek Lisans Tezi
2008	Türkçe Metinler İçin Hece Tabanlı Konuşma Sentezleme Sistemi	Rıfat Aşlıyan, Korhan Günel	Makale

Tablo 3.2 (devam): Türkçe MKS çalışmaları

Yıl	Çalışmanın Adı	Yazar	Çalışma Türü
2009	Makine Öğrenme Algoritmalarıyla Türkçe Metin Seslendirme Sistemi Yazılımı	Zeliha Görmez	Yüksek Lisans Tezi
2009	Türkçe Metin Seslendirme	Kenan Güldalı	Yüksek Lisans Tezi
2010	Türkçe Metin Seslendirme	Tuncay Şentürk	Yüksek Lisans Tezi
2010	Metinden Konuşma Sentezleme	İbrahim Baran Uslu	Makale
2010	Türkçe Metin Seslendirme İçin Doğal Konuşma Sentezleme	Cavit Erdemir	Yüksek Lisans Tezi
2010	Türkçe Metinden Konuşma Sentezleme	Yücel Bicil	Yüksek Lisans Tezi
2011	Beyin Bilgisayar Arayüzleri İçin Türkçe Metinden Konuşma Sentezleme Sistemi	İlhami Sel, Davut Hanbay, Murat Karabatak	Sempozyum Bildirisi
2011	Görme Engelliler İçin Bilgisayar Kullanımının Etkinleştirilmesi, Erişilebilirlik ve Türkçe Hece Tabanlı Konuşma Sentezleme Sisteminin Geliştirilmesi	Güray Arık	Yüksek Lisans
2012	Görme Engelliler İçin Türkçe Metinden Konuşma Sentezleme Yazılımı Geliştirilmesi	Benian Tekindal, Güray Arık	Makale
2013	Türkçe Metinler İçin Hece Tabanlı Metinden Konuşma Sentezleme Sistemi	İlhami Sel	Yüksek Lisans Tezi
2013	Morfolojik Olarak Zengin Diller İçin Melez İstatistiksel/Birim Seçmeli Metinden Konuşma Sentezleme Sistemi	Ekrem Güner	Yüksek Lisans Tezi
2015	RC8660 Ses Sentezleyici ile Türkçe Metinden Konuşma Sentezleme	Timur Karamehmet	Yüksek Lisans Tezi
2016	Türkçe Metin Seslendirme	Tuncay Şentürk, Eşref Adalı	Makale
2016	DSP Kartı Kullanarak Türkçe Metinden Konuşma Sentezleme	Uğur Ayaz	Yüksek Lisans Tezi

3.4.1. Doğal Dil İşleme Modülü Kullanan Metinden Konuşma Sentezleme Çalışmaları

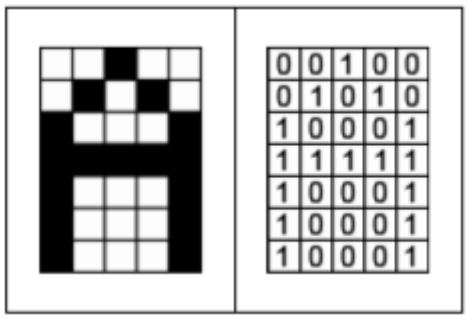
Bu grupta incelenen çalışmalar, metin analizi aşamasında DDİ modülünden faydalanmış, çalışmanın test aşamasında DDİ'den faydalanılması gerektiği sonucunu bulan veya yapay zekâ tekniklerini kullanan çalışmalardır. 19 çalışmadan 7'si bu grup kapsamına girmektedir.

- Uzun (2007), “Görme Engelliler İçin Basılı Doküman Yorumlama ve Seslendirme Sisteminin Gerçekleştirilmesi” isimli bir çalışma yapmıştır. Geliştirilen sistemde kullanılan yöntem ve algoritmaların tamamı C++ Builder ortamında uygulanmıştır. Geliştirilen sistem ile metin, grafik ve tablo gibi görüntülerin de sesli olarak aktarımı sağlanmıştır. Bunun için de Yapay Sinir Ağları (YSA) teknolojisine başvurulmuştur. Yapay zekâ algoritmalarına başvurulması açısından, bu çalışma diğer çalışmalardan farklılık oluşturmaktadır.

  Sistemin işleyişi şu şekildedir:

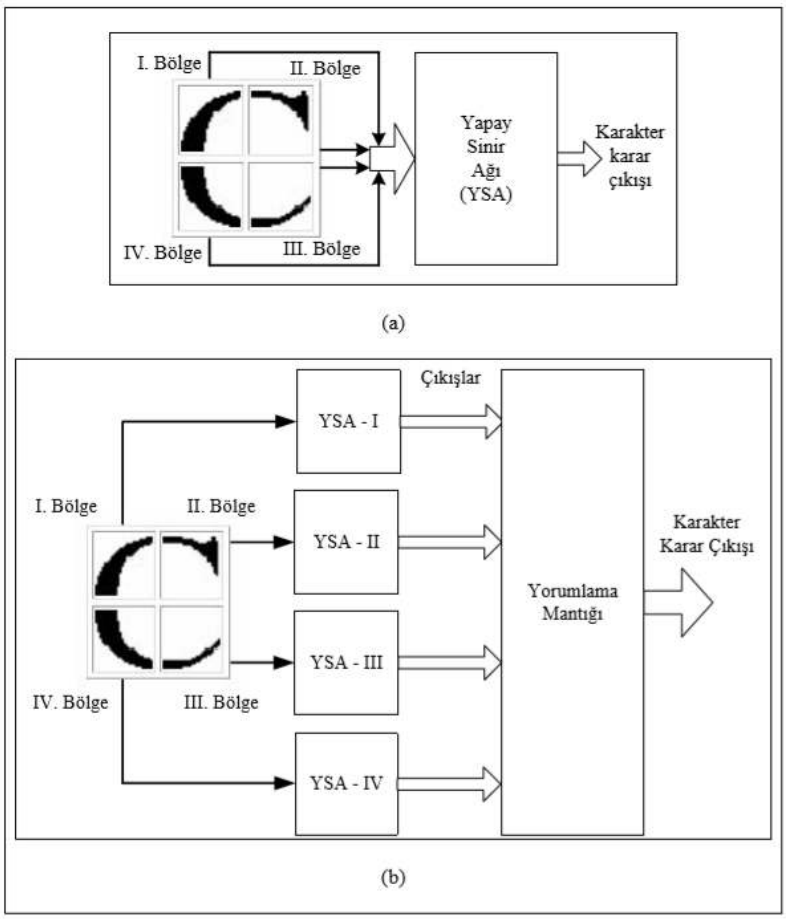
1. Tarayıcı ile doküman görüntüsünün alınması
2. Doküman görüntüsü analizi ve yorumlanması
3. Türkçe karakter tanıma
4. Türkçe metni seslendirme

Verileri modellemenin ve kullanmanın en kolay yolu dizi veya matris biçiminde ifade olduğu için bu yönteme başvurulmuştur. Modellemenin bir örneği aşağıda gösterilmiştir.



Şekil 3.4: Karakter Matrisi ve Veri Düzeni (Uzun, 2007)

Verilerin eğitilmesi kısmı için kullanılan dataset ile girdi metninin modellenmesi ve tanınması aşaması da görsel olarak aşağıda örneklenmiştir.

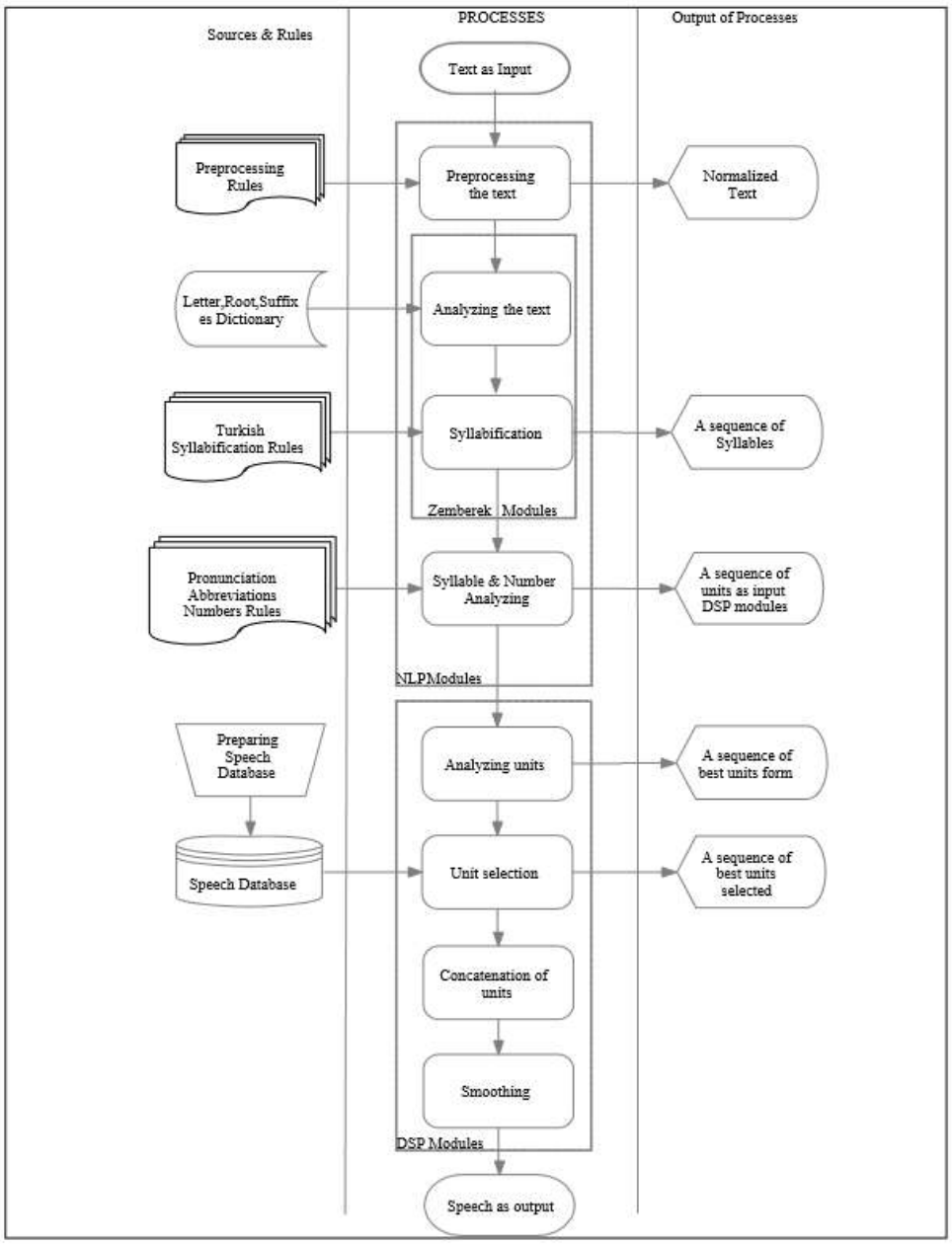


Şekil 3.5: Benzer Bölümlerin Belirlenmesi, Eğitimin Düzenlenmesi ve Karakter Görüntülerinin Farklı Sinir Ağları İle Tanınması (Uzun, 2007)

Çalışmaya benzer olarak geliştirilen diğer yazılımlardan farklı olarak bu sistemin grafik ve tablo tanıma ve yorumlama avantajının olması, bir adım önde olmasını sağlamıştır. Gelecekteki çalışmalar için, bu sistemin yardımcı donanım sistemleri ile birlikte tasarlanması gerektiği, bu sayede bireylerin sistemden daha kolay faydalanabileceği belirtilmiştir.

- Görmez (2009), gerçekleştirdiği "Makine Öğrenme Algoritmalarıyla Türkçe Metin Seslendirme Sistemi Yazılımı" isimli çalışmada; metni sese çeviren bir yazılım geliştirmiştir. Eklemeli sentezleme yöntemi kullanılmıştır. Birleştirme aşamasında yumuşak geçişlerin sağlanabilmesi için PSOLA algoritmasından yararlanılmıştır.

Araştırmacı, metin sentezlemenin bilgisayar bilimleri taksonomisinde DDİ altında olduğunu ve MKS'yi anlamak için DDİ'yi anlamak gerektiğini savunmuştur. Bu yüzden çalışmayı DDİ modülü ve dijital sinyal işleme modülü olarak iki modülde ele almıştır. Bu modüller aşağıdaki gibi ifade edilmiştir.

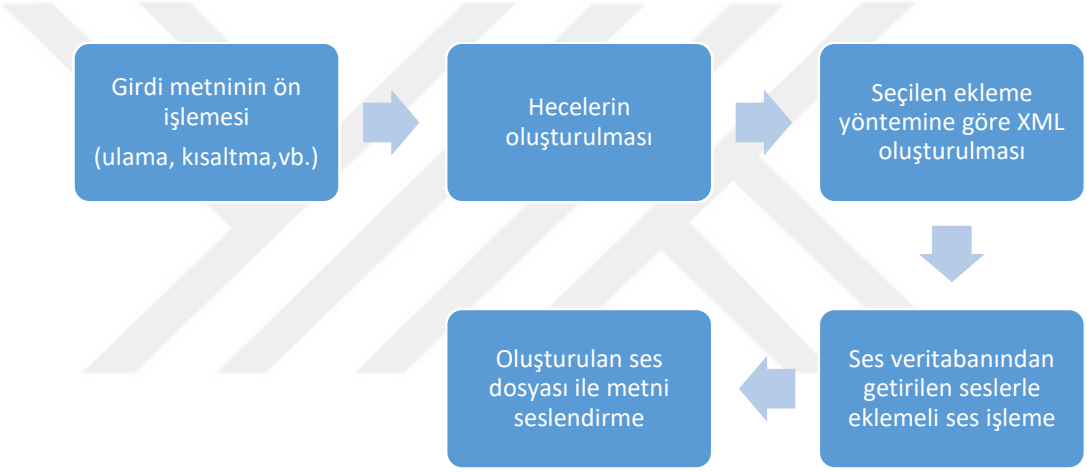


Şekil 3.6: Türkçe TTS Yapısı (Görmez, 2009)

Makine öğrenme algoritması olarak K * (Kstar) algoritması kullanılmıştır. Benzer eğitim örnekleri sınıfına dayandığı için bu algoritma tercih edilmiştir. Aynı zamanda veri madenciliği için makine öğrenme algoritmaları topluluğu olan WEKA (Waikato Environment for Knowledge Analysis) kullanılmıştır. Bu algoritmalar doğrudan bir veri kümesine uygulanabilmekte ve kendi Java kodumuzdan çağrılabilmektedir.

Bu yazılım, anlaşılabilirlik ve doğallık açısından ayrı ayrı testlere tabi tutulmuştur ve ümit verici sonuçlar alındığı belirtilmiştir. İlerideki çalışmalar için de ayrı ayrı değil, tek bir teste tabi tutulup sonuçların gözlenmesi gerektiği söylenmiştir.

- Şentürk (2010), gerçekleştirdiği "Türkçe Metin Seslendirme" isimli çalışmada hece eklemeli yöntem kullanmıştır. Sistemin akışı aşağıdaki gibidir.



Şekil 3.7: Sistemin Akış Şeması (Şentürk, 2010)

Oluşturulan XML dosyasında birleştirilen hecelerin uzunluk ve duraksama sürelerinin ayarlanabilmesi sağlanmıştır. Aynı çalışmada, görme engelliler için klavyeden girilen ifadeleri seslendiren metin düzenleme programı da hazırlanmıştır.

Geliştirilen sistemde anlaşılırlık ve ses seviyesinde olumlu yönde bulgular elde edilmiştir. Aynı çalışmada, görme engelliler için metin düzenleme programı hazırlanmıştır. DDİ çalışmalarının bu sisteme eklenmesiyle daha doğal seslerin elde edilebileceği vurgulanmıştır. Hecelerin kelimedeki yerine göre vurgusunun değişmesi konusuyla alakalı olarak frekans alanı üzerinde çalışmalar yapılması gerektiği de belirtilmiştir.

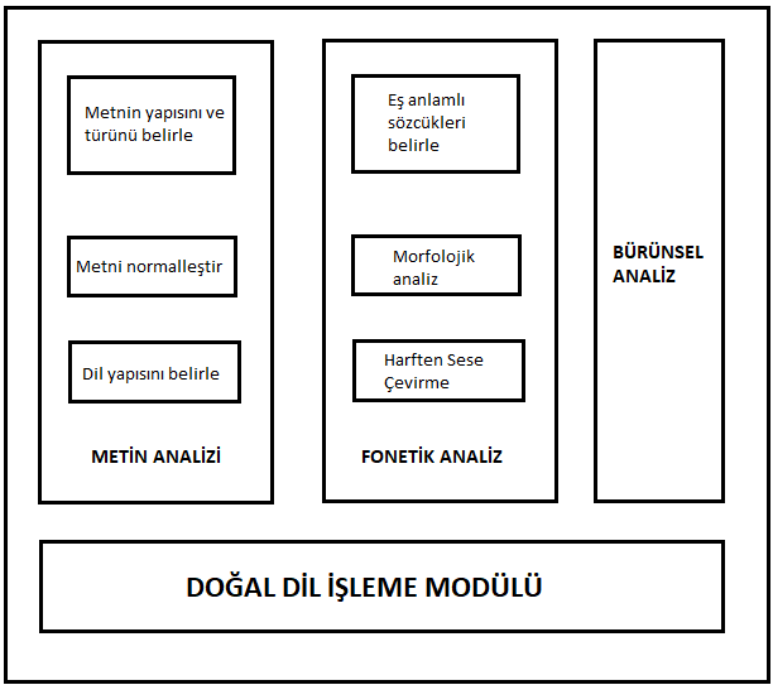
- Bicil (2010), gerçekleştirdiği "Türkçe Metinden Konuşma Sentezleme" isimli çalışmada; Türkçe ses kümesi oluşturmuş ve ikili ve üçlü sesleri kullanan eklemeli sentezleyici ile bir sentezleme sistemi geliştirmiştir. Çalışmada, MKS sistemlerinin, sinyal işlemeden önceki metin ön işleme ve normalizasyon aşamalarına ihtiyaç duyulduğundan, DDİ'nin kullanım alanlarından biri olarak kabul edilebileceği belirtilmiştir.

  Araştırmacı; MKS sistemini üç aşamada gerçekleştirmiştir:

  1. Ses veri tabanının hazırlanması
  2. DDİ modülü (dil ve metin analizi)
  3. Sentezleme modülü (konuşma üretimi)

  Ses veri tabanı oluşturulurken parça olarak üçlü seslerin kullanılması öngörülmüştür ve bunun için Türkçede en fazla geçen üçlü heceler tespit edilip veri tabanı hazırlanmıştır.

  Daha sonra DDİ modülü gerçeklenmiştir. Bu modülün şematik gösterimi aşağıda verilmiştir.

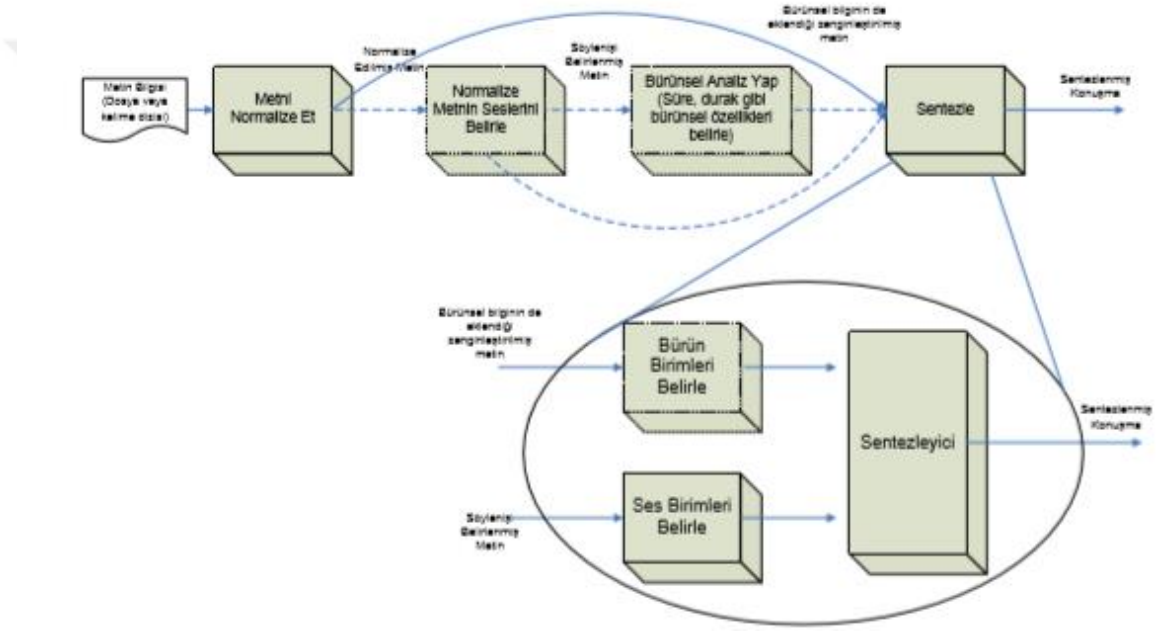


Şekil 3.8: DDİ Modülünün Şematik Gösterimi

Bürünsel analiz kısmında, metin içindeki hecelerin hangisine vurgu verileceği, hangi kısmının uzatılacağı, cümle tonlamasının nasıl olacağı belirlenir.

Sentezleme modülü ise ses veri tabanından çekilen uygun hecelerin birbirine eklenerek konuşmaya çevrildiği kısımdır. Tek başına kullanıldığında tonlamasız bir konuşma elde edilir fakat bürün üretimi ile beraber kullanıldığında bürünsel bir sentez elde etmek mümkün olmuştur. Seslerin birleştirilmesi aşamasında TD-PSOLA algoritması kullanılmıştır.

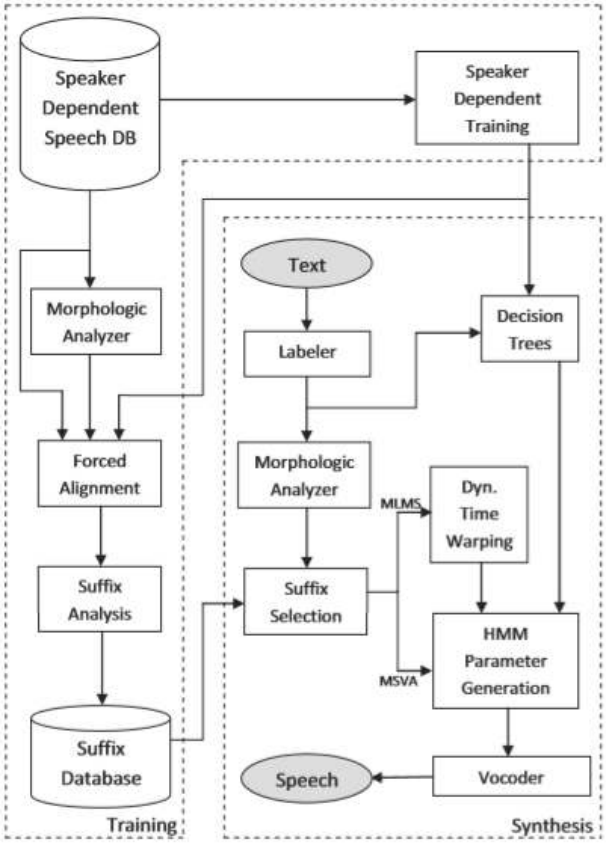
Sistemin tamamı düşünüldüğünde MKS'nin veri akış diyagramı aşağıdaki gibidir.



Şekil 3.9: MKS Veri Akış Diyagramı (Bicil, 2010)

Testler sonucunda sentezlenen konuşmanın anlaşılırlık oranı ortalamanın üzerinde tespit edilmiştir. Geliştirilen sistemin kalitesinin arttırılması için daha büyük bir ses veri tabanı kullanılması gerektiği ve seslerin profesyonel konuşmacılar tarafından kaydedilmesi gerektiği belirtilmiştir.

- Güner (2013), "Morfolojik Olarak Zengin Diller İçin Melez İstatistiksel/Birim Seçmeli Metinden Konuşma Sentezleme Sistemi" isimli çalışmasında; birim seçmeli MKS (BMKS) ve saklı markov modeli tabanlı MKS (SMKS) tekniklerini karşılaştırmış ve BMKS'de ortaya çıkan seslendirmedeki süreksizlik probleminin SMKS sistemlerinde yaşanmadığı tespit etmiştir. BMKS'de daha büyük veri tabanı kullanıldığı için de ses kalitesinin daha iyi olduğu belirtilmiştir. Bu yüzden morfolojik olarak zengin diller için bu iki sistemden de özellikler alan bir melez MKS sistemi önerilmiştir. Bu sistemin akış şeması aşağıdadır.



Şekil 3.10: Önerilen Sistemin Yapısı (Güner, 2013)

Geliştirilen sistemin literatürde İngilizce için yapılan sistemlerde raporlanan değerlerle benzer sonuçlar verdiği gözlemlenmiştir.

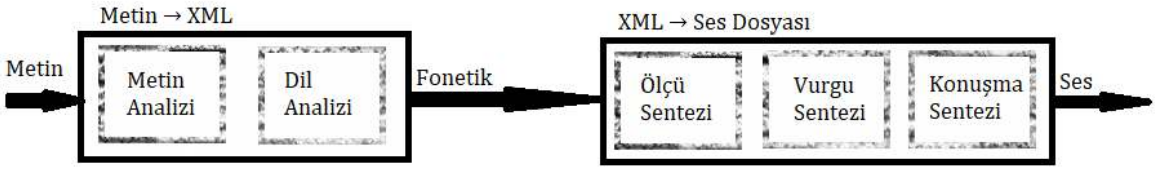
- Şentürk ve Adalı (2016), yazdıkları “Türkçe Metin Seslendirme” isimli makalede; iki farklı yöntem incelemiştir. Bu sistemde, çift-ses eklemeli yöntemde seslendirilen metnin doğallıktan uzak olduğu tespit edilince hece eklemeli yöntem kullanılmıştır. Aynı zamanda daha önce yapılan çalışmalarla karşılaştırılabilmesi ve yapılara müdahale edilebilmesini sağlamak amacıyla XML biçiminde kullanıcıya sunulmuştur. Özellikle bu nokta, makalenin önemli bir katkı sunmuş olduğu noktalardandır. Tablo 3.3’te metin işleme sürecinin son aşamasında üretilen XML dosyası gösterilmiştir.

Tablo 3.3: Üretilen XML Dosyası (Şentürk ve Adalı, 2016)

```
<?xml version="1.0" encoding="ISO-88599"?>
<metin>
  <cumle>
   <kelime vurguKatsayisi="1">
    <hece vurguKatsayisi="1">
     <ses sampa="j" sure="60"/>
     <ses sampa="e" sure="90"/>
    </hece>
    <hece vurguKatsayisi="1">
     <ses sampa="S" sure="60"/>
     <ses sampa="I" sure="90"/>
    </hece>
    <hece vurguKatsayisi="1">
     <ses sampa="l" sure="60"/>
     <ses sampa="a" sure="90"/>
    </hece>
    …
    <hece vurguKatsayisi="1.4">
     <ses sampa="d" sure="90"/>
     <ses sampa="u" sure="120"/>
    </hece>
   </kelime>
   <bekle sure="50"/>
  </cumle>
  <bekle sure="100">.</bekle> </metin>
```

Bu XML yapısında hangi hecenin ne kadar vurgulanacağı, kelimeler veya heceler arasında ne kadar bekleme süresi olacağı ve hangi hecenin ne kadar uzunlukta okunabileceği belirlenebilmektedir.

Bu çalışmada, metinden ses elde etmek için iki bileşen tasarlanmıştır. Metinlerden ses doyasına dönüşme işlemi esnasında yine XML kullanılmıştır. Bu bileşenler aşağıdaki şekilde gösterilmiştir.



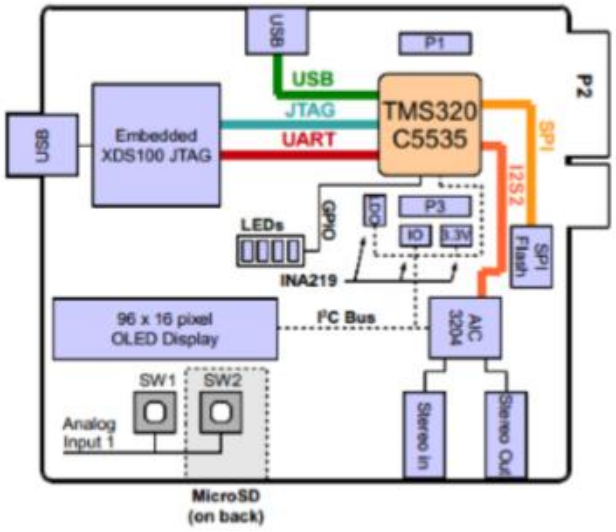
Şekil 3.11: Metinden Ses Elde Etme Bileşenleri

Aynı çalışmada; görme engelliler için metin düzenleyici program da geliştirilmiştir. Klavyeden girilen her ifadenin sesli olarak iletilmesi sağlanmıştır.

Farklı yaş gruplarından dinleyicilerle yapılan deneylerde hece eklemeli yöntemin kullanıldığı sistemin anlaşılırlık oranı %96.1 olarak ölçülmüştür.

Geliştirilen yazılıma ek olarak; hecenin cümledeki yerine göre farklı okunması konusuyla alakalı iyileştirici bir yol izlenmesi gerektiği, doğal bir konuşma elde edilmesi için de frekans alanı üzerinde çalışmalar yapılması gerektiği ve bunların çözümü için DDİ alanından destek alınması gerektiği belirtilmiştir.

- Ayaz (2016), "DSP Kartı Kullanarak Türkçe Metinden Konuşma Sentezleme" isimli çalışmasında; DSP kartı ile metinden konuşma üreten bağımsız bir sistem geliştirmiştir. DSP kartı geliştirme kitinin blok diyagramı aşağıda gösterilmiştir.



Şekil 3.12: Geliştirme Kitinin Blok Diyagramı (Ayaz, 2016)

Bu kit, bilgisayara entegre edilmiş ve seslendirilecek metin direkt kart üzerinden aktarılmıştır.

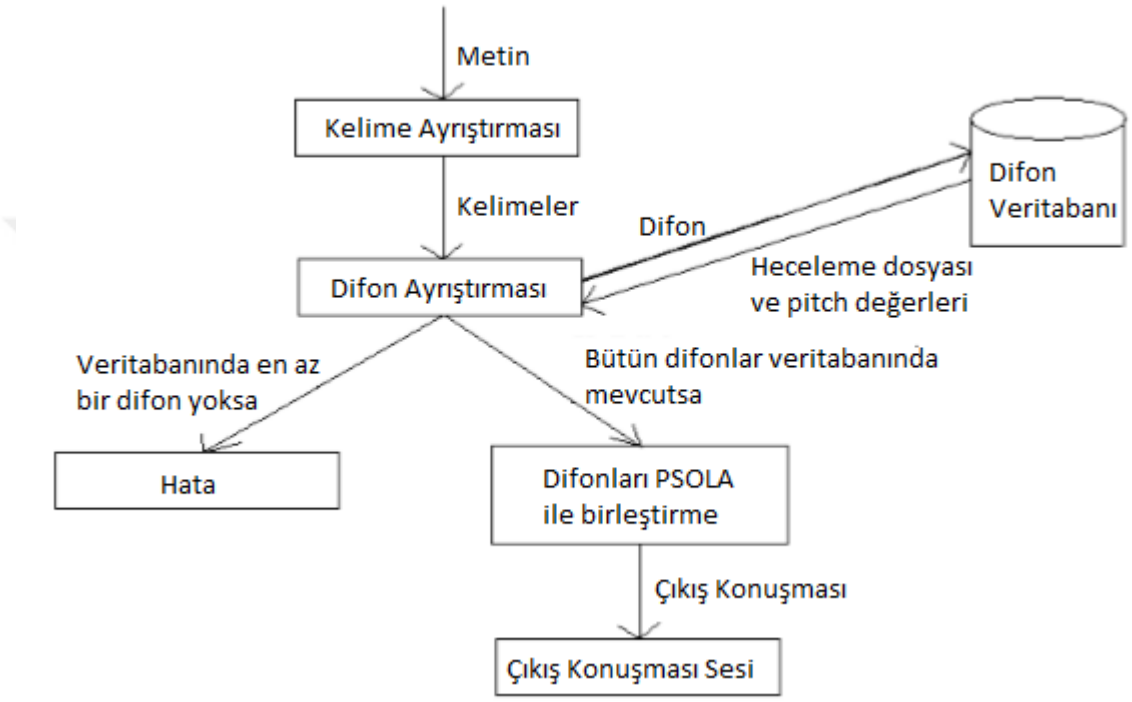
Yazılımın geliştirilmesi Matlab ortamında sağlanmıştır. Girdi olarak alınan metin, metin çözümlemesi işleminden geçirilip hecelenmiş, sentezlenmek üzere DSP kartına gönderilmiştir.

Anlaşılırlık ve doğallık açısından ortalama bir sonuç alındığı belirtilmiştir. Daha az gürültüsüz ortamda kaydedilen seslerin kullanılması gerektiği, konuşmaya vurgu ve tonlamanın da eklenmesi gerektiği, konuşma sentezi ve DDİ çalışmalarının sayıca artması gerektiği vurgulanmıştır.

3.4.2. Klasik Yöntemleri Kullanan Metinden Konuşma Sentezleme Çalışmaları

Bu grupta incelenen çalışmalar daha önce literatürde kullanılan MKS yöntemlerini kullanan ve bu yöntemlere ek olarak iyileştirme algoritmaları geliştiren çalışmalardır. Ses veri tabanında geliştirmeler yaparak verimi arttırmayı hedefleyen çalışmalar da bu grupta incelenmiştir. Araştırma kapsamındaki 19 çalışmadan 12'si bu grup kapsamındadır. Klasik yöntemlerdeki sonuçlar incelendiği zaman araştırmacıların çoğunluğu vurgu ve tonlama konularında iyileştirme yapılması gerektiği sonucuna ulaşmışlardır.

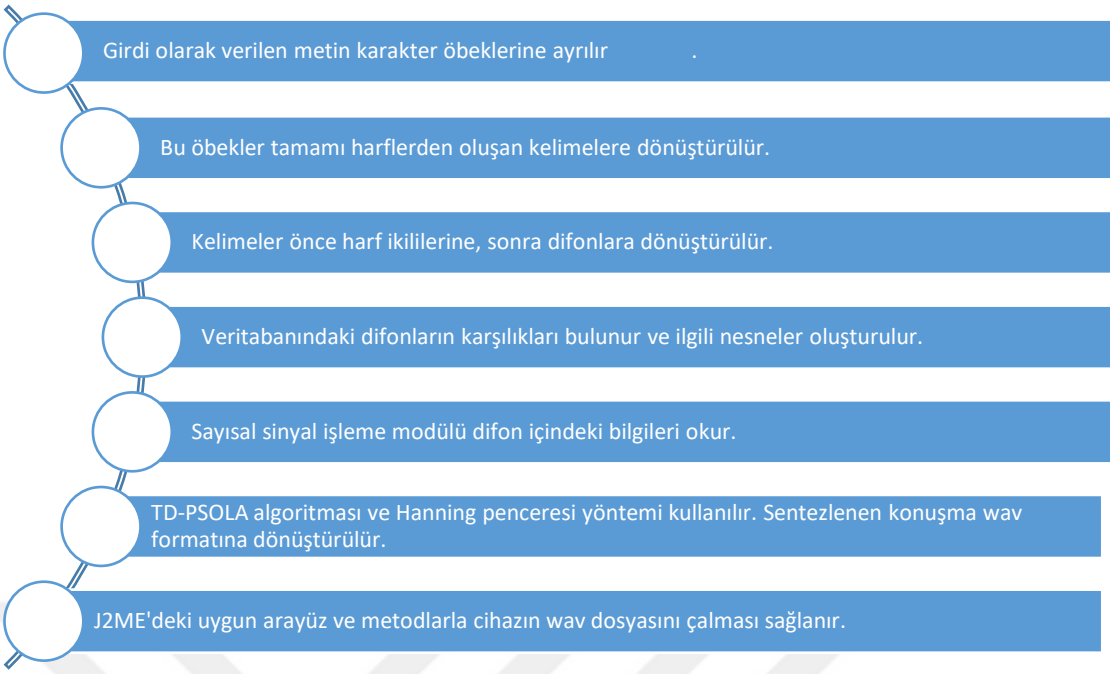
- Eker (2002), "Türkçe Metin Seslendirme" isimli çalışmasında; girdi olarak metni alan ve çıktı olarak bu metne karşılık gelen seslerin verildiği bir sistem tasarlamıştır. En küçük birim olarak difonlar kullanılmıştır. İkili fonem birleştirme tekniği kullanılmıştır. Geliştirme ortamı olarak Matlab tercih edilmiştir. Seslerin birleştirilmesi için de PSOLA algoritması kullanılmıştır. Sistemin genel işleyişi aşağıdaki gibidir.



Şekil 3.13: Sistemin Genel Yapısı (Eker, 2002)

Test aşamasında, uzun cümlelerin seslendirilmesinde karşılaşılan performans zayıflığından dolayı sistemin daha hızlı olması gerektiği ve konuşmanın daha doğal olması için iyi bir metin işlemcisine ihtiyaç duyulduğu belirtilmiştir.

Ünaldı (2007), gerçekleştirdiği "Taşınabilir Cihazlar İçin Türkçe Metinden Konuşma Sentezleme Sistemi" isimli çalışmada; J2ME platformunu destekleyen bir mobil cihaz üzerinde çalışabilen ve kısa mesajları sesli olarak okuyabilen bir yazılım geliştirmiştir. Difonları birleştirerek sentezleyen TD-PSOLA algoritması ile Java 2 Platformu kullanılmıştır. J2ME açık bir standart olmasından dolayı avantajlıdır. Sistemin çalışma mantığı aşağıda gösterilmiştir.



Şekil 3.14: Çalışmanın Blok Şeması

Sistemin daha hızlı çalışabilmesi için daha yüksek kapasiteli bir işlemciye gereksinim duyulduğu ve bu işlemci sayesinde daha büyük veri tabanları ile çalışılabileceği belirtilmiştir.

- Aşlıyan ve Günel (2008), tarafından yazılan "Türkçe Metinler İçin Hece Tabanlı Konuşma Sentezleme Sistemi" isimli makalede; metni konuşmaya çeviren bir sistem tasarlanmıştır. Geliştirilen sistemde en küçük ses birimi olarak heceler kullanılmıştır. Tasarlanan heceleme algoritması C++ dili ile kodlanmıştır. Veri tabanına kaydedilen seslerin birleştirme aşamasında ise TD-PSOLA algoritması kullanılmıştır.

  Hece tabanlı konuşma sentezleme sistemi üç ana işlemden geçirilir. İlk aşama, girdi metninin ön işlemden geçirilmesi ve hecelere ayrılmasıdır. İkinci işlem, hece ses sinyallerinin sırasıyla veritabanından alınması; son aşama ise ses sinyallerinin TD-PSOLA algoritmasıyla istenilen hızda okutulmasıdır. Sistemin kullanıcı arayüzü aşağıda gösterilmiştir.

Şekil 3.15: Türkçe Metin Okuma Programının Arayüzü (Aşlıyan ve Günel, 2008)

Bu arayüz sayesinde kullanıcılar zaman ölçeği çarpanını ve ses yüksekliğini ayarlayabilmektedir.

Test edilen sistem ortalama bir puan almıştır. Seslendirilen metnin iyi anlaşıldığı fakat vurgu ve tonlamalarda iyileştirmeler yapılması gerektiği vurgulanmıştır.

- Yılmaz (2009), tarafından yapılan "Türkçe Metinden Konuşma Sentezleme Uygulamaları İçin Bir Veri Sözlük Seti ve Yazılım Çerçevesi" isimli çalışmada; MKS sistemleri için altyapı niteliğinde bir veri sözlük seti ve yazılım çerçevesi tanıtılmıştır. Bu altyapı; metin içerisindeki kısaltmaların, sayıların vs. doğru bir şekilde okunabilmesini sağlamaktadır.

  Konuşma dilinde, aynı harflerin farklı uzunlukta okunabilmesinden dolayı, hece veri tabanı oluşturulurken yeni semboller tanımlanması ve veri tabanının ona göre oluşturulması gerektiği savunulmuştur. Önerilen yeni sembollerin kullanım örnekleri Tablo 3.4'te gösterilmiştir.

Tablo 3.4: Önerilen Sembollerin Kullanım Örnekleri (Yılmaz, 2009)

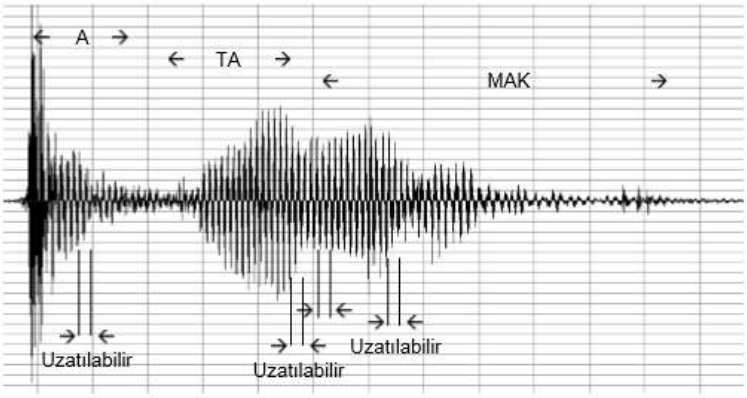
	a	**e**	**i**	**o**	**u**	**ü**
Normal	aba	elek	inat	otomobil	uzun	ülkü
Uzun	âşık	têmin	îkaz	limônî	ûdî	mÿmin
İnce	láma, káğıt, gávur	-	-	lómboz	billúr, sükút	-
Uzun ve ince	lãle, kãbus, yegãne, mânã	-	-	-	ulýfe, sükýnet	-
Yumuşak	ihmàl,itaàt	-	-	gòl	kabùl	-
Geniş	-	dirhëm	-	-	-	-

Ses veri tabanı, önerilen şekilde kullanılırsa, bütün kelimeler uygun tonlamada okunmuş olacağından anlam kargaşalarının önüne geçilmiş olacaktır.

Geliştirilen ürünün gelecekteki araştırmacıların hizmetine sunulması hedeflendiği belirtilmiştir.

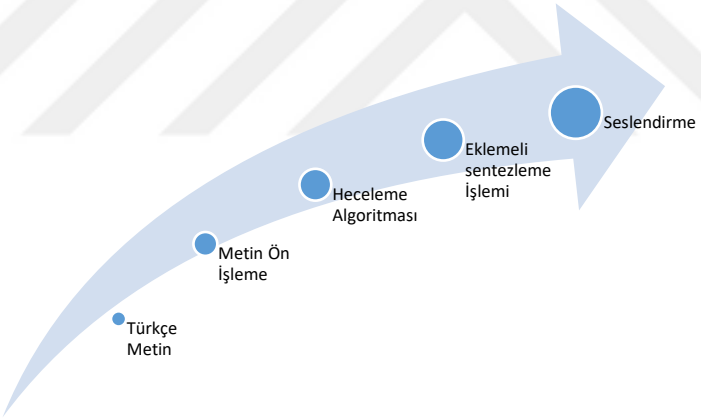
Güldalı (2009), tarafından yapılan Türkçe Metin Seslendirme isimli çalışmada; ses birimi olarak heceyi kullanan ve eklemeli sentezleme tekniği kullanan bir sistem geliştirilmiştir.

Bu çalışmada, Türk Dil Kurumu sesli sözlük ses dosyalarına uygun şekilde kesimlemeler yapılmıştır. Kaydedilecek hecelerin parametre dosyaları için hece başlangıç-bitişleri ve uzatılabilen seslerin periyodunun başlangıç-bitişleri için işaretleme yapılır. Uzatılabilir periyotlu sesler, sesli harfler ile bazı nefesli sesler olabilmektedir. İşaretleme işleminin bir örneği aşağıda gösterilmiştir.



Şekil 3.16: İşaretleme Örneği (Güldalı, 2009)

Girdi metni alınıp boşluklardan ve satır başlarından arındırılmıştır. Tek tek kelimeler tespit edilmiştir ve bu kelimeler hecelerine ayrılıp liste yapısında tutulmuştur. Kelimedeki her hece için veri tabanından ilgili ses dosyası getirilmiştir. Eklemeli sentezleme algoritmasıyla da oluşturuşan ses dosyasının seslendirme işlemi yapılmıştır. Bu akışın görsel olarak anlatımı aşağıdadır.



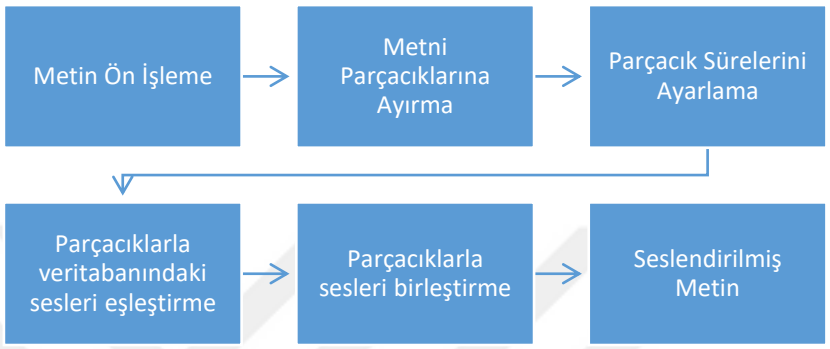
Şekil 3.17: Hece Birleştirmeli Sistem

Yapılan testlerde Anlaşılırlık açısından ortalamanın üzerinde sonuçlar elde edilmiştir. Ancak vurgu ve tonlama açısından başarının çok yüksek olmadığı ve ses veri tabanının sabit ses seviyesi ile hazırlanmış olması gerektiği belirtilmiştir.

Erdemir (2010), tarafından yapılan "Türkçe Metin Seslendirme İçin Doğal Konuşma Sentezleme" isimli çalışmada; metni seslendirmeye yarayan bir çalışmanın adımları açıklanmıştır.

Parçacık birimi olarak “en çok iki harfli hece” mantığı kullanılmıştır ve bu parçacıkların art arda eklenmesiyle sentezleme işlemi gerçekleştirilmiştir. Özel parçacık süre belirleme algoritmalarıyla hecelerin sürelerinin ayarlanabilmesi hedeflenmiştir. Her parçacık türüne karşılık gelen bir ses veri tabanı oluşturulmuştur.

Sistem C++ ve Java dilleriyle Matlab üzerinde geliştirilmiştir. Sistemin çalışma mantığı şu şekildedir:



Şekil 3.18: Türkçe Metin Seslendirme Sisteminin Çalışma Mantığı

Oluşturulan sistem hazır paket kodlarla hazırlanmadığı için Matlab yazılımından bağımsız olarak çalışan Java ve C++ dilleriyle yazılmış bir programa dönüştürülebileceği ve çoğu platformda kullanılabileceği belirtilmiştir. Ayrıca bu çalışmanın Matlab kodları tez içerisinde verilmiş ve araştırmacıların yararlanabilmesi sağlanmıştır.

Yapılan testlerin sonuçlarına göre anlaşılırlık %91, doğallık ise %74 olarak belirlenmiştir. Doğallığın arttırılması için de açıklanan kurallar çerçevesinde çok gelişmiş ses kütüphanelerinin oluşturulması gerektiği belirtilmiştir (Erdemir, 2010).

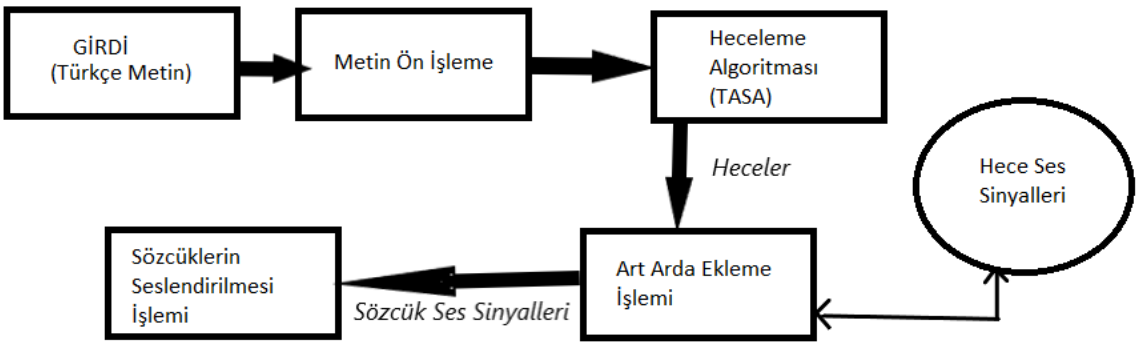
Uslu (2010), tarafından yazılan “Metinden Konuşma Sentezleme” isimli çalışmada; temel bir MKS sisteminin yapısı incelenmiş ve hayata geçirilmesi gereken uygulamalar tespit edilmiştir.

MKS sistemleri dil çözümleme ve sinyal işleme olarak iki modülde ele alınmıştır. Dil çözümlemesinde metin ön incelemesi ve metnin parçalara ayrılması, sinyal işleme modülünde ise sayısal sinyal işleme işlemi ve seslendirme kuralları doğrultusunda seslendirme işlemi yapıldığından bahsedilmiştir.

Elde edilen konuşmanın kalitesinin ve doğallığının önemli bir faktör olduğu belirtilmiş ve ileride yapılacak olan çalışmalarda özellikle bu konulara eğilim gösterilmesi gerektiği vurgulanmıştır.

Sel ve diğ. (2011), tarafından yapılan "Beyin Bilgisayar Arayüzleri İçin Türkçe Metinden Konuşma Sentezleme Sistemi" isimli çalışmada; konuşma veya hareket etme yeteneği olmayan engellilere yönelik bir Türkçe MKS sistemi geliştirilmiştir.

Bu sistem, seslendirmeyi yapabilmek için iki tip metin girişi almaktadır. Kullanıcı kelimeleri klavyeden kendisi yazabildiği gibi, listbox içerisindeki hazır komutları da kullanabilmektedir. Bu sistemin yazılımı için Microsoft.NET 2.0 ve MS Visual C# ortamı ve eklemeli sentezleme yöntemi kullanılmıştır. Sistemin işleyiş şeması aşağıdaki gibidir.



Şekil 3.19: Sistemin İşleyiş Mantığı

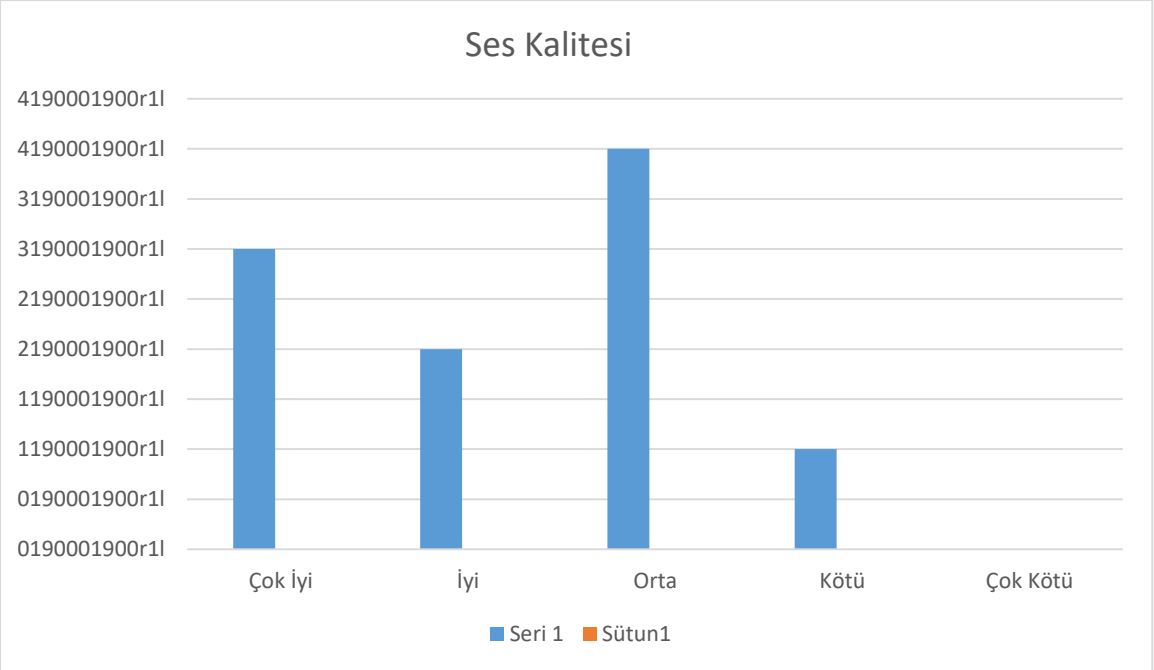
Girdi olarak alınan metin ön işlemeden geçirilerek, heceleme sistemine gönderilir. Bu algoritmanın %100 doğru çalıştığı gözlemlenmiştir. Oluşturulan hecelerin ses sinyalleri veri tabanından alınır ve art arda ekleme işlemi ile sözcüğün ses sinyali oluşturulur.

Bu sistemde, girdi olarak alınan metnin anlamsal olarak incelenmediği ve bu yüzden Türkçeye özgü ses değişimlerinin kaliteli bir şekilde okunmadığı saptanmıştır. Mevcut MKS sistemlerinde nümerik ifade, kısaltmalar, ses değişimleri gibi farklı telaffuz edilen ifadelerin pek dikkate alınmadığı da eklenmiştir. Gelecekte bu konulara dikkat edilmesi gerektiği ve engellilere yönelik yapılan araçların maliyetinin düşürülmesi gerektiği belirtilmiştir.

- Arık (2011), "Görme Engelliler İçin Bilgisayar Kullanımının Etkinleştirilmesi, Erişilebilirlik ve Bir Türkçe Hece Tabanlı Konuşma Sentezleme Sisteminin Geliştirilmesi" isimli çalışmasında; eklemeli sentezleme yöntemi kullanarak hece tabanlı bir MKS sistemi geliştirmiştir.

Kesimleme işlemi sırasında seslerin kaydı ve düzenlenmesi için Audacity yazılım paketi kullanılmıştır. Bu yazılım açık kaynak kodlu olduğu için avantajlıdır.

Test aşamasında, 10 kullanıcıya programdaki konuşmanın kalitesi sorulmuştur. Alınan sonuçlar aşağıdaki grafikteki gibidir.



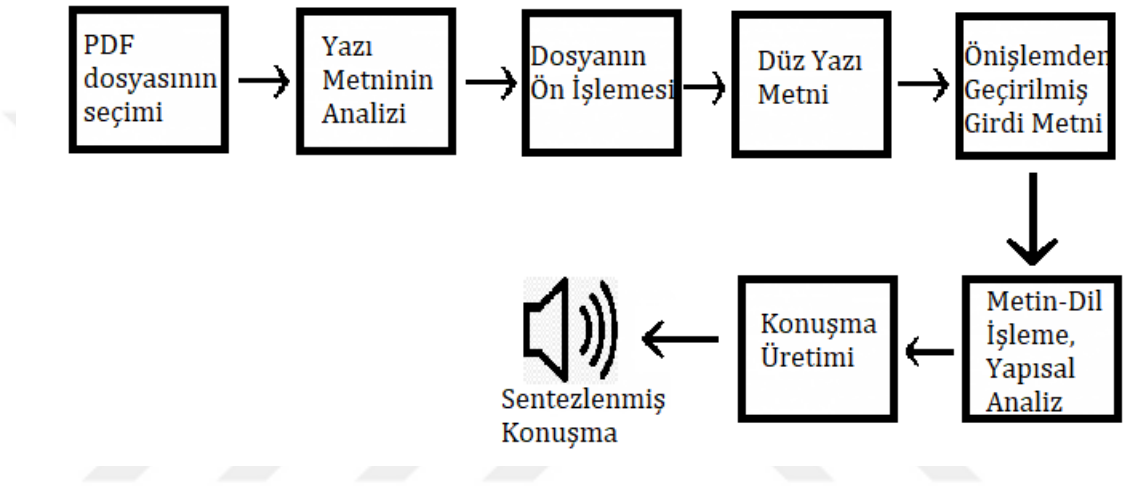
Şekil 3.20: Programdaki Konuşmanın Kalite Grafiği

Yapılan testlerde kalitenin ortalamanın üzerinde olduğu saptanmıştır. Kullanıcı üzerinde yapılan testlerde ise ses kalitesinin daha iyi olması gerektiği, seslerin diksiyonu düzgün bir konuşmacı tarafından seslendirilmesi gerektiği ve vurgusal anlamda iyileştirilmeler yapılması gerektiği sonuçları elde edilmiştir.

Tekindal ve Arık (2012), tarafından yapılan "Görme Engelliler İçin Türkçe Metinden Konuşma Sentezleme Yazılımı Geliştirilmesi" isimli çalışmada; eklemeli sentezleme yöntemini kullanan bir metin okuyucu yazılımı geliştirilmiştir. Sistemde Visual C#.NET programlama dili ile Visual Studio 2008 Professional kullanılmıştır.

Çalışma, elektronik ortamda bulunan bir dokümanı insan sesine yakın bir ses ile konuşmaya çevirebilmektedir. Bunun için yazılım içerisine entegre edilebilen; sesli harfle başlayan kelimeler ve sessiz harfle başlayan kelimeler için kullanılabilen iki ayrı algoritma geliştirilmiştir. Aynı heceleme işlemi, diğer MKS çalışmalarında tek bir algoritmayla yapılabilmektedir. Bu durum bu çalışma için bir dezavantajdır.

Elektronik ortamdaki PDF uzantılı dosyanın ön işleme süreci ve devamındaki konuşma sentezleme süreci, aşağıdaki diyagramda gösterilmiştir.

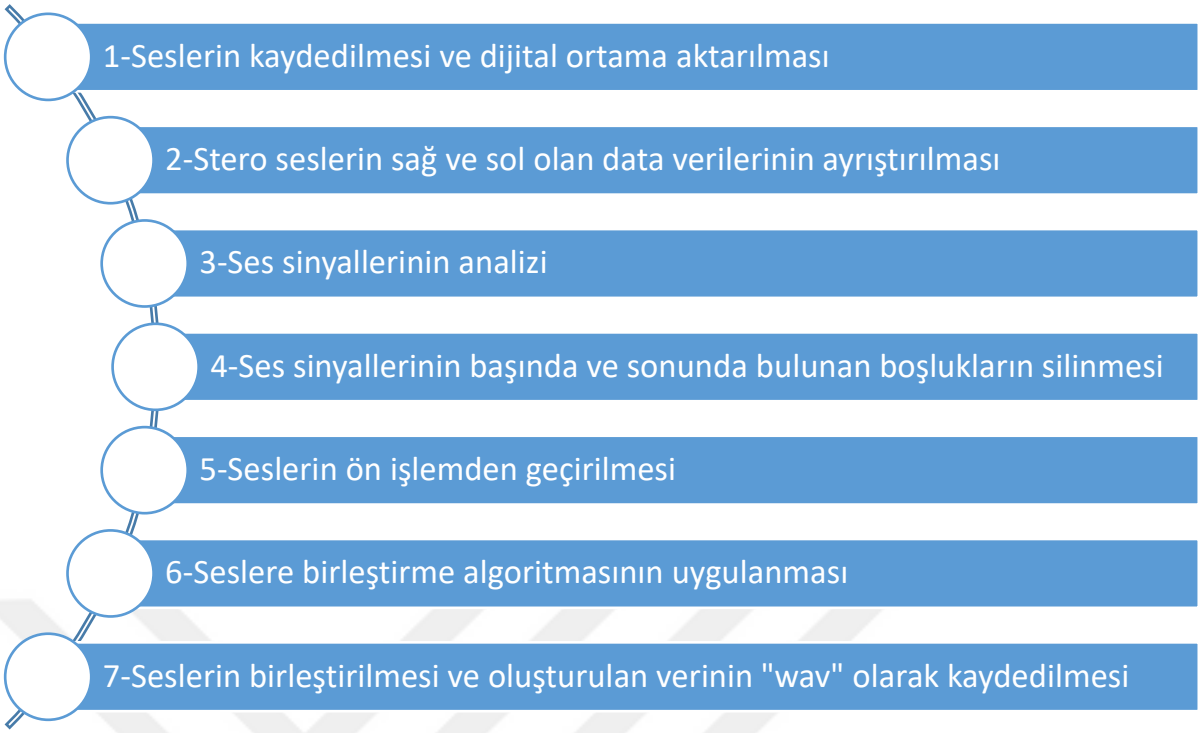


Şekil 3.21: Ön İşleme ve Konuşma Sentezleme Süreci

Görme kaybı olan 10 kişi üzerinde testler yapılmış ve vurgunun daha fazla iyileştirilmesi gerektiği, aynı zamanda hece ses dosyalarının kayıtlarının daha kaliteli bir ses kayıt ortamında yapılması gerektiği saptanmıştır.

- Sel (2013), tarafından yapılan "Türkçe Metinler İçin Hece Tabanlı Metinden Konuşma Sentezleme Sistemi" isimli çalışmada; eklemeli sentezleme yöntemini kullanan bir MKS yazılımı geliştirilmiştir. Temel olarak iki sistem oluşturulmuştur. İlk sistemde sentezleme işlemi hece tabanlı olarak gerçekleştirilmiştir. İkinci sistem ise nesne tabanlı hazırlanmış ve seslendirilecek bölüm olarak difon veritabanı kullanılmıştır.

  İlk sistem Matlab ile hazırlanmıştır. Birleştirilecek parça olarak heceler kullanılmıştır ve bu heceler önceden seslendirilerek kaydedilmiş, Hanning penceresinden geçirilen seslerin frekans olarak uyumlu olması sağlanmıştır. Birleştirme algoritması olarak ise SOLA algoritması kullanılmıştır. Bu sayede doğal konuşmaya yakın bir konuşma elde edilmeye çalışılmıştır. Bu sistemin blok şeması aşağıda gösterilmiştir.

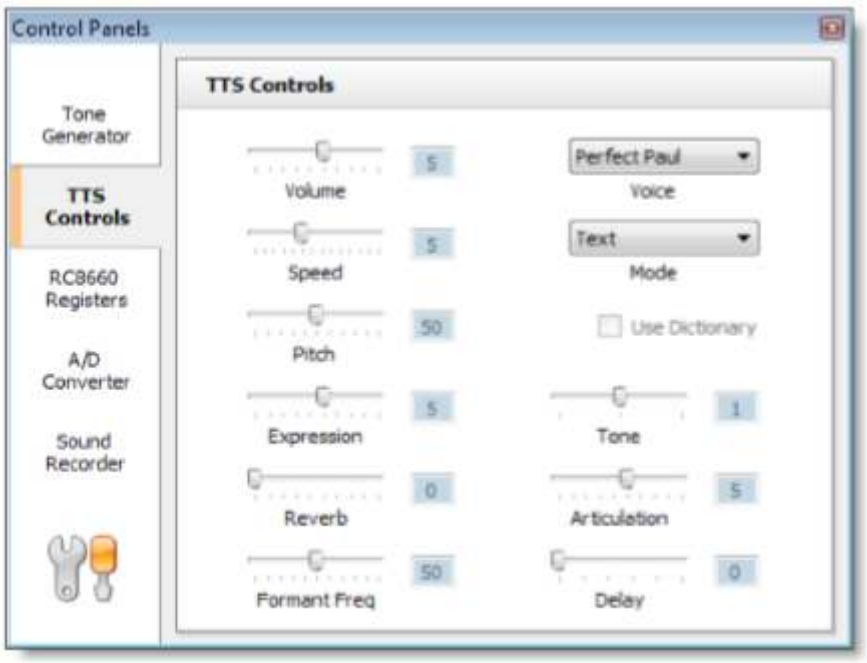


Şekil 3.22: Sistemin Blok Şeması

İkinci sistem ise C# ile hazırlanmıştır. İlk sisteme göre daha kullanışlı olduğu fakat robotik bir ses sentezlendiği belirtilmiştir. Ulama, ünsüz benzeşmesi, ünsüz yumuşaması gibi Türkçeye özgü kuralların sisteme dahil edilmesi gerektiği, vurgu ve tonlamalarda eksiklikler olduğu ve bunların giderilmesi gerektiği belirtilmiştir. Bu konuların çözüme kavuştuğu bir sistemin henüz geliştirilmediği de özellikle vurgulanmıştır.

- Karamehmet (2015), tarafından yapılan "RC8660 Ses Sentezleyici ile Türkçe Metinden Konuşma Sentezleme" isimli çalışmada; yeni bir harici sözlük tanımlanmış ve doğru heceleme için gereken kurallar bu sözlüğe entegre edilmiştir. Aynı zamanda vurgu ve tonlama için gereken kurallar da tanımlanmıştır.

  RC 8660 kartı, çıkış sesini kontrol edebilen bir özelliğe sahiptir. Ses seviyesi, hız, karakter aralığı, yankı, ton gibi etkenler istenilen düzeyde ayarlanabilmektedir. Aşağıda, RC8660'ın kontrol panelleri gösterilmiştir.



Şekil 3.23: TTS Kontrol Paneli (Karamehmet, 2015)

Uygulama için RCStudio ortamı kullanılmıştır. Metin kutucuğuna yazılan girdinin sesli olarak iletimi sağlanmaktadır. TTS kontrol paneli RCStudio ortamının içerisinde bulunmaktadır.

Geliştirilen sistemin engellilere yardımcı olabilmesi için diğer sistemlere (e-kitap okuyucuları, ATM makineleri, navigasyon sistemleri vs.) de uygulanması gerektiği belirtilmiştir.

3.5. Verilerin Karşılaştırılması

Çalışma yapıları, kullandıkları algoritmalar, heceleri/sesleri birleştirme ve kesimleme yöntemleri vb. açıklanan çalışmaların tablo halinde gösterimi bu bölümde tablo 3.5'te verilmiştir. Verilerin tabloya dökülmesiyle araştırma kapsamındaki 19 çalışmanın karşılaştırmasının daha kolay yapılması sağlanmıştır. Çalışma içerisinde belirtilmeyen veya kullanılmayan ifadeler "-" sembolü ile gösterilmiştir.

Tablo 3.5: Çalışmaların Karşılaştırılması

Yıl	Yazar	İlk Gupta mı? (DDİ)	İkinci Grupta mı? (Klasik)	Sentezleme Yöntemi	Birleştirme Algoritması	Programlama Dili ve Geliştirme Ortamı
2002	Barış Eker	-	Evet	İkili fonem birleştirme	PSOLA	MATLAB
2007	İlker Ünaldı	-	Evet	Difonları birleştirme	TD-PSOLA	Java 2 – J2ME
2007	Emre Uzun	Evet	-	Heceleri birleştirme	-	C++ Builder
2008	Rıfat Aşlıyan, Korhan Günel	-	Evet	Heceleri birleştirme	TD-PSOLA	C++
2009	Asım Egemen Yılmaz	-	Evet	Yeni hece türleri tanımlama	-	-
2009	Zeliha Görmez	Evet	-	Eklemeli sentezleme	PSOLA	K * (Kstar) algoritması
2009	Kenan Güldalı	-	Evet	Eklemeli sentezleme	-	C
2010	Tuncay Şentürk	Evet	-	Eklemeli sentezleme	-	Eclipse SpringSource Tool Suite – Visual C++
2010	Cavit Erdemir	-	Evet	Eklemeli sentezleme	Özel parçacık süre belirleme algoritması	C++, Java - MATLAB
2010	İbrahim Baran Uslu	-	Evet	-	-	-
2010	Yücel Bicil	Evet	-	Eklemeli Sentezleme	TD-PSOLA	-
2011	İlhami Sel, Davut Hanbay, Murat Karabatak	-	Evet	Eklemeli Sentezleme	-	Microsoft.NET 2.0 - MS Visual C#
2011	Güray Arık	-	Evet	Eklemeli sentezleme	-	Visual C#.NET
2012	Benian Tekindal, Güray Arık	-	Evet	Eklemeli sentezleme	-	Visual C#.NET - Visual Studio 2008

Tablo 3.5 (devam): Çalışmaların Karşılaştırılması

Yıl	Yazar	İlk Gupta mı? (DDİ)	İkinci Grupta mı? (Klasik)	Sentezleme Yöntemi	Birleştirme Algoritması	Programlama Dili ve Geliştirme Ortamı
2013	İlhami Sel	-	Evet	Eklemeli Sentezleme	SOLA	Matlab, C#
2013	Ekrem Güner	Evet	-	-	-	saklı markov modeli tabanlı
2015	Timur Karamehmet	-	Evet	Eklemeli sentezleme	-	RCStudio
2016	Tuncay Şentürk, Eşref Adalı	Evet	-	Hece eklemeli yöntem	XML dosyası	-
2016	Uğur Ayaz	Evet	-	DSP Kartı	-	Matlab

4. BULGULAR

Bu bölümde, yapılan araştırma sonucu elde edilen bulgular verilmiştir. Belirtilen araştırma sınırlılıklarına göre ülkemizde yapılmış olan Türkçe MKS sistemlerinden 19 tanesinin verileri, sonuçları, eksiklikleri ve DDİ ile olan bağlantıları tablolarla açıklanmıştır. Aynı zamanda bu çalışmaların kaç tanesinin direkt olarak engellilere yönelik geliştirilmiş olduğu da yüzdelerle ifade edilmiştir.

4.1. Çalışmaya Ait Betimleyici Veriler

Araştırmaya dâhil edilmiş olan çalışmaların frekans ve yüzde değerleri, yıllara göre tablo 4.1'de verilmiştir.

Tablo 4.1: Çalışmaların Yıllara Göre Frekans ve Yüzde Değerleri

Yıl	Frekans	Yüzde Değeri
2002	1	5.263 %
2007	2	10.526 %
2008	1	5.263 %
2009	3	15.789 %
2010	4	21.052 %
2011	2	10.526 %
2012	1	5.263 %
2013	2	10.526 %
2015	1	5.263 %
2016	2	10.526 %

Araştırma kapsamına alınan 19 çalışmanın yıllara göre dağılımı yukarıdaki gibidir. Tablo 4.1'e göre; 2002 ve sonrasında yapılan çalışmaların en fazla yapıldığı yıl; 4 çalışma ile 2010 yılıdır. 2010 yılındaki çalışmalar Türkçe MKS sistemlerine en fazla katkıda bulunan yıl (%21.052) olarak görülmektedir.

Tablo 4.2: Çalışmaların Yayın Türüne Göre Frekans ve Yüzde Değerleri

Yayın Türü	Frekans	Yüzde Değeri
Tez	13	68.421 %
Makale	5	26.315 %
Bildiri	1	5.263 %

Tablo 4.2'de görüldüğü gibi Türkçe MKS sistemleriyle alakalı yapılan çalışmaların %68'lik bir kısmını 13 çalışma ile tez çalışmaları oluşturmaktadır.

Araştırma kapsamındaki 19 tezin hiçbirinde direkt olarak DDİ tekniklerinin kullanıldığı görülmemiştir. 19 çalışmanın 7'sinde sadece metin işleme sürecinde DDİ modülü kullanılmış veya diğer yapay zekâ teknolojilerinden yararlanılmıştır. Çalışmaların öneriler kısmında DDİ konusunun Türkçe MKS sistemlerine entegre edilmesi gerektiğini savunan araştırmacılar vardır.

Araştırmaya dâhil edilen 19 çalışmanın 10'u elde edilen sonuçlara göre konuşmadaki vurgu ve tonlamanın iyileştirilmesi gerektiğini vurgulamıştır. Bu iyileştirmenin ise DDİ teknikleriyle mümkün olduğunu savunmuşlardır.

Türkçe yapısı gereği sondan eklemeli bir dil olduğu için eklemeli sentezleme ve hece birleştirme yöntemleri kullanılan çalışma sayısının çok fazla olduğu görülmüştür. 19 çalışmanın 13'ü; yani araştırmaya dâhil edilen çalışmaların yaklaşık %68'i bu sentezleme yöntemini kullanmıştır. 2007 yılı ve sonrasındaki çalışmalarda neredeyse sadece bu yöntemin kullanıldığı görülmüştür.

Çalışmalarda en fazla TD-PSOLA algoritması kullanılmıştır. Kullandığı birleştirme algoritmasını belirten çalışmalardan 3 tanesinde TD-PSOLA algoritmasının kullanıldığı görülmüştür. Bunları 2 adet PSOLA ve 1 adet SOLA algoritması takip etmiştir.

Araştırmaya dâhil edilen çalışmalarda en çok kullanılan programlama dillerinin MATLAB, C/C++, Java ve C# olduğu gözlemlenmiştir. Tablo 4.3'te kullanılan programlama dilleri ve yüzdeleri verilmiştir. En fazla kullanılan programlama dilinin %26'lık bir oranla C ve C++ olduğu görülmüştür.

Tablo 4.3: En Çok Kullanılan Programlama Dilleri

Kullanılan Programlama Dili	Sayı	Yüzde
C/C++	5	%26.3
MATLAB	4	%21
C#	4	%21
Java	2	%10.5

Tablo 4.4'te en çok kullanılan sentezleme yöntemi, birleştirme algoritması ve programlama dili, adet ve yüzde bazında gösterilmiştir.

Tablo 4.4: Çalışmaların Genel Özeti

En çok kullanılan sentezleme yöntemi	En çok kullanılan birleştirme algoritması	En çok kullanılan programlama dili
Eklemeli Sentezleme	TD-PSOLA	C/C++
(13 Adet - %68)	(3 adet - %15.7)	(5 adet - %26.3)

Tablo 4.4'e bakıldığında en çok kullanılan sentezleme yönteminin %68 ile eklemeli sentezleme, en çok kullanılan birleştirme algoritmasının %15 ile TD-PSOLA, en çok kullanılan programlama dilinin ise %26 ile C/C++ olduğu görülmektedir.

5. TARTIŞMA VE SONUÇ

DDİ, dillerin bilgisayar vasıtasıyla işlendiği ve bilgisayar-insan etkileşimini kolaylaştırmayı amaçlayan bir yapay zekâ dalıdır. MKS sistemleri ise en basit tabiriyle, bilgisayardaki yazılı metni sesli olarak okumayı sağlamaktadır. MKS sistemleri günümüzde birçok cihazda kullanılmaktadır. Türkçe sondan eklemeli bir dil olduğu için genellikle bu çalışmalarda eklemeli sentezleme yöntemi tercih edilmiştir. MKS sistemlerinin ilk aşaması Türkçede var olan hece tiplerini belirleyip, bu hecelerin bir konuşmacı yardımıyla seslendirilerek ses veri tabanına kaydedilmesidir. Daha sonra bir heceleme algoritması ile metin hecelere ayrılır ve kaydedilen seslerle eşleştirilip metin seslendirilir. Bu kısma kadar bütün MKS çalışmaları birbirine benzemektedir. Ancak üretilen konuşma, insanın doğal konuşmasındaki vurgu ve tonlamalara ne kadar benzetilmeye çalışılırsa çalışılsın tamamen aynı olmamaktadır. Bu durum bazen anlam kargaşalarına sebep olmaktadır. İncelenen çalışmalardan elde edilen sonuçlar da bu durumu doğrulamaktadır. Cümlenin anlamını en iyi şekilde anlamamızı sağlayan şey konuşmadaki vurgu ve tonlamalardır. Bu ise, DDİ teknikleri ile sağlanabilmektedir.

Tez kapsamında, Türkçe MKS sistemleri hakkındaki literatür incelenmiştir. Türkçe MKS sistemlerinde genel olarak karşılaşılan sorunlar ve eksiklikler saptanmıştır. Türkçe MKS konusunda yapılan literatür taraması ve araştırmanın sınırlılıkları doğrultusunda çalışma kapsamına dahil edilen tez, makale ve bildiri sayısı 19'dur.

Türkçe yapısı gereği sondan eklemeli bir dil olduğu için sentezleme yöntemi olarak büyük çoğunlukla eklemeli sentezleme yöntemi kullanılmıştır. İncelenen çalışmaların büyük çoğunluğunda bu yöntemin kullanılması, doğru olan tekniğin bu olduğunu göstermektedir. Eklemeli sentezleme/hece eklemeli yöntem kullanan araştırmacılar, sentezlenen cümlenin anlaşılabilirliğinin iyi seviyede olduğunu kaydetmişlerdir.

İncelenen 19 çalışmanın %68'i eklemeli sentezleme yöntemini kullanmıştır. Türkçe MKS sistemlerinde kullanılacak olan doğru sentezleme yönteminin eklemeli sentezleme olduğu da bu rakamla kanıtlanmıştır. Çalışmaların %26'sı C ve C++ dillerini, %15'i ise TD-PSOLA birleştirme algoritmasını kullanmıştır.

Türkçe MKS sistemlerinde kullanılan programlama dilleri değişkenlik göstermektedir. Bu durum bize, çalışmadan elde edilen sonucun, kullanılan programlama dili ile doğrudan bir

bağlantısı olmadığını ve farklı programlama dillerinin aynı sonucu verebileceğini göstermiştir.

Araştırmaya dâhil edilen 19 çalışmanın 7'sinde sadece metin işleme sürecinde DDİ modülü kullanılmış veya diğer yapay zekâ teknolojilerinden yararlanılmıştır. DDİ modülünden faydalanan çalışmacıların elde ettiği başarıları incelendiği zaman, sentezlenen konuşmanın anlaşılırlığı yüksek, vurgu ve tonlamalarının ise yeterli seviyede olduğu sonucuna varılmıştır. 19 çalışmanın 10'u elde edilen sonuçlara göre konuşmadaki vurgu ve tonlamanın iyileştirilmesi gerektiğini vurgulamıştır. Standart bir MKS sisteminde genel olarak anlaşılırlığın iyi fakat insan konuşmasına yakın vurgu ve tonlamaların eksik olduğu belirtilmiştir. Çalışmacıların 12'si; yani yaklaşık %63'ü, vurgu ve tonlamalarda eksiklikler olduğunu, bu durumun iyileştirilmesinin; DDİ bilimi ile mümkün olduğunu savunmaktadır. Bu anlamda; gelecekteki çalışmalarda, MKS çalışmalarının DDİ teknikleriyle yapılması önerilmektedir.

Sözcük anlamını belirginleştirme çalışmaları, Türkçede anlam kargaşasına sebep olabilecek bazı durumların giderilmesini amaçladığı ve gerçekleştirdiği için, bu konuda örnekleri verilen derleme metinlerinin sayısı arttırılmalı ve kullanıma açılmalıdır. Daha verimli ve doğal konuşmaya yakın sentezlenen sonuçlar alınabilmesi için şimdiki derleme metinlerinin sözcük sayısı arttırılmalı ve biçimbilimsel analiziyle çözümlemesi yapılmış daha geniş yelpazede bir sözlük oluşturulmalıdır. Çünkü şu anki derleme metinlerinde geçen sözcük dağarcığının belli sayıda olduğu görülmüştür. Geliştirilen derleme metinlerinin Türkçe MKS sistemleri ile entegresi de bir başka iyileştirme yolu olarak gösterilebilir.

Gelecekteki MKS çalışmaları için, sistemlerde DDİ teknikleri kullanılmalıdır. Türkçedeki yazılışları aynı fakat anlamları ve okunuşları farklı sözcükler tespit edilmelidir. Geliştirilen sözcük anlamını belirginleştirme derlemleri ile birlikte bu tip sesteş sözcüklerin her okunuşu kaydedilip ses veri tabanında saklanmalıdır. Sözcüklerin cümlede kullanıldıkları anlamına göre doğru okunuşları ise geliştirilen veri tabanından getirilip seslendirilmelidir. Anlaşılabilirliğin iyi olduğu çalışmalarda vurgu ve tonlamanın da iyileştirilmesinin bu yollarla sağlanabileceği görülmüştür. Bu sayede insan konuşmasına en yakın seslendirme elde edilmiş olacaktır.

KAYNAKLAR

Adalı, E., 2016, Doğal Dil İşleme, *Türkiye Bilişim Vakfı Bilgisayar Bilimleri ve Mühendisliği Dergisi,* 5 (2).

Altan, Z. ve Orhan, Z., 2005, Anlam Belirsizliği İçeren Türkçe Sözcüklerin Hesaplamalı Dilbilim Uygulamalarıyla Belirginleştirmesi, *Ulusal Dilbilim Kurultayı,* 2005 İstanbul.

Arık, G., 2011, *Görme Engelliler İçin Bilgisayar Kullanımının Etkinleştirilmesi, Erişilebilirlik ve Bir Türkçe Hece Tabanlı Konuşma Sentezleme Sisteminin Geliştirilmesi,* Yüksek Lisans Tezi, Ankara: Gazi Üniversitesi Bilişim Enstitüsü.

Aşlıyan, R. Ve Günel, K., 2008, Türkçe Metinler için Hece Tabanlı Konuşma Sentezleme Sistemi, *Akademik Bilişim 2008.*

Ayaz, U., 2016, *Text-To-Speech Synthesıs For Turkısh Usıng A DSP Board,* Yüksek Lisans Tezi, İzmir: Dokuz Eylül Üniversitesi.

Aydın, E. A., 2011, *Görme Engelli Üniversite Öğrencilerinin Bilgiye Erişim Sorunları,* Yüksek Lisans Tezi, Ankara: Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü.

Bicil, Y., 2010, *Türkçe Metinden Konuşma Sentezleme,* Yüksek Lisans Tezi, Sakarya: Sakarya Üniversitesi.

Delibaş, A., 2008, *Doğal Dil İşleme İle Türkçe Yazım Hatalarının Denetlenmesi,* Yüksek Lisans Tezi, İstanbul: İTÜ Fen Bilimleri Enstitüsü.

Dutoit, T., 1999, *A short introduction to text-to-speech synthesis*.

Eker, B., 2002, *Turkish Text To Speech System*, Yüksek Lisans Tezi, Ankara: Bilkent Üniversitesi Fen Bilimleri Enstitüsü.

Elmas, Ç., 2011, *Yapay Zeka Uygulamaları*, Seçkin Yayıncılık, Ankara.

Erdemir, C., 2010 *Türkçe Metin Seslendirme İçin Doğal Konuşma Sentezleme,* Yüksek Lisans Tezi, İstanbul: İstanbul Üniversitesi.

Eryiğit, G., 2014, *ITU Turkish Natural Language Processing Pipeline,* http://tools.nlp.itu.edu.tr/, [Ziyaret Tarihi: 9 Ocak 2019].

Eryiğit, G., Adalı, E. ve Oflazer, K., 2006, Türkçe Cümlelerin Kural Tabanlı Bağlılık

Analizi, *In Proceedings of the 15th Turkish Symposium on Artificial Intelligence and Neural Networks,* 21-24 Haziran 2006 Muğla, 17–24.

Eryiğit, G. ve Torunoğlu Selamet, D., 2017, Social media text normalization for Turkish, *Natural Language Engineering,* 23 (6), 835-875.

Görmez, Z., 2009*, Makine Öğrenme Algoritmalarıyla Türkçe Metin Seslendirme Sistemi Yazılımı,* Yüksek Lisans Tezi, İstanbul: Fatih Üniversitesi.

Güldalı, K., 2009, *Türkçe Metin Seslendirme*, Yüksek Lisans Tezi, İstanbul: İTÜ Fen Bilimleri Enstitüsü.

Güner, E., 2013, *A Hybrıd Statıstıcal/Unıt-Selectıon Text-To-Speech Synthesıs System For Morphologıcally Rıch Languages,* Yüksek Lisans Tezi, İstanbul: Özyeğin Üniversitesi.

Karamehmet, T., 2015, *Implementatıon Of Turkısh Text To Speech Synthesıs,* Yüksek Lisans Tezi, Ankara: Atılım Üniversitesi

Kaya, T. ve İnce, M. C., 2010, Pencere Fonksiyonu Aileleri ve Uygulama Alanları, *Erciyes Üniversitesi Fen Bilimleri Enstitüsü Dergisi,* 26 (3), 291-306.

Klatt D., 1987, Review of Text-to-Speech Conversion for English, *Journal of the Acoustical Society of America, JASA*, 82(3), 737-793.

Nabiyev, V. V., 2012, *Yapay Zeka,* Seçkin Yayıncılık, Ankara.

Oflazer, K., 2016, Türkçe Doğal Dil İşleme / Turkish Natural Language Processing, *Türkiye Bilişim Vakfı Bilgisayar Bilimleri ve Mühendisliği Dergisi*, 5 (2).

Oflazer, K., 2018, Türkçe Doğal Dil İşleme, *Boğaz'da Yapay Öğrenme İsmail Arı Yaz Okulu*, 2 -5 Temmuz 2018 Boğaziçi Ünversitesi İstanbul.

Orhan, Z., 2006, *Türkçe Metinlerdeki Anlam Belirsizliği Olan Sözcüklerin Bilgisayar Algoritmaları İle Anlam Belirginleştirmesi,* Doktora Tezi, İstanbul Üniversitesi Fen Bilimleri Enstitüsü.

Schroeder, M. R., 1972, Computer speech recognition, compression, synthesis, *Berlin: Springer*, (2), 26-27.

Sel, İ., 2013, *Türkçe Metinler İçin Hece Tabanlı Metinden Konuşma Sentezleme Sistemi*, Yüksek Lisans Tezi, Elazığ: Fırat Üniversitesi Fen Bilimleri Enstitüsü.

Sel, İ., Hanbay, D. ve Karabatak, M., 2011, Beyin Bilgisayar Arayüzleri İçin Türkçe Metinden Konuşma Sentezleme Sistemi, *Fırat Üniversitesi Elektrik-Elektronik Bilgisayar Sempozyumu, Bildiri Kitabı II,* 2011 Elazığ, 273-276.

Şeker, G. ve Eryiğit, G., 2016, Extending a CRF-based Named Entity Recognition Model for Turkish Well Formed Text and User Generated Content, *Semantic Web,* 8 (5) , 1-18.

Şentürk, T., 2010*, Türkçe Metin Seslendirme,* Yüksek Lisans Tezi, İstanbul: İstanbul Teknik Üniversitesi.

Şentürk, T. ve Adalı, E., 2016, Türkçe Metin Seslendirme, *Türkiye Bilişim Vakfı Bilgisayar Bilimleri ve Mühendisliği Dergisi,* 4 (1).

Tekindal, B. ve Arık, G., 2012, Görme Engelliler için Türkçe Metinden Konuşma Sentezleme Yazılımı Geliştirilmesi, *Bilişim Teknolojileri Dergisi,* 5 (3), 9-18.

Topbaş, S. ve Kopkallı, H., 1994, "Sesbilim" ve "Sesbilgisi" Terimleri Üzerine, *Dilbilim Araştırmaları Dergisi,* 5 (), 310-322.

Türk Dil Kurumu, 2016, http://www.tdk.gov.tr/. [Ziyaret Tarihi: 25 Aralık 2018].

Tüysüz, M. A. A. ve Güvenoğlu, E., 2014, Türkçe İçin Karşılaştırmalı Bir Kelime Anlamı Belirginleştirme Uygulaması, *XVI. Akademik Bilişim Konferansı,* 5-7 Şubat 2014 Mersin, 721-726.

Uslu, İ. B., 2010, Metinden Konuşma Sentezleme, TMMOB Elektrik Mühendisleri Odası Ankara Şubesi Haber.

Uzun, E., 2007*, Görme Engelliler İçin Basılı Doküman Yorumlama ve Seslendirme Sisteminin Gerçekleştirilmesi,* Yüksek Lisans Tezi, Trabzon: KTÜ Fen Bilimleri Enstitüsü.

Ünaldı, İ., 2007, *Taşınabilir Cihazlar İçin Türkçe Metinden Konuşma Sentezleme Sistemi*, Yüksek Lisans Tezi, Trabzon: KTÜ Fen Bilimleri Enstitüsü.

Yılmaz, A. E., 2009, Türkçe Metinden Konuşma Sentezleme Uygulamaları İçin Bir Veri Sözlük Seti ve Yazılım Çerçevesi, *Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi,* 24 (4), 735-744.

ÖZGEÇMİŞ

Kişisel Bilgiler	
Adı Soyadı	Gamze YILMAZ
Doğum Yeri	Kırşehir
Doğum Tarihi	08.11.1991
Uyruğu	☑ T.C. ❒ Diğer:
E-Posta Adresi	gamzeyilmaz08@gmail.com

Eğitim Bilgileri Lisans	
Üniversite	Erciyes Üniversitesi
Fakülte	Mühendislik Fakültesi
Bölümü	Bilgisayar Mühendisliği
Mezuniyet Yılı	2015

Yüksek Lisans	
Üniversite	Ahi Evran Üniversitesi
Enstitü Adı	Fen Bilimleri Enstitüsü
Anabilim Dalı	İleri Teknolojiler Anabilim Dalı
Mezuniyet Yılı	2019