

**KARADENİZ TEKNİK ÜNİVERSİTESİ  
FEN BİLİMLERİ ENSTİTÜSÜ**

**ELEKTRİK-ELEKTRONİK MÜHENDİSLİĞİ ANABİLİM DALI**

**KONUŞMACI AŞINALIĞININ EEG TABANLI BELİRLENMESİ**

**YÜKSEK LİSANS TEZİ**

**Elektrik-Elektronik Mühendisi Zübeyir ÖZCAN**

**HAZİRAN 2017  
TRABZON**



**KARADENİZ TEKNİK ÜNİVERSİTESİ**  
**FEN BİLİMLERİ ENSTİTÜSÜ**



**Karadeniz Teknik Üniversitesi Fen Bilimleri Enstitüsünce**

**Unvanı Verilmesi İçin Kabul Edilen Tezdir.**

**Tezin Enstitüye Verildiği Tarih : / /**

**Tezin Savunma Tarihi : / /**

**Tez Danışmanı :**

**Trabzon**

**KARADENİZ TEKNİK ÜNİVERSİTESİ  
FEN BİLİMLERİ ENSTİTÜSÜ**

**Elektrik-Elektronik Mühendisliği Anabilim Dalında  
Zübeyir ÖZCAN Tarafından Hazırlanan**

**KONUŞMACI AŞINALIĞININ EEG TABANLI BELİRLENMESİ**

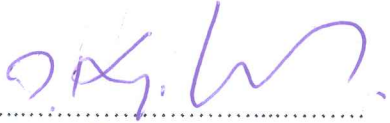


başlıklı bu çalışma, Enstitü Yönetim Kurulunun 16/05/2017 gün ve 1702 sayılı kararıyla oluşturulan jüri tarafından yapılan sınavda  
**YÜKSEK LİSANS TEZİ**  
olarak kabul edilmiştir.

**Jüri Üyeleri**

**Başkan : Prof. Dr. Temel KAYIKÇIOĞLU**

**Üye : Prof. Dr. Ali GANGAL**

**Üye : Yrd. Doç Dr. Yasin OĞUZ**

  
.....  
  
.....  
  
.....

**Prof. Dr. Sadettin KORKMAZ**

**Enstitü Müdürü**

## ÖNSÖZ

Bu tez, Karadeniz Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Elektrik-Elektronik Mühendisliği Anabilim Dalı, Elektronik Mühendisliği Bilim Dalı Yüksek Lisans Programı'nda hazırlanmıştır. Çalışmada konuşmacı aşinalığının EEG tabanlı belirlenmesi amaçlanmış ve bu doğrultuda ilgili adımlar sunulmuştur. Tez çalışmam süresince bilimsel desteği ve değerli düşünceleriyle bana her aşamada sabır, hoşgörü ve dikkatle yardımcı olan danışmanım Prof. Dr. Temel KAYIKÇIOĞLU'na teşekkür ederim.

Tüm eğitim-öğretim hayatım boyunca benden maddi ve manevi desteğini esiremeyen; bana anadilim Türkçeyi öğreten değerli aileme, eğitim hayatımın başlangıcında bana okuma ve yazma öğretmek bu tezi yazabilme yetisini kazandıran ilkokul öğretmenim Sevinç TOPRAK'a, çalışmalarımda takip ettiğim literatürleri anlamamda bana en çok katkıyı sağlayan İngilizceyi bana öğreten lise öğretmenim Nejat TÖNGÜR'e, çalışmam süresince bana güven aşılayan ve bana başarabileceğimi söyleyen, her zaman arkamda hissettiğim Seniha KETENCİ'ye ve çalışmam süresince aldığım verilerde emeği geçmiş deneylerimi birlikte yaptığım değerli arkadaşlarıma teşekkür ederim. Bu tezin sonraki çalışmalara katkı sağlamasını temenni ederim.

Zübeyir ÖZCAN

Trabzon 2017

## **TEZ ETİK BEYANNAMESİ**

Yüksek Lisans Tezi olarak sunduğum “Konuşmacı Aşinalığının EEG Tabanlı Belirlenmesi” başlıklı bu çalışmayı baştan sona kadar danışmanım Prof. Dr. Temel KAYIKÇIOĞLU’nun sorumluluğunda tamamladığımı, verileri/örnekleri kendim topladığımı, deneyleri/analizleri ilgili laboratuvarlarda yaptığımı/yaptırdığımı, başka kaynaklardan aldığım bilgileri metinde ve kaynakçada eksiksiz olarak gösterdiğimi, çalışma sürecinde bilimsel araştırma ve etik kurallara uygun olarak davrandığımı ve aksinin ortaya çıkması durumunda her türlü yasal sonucu kabul ettiğimi beyan ederim. 12/06/2017

Zübeyir ÖZCAN

## İÇİNDEKİLER

|  | <u>Sayfa No</u> |
|--|-----------------|
| ÖNSÖZ.....                                     | III             |
| TEZ ETİK BEYANNAMESİ.....                      | IV              |
| İÇİNDEKİLER.....                               | V               |
| ÖZET.....                                      | VIII            |
| SUMMARY .....                                  | IX              |
| ŞEKİLLER DİZİNİ.....                           | X               |
| TABLolar DİZİNİ.....                           | XI              |
| SEMBOLLER DİZİNİ.....                          | XII             |
| 1. GENEL BİLGİLER.....                         | 1               |
| 1.1. Giriş.....                                | 1               |
| 1.1.1. Amaç.....                               | 1               |
| 1.1.2. Konuşmacı Tanımda Ayırıcı Etkenler..... | 3               |
| 1.1.2.1. Cinsiyet.....                         | 3               |
| 1.1.2.2. Bölgesel Ağız .....                   | 3               |
| 1.1.2.3. Yaş.....                              | 4               |
| 1.1.2.4. Duygu .....                           | 4               |
| 1.1.2.5. Konuşmacı Aşinalığı .....             | 4               |
| 1.1.3. Konuşmacı Tanımda Öznitelikler.....     | 5               |
| 1.1.3.1. Kısa Zamanlı Öznitelikler .....       | 5               |
| 1.1.3.2. Yüksek Seviyeli Öznitelikler.....     | 6               |
| 1.2. Konuşmacı Tanımaya Genel Bakış.....       | 6               |
| 1.2.1. Konuşmacı Tanıma Kolları.....           | 7               |
| 1.2.1.1. Konuşmacı Doğrulama.....              | 7               |
| 1.2.1.2. Konuşmacı Teşhis Etme .....           | 8               |
| 1.2.1.3. Konuşmacı ve Olay Sınıflandırma ..... | 8               |
| 1.2.1.4. Konuşmacı Bölümleme .....             | 9               |
| 1.2.1.5. Konuşmacı Saptama .....               | 9               |
| 1.2.2. Konuşmacı Tanıma Çeşitleri.....         | 9               |
| 1.2.2.1. Metne Bağlı Konuşmacı Tanıma.....     | 10              |

|            |  |    |
|------------|--|----|
| 1.2.2.2.   | Metinden Bağımsız Konuşmacı Tanıma .....                       | 10 |
| 1.2.2.3.   | Metin Yönlendirmeli Konuşmacı Tanıma .....                     | 10 |
| 1.2.3.     | Konuşmacı Tanıma Aşamaları .....                               | 11 |
| 1.3.       | Konuşma, Dil ve İşitmeye Dair Fizyolojik Süreçler .....        | 11 |
| 1.3.1.     | Beyinde Süreçler .....   | 14 |
| 1.3.1.1.   | Broca Bölgesi .....  | 15 |
| 1.3.1.2.   | Wernicke Bölgesi .....   | 15 |
| 1.3.2.     | Konuşma Üretim Sistemi .....                                   | 15 |
| 1.3.3.     | İşitme Sistemi .....   | 17 |
| 1.3.3.1.   | Kulak .....  | 17 |
| 1.3.3.2.   | İşitsel Korteks .....  | 18 |
| 1.3.4.     | Konuşma, Konuşmacı ve Müzik Algısı .....                       | 19 |
| 1.4.       | Elektroensefalogram .....                                      | 20 |
| 1.4.1.     | Sinir Hücresinden EEG'ye .....                                 | 20 |
| 1.4.2.     | EEG Alt Bantları .....   | 22 |
| 1.4.3.     | EEG Veri Toplama Yöntemi .....                                 | 23 |
| 1.4.4.     | İşitsel Olaya İlişkin Potansiyeller .....                      | 24 |
| 1.4.4.1.   | Çok Erken İşitsel Bileşenler .....                             | 25 |
| 1.4.4.2.   | Uyumsuzluk Negativitesi .....                                  | 25 |
| 1.5.       | Literatür Araştırması .....                                    | 25 |
| 2.         | YAPILAN ÇALIŞMALAR, BULGULAR VE İRDELEME .....                 | 30 |
| 2.1.       | Ön Çalışma: Bilgisayar Tabanlı Otomatik Konuşmacı Tanıma ..... | 30 |
| 2.1.1.     | Materyal .....   | 30 |
| 2.1.2.     | Yöntem .....   | 30 |
| 2.1.2.1.   | Önişleme .....   | 32 |
| 2.1.2.2.   | Mel Frekansı Kepstrum Katsayıları .....                        | 33 |
| 2.1.2.2.1. | Çevçeveleme ve Pencereleme .....                               | 33 |
| 2.1.2.2.2. | Ayrık Fourier Dönüşümü .....                                   | 33 |
| 2.1.2.2.3. | Mel Frekansı Bükme .....                                       | 35 |
| 2.1.2.2.4. | Üçgen Filtre Bankaları ve Enerjileri .....                     | 36 |
| 2.1.2.2.5. | Ayrık Kosinüs Dönüşümü .....                                   | 37 |
| 2.1.2.3.   | Sınıflandırma: Doğrusal Destek Vektör Makineleri .....         | 37 |

|            |  |    |
|------------|--|----|
| 2.1.3.     | Öznitelik Çıkarma ve Sınıflandırma .....                           | 40 |
| 2.1.3.1.   | Uygulanan Adımlar .....  | 40 |
| 2.1.3.2.   | Hata Analizi .....   | 44 |
| 2.2.       | Ana Çalışma: Konuşmacı Aşinalığının EEG Tabanlı Belirlenmesi ..... | 45 |
| 2.2.1.     | Materyal .....   | 45 |
| 2.2.1.1.   | Uyarım .....   | 46 |
| 2.2.1.2.   | Deney Senaryosu .....  | 46 |
| 2.2.1.3.   | EEG Verisi .....   | 48 |
| 2.2.2.     | Yöntem .....   | 49 |
| 2.2.2.1.   | Filtreleme .....   | 51 |
| 2.2.2.2.   | Öznitelik Çıkarma .....  | 51 |
| 2.2.2.2.1. | Hjorth Değişkenleri .....  | 51 |
| 2.2.2.2.2. | Güç Yüzde Artışı .....   | 52 |
| 2.2.2.3.   | Elektrot Seçimi .....  | 53 |
| 2.2.2.4.   | Sınıflandırma: Doğrusal Ayırıcı Analizi .....                      | 53 |
| 2.2.3.     | Öznitelik Çıkarma ve Sınıflandırma .....                           | 55 |
| 2.2.3.1.   | Uygulanan Adımlar .....  | 55 |
| 2.2.3.2.   | Davranışsal Bulgular .....   | 57 |
| 2.2.3.3.   | Sınıflandırma ve Bulgular .....                                    | 59 |
| 3.         | SONUÇLAR VE TARTIŞMA .....   | 63 |
| 3.1.       | Beyin Konuşmacı Tanıma ve Otomatik Konuşmacı Tanıma .....          | 63 |
| 3.2.       | Konuşmacı Aşinalığı .....  | 65 |
| 4.         | ÖNERİLER .....   | 68 |
| 5.         | KAYNAKLAR .....  | 69 |
| ÖZGEÇMİŞ   |  |    |



Yüksek Lisans Tezi

ÖZET

KONUŞMACI AŞINALIĞININ EEG TABANLI BELİRLENMESİ

Zübeyir ÖZCAN

Karadeniz Teknik Üniversitesi  
Fen Bilimleri Enstitüsü  
Elektrik-Elektronik Mühendisliği Anabilim Dalı  
Danışman: Prof. Dr. Temel KAYIKÇIOĞLU  
2017, 75 Sayfa

Adli bilimlerde, organize suç örgütlerine mensup kişilerin birbirleriyle bağlantısının çözülmesi çok büyük önem arz etmektedir. Bir insan grubunun örgütlenerek işlediği düzenli ve organize suçlar, tekil kişilerin işlediği suçlara göre nispeten daha tehlikeli görülmektedir. Bu tip organize suç örgütlerinin bireylerinin belirlenmesi ve hiyerarşik yapısının çözülmesi, örgüt faaliyetlerinin ortaya konabilmesi ve örgüt elemanlarının yakalanabilmesi için ilk adım olarak görülmektedir. Emniyet birimleri tarafından yakalan örgüt üyeleri bağlı oldukları suç topluluğu hakkındaki bilgileri saklı tutma çabasında olabilmektedir.

Bu soruna yardımcı bir uygulama olarak önerilen bu yüksek lisans tezinin amacı organize suç örgütlerine mensup kişilerin örgüt içi kişilerle bağlantısının, ilgili örgütte etkin olarak irtibatta bulunduğu kişilerin belirlenebilmesinde alternatif bir yol sunmaktır. İnsanların kişi seslerine olan aşinalığını uzun süreli ve sık birliktelik ile oluştuğunu ortaya koyan çalışmalardan yola çıkılarak bu tez konusu belirlenmiştir.

Literatürde öncü bir uygulama olan bu çalışmada altı kişi üzerinden kaydedilen EEG ölçümlerinde kişilerin dinledikleri kişi seslerini tanıyıp tanıyamama durumu ortalama %80,5 ve standart sapması 3,7 olmak üzere belirlenebilmiştir.

**Anahtar Kelimeler:** Konuşmacı Tanıma, Konuşmacı Aşinalığı, Elektroensefalogram, Sınıflandırma, Öznitelik Çıkarma

Master Thesis

SUMMARY

EEG BASED DETECTION OF SPEAKER FAMILIARITY

Zübeyir ÖZCAN

Karadeniz Technical University  
The Graduate School of Natural and Applied Sciences  
Electrical and Electronics Engineering Graduate Program  
Supervisor: Prof. Dr. Temel KAYIKÇIOĞLU  
2017, 75 Pages

In forensic sciences, finding the connection between members of a criminal organization is quite an important task. The regular and organized crimes which are committed by a group of people are accepted more dangerous than the crimes which individuals did. Detecting the members and solving the hierarchical structures of these type organizations are accepted as the first step to define the organization and catch the members. The members of these type organizations which was caught by security units could be in an effort to keep the secrets of the organizations.

The purpose of this master thesis which is offered as an assistive application to this task is finding whether a connection of a member with another member of an organization is existed. This thesis is presented as an alternative way to assist this task. The studies which shows that the speaker familiarity could be built by a long and closed relationship, led us to choose the thesis topic.

As pioneer application in the literature, whether someone could recognize a person from his voice or not, could be detected with the mean accuracy of %80,5 and the standard deviation of 3,7 on EEG recorded from scalp of six subjects.

**Key Words:** Speaker Recognition, Speaker Familiarity, Electroencephalogram, Classification, Feature Extraction

## ŞEKİLLER DİZİNİ

|  | <u>Sayfa No</u> |
|--|-----------------|
| Şekil 1.1. Konuşmacı tanıma aşamaları.....   | 12              |
| Şekil 1.2. Konuşma üretim ve algısı.....   | 13              |
| Şekil 1.3. Beyin sol yarımküresi ve beyin bölgeleri.....                             | 14              |
| Şekil 1.4. Konuşma üretim sistemi.....   | 16              |
| Şekil 1.5. Kulağın yapısı .....  | 17              |
| Şekil 1.6. Sinir hücresi yapısı .....  | 21              |
| Şekil 1.7. 10-20 EEG elektrot sistemi .....  | 24              |
| Şekil 2.1. Bir kontrol sistemi olarak konuşma akışı.....                             | 31              |
| Şekil 2.2. Konuşma işaretinin önışleme ve MFKK hesaplanmasına ait yöntem akışı ..... | 32              |
| Şekil 2.3. Mel-Frekans deęişimi.....   | 36              |
| Şekil 2.4. Ham konuşma işareti.....  | 40              |
| Şekil 2.5. Ham konuşma işaretinden alınmış işaret parçası .....                      | 41              |
| Şekil 2.6. İşaret parçasının Ayırık Fourier Dönüşümü karesi .....                    | 41              |
| Şekil 2.7. Melodi filtre bankalarının enerjileri .....                               | 42              |
| Şekil 2.8. Bir parça işarete ait MFKK .....  | 42              |
| Şekil 2.9. Bir denemeye ait MFKK .....   | 43              |
| Şekil 2.10. EEG kayıt aşaması deney senaryosu .....                                  | 47              |
| Şekil 2.11. Deneme verisi içerięi.....   | 49              |
| Şekil 2.12. EEG verisi yöntem akışı.....   | 50              |
| Şekil 2.13. Ham EEG işareti.....   | 56              |
| Şekil 2.14. Delta bandında elde edilen işaret.....                                   | 56              |
| Şekil 2.15. Teta bandında elde edilen işaret .....                                   | 56              |
| Şekil 2.16. Alfa bandında elde edilen işaret .....                                   | 57              |
| Şekil 2.17. Beta bandında elde edilen işaret.....                                    | 57              |
| Şekil 2.18. Gama bandında elde edilen işaret.....                                    | 57              |
| Şekil 3.1. Beyin konuşmacı tanıma akışı.....   | 64              |

## TABLolar DİZİNİ

|   | <u>Sayfa No</u> |
|---|-----------------|
| Tablo 2.1. Otomatik konuşmacı tanıma eğitim ve test sınıflandırma sonuçları .....           | 43              |
| Tablo 2.2. Yapılan hataların cinsiyete bağlı yüzdeler dağılımı .....                        | 44              |
| Tablo 2.3. Davranışsal bulgular ve sınıf dağılımları .....                                  | 58              |
| Tablo 2.4. En etkin elektrot için bant ve kişiye bağlı Eğitim ve test başarıları .....      | 59              |
| Tablo 2.5. En etkin iki elektrot için bant ve kişiye bağlı Eğitim ve test başarıları .....  | 60              |
| Tablo 2.6. En etkin üç elektrot için bant ve kişiye bağlı eğitim ve test başarıları .....   | 61              |
| Tablo 2.7. En etkin dört elektrot için bant ve kişiye bağlı Eğitim ve test başarıları ..... | 62              |
| Tablo 3.1. En yüksek başarılar ve elektrot bölgeleri .....                                  | 66              |

## SEMBOLLER DİZİNİ

**EEG:** Elektroensefalogram

**ERP:** Olaya İlişkin Potansiyeller

**BERs:** Beyin Sapı Uyarılmış Tepkiler

**ABRs:** İşitsel Beyin Sapı Tepkileri

**DVM:** Destek Vektör Makineleri

**DAA:** Doğrusal Ayırıcı Analizi

**MFKK:** Mel Frekanslı Kepstrum Katsayıları

**AFD:** Ayrık Fourier Dönüşümü

**AKD:** Ayrık Kosinüs Dönüşümü

**Mel:** Melodi Frekansı

## 1. GENEL BİLGİLER

### 1.1. Giriş

Konuşmalar birincil olarak içeriğindeki iletiyi taşıyor olsa da ikincil olarak da konuşmacının kimliğini barındırır. Bir konuşma içerisinden çıkarılabilecek ilk bilgi barındırdığı iletidir. Buna ek olarak konuşmalar, konuşan kişiyi ele veren özelliklere de sahiptir. İstemsizce kulak misafiri olunan bir konuda, aktarılan ileti ne olursa olsun, dinleyici konuşmacının kim olduğunu görmese dahi, konuşmacının ilk olarak yaşı, cinsiyeti, hangi dili konuştuğu, duygusal durumunu hatta tanıyıp tanımadığını söyleyebilir. Dinleyici kimlik ile ilgili bu kadar saptamanın yanında, konuşmacıyı tanıması halinde konuşmacının kimliğini de tahmin edebilir. Bu da konuşmaların konuşmacı kimliğini de barındırdığını ortaya koymaktadır.

İnsan sesi kişiye özel bir nitelik olması nedeniyle biyometrik özellikler arasında gösterilir. Gerek ses tellerinin gerekse ses yolu şeklinin kişiye özel olması hiç kimsenin sesinin bir diğeri ile aynı olmamasına neden olur. Tüm bu fiziksel özelliklerin yanında bölgesel ağız, aksan, kelime dağarcığı ve telaffuz gibi davranışsal farklılıklar da konuşmacı kimliğinin çeşitlenmesine olanak sağlar [1].

İnsanlar hayatlarının daha ilk evrelerinde konuşmacı ile ilgili farklılıkları ayırt etmeye başlarlar. Anne karnındaki fetüsler üzerinde yapılan bir çalışma, fetüslerin kendi annelerinin seslerini diğer seslerden ayırabildikleri gösterilmiştir [2]. Yeni doğmuş bebekler üzerinde yapılan bir çalışma ise bu görüşü destekler niteliktedir [3]. İlerleyen yaşlarda 4 ila 7 aylık bebekler de konuşmacıları birbirinden ayırmada ustalaşmaya başlamaktadır [4].

#### 1.1.1. Amaç

Bu çalışmanın yapıldığı yıl itibariyle Türkiye Cumhuriyeti sınırları içinde gerek terör örgütleri olsun gerekse organize suç örgütleri olsun ülke çapında eylemlerini artırmışlardır. Hem bulunduğu bölgenin siyasi durumu hem de jeopolitik konum ve önemi sebebiyle Türkiye Cumhuriyeti'nde birçok terör ve suç örgütünü hem çeşitliğini hem de mevcudiyetini artırmıştır. Bunun yanında Ortadoğu, Orta Asya ve Afrika merkezli savaş, kıtlık veya işsizlik gibi nedenlerden dolayı göç etmek zorunda kalan mülteciler ülkemizi yolculuk güzergahı

olarak kullanılmaktadır. Bu da ülke nüfusuna kayıtlı olmayan birçok çocuk kayıplarının olduğu vakaları ülkenin bir gerçeği olarak ortaya koymaktadır. Gerek dil konusunda zorluklar gerekse yaş olarak çok küçük olan bazı çocukların dil gelişiminin tamamlanmamış olması bu kayıp çocuklarının ailesinin bulunmasını zorlaştırmaktadır. Motivasyonu bu gibi durumlara yardımcı bir yöntem sunmak olarak belirlenen bu çalışmanın amacı konuşmacı aşinalığı üzerinden adli birimlerce araştırılan terör veya suç örgütlerinin yapısının çözülmesini sağlamak ve kayıp çocukların ailelerini bulmaktır.

Yapılan araştırmalar göstermiştir ki bebekler doğumlarını hemen ardından annelerinin seslerini tanıyabilmekte ve diğer kişilerin sesini tanıma yetisi de 4 ila 7 aylıkken başlamaktadır [4]. Bunun yanında dil anlama ve konuşma yetisi ortalama 18 aylıkken gelişimini tamamlamaktadır [5]. Mülteci göçlerinin güzergahı üzerinde bulunan ülkemizde birçok mülteci çocuk kayıpları meydana gelmektedir. Yetkili birimlerce bulunan bu çocukların dil bilmemesi veya dil gelişimini henüz tamamlamamış olmasından dolayı bu çocuklar ailelerine ulaştırılamamakta ve devlete bağlı çocuk yetiştirme kurumlarında yaşamlarına devam etmektedir. Tüm bu gelişmeler meydana gelirken çocuklarını kaybeden yabancı uyruklu aileler çocuklarından haber alamamakta ve yolculuklarına devam etmek zorunda kalmaktadır. Bu da aile ve çocuğunun ayrılığıyla sonuçlanmakta ve devlet ekonomisine bir yük olarak yansımaktadır. Bu çalışma sayesinde bu çocuklar başvuran mülteci ailelerin ses kayıtlarını dinleyerek tanıdıkları kişi veya kişilere ulaştırılabilir.

Bir başka açıdan bakıldığında yükselen terör ve suç örgütü sayısı ve çeşitliliği bu örgütlerin yapısının ve faaliyetlerinin ortaya koyulmasını gerekli hale getirmiştir. Konuşmacı aşinalığının ancak uzun süreli birliktelik ile ortaya çıkacağını ortaya koyan çalışmalara bakıldığında bir örgüt üyesinin örgüt içinde henüz yakalanmamış diğer bireylerini seslerinden tanıyabiliyor olması o örgüt üyesi ile uzun bir süre birlikteliği olduğunu ortaya koymaktadır. Bu tip örgütlerde hiyerarşik bir yapının olduğu göz önünde bulundurulduğunda bu yapılarda birbirleri ile yakın ilişkilerde olan kişilerin hiyerarşik yapıda birbirlerine yakın düzeylerde olduğu ortaya koyulabilir. Bu da emniyet güçlerinin örgütün bu seviyedeki bireyleri hakkında istihbarat alabileceklerini göstermektedir. Yapılan sorgulamalarda yakalanan örgüt üyelerinin konuşmayı reddetmeleri veya örgüt içi hiyerarşik yapıdaki seviyelerini saklamaları verimli bilgi akışını zorlamaktadır. Bu çalışma kişilerin örgüt içindeki hiyerarşik yapıdaki yerlerinin belirlenmesine yardımcı olma niteliğindedir. Ayrıca birbirlerini tanıdıklarını reddeden kişilerin arasındaki bağın sorgulanmasını

sağlamaktadır. Bu açıdan bakıldığında adli bilimlerde büyük bir sorun olan örgütsel bağlılığın belirlenmesine yardımcı bir uygulama olarak sunulmuştur.

### **1.1.2. Konuşmacı Tanımda Ayırıcı Etkenler**

Konuşma içinde kişiye ait birçok özellik bulunabilir. Dinleyici konuşmacının kimliğini bilmesede dahi ilk etapta birçok saptama yapabilir. Bunlar arasında konuşmacının cinsiyeti, yaşı, memleketi veya duygusal pozisyonu sayılabilir. Konuşmacının hakkında fazladan bilgi edinilebilmesini sağlayan bu özelliklerin yanında konuşmacı bir başkasını taklit ederek de konuşmacının tanınmasını zorlaştırabilir. Dinleyici eğer konuşmacının sesine yeterince aşına ise taklit konuşmacının kendini saklama şansını azaltacaktır. Bu gibi etkenlerin konuşmacı tanımda bu kadar etkili olmasıyla bu alanda da bilimsel araştırmalar yapılmıştır.

#### **1.1.2.1. Cinsiyet**

Cinsiyet konuşmacıların sınıflandırılmasında bir hayli etken bir özelliktir. Lass ve diğerleri [6] ana frekansların cinsiyet tanımaya etkisini araştırdıkları çalışmalarında, konuşmacıların doğal konuşmaları ve fısıldamalarının kayıtlarını dinleyicilerine sunmuşlardır. Daha sonra sesli kayıtları yüksek geçiren süzgeçten geçirerek ana frekansları ortadan kaldırmışlardır. Alınan sonuçlara göre, dinleyiciler en iyi sınıflandırmayı üzerinde işlem yapılmamış sesli konuşma kayıtlarında, en kötü sınıflandırmayı fısıltı kayıtlarında ve nispeten daha iyi sınıflandırmayı ise ana frekansların süzüldüğü kayıtlarda yakalamışlardır.

Murry ve Singh [7] konuşmacı seslerinin benzerliğini araştırdıkları çalışmalarında, dinleyicilere kadın ve erkek sesi kayıtlarından tam bir cümle ve ünlü sesler olmak üzere iki farklı kayıt dinletmişlerdir. Çalışmalarında hiçbir dinleyici bir cinsiyete ait ses kaydını bir diğer cinsiyetle eşleştirmemişlerdir.

#### **1.1.2.2. Bölgesel Ağız**

Bölgesel ağız konuşmacıların tanınmasında önemli olan bir başka etkidir. Bu konu ile yapılmış çalışmalar göstermiştir ki dinleyiciler konuşmacının yaşadığı bölgeyi, bölgesel



çözünürlük düşük olmak kaydı ile saptayabilirler [8], [9]. Preston [10] çalışmasında kişilerin ancak hakkında bilgi sahibi oldukları bölgelerden gelen kişileri sınıflandırabildiklerini göstermiştir. Remez ve diğerleri [11] aynı bölgeden ve farklı bölgelerden gelen kişilerin ses benzerliği ile ilgili çalışmasında bunu destekler nitelikte sonuçlara ulaşmıştır. Dinleyiciler farklı bölgeden konuşmacı kayıtlarında daha yüksek başarıya ulaşmışlardır. Ayrıca hakkında herhangi bilgileri olmadığı bölgelerin konuşmacılarında hata yapmaya daha meyilli olmuşlardır.

### **1.1.2.3. Yaş**

Abercrombie [12] yaşın kişinin sesi üzerinde büyük etki sahibi olduğu ve bu yüzden de sesler üzerinde kişilerin yaşlarının tahmin edilebileceğini öne sürmüştür. Cerrato ve diğerlerinin [13] yaptıkları çalışma göstermiştir ki kişiler yaş tahminini ancak geniş aralıklarda gruplandırma yapıldıysa yüksek başarı ile gerçekleştirebilmektedirler. Bu da tasarlanan deneylere bir hayli bağımlı bir durum olduğunu ortaya koymuştur.

### **1.1.2.4. Duygu**

Konuşmacının tanınmasını etkileyen faktörlerden biri de konuşmacının duygusal pozisyonudur. Read ve Craik [14] bazı aktörler tarafından duygusal ve duygusal olmayan ifadeleri kaydetmişlerdir. Çalışmalarında ortaya çıkan sonuç, duygusal pozisyonun dinleyicilerin konuşmacıyı tanıma başarılarını etkilememiştir. Buna rağmen otomatik konuşmacı tanıma uygulamalarında konuşmacının duygusal pozisyonunu saptama çok önemli bir alanı kapsar. Dinleyiciler duygusal konuşmacıyı tanımada duygusal pozisyonun etkilenmeseler de duygular konuşmacı sesinin akustik olarak etkilemekte ve konuşmacının doğru tanınmasını zorlaştırmaktadır. Bunun nedeni duygu ve konuşmacı kimliğini ön plana çıkartan özneliliklerin birbirine girişim yapmasıdır [15], [16].

### **1.1.2.5. Konuşmacı Aşinalığı**

Yarmey ve diğerleri [17] dinleyicilerin konuşmacıları tanıyıp tanımamasının konuşmacı tanıma işlemine doğrudan etki ettiğini göstermiştir. Bulgularına göre dinleyiciler

tanıdıkları kişilere ait olan konuşmalardan daha hızlı ve başarılı tanıma göstermişlerdir. Ayrıca dinleyicilerin konuşmacıları tanımada aşına oldukları kişiler için sadece kısa ses kayıtlarına ihtiyaç duymuş, tanımadıkları kişiler için ise kayıt uzunluğu bir hayli uzun olması gerekmiştir.

Cook ve Wilding [18] yaptıkları çalışmalarında dinleyicilerin tanımadıkları konuşmacıları bir cümle uzunluğundaki kayıtlarla aşinalığını sağlamayı amaçlamışlardır. Fakat çalışmalarını benzer bir yüz tanıma çalışmasıyla karşılaştırdıklarında başarılarının bir hayli düşük olduğunu saptamışlardır. Bu sonuçlardan yola çıkarak dinleyicilerin ancak çok daha uzun süreli eğitim sürecinden geçirilme durumunda konuşmacılara olan aşinalığının artırılabilceğini savunmuşlardır.

### **1.1.3. Konuşmacı Tanıma Öznitelikleri**

Öznitelik çıkarma, konuşmacı tanıma uygulamalarında konuşmacıların modellenmesi için konuşmacıların karakteristiklerinin en ayırt edici özelliklerin vurgulandığı kısımların ortaya koyulmasını amaçlar. Konuşma işaretleri konuşmacı tanıma için çok önemli olmayan birçok öznitelik içerir. Bu yüzden konuşmacı tanıma için en uygun özniteliklerin özellikleri:

1. Konuşmacılar arası yüksek, konuşmacı içinde düşük değişkenlik göstermelidir.
2. Gürültü ve bozulmalardan etkilenmemelidir.
3. Konuşma içinde sık ve doğal olarak ortaya çıkmalıdır.
4. Kolay ölçülebilmelidir.
5. Taklidi zor olmalıdır.
6. Konuşmacının sağlığından veya uzun süreli yaşsal değişimlerinden etkilenmemelidir.

olarak sıralanmıştır [19].

#### **1.1.3.1. Kısa Zamanlı Öznitelikler**

Konuşma işaretleri iletilen mesajın devamlı değişmesi nedeniyle kısa zamanlarda değişkenlik gösterir. Bu yüzden konuşma işaretleri 20 ila 30 milisaniyelik parçalara bölünerek incelenir. Öznitelikler de bu parçalardan ayrı ayrı çıkarılır. Genelde bu parçalar bir önvurgulama ve pencereleme işleminden geçirilir. Önvurgulama parçaların yüksek

frekans bileşenlerinin genliklerinin gırtlak üzerinden geçerken düşmesi nedeniyle bu frekans bileşenlerinin yükseltilmesi amacıyla yapılır [20]. Bunun yanında pencereleme işlemi de Ayırık Fourier Dönüşümünün (AFD) sonsuz uzunluk etkisi nedeniyle yapılır [21].

Ayrık fourier Dönüşümü işaretlerin frekans bileşenlerinin hesaplanmasına yarayan basit ve etkin bir yöntemdir. Genelde frekans bileşenlerinin genlikleri ile ilgili çalışmalar yapılsa da Paliwal ve Alsteris [22], Hegde ve diğerleri [23] frekans bileşenlerine ait faz bilgisinin de doğruluğa etki ettiğini çalışmalarında göstermişlerdir.

### 1.1.3.2. Yüksek Seviyeli Öznitelikler

Konuşmacılar sadece konuşma işaretlerindeki spektral özellikleri ile birbirinde ayrılmazlar. Ayrıca kişilerin kelime dağarcığı, konuşma hızı, konuşmalarında sık kullandıkları kelimeler gibi özellikler de konuşmacıların birbirinde ayrılmasına olanak sağlar [24]. Kişi konuşmasında diğerlerinde farklı olarak “Eeee...”, “Evet”, “Efendime söyleyeyim” gibi kelime öbeklerini sık kullanıyor olması ayırıcı bir özellik olarak kullanılabilir. Bunun yanında kişinin diğer konuşmacılardan daha hızlı veya yavaş konuşuyor olması da konuşmacıyı ele veren özelliklerden sayılır.

## 1.2. Konuşmacı Tanımaya Genel Bakış

Konuşmacı tanıma, bir kişinin sesi üzerinden kimlik bilgisinin tespit edilmesi üzerine yapılan çalışmalara verilen isimdir. Konuşmacıların doğrulanması, teşhis edilmesi, izlenmesi veya tespit edilmesi gibi uygulamalarda kullanılabilir. Süregelen tarih boyunca bu alana farklı isimlendirmeler de yapılmıştır. Bunların arasında, insan sesi biyometrisi, konuşma biyometrisi, biyometrik konuşmacı tanıma ve sez izi teşhisi gibi isimlendirmeler bulunur [25]–[28]. Sayılan birçok isimlendirme alana özgün birçok kaynağın parça parça sunulmasına yol açmıştır. Konuşmacı tanıma sistemleri kişilerin ses yollarının karakteristiklerini modellemeyi amaçlar. Bu insan konuşma üretme sisteminin fizyolojik yapısının matematiksel olarak modellenmesi olarak yapılabilirken ses yolunun istatistiksel olarak modellenmesi olarak da yapılabilir [25], [29], [30]. Temel olarak konuşmacı tanıma bir konuşmacının sesi modellendirildikten sonra, daha önce modellendirilmiş konuşmacı modelleri ile karşılaştırılması ile yapılır. Bu konuşmacı tanıma sisteminin altında yatan temel

yöntemdir. Mobil iletişim ve kişisel ses kayıt cihazlarının oldukça yaygın olduğu son zamanlarda, konuşmacı tanıma sesin kolay elde edilebilmesi ve kolay işlenebilmesi bakımından diğer biyometrik sistemlerin yanında daha önemli hale gelmeye devam etmektedir.

### **1.2.1. Konuşmacı Tanıma Kolları**

Konuşmacı tanıma sistemleri, uygulama tiplerine göre kollara ayrılmıştır. Temel olarak bu dallar üçü basit ikisi de bileşik olmak üzere beş farklı kola ayrılabilir. Basit kollar arasında Konuşmacı Doğrulama, Konuşmacı Teşhis Etme ve Konuşmacı Sınıflandırma bulunurken, bileşik kollar arasında Konuşmacı Bölümleme ve Konuşmacı Saptama sayılabilir.

#### **1.2.1.1. Konuşmacı Doğrulama**

Konuşmacı doğrulamada bir ses işaretinin bir kişiye ait olup olmadığı araştırılır. Bir kişi sunulan bir ses işaretinin kendine ait olduğu veya olmadığı konusunda beyanat verir. Bu beyanat sonunda kişiden alınan ses işareti bir evrensel model yardımı ile karşılaştırılır ve işaretin o kişiye ait olup olmadığı tespit edilir. Uygulamada temel iki farklı yöntem mevcuttur.

Birinci yöntem “Evrensel Arka Plan Modeli” olarak adlandırılır [31]. Bu yöntemde, hedef kişiden alınan ses işareti eğitim işareti olarak daha önce oluşturulmuş ve çok sayıda kişiye ait ses modellerinin bulunduğu bir evrensel model içinde yeni bir model olarak eklenir. Bu modelleme sonunda test işareti aynı şekilde modellenerek yeniden oluşturulmuş bu model içinde kıyaslanır. Bu kıyas sonucunda test işareti evrensel model içinde hedef kişiye ait eğitim işaretine benzerlik gösterirse doğrulanmış, evrensel modeldeki diğer işaretlere benzerlik gösterirse de yalanlanmış olur.

İkinci yöntem ise “Kohort model” olarak adlandırılan yöntemdir [25]. Birinci yöntemde çok fazla kişinin sesinin modellenmesi gerektiği ve bu yüzden çok fazla işlem yapılacağı savunulması ile ortaya çıkmış bu yöntemde, kişiden alınan eğitim işaretine çok benzeyen kişilerin sesleri kıyaslanır. Eğitim işareti ile yüksek benzerlik gösteren işaretlerden oluşan ve evrensel model yöntemine kıyasla çok daha az model sayısına sahip nispeten

küçük bir model grubu oluşturulur. Bu kıyas sonucunda test işareti kohort model içinde hedef kişiye ait eğitim işaretine benzerlik gösterirse doğrulanmış, kohort modeldeki diğer işaretlere benzerlik gösterirse de yalanlanmış olur.

### **1.2.1.2. Konuşmacı Teşhis Etme**

Konuşmacı teşhis etme bir test kişinin bir modeller topluluğu içinde hangi modele benzediğinin araştırıldığı konuşmacı tanıma koludur. Temel olarak açık-küme konuşmacı teşhis etme ve kapalı-küme konuşmacı teşhis etme olarak iki farklı yöntem uygulanır.

Kapalı-küme konuşmacı teşhis etme bu iki yöntem arasında nispeten en kolay olanıdır. Bir test işareti daha önce modellenmiş bir topluluk arasında kıyaslanır ve benzerliğin en yüksek olduğu model ile ilişkilendirilir. Sonuç olarak test kişinin kimlikleri belirli olan modeller içinden biri olduğu ortaya koyulur. Bu yöntemin dezavantajlarından biri test kişinin gerçekte modelin içinde bulunmaması durumudur. Bu yöntemde test kişisi model içinde bulunmasa dahi yapılan işlemlerde test kişisine en çok benzeyen modelin kimliği test kişisine atanacak ve yanlış bir teşhis yapılmış olacaktır. Bu nedenle ikinci bir yöntem olarak açık-küme konuşmacı tanıma kullanılmaktadır.

Açık-küme konuşmacı tanıma, test kişinin model içinde bulunmaması durumuna karşı ortaya koyulmuştur. Modeller içinde en çok benzerlik gösterdiği model ile arasında bir benzerlik eşiği belirlenir. Bu benzerlik eşiğinin altında bir benzerlik parametresi çıkması durumunda, modellerden hiçbirinin test kişisi olmadığı sonucuna varılır. Açık-küme ve kapalı-küme konuşmacı teşhis etme yöntemleri arasındaki temel fark budur.

### **1.2.1.3. Konuşmacı ve Olay Sınıflandırma**

Konuşmacı ve olay sınıflandırma, konuşmacılara ait cinsiyet veya yaş gibi özelliklerin sınıflandırıldığı konuşmacı tanıma koludur. Bunun yanında patlamalar, silah sesler, kişi çığlıkları ve müzik gibi özellikler de bu kolda sınıflandırılır.

Cinsiyetin sınıflandırılması insan konuşmacı tanıma yeteneklerine kıyasla ne kadar kolay görünse de çocukların cinsiyetlerinin sınıflandırılması itibariyle birçok zorlukları bünyesinde barındırır. Konuşmacıların seslerindeki kişisel farklılıklar ses yollarının yapısından meydana gelmesi ve çocuklarda birbirine çok yakın özelliklere sahip olması bu

kolun en önemli zorluklarından biridir [32]. Genel itibariyle erkek, kadın ve çocuklara ait seslerin ortalama temel frekansları sırasıyla 130 Hz, 220 Hz ve 265 Hz civarındadır [33].

#### **1.2.1.4. Konuşmacı Bölümleme**

Konuşmacı bölümleme, birden fazla kişinin konuşmasının bulunduğu ses işaretlerinde bu kişilere ait ses kısımlarının bölünerek ayrıştırılmasını amaçlayan konuşmacı tanıma koludur. Ayrıca bu ses işaretleri içinde konuşmanın olmadığı, konuşma tabanlı olmayan sesler veya müzik gibi kısımlarında bölünmesi amaçlanır. Konuşma haricinde bulunan kısımların varlığı bu kolun en büyük zorlayıcı etkenlerindedir.

Bir ses işareti içinde konuşmacıların kim olduğu bilinmese dahi konuşmacıların sayısının bilinmesi de konuşmanın bölünmesi için yeterli olacaktır. Kümeleme kullanılarak sağlanacak bu işlemde kimlikleri bilinmeyen kişilerin sayısının bilinmesi ile her bir kişiye ait konuşma kısımlarının bölünmesi sağlanabilir.

#### **1.2.1.5. Konuşmacı Saptama**

Konuşmacı saptama, bir veya daha fazla kişinin birden fazla kişiye ait sesin olduğu bir işarete konuşma zaman ve uzunluklarının saptandığı konuşmacı tanıma koludur. Genelde daha önceden kişilere ait kaydedilmiş seslere ihtiyaç duyulur. Bir konferans konuşması içinde bir veya daha fazla kişinin konuştuğu akışta, istenen kişilerin konuşma zamanlarının ve uzunluklarının saptanmak istediği durumda kullanılır. Konuşmacı bölümleme kısmında da açıklandığı üzere bu tip akışlarda içerikte olabilecek konuşmanın olmadığı kısımlar, konuşma tabanlı olmayan sesler ve müzik karşılaşılabilecek en büyük zorluklardandır.

#### **1.2.2. Konuşmacı Tanıma Çeşitleri**

Bir konuşmacı tanıma uygulamasında veriler, konuşulan dil, kişinin aksanı, test işaretinin bilgisayar tabanlı olması veya konuşulan metnin aynı veya farklı olmasından etkilenebilir. Farklı dillerin farklı fonetik yapıya sahip olması, aynı dilde farklı fonetik yapıya sahip aksanların mevcut olması, sınıflandırılmak istenen test işaretinin bilgisayar tabanlı bir sistem ile taklit edilerek kullanılabilme olasılığı, başka konuşmacıların taklit

yapmaları ve konuşmacıların söyledikleri sözlerin benzerliği; doğrulama, teşhis etme ve sınıflandırma gibi işlemlerde başarısızlık kaynağı olabilir. Konuşmacıların bir bilgisayar sistemi ile kandırılabilmesi konuşmacı tanıma sistemleri için en büyük sorunlardan biridir. Bu soruna canlılık değerlendirmesi denir. Bu yüzden konuşmacı tanıma çeşitli şekillerde ortaya çıkmıştır.

#### **1.2.2.1. Metne Bağlı Konuşmacı Tanıma**

Metin bilgisi ve konuşmacı bilgisinin içi içe bulunduğu bu konuşmacı tanıma çeşidinde modeller, konuşmacılara aynı metinden kesitler okutulur. Daha sonra teşhis edilmesi veya doğrulanması istenen konuşmacı aynı metni tekrar eder ve işleme alınır. Diğer konuşmacı tanıma çeşitleri ile kıyaslandığında nispeten en yüksek başarıyla doğrulama ve teşhis etme yapılabilir. Fakat metin bilgisine sahip olan başka bir konuşmacı veya bir bilgisayar programı hedef kişiyi taklit ederek oluşturulan sistemi kandırabilir. Canlılık değerlendirmesi bu uygulamaların en büyük problemidir. Bu yüzden bu tür konuşmacı tanıma sistemleri yüksek güvenlik gerektirilen yerlerde tercih edilmez.

#### **1.2.2.2. Metinden Bağımsız Konuşmacı Tanıma**

Metne bağlı konuşmacı tanımadaki canlılık değerlendirilmesi ile ilgili problem sonucu ortaya çıkarılmış olan bu konuşmacı tanıma çeşidinde, modeller konuşmacıların rastgele konuşmalarının kaydedilmesi ile oluşturulur. Hedef kişinin herhangi bir konuşmasından doğrulama veya teşhis etme yapılabilme imkanı sunar. Konuşmacının kişisel ses kimlikleri bu tip uygulamalarda daha öne çıkacağından diğer çeşitlerin arasında metinden bağımsız konuşmacı tanıma daha güvenilirdir. Eğitim aşamasındaki zorluklar, eğitim kısmının çok uzun sürmesine neden olabilir. Çünkü yalnız kişinin ses yolu karakteristiğine bağlı olan eğitim işaretlerinin yüksek başarı ile sınıflandırılması uzun süre ve çaba gerektirmektedir.

#### **1.2.2.3. Metin Yönlendirmeli Konuşmacı Tanıma**

Canlılık değerlendirmesi konusunda başka bir çözüm olarak ortaya çıkarılmış olan metin yönlendirmeli konuşmacı tanıma da modeller birden fazla aynı metnin okutulduğu

konuşmacılardan oluşturulur. Her bir metin için eğitim metne bağlı konuşmacı tanıma tabanlı yapılır. Test aşamasında metinlerden rastgele bir tanesi hedef kişiye sunulur ve okuması istenir. Kaydedilen işaretin doğrulanması veya teşhis edilmesi ait olduğu metin modelinde yapılır. Canlılık değerlendirilmesinde kendini hedef kişi olarak gösteren bir bilgisayar sistemi ise gösterilen metni bilemeyeceğinden konuşmacı tanıma sistemi kandırılmaz olur. Bu metin yönlendirmeli konuşmacı tanıma yöntemlerinden biridir.

Bir diğer yöntemde ise modeller metinden bağımsız şekilde oluşturulur. Diğer yönteme benzer şekilde rastgele bir metin hedef kişiye okutulur ve test işaretinde hem metinden bağımsız konuşmacı tanıma yapılır hem de konuşma tanıma yapılarak hedef kişinin söylediklerinin metne uygunluğu karşılaştırılır. Sonuçta hedef kişinin metinden bağımsız konuşmacı tanımada doğrulanması veya teşhis edilmesi ve aynı zamanda söylediklerinin metne birebir uyması durumunda hedef kişi doğrulanmış veya teşhis edilmiş olur.

### **1.2.3. Konuşmacı Tanıma Aşamaları**

Herhangi bir kişinin sesi bir başka kişinin sesi ile birebir aynı olamaz. Çünkü her bir kişinin ses yolu şekli, gırtlak boyutu ve ses üretmeye yardımcı diğer organları bir başkasından farklıdır. Ayrıca bu fiziksel özelliklere ek olarak aksan, ritim, tonlama şekli, kullanılan kelime dağarcığı ve telaffuz gibi farklılıklar kişilerin seslerinin bir başkası ile aynı olmasına engel olur [1].

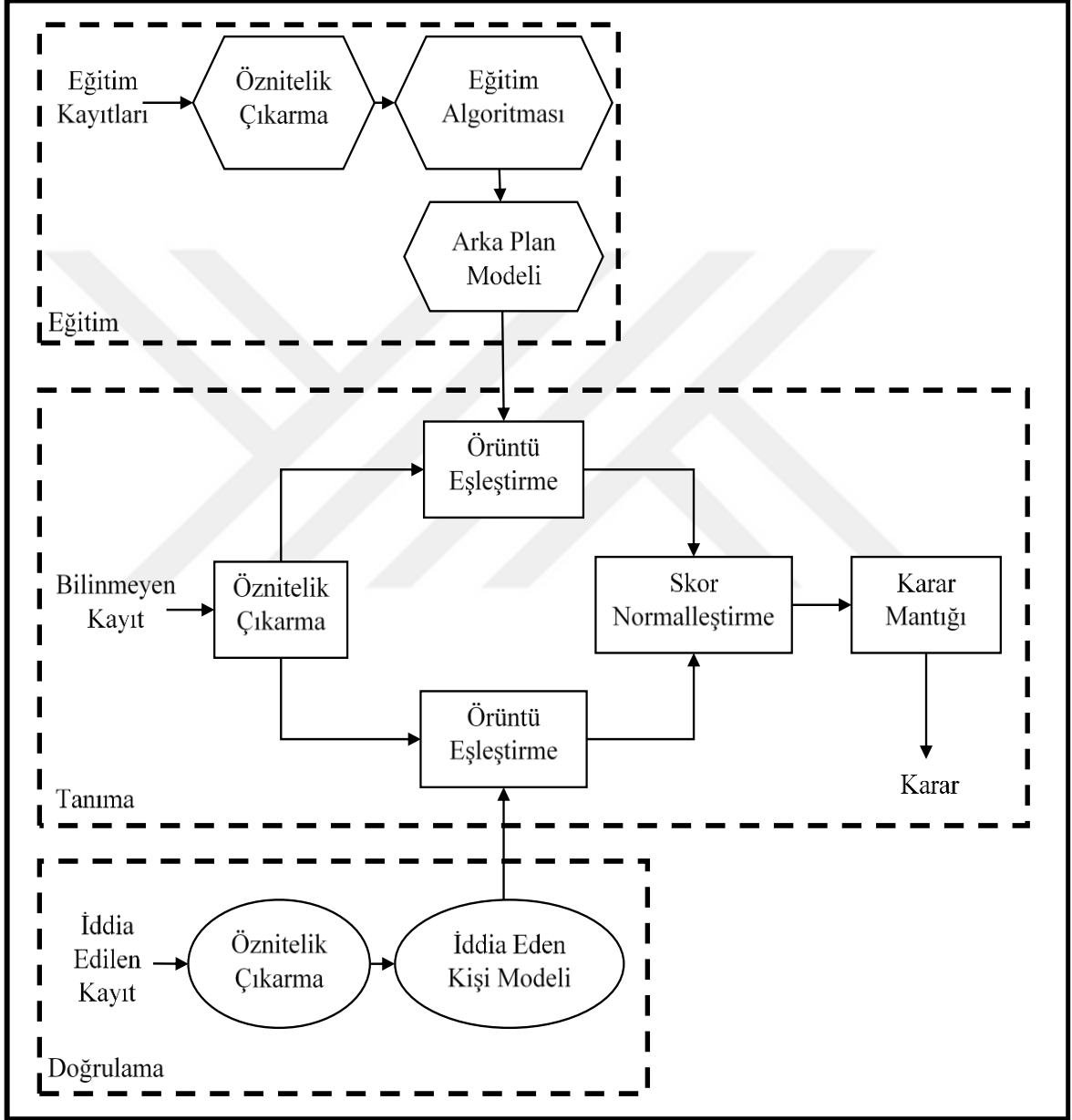
Bu gibi farklılıklar konuşmacı tanımının en önemli motivasyonudur. Konuşmacı tanıma, yapılan uygulamalara göre farklı akışlar izleyebilir. Temelde sırasıyla öznitelik çıkarma, konuşmacı modelleme, test işaretlerinden elde edilen modellerin eğitim modelleri ile karşılaştırılması ve karar aşaması olarak süregelen aşamalardan meydana gelir. Bu aşamalarda uygulamalara göre farklı aşamalar veya sıra değişimi ile karşılaşılabılır. En genel hali ile konuşmacı tanıma uygulama aşamaları Şekil 1.1’de gösterilmiştir.

### **1.3. Konuşma, Dil ve İşitmeye Dair Fizyolojik Süreçler**

Beyin günümüzde gittikçe daha da anlaşılmakta olan vücut fonksiyonlarını yöneten organdır. Karmaşıklığı ve görevleri bilim dünyasının ilgisini her geçen gün daha da



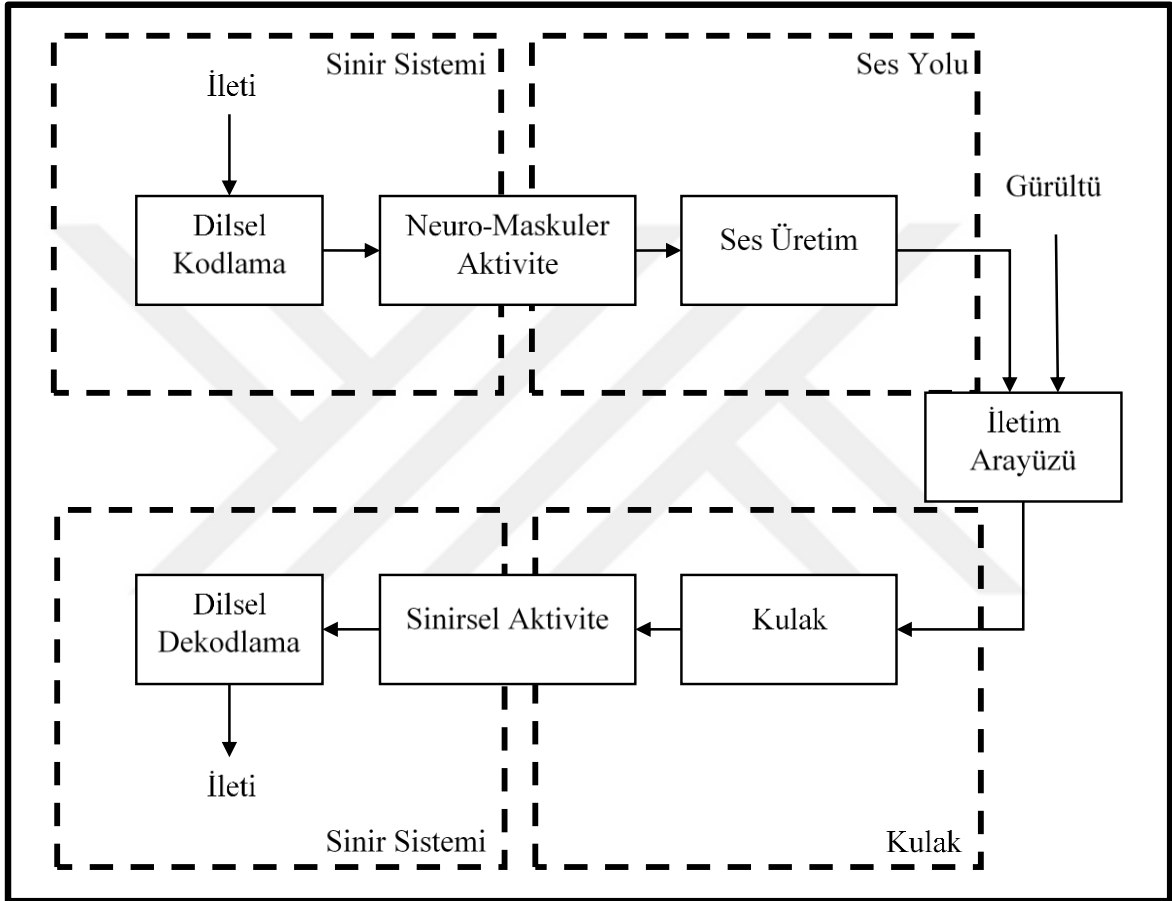
çekmektedir. Tarih boyunca vücudun fonksiyonlarından sorumlu organın kalp olduğunun düşünüldüğü asırlar vardır. Aristo tavukların başının kesildiğinde koşabildiklerini kanıt göstererek beynin çok önemli bir organ olmadığını, kanı soğutmaya yarayan bir organ olduğunu savunmuştur [34].



Şekil 1.1. Konuşmacı tanıma aşamaları [1]

Tarihte beyin ile ilgili yazılmış en eski kaynak M.Ö. 1700'lere tarihlenen Edwin Smith Papirüsü'dür. Antik Mısır zamanından kalan bu ilk yazılı tıbbi belge, Breasted [35]

tarafından incelenmiş ve genellikle kafa yaralanmaları ile ilgili kırk sekiz farklı vaka üzerinde bilgiler barındırmaktadır. Belgede geçen vakalardan birinde şakaktan alınan bir darbe sonucunda konuşma yetisini kaybeden bir vakadan bahsedilir. Bugünkü bilimsel ilerleme ile anlaşılıyor ki şakak kısmında bulunan Broca bölgesinden muhtemelen bu belgede bahsedilmiştir.

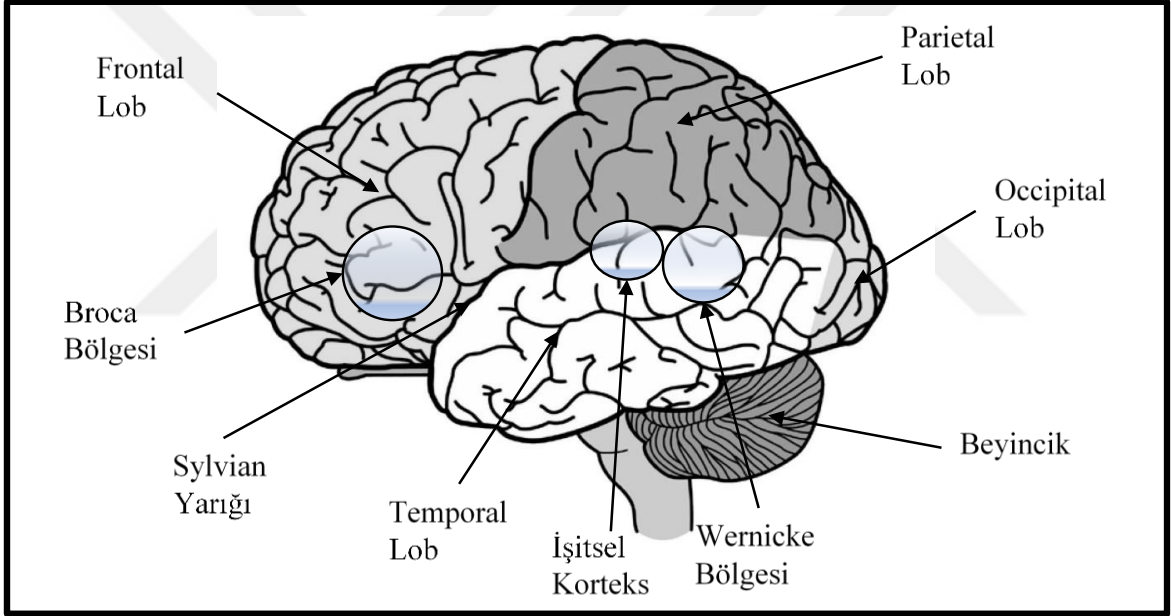


Şekil 1.2. Konuşma üretim ve algısı

İnsan vücudunun diğer tüm fonksiyonları gibi konuşma ve işitme fonksiyonları da beyin tarafından yönetilir. Kişilerin dil öğrenme, konuşulan iletileri anlama, konuşmacıları tanıma, duyma ve konuşma gibi fonksiyonları beyin tarafından yönetilir. Konuşma fonksiyonu insan ses üretim sistem, duyma fonksiyonu da işitme sistemi tarafından gerçekleştiriliyor olsa da bilgilerin işlendiği ve anlamlandırıldığı yer yine beyindir. Bu bölümde konuşma ve işitme olaylarının nasıl gerçekleştirildiğinden bahsedilecektir. Şekil 1.2’de konuşma üretim ve algısına ait akış diyagramı verilmiştir.

### 1.3.1. Beyinde Süreçler

Merkezi sinir sistemi 2 büyük temel parçaya ayrılır: Omurilik ve Beyin. Beyin ise 3 ana bölüme ayrılır. Bunlar: Ön beyin, orta beyin ve arka beyindir. Ön beyin *Cerebrum*'u içerir ve *Cerebrum* kontrolün en üstünde bulunur. Beynin tüm bölgeleri *Cerebrum* hariç beyaz maddeden, *Cerebrum* ise gri maddeden oluşur. Beynin bölgelerinde bulunan nöronlarda miyelin kılıf olup olmamasına göre o bölgenin rengi beyaz veya grimsi bir renk alır. Bu renginden dolayı beyaz madde ve gri madde olarak ayrılırlar. Beyaz madde genelde düşük seviye fonksiyonlardan sorumluyken gri madde daha karmaşık fonksiyonlardan sorumludur.



Şekil 1.3. Beyin sol yarımküresi ve beyin bölgeleri [36]

*Cerebrum*'un yüzeyine *Cerebral korteks* denir. *Cerebral korteks* beyinde düşünme, konuşma üretme ve anlama, görme ve işitme gibi yüksek seviyeli fonksiyonların gerçekleştirildiği bölgedir. *Cerebral korteks* ise 4 ana loba ayrılır. *Cerebral korteks*in arka kısmında bulunan *Occipital lob*, daha çok görme ile ilgili, yan kısımda bulunan *Temporal lob* görme olayının yüksek seviyeli işlemlerini, hafıza olaylarını, ses işleme fonksiyonlarını ve seslerin dilsel ve müziksel anlamlarının işlendiği bölümdür. Tepe kısmında konuşlanmış olan *Parietal lob* işitme ve görme ile ilgili işlemlerin beraber ve senkronize yapıldığı

bölümdür. Ön kısımda bulunan *Frontal lob* ise karar verme ve motor kontrolden sorumludur. Konuşma ve konuşmacı tanımada önemli olan kısım *Frontal* ve *Parietal lobları* birbirinden ayıran *Sylvian yarığdır* [36]. Bahsedilen bölgelerin beyin üzerindeki pozisyonları Şekil 1.3'te gösterilmiştir.

### 1.3.1.1. Broca Bölgesi

Fransız bir bilim adamı olan Pierre-Paul Broca, 1865'te dil ile ilgili rahatsızlıkların beyinin sol yarımküresinin Frontal lobunda bulunduğunu ve simetrik olan diğer yarım kürede bulunmadığını ortaya atmıştır. Broca bu bölgenin hasarlı olduğu hastaların konuşmalarında kullandıkları dilde sıkıntı yaşadıklarını fakat söylenenleri gayet iyi anlayabildiklerini ortaya koymuştur. Broca'nın buluşu ile birlikte *Broca bölgesi* adı ile anılan bu bölge konuşma üretiminden sorumlu olduğu düşünülmüştür [37].

### 1.3.1.2. Wernicke Bölgesi

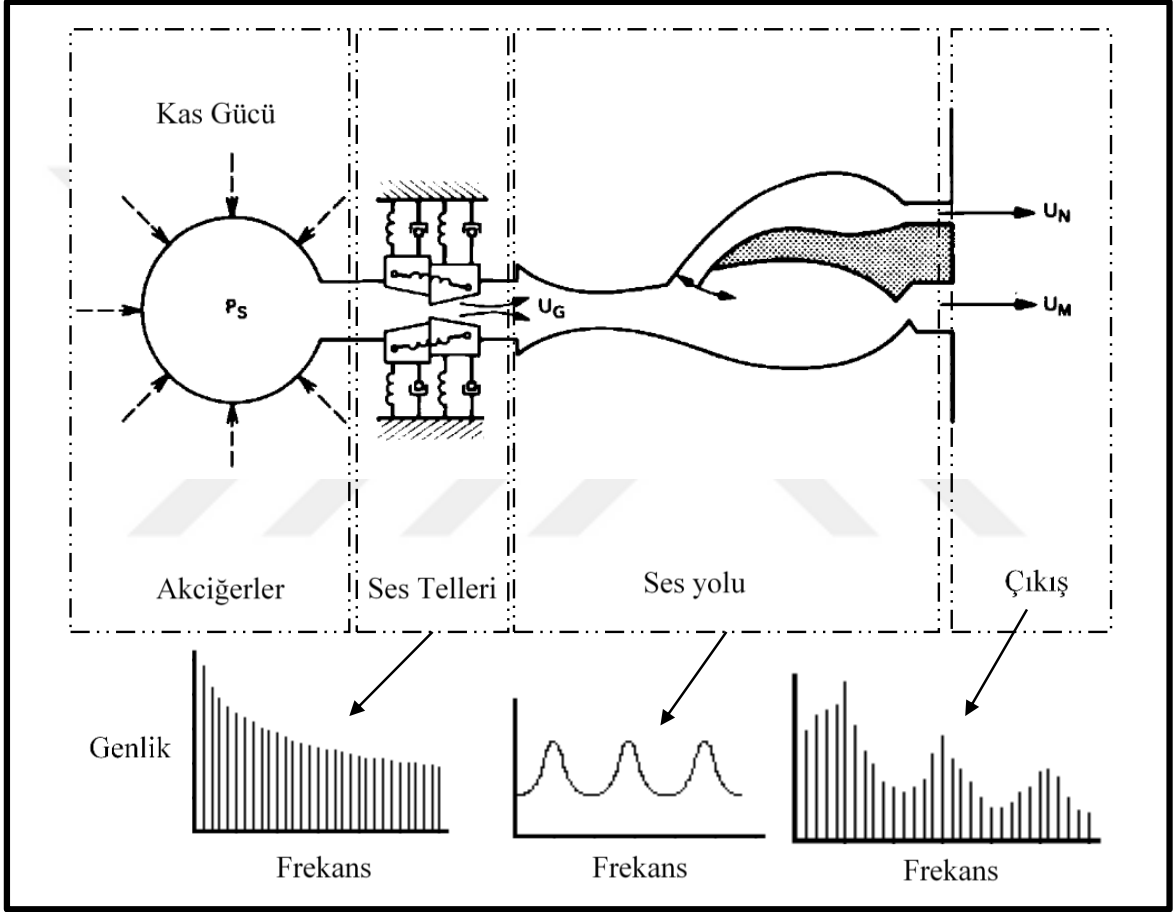
Alman bir bilim adamı olan Karl Wernicke bazı hastalarının duyabildiklerini kelimeleri anlayabildiklerini hatta çok rahat bir şekilde konuşabildiklerini ama söylenen cümleleri anlayamadıklarını keşfetmiştir. Dil ile ilgili hiçbir girişin anlamlandırılmaması üzerine Wernicke bu hastalarının ölümlerinden sonra yaptığı otopsilerinde Sylvian yarığı ile Temporal lob arasındaki bölgenin hasarlı olduğunu saptamıştır. Bu bölge *Wernicke bölgesi* olarak anılır ve dilin anlaşılmasından sorumludur [37].

## 1.3.2. Konuşma Üretim Sistemi

Konuşma üretim sistemi Akciğerlerden başlayarak, ağız ve burunda son bulur. Akciğerler ses üretme aşamasında doğrudan bir etkide bulunmasa da burada üretilen hava basıncı ses tellerinde sese dönüştürülür. Bu yüzden kas hareketleri ile hava basıncının üretildiği akciğerler konuşma üretim sisteminin başlangıcı olarak kabul görür [36].

Akciğerlerde başlayan konuşma üretim mekanizması, bu bölgeden gelen hava basıncının ses tellerini (ses kıvrımlarını) titreştirmesi ile ilk ses işareti üretilmiş olur. Ses kıvrımları olarak da adlandırılan ses telleri bir telden çok mukozadan oluşmuş kıvrımlardır.

Yüksek hava basıncı bu kıvrımları titretir. Konuşmacı kıvrımlar etrafında bulunan kaslar sayesinde farklı frekansta sesleri oluşturabilir. Burada üretilen ses işaretinin frekans bileşenleri incelendiğinde bir  $F_0$  ana frekansı ve bu frekansın tam katlarında oluşan armoniklere (doğuşkanlara) rastlanır. Konuşmacının burada ürettiği sesin frekansı bu  $F_0$  ana frekansı olarak görülür.  $F_0$  frekansının tam katlarında konuşlanmış armonikler ise konuşmacıların birbirinde ayrılmasında önemli etkenlerden biri olma adaylarıdır [38].



Şekil 1.4. Konuşma üretim sistemi [19]

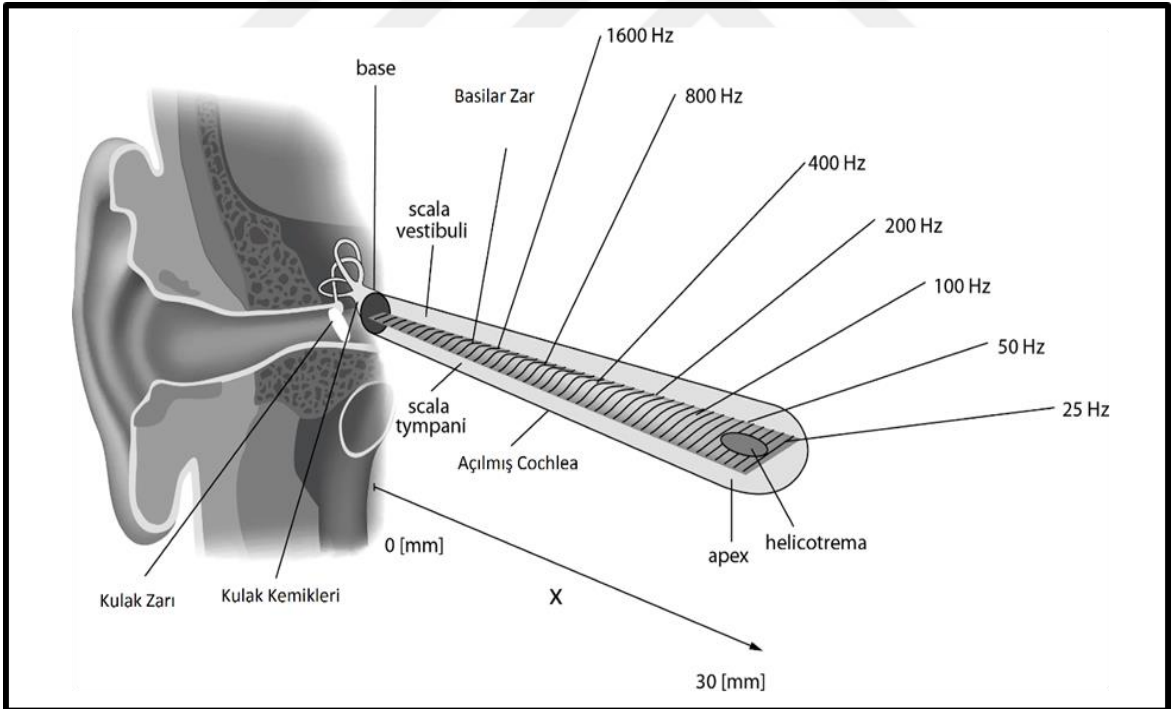
Ses tellerinde üretilmiş olan işarete Şekil 1.4'te görülebilir. Bu işaret ses yoluna aktarılır. Ses yolunun kendine ait bir süzgeç etkisi vardır. Konuşmacıların seslerinin ayrılmasında en önemli farklılıklardan olan ses yolu şekilleri kişiye özeldir. Her insanın ses yolunda farklı bir süzgeç etkisi görülür ve bu yüzden konuşmacı tanıma uygulamalarının önem verdiği kısımlardan biridir. Ses tellerinde üretilen ses işareti ses yolunun süzgeç etkisinde geçtikten sonra çıkışa Şekil 1.4.'te görüldüğü gibi yansır [39].

### 1.3.3. İşitme Sistemi

İşitme sistemi mekanik sistem ve sinirsel sistem olarak ikiye ayrılır. Ses işaretlerini ilk olarak karşılayan kulak Mekanik sistemi, kulaktan işaretleri beyne aktaran işitsel sinir demetleri ve beyinde işleyen işitsel korteks ise sinirsel sistem olarak adlandırılır.

#### 1.3.3.1. Kulak

Kulak üç kısımdan oluşur: Dış kulak, orta kulak ve iç kulak. Dış kulak, kulak kepçesi ve kulak kanalından oluşur. Dış ortamda arayüz içinde serbest dolaşan ses işaretleri kulak kepçesi tarafından yakalanır ve kişiye özel olan kulak kıvrımları sayesinde kulak kanalına aktarılır. Kulak kanalında ses işareti yol olarak orta kulağın başlangıcı olan kulak zarını titretir.



Şekil 1.5. Kulağın yapısı [36]

Kulak zarı gelen ses işaretinin yoğunluğuna göre titreşir ve bu titreşim orta kulak kemiklerine aktarılır. Orta kulak kemikleri kulak zarından alınan ses işaretlerinin genliklerini yükseltmekten sorumludur. Kulak içinde mekanik bir amplifikatör görevi olan bu kemikler

yükseltilen bu işaretleri iç kulağa aktarılmak üzere *kokleanın* girişi durumundaki oval pencereye iletir.

İç kulağın en önemli iki organı olan yarıdaire kanalları ve koklea (salyangoz) duyma işleminin sinirsel aktiviteye dönüştürülmesinde etkindir. Yarıdaire kanalları daha çok vücudun dengesinden sorumludur. Koklea ise içinde bulunan sıkıştırılmaz sıvı sayesinde sinirsel aktivite üretiminde etkindir. Oval pencere üzerinden kokleaya iletilen ses işareti koklea içindeki sıkıştırılmaz sıvıyı titreştirir. Bu titreşim bir salyangoza benzeyen ve Şekil 4'te açılmış hali gösterilen koklea üzerinde yol alır. Koklea üzerinde bulunan kıl hücreleri yol boyunca iletilen titreşimin içinde bulunan frekans bileşenlerini saptar. Her bir frekans değerinden sorumlu kıl hücreleri farklıdır ve Şekil 1.5'te görüleceği üzere yüksek frekansa duyarlı kıllar önce, düşük frekansa duyarlı kıllar sonra olacak şekilde dizilmişlerdir. Her bir kıl hücresi titreşimde saptadıkları frekans ve genlik bilgisine karşılık gelen sinirsel aktiviteyi beyne iletir. Böylece beyin eş zamanlı olarak işitme işlemini alan ses işaretlerinin frekans bileşenlerini göz önünde bulundurarak gerçekleştirir. Beyinde işitme olayında sorumlu bölge ise işitsel kortekstir.

### 1.3.3.2. İşitsel Korteks

İşitme sisteminin sinirsel sistemi kokleadan beyne giden sinir demetleri ve işitsel korteksten meydana gelir. Sinir demetleri sadece iletim görevinden sorumluyken ses işaretinin anlamlandırılmasıyla alakalı değildir. İşitsel korteks Birincil İşitsel Korteks, İkincil İşitsel Korteks ve Üçüncül İşitsel Korteks olarak üç ana kısımdan meydana gelir.

Birincil işitsel korteks kokleadan sinir demetleri ile aktarılan ses işaretlerinin ilk uğradığı yerdir. Bu bölge kokleada farklı hücre kıllarının farklı sinir hücresi gruplarına bağlanması nedeniyle iletilen tonlar ile ilişkilidir. Yani her bir tona karşılık gelen farklı bir birincil işitsel korteks sinir hücresi grubu vardır. Bu nedenle Frekans-Beyin ilişkisinin en net gözlemlendiği bölgedir. Ayrıca bu bölgeden kokleaya giden geribildirim demetleri de bulunmaktadır [40].

İkincil işitsel korteks, beynin her iki yarıküresinde bulunsa da birincil işitsel korteksten alınan bilgilerden, sol beyinde konuşma ile ilgili fonetik bilgilerin çıkartılmasından sorumluyken, sağ beyinde müzikle ilgili melodi ve ritmik örüntülerin çıkartılmasından sorumludur.

Üçüncül işitsel korteks, ikincil işitsel bölgeden alınan bilgilerin daha derinlemesine işlenmesinden sorumludur. İkincil işitsel kortekste olduğu gibi üçüncül işitsel kortekste de sol yarıkürede müzikle ilgili özellikler değerlendirilirken, sağ yarıkürede konuşma ve dil ile ilgili özellikler değerlendirilir.

Daha önce anlatıldığı gibi Wernicke bölgesi hem sözlü hem yazılı dilin anlamlandırılmasında sorumludur. Bu bölgenin sağ yarıküredeki simetriği olan bölge ise tempo ritim ve tonlama gibi müziksel özelliklerin belirlenmesinden sorumludur. İşitsel korteks ve konuşma bölgelerinin her birinden anlaşılacağı üzere beynin sol yarıküresi dil ve konuşma, sağ yarı küresi ise müzik ile ilgili kısımlardan sorumludur [41]. Birbirinden ayrı çalışan iki yarıküre elemanları arasındaki bağlantı ise beynin iki yarıküresini birbirine bağlayan *Corpus Callosum* sayesinde sağlanır.

Burada dikkat çekilmesi gereken bir nokta, bu gibi özelliklerin beynin baskın yarıküresine bağlı olmasıdır. Sağ elini kullanan kişilerin %90'ı, sol elini kullanan kişilerin ise %60'ı için beynin baskın yarıküresi sol yarıküredir. Yani sol yarıküre olarak anlatılan yerden kastedilen, aslında beynin baskın yarıküresidir [42].

#### **1.3.4. Konuşma, Konuşmacı ve Müzik Algısı**

İşitsel korteksin görevi, alınan ses işaretlerin işlenmesi ve daha yüksek seviyeli işlemler için Wernicke bölgesine iletmektir. Wernicke bölgesi böylece iletilen bu ses bilgilerinin içinden sol yarıkürede dil ile ilgili kısımların anlamlandırılması, sağ yarı kürede ise bu bilgilerin müziksel özelliklerinin çıkarılması görevini yapar. Wernicke bu bölgeyi belirlediğinde görsel kortekste okuma yolu ile elde edilen dilsel bilgilerinde bu bölgeye aktarılarak işlendiğini ortaya koymuştur. Ayrıca Wernicke ve Broca bölgelerinin beynin iç kısmında birleştiğini gösteren sinirsel ağları da saptamıştır. Bu ağların dil üretim ve anlama kısımları arasında çift taraflı geribildirim sağladığını göstermiştir. Ayrıca bu ağların sol yarıkürede sağ yarıküreye göre çok daha sağlam olduğunu görmüştür.

Beyinde konuşmacı tanıma ile ilgili çalışmalar göstermiştir ki beynin konuşmacıları birbirinde ayırması çok fazla karmaşık bir yapıya sahiptir. Konuşmacıların tanınmasında ayırıcı etken olan dil, tonlama, konuşma hızı, aksan, bölgesel ağız, spektral özellikler gibi özniteliklerin ayrılması beynin farklı bölgelerinin sorumluluğu altındadır. Gougoux ve diğerleri [43], körlerin insan sesi algısı üzerine yaptıkları İşlevsel Manyetik Rezonans Görüntüleme (fMRI) çalışmasında, kişilerin insan seslerini birbirinden ayırmaya



çalıştıklarında bölgenin Temporal lob üzerinde yukarıda anlatılan bölgeler ve yakınındaki bölgelerin her iki yarıkürede de etkinlik gösterdiğini saptamıştır.

#### **1.4. Elektroensefalogram**

Hans Berger [44], İnsan beyninden ilk işaret kaydını yapan kişi olarak bu çalışmasıyla ilk zamanlar bilim dünyasında büyük şüpheyle karşılanmıştır. Berger, alın bölgesi ve Occipital lob üzerine yerleştirdiği iki elektrot üzerinden aldığı kayıtlarda yaklaşık 10 Hz düzeyinde salınım yapan işaretlere rastlamıştır. Özellikle duyuşal herhangi bir uyarımın bulunmadığı, zihinsel aktivitenin olmadığı ve uyanıklık halinde gözlenen bu salınımlar sonraları “Alfa Dalgaları” olarak isimlendirilmiştir. Bu çalışmasıyla Berger İnsan beyin işaretlerinin zihinsel işlevlerden, hafızaya birçok aktiviteyle ilişkili olduğunu ortaya atan ilk kişi olmuştur. Daha sonraları veri kaydı ve analiz yöntemlerinin gelişmesiyle “Elektroensefalogram” (EEG) ismini alan bu işaretler insan beyninin incelenmesinde en önemli tekniklerden biri haline almıştır.

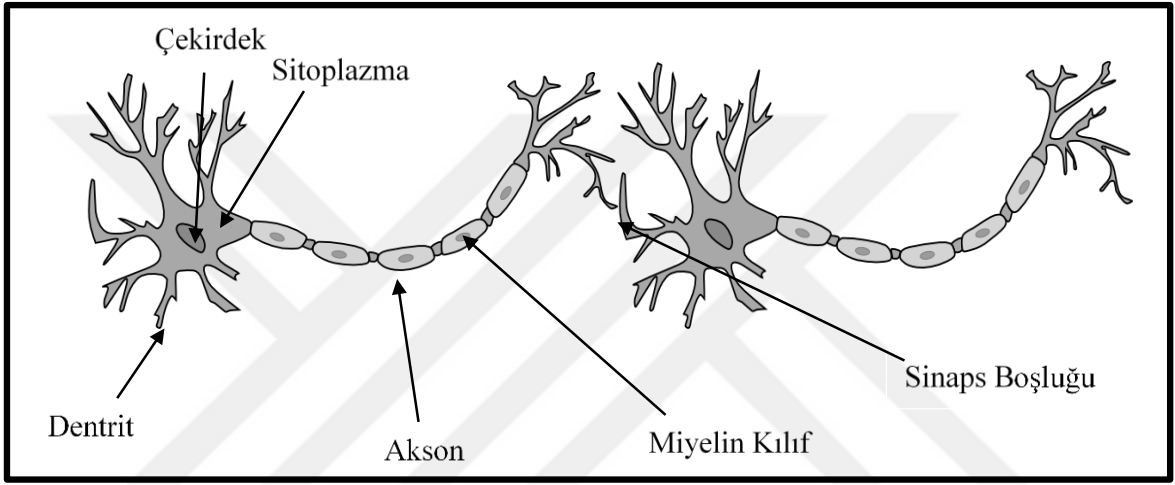
EEG işaretleri üzerinde Fourier analizi yapan ilk kişi olan Dietsch'den [45] sonra Gibbs ve diğerleri [46] EEG'nin epilepsi hastalarında kullanılmasının, Loomis ve diğerleri [47] ise uykuda EEG'nin incelenmesinin öncüleri olmuşlardır. Bir yıl sonra Walter [48] tümör varlığında “Delta” bandında aktivite olduğunu ortaya koymuştur. 1947 yılında Amerikan EEG Topluluğunun kurulması, 1949 yılında ise ilk EEG dergisinin çıkartılmasıyla EEG'ye olan ilgi daha da artmıştır. Jasper [49] uluslararası 10-20 sistemini uluslararası federasyon adına tanıtmıştır. Bu tarihten itibaren EEG'ye olan bilimsel ilgi giderek artmış ve günümüzde EEG birçok alanda önemli bir teknik olarak kullanılmaya devam etmektedir.

EEG'nin kullanımının hızla artmasının nedenleri arasında ölçüm maliyetlerinin ucuz olması, zamansal çözünürlüğün yüksek olması ve noninvasif (acı vermeyen) olması sayılabilir. Bu özellikleri diğer beyin görüntüleme tekniklerine göre daha çok kullanılabilir olmasını sağlamıştır.

##### **1.4.1. Sinir Hücresinden EEG'ye**

Sinir hücreleri; bir hücre gövdesi, dentritler, akson ve akson terminallerinden meydana gelen; vücuttaki sinirsel elektrik aktivitelerin iletiminden sorumlu ve bu özelliği itibarıyla

sinir sisteminin temel yapıtaşı olma görevini üstlenmiş hücrelerdir. Hücre gövdesinde çekirdek ve sitoplazma bulunur. Akson terminalleri ve dentritler sinir hücrelerinin birbirine bağlanmasını sağlayan kısımlardır. Bir sinir hücresinin akson terminali başka bir sinir hücresinin dentriti ile bağlanır. Bir hücrenin akson terminali ve dentrit sayısının miktarına bağlı olarak bir sinir hücresi binlerce sinir hücresine bağlanmış olabilir. Bu nedenle sinir sistemi karmaşık bir yapı haline gelir. Şekil 1.6’da iki sinir hücresinin bağlantısı ve anatomik yapısı gösterilmektedir.



Şekil 1.6. Sinir hücresi yapısı [36]

Akson sinir hücresinde iletimi sağlanacak olan “Aksiyon Potansiyeli” adlı özel bir elektriksel işaretin iletiminde görev alır. Bu iletim hücre gövdesinden akson terminallerine doğrudur. Birçok sinir hücresinde akson etrafına sarılı halde miyelin kılıf bulunur. Bu kılıfın görevi akson üzerinde iletilen elektriksel işaretin zayıflamasını önlemek ve bu işaretin kaliteli bir şekilde iletilmesini sağlamaktır. Bu özelliği itibariyle miyelin kılıf için sinir hücresinin yükselteci tabiri kullanılabilir. İki sinir hücresinin birbirine bağlantısı doğrudan değildir. Bağlantı noktalarında iki sinir hücresi arasında sinaps boşlukları bulunur. Bu boşluklar iletim yapılması ve yapılmamasına bağlı olarak aralarında nörotransmitter adlı proteinler bulundurulur.

Aksiyon potansiyelleri hücre içi ve dışı arasındaki ani gerilim veya elektriksel potansiyel değişimleridir. Bir sinir hücresinin dinlenme durumunda hücre içi ve dışı arasında içi negatif ve dışı pozitif olmak üzere  $-60$  mV’luk elektriksel potansiyele sahiptir. Bu potansiyele dinlenme potansiyeli ismi verilir. Aksiyon potansiyelinin maksimum olduğu

noktada +50 mV'luk elektriksel potansiyele sahiptir ki bu da yaklaşık 110 mV'luk bir farka denk gelir. Hücre içi ve dışı arasındaki bu depolarizasyonu hızlı bir repolarizasyon izler ve yeniden dinlenme değerine düşer. Aksiyon potansiyelleri akson üzerinde hızlı hareket eder. Bu potansiyeller 100 m/s hızıyla ilerleyebilirler. Bazı aksonların uzunluğunun bir metreden fazla olduğu düşünülürse bu hızla birkaç milisaniyede iletimin sağlanacağı anlamına gelir.

Sinapslar sinir hücrelerinin birbirleri ile iletişimi sağlamakla yükümlü özelleşmiş alanlardır. Kimyasal sinapslar ve nadir bulunsa da elektriksel sinapslar olmak üzere iki çeşittir. Kimyasal sinapslar nörotransmitter isimli kimyasallar yardımı ile hücreler arası bağlantıyı sağlar. Bir hücrenin aksonundan gelen aksiyon potansiyeli belirlenmiş eşik değerinin üzerine çıkarsa sinaps boşluğuna nörotransmitterler dolar ve iletim sağlanır [50].

EEG sinir hücrelerinin aktivasyonu ile oluşan aksiyon potansiyellerinin cebirsel toplamıdır. Bir bölgede senkron olarak aktif olan sinir hücrelerinin aksiyon potansiyelleri, o bölgede kayıt yapmakta olan EEG kanalına toplamsal olarak yansır. Beyin yüzeyinde alınan ECoG ölçümleri genlik ve frekans aralığı olarak çok daha yüksek olsa da kafatası ve kafa derisinin; genliği zayıflatma ve alçak geçiren süzgeç etkisinden dolayı EEG işaretleri daha zayıflamış şekilde ölçülür.

#### 1.4.2. EEG Alt Bantları

Kişilerin farklı zihinsel ve hafıza aktiviteleri EEG işaretlerinin farklı frekans bileşenlerinde gözlemlenebilir. Gerçekleştirilen aktiviteler farklı bantlarda aktivasyona sebebiyet verir. EEG işaretleri de bu yüzden Delta (1-4 Hz), Teta (4-8 Hz), Alfa (8-13 Hz), Beta (13-30 Hz) ve Gama (30-üstü Hz) olmak üzere beş ana alt banda ayrılır [51].

1. Delta Bandı: Delta salınımları, sağlıklı kişilerde uyku esnasında gözlenen dalgalarıdır. Yetişkinlerde delta dalgalarının tümör [52] ve doku bozulması [53] olan bölgelerde delta aktivitesinin arttığı gözlenmiştir. Delta dalgaları ayrıca çocuklarda hayatlarının ilk iki yılında oldukça aktiftir. John ve diğerleri [54] yaş ilerlemesiyle birlikte delta dalgalarının azaldığını bildirmiştir.
2. Teta Bandı: Teta salınımları belirgin olarak uykuda görülmektedir. Uyanıklık durumunda iki farklı teta dalgalarına rastlanabilir [55]. Bunlardan biri uykulu olma durumudur. Bir diğeri ise beynin ön kısımlarında gözlenen odaklanmaya ve karmaşık zihinsel süreçlere bağlanan teta dalgalarıdır.

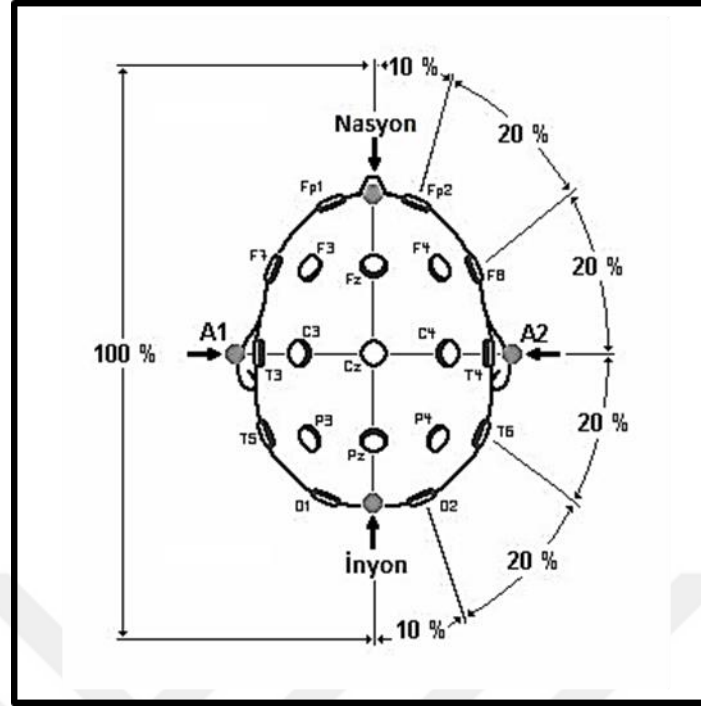
3. Alfa Bandı: Alfa salınımları diğer bantlar arasında gözlenmesi en kolay dalgalardır. Genelde rahat durumdaki uyanık kişilerde kolaylıkla gözlenir [56]. Ayrıca uyanıklık durumunda gözlerini kapamış kişilerde de rastlanır. Gözlerin açıldığı zamanın hemen sonrasında ise yüksek olan alfa dalgaları düşer. Özellikle bilişsel araştırmalarda Düşük Alfa (8-10 Hz) ve Yüksek Alfa (10-12 Hz) olarak iki alt banda ayrılır.
4. Beta Bandı: Beta salınımları bilişsel aktivitelerde yüksek olan alfa salınımlarının düşmesiyle ortaya çıkar. Bu bakımdan beta dalgaları dikkat atışı ile arttığı savunulmuştur [57].
5. Gama Bandı: Gama salınımları yüksek dikkat, nesne tanıma, problem çözme ve algısal bağlantılarla ilişkilendirilir (Engel vd., 2001). Bazı çalışmalarda algı ile ilgili [58] ve öğrenme ile ilgili [59] işlemlerde gama aktivitesinin arttığı gözlemlenmiştir. Anestezi uygulanan kişilerde gama bandında doza bağlı olarak düşüş gözlemlenmiştir [60]. Gross ve diğerleri [4] ise Gama salınımlarının uyku evrelerinde sistematik olarak değiştiğini göstermiştir.

### 1.4.3. EEG Veri Toplama Yöntemi

EEG işaretleri iki nokta arasındaki potansiyel farkı temsil eder. Bu bağlamda alınan işaret için bir aktif elektrot ve bir referans söz konusudur. Yüz yıla yakın EEG kayıt tecrübeleri göstermiştir ki alınan işaretlerin kalitesi seçilen referans ve kullanılan elektrotların kalitesine bağlıdır [61].

Elektrotlar EEG işaretinin karakteristiği gereği 0 ila 100 Hz arasında kalan bantları zayıflatmamalıdır. Ayrıca yapılan çalışmaların tutarlılık göstermesi amacıyla ortak bir elektrot yerleşim sistemine de ihtiyaç doğmuştur. Jasper [49] ortak bir elektrot yerleşim sistemi olarak uluslararası 10-20 elektrot sistemini sunmuştur. Bu isim kafatasının 4 noktasında %10 ve %20'lik bölümler ayırarak kişiden kişiye fark yaratmayan bir sistem olmasından gelir. Bu sistem sayesinde 19 elektrot sistematik bir şekilde konumlandırılabilir. Uluslararası 10-20 sistemi Şekil 1.7'de gösterilmiştir.

Uluslararası 10-20 sistemi üzerinde daha çok kanal elde etme amacıyla sonraları 10-10 sistemi [62] ve 5-5 sistemi [63] türetilmiştir. Bu sistemler birbirinden türetilmiş olsa da elektrotların kafatasına yerleştirilme süreleri, kaydı alınan kişinin rahatlığı ve enfeksiyon riski gibi nedenlerden ötürü uluslararası 10-20 sistemi en çok kullanılan sistem olmuştur.



Şekil 1.7. 10-20 EEG elektrot sistemi [61]

#### 1.4.4. İşitsel Olaya İlişkin Potansiyeller

Berger [44] kafa derisi üzerine konuşlandırılmış elektrotlar sayesinde beyin dalgalarını ölçülebildiğini gösterdikten sonra bu işaretlerle beynin nörofizyolojik aktivitelerin arasındaki bağlantının araştırılması hız kazanmıştır. Berger ve takipçileri bu işaretlerin daha çok salınımları üzerinde durmuş olsalar da daha sonraları kafa üzerinde belli olaylara bağlı potansiyel değişimler konusunda araştırmalar da yapılmaya başlanmıştır.

Olaya ilişkin potansiyeller özel bir olaya ilişkin beynin belirli bölümlerindeki potansiyel değişimlerinin incelendiği bir teknik olarak ortaya çıkmıştır. Bu teknik temel olarak aynı uyarımın defalarca kes dışarıdan deneğe uygulanması ve bu olayın beynin belirli bölgelerindeki bu olaya bağlı değişimleri bulmayı amaçlar. Birçok kez tekrarlanan uyarımlara ait EEG kayıtlarının ortalamaları alınarak olaya ilişkin potansiyeller elde edilir.

Bir uyarım anında beyin sadece o uyarıma karşılık gelen tepkiyi vermez. Eş zamanlı olarak devamlı yapılmakta olan aktiviteler istenen uyarıma ait işaretin elde edilmesinde bir problem yaratır. Beynin belirli bölgesindeki toplamsal aktiviteyi temsil eden bir kanal kaydı içerisinde birçok bileşeni bulundurmış olur. Sadece olaya ilişkin potansiyellerin elde edilmesi amacıyla ilgili kayıtların ortalaması işaretler zamana kilitlenerek alınır ve beynin istenen olaya ilişkin tepkisi yüksek zaman çözünürlüğünde elde edilir.

Ortalaması alınan işaretlerin miktarının fazlalığı elde edilen olaya ilişkin potansiyellerin güvenilirliği açısından önemlidir. Fazla verinin yüksek temsil anlamına gelmesi nedeniyle ortalaması alınan verinin sayısının artması ile olaya ilişkin potansiyelin İşaret-Gürültü oranı yüksek olacaktır. Bu yüzden bu teknikte olabildiğince çok miktarda kayıt yapılmaya çalışılır.

#### **1.4.4.1. Çok Erken İşitsel Bileşenler**

Uygun koşullar oluşturulduğunda, işitsel bir uyarımın yapılmasında 10 ms sonra olaya ilişkin potansiyele ait bir tepe görmek mümkündür. Çeşitli kaynakların iddia ettiğine göre bu tepeler beyin sapının işitsel yolları ile ilgilidir. Bu yüzden bu tepelere beyin sapı uyarılmış tepkiler (BERs) ya da işitsel beyin sapı tepkileri (ABRs) de denir. Bu tepeler özellikle çocuklarda işitsel rahatsızlıkların tespitinde kullanılır [64]. Beyin sapı uyarılmış tepkileri 10-50 ms aralığında orta gecikme bileşenleri takip eder. Bu bileşenler işitsel korteks ile ilişkilendirilir.

#### **1.4.4.2. Uyumsuzluk Negativitesi**

Uyumsuzluk negativitesi, denekler araka arkaya hep aynı uyarımı dinledikleri deneyde nadir olarak dinledikleri ikinci uyarımda görülür. Uyumsuzluk negativitesi negatif potansiyele sahiptir ve kafa derisinin orta çizgisinde en yüksek mutlak potansiyele sahiptir. Genelde 160- 220 ms arasında tepeye ulaşır. Kitap okurken işitsel uyarım verilen deneklerde de uyumsuzluk negativitesine rastlanmıştır. Uyumsuzluk negativitesi benzer olarak devamlı tek bir kulaktan uyarım verilen deneye nadir diğer kulaktan uyarım verildiğinde de ortaya çıkmaktadır [65].

### **1.5. Literatür Araştırması**

Beyinde konuşmacı tanıma sinirsel temelde henüz tam anlaşılabilmiş bir konu değildir. Konuşma üretim, konuşmanın anlaşılması ve dilsel kodlamada yapıldığı yerler beyin üzerinde konumlandırılabilse [37] de konuşmacı kimliğinin belirlendiği bölge hala netlik kazanmış değildir [66]. Bazı klinik çalışmalar beynin bazı bölgelerinde hasar oluşmuş

hastalarda konuşmacı tanıma ve ayırma süreçlerinin sekteye uğradığını bildirmiştir [67]–[70]. Bu çalışmalar kişi seslerinin birbirinden ayrılması için sorumlu bölgelerin daha çok beynin sağ yarı küresinde bulduklarını ortaya atmışlardır. Bununla birlikte bazı çalışmalar beyinde kişi seslerine duyarlı olan bölgelerin mevcut olduğunu bildirmiştir. Belin ve diğerleri [71] vokal ve doğal sesler olacak şekilde iki farklı sınıflı uyarımlı fMRI sunduğu çalışmasında beyin üzerinde insan sesine duyarlı bölgeleri tespit etmeyi amaçlamıştır. Birçok bölgede herhangi bir aktiviteye rastlanmasa da her kişi için farklı yerler olmak üzere işitsel kortekste insan sesi uyarımlarına yüksek tepki gösteren bölgeler gözlemlenmiştir. Bunun yanında sunulan doğal seslere tepki veren özel bir alan gözlemlenmemiştir.

Kişiler insan seslerini birbirinden ayırmaya da anne karnındayken başlar. Kisilevsky ve diğerleri [2] anne karnındaki fetüsler üzerinde yaptığı çalışmasında fetüslerin annelerinin sesini tanıyabildiklerini ortaya koymuştur. Anne adaylarının çocuklarına sözel olarak seslenmelerini kaydeden araştırmacılar, bu kayıtları daha sonra anne karnındaki fetüslere dinletmişlerdir. Bütün anne kayıtlarını bütün fetüslere dinleten araştırmacılar fetüs kalplerin artışları üzerinde belirgin değişiklikler gözlemlenmiştir. Buna göre kendi anne sesini duyan fetüslerin kalp atım oranı, diğer annelerin seslerini dinlediği duruma göre daha yüksek çıkmıştır.

Anne karnında fetüsler kendi annelerini tanıyor olsa da diğer kişilerin sesleriyle ilgili bir aşinalık göstermemişlerdir. Kisilevsky'den bağımsız olarak yürütülen Beauchemin ve diğerlerinin yürüttüğü çalışmada bu durum henüz doğmuş bebekler üzerinde araştırılmıştır [3]. Bu çalışmada yeni doğmuş bebeklere yine annelerinin ve başka annelerin sesleri dinletilmiş ve bebeklerden EEG kaydı alınmıştır. Bebekler annelerinin seslerini duyduklarında merkezi elektrot bölgelerinde uyumsuzluk negativitesi gözlenmiştir. Bu da aşinalığın bebeklerin beyinlerinde nasıl etki gösterdiğine ışık tutmuştur.

Kişilerin hayatlarının erken yaşlarında konuşmacı aşinalığının ne zaman başladığını ortaya koymak için Belin ve diğerleri 4 ila 7 aylık bebekler üzerinde çalışmışlardır. Bebeklere tanıdığı ve tanımadığı kişilere ait sesler dinletilmiştir. bu grup içinde 7 aylık olan bebeklerde tanıdık seslerde beynin sağ ön temporal bölgesinde etkinlik görülmüş bunun yanında 4 aylık bebeklerde herhangi bir farklılık görülmemiştir. Bu bulgular ışığında Belin ve diğerleri kişi ses tanıma yetisinin 7 aydan önce 4 aydan sonra gelişimini tamamladığı sonucuna varmıştır [4].

Kişi aşinalığının beynin hangi bölgesinde yoğunlaştığı da önemli bir bilgi olarak göze çarpmaktadır. Barret ve diğerleri 1988 yılında yaptıkları çalışmalarında kişilere tanıdıkları

ve tanımadıkları yüzler olmak üzere iki farklı uyarım göstermişlerdir. Deneysel süresince kaydedilen EEG kayıtlarından elde edilen ERP çalışmasında kişilere tanıdık yüzler sunulduğunda beyin sağ yarı küresinde sol yarı küreye göre daha negatif takip eden bileşenlere rastlanmıştır [72].

Schweinberger ve diğerlerinin yaptığı çalışmada Barret ve diğerlerinden farklı olarak kişilere tanıdık ve tanımadık sesler olmak üzere iki farklı uyarım dinletmişlerdir. EEG kayıtlarında elde edilen ERP'ler göstermiştir ki Barret ve diğerlerinin çalışmalarına benzer olarak bu çalışmada da tanıdık kişilerin sesleri sunulduğunda beyin sağ yarı küresinde bir aktivasyon artışı meydana gelmiştir [73].

Schweinberger ve diğerleri bir başka çalışmalarında kişilere insan yüzleri, ters çevrilmiş insan yüzleri, arabalar ve maymunlar sunmuşlardır. Bir ERP çalışması olarak yapılan bu çalışma sonunda kişilerin beyinlerinde 250 ms sonra tepki görmüşlerdir. Bu tepkilerde genliğin en yüksek olduğu sınıf insan yüzleri olurken, ters çevrilmiş insan yüzlerinde önemli bir değişime rastlanamamıştır [74].

Beauchemin ve diğerlerinin yaptığı bir başka ERP çalışmasında kişilere tanıdıkları ve tanımadıkları sesler dinletilmiştir. ERP sonuçlarında beyin ön bölgesinde uyumsuzluk negativitesi ve P3a bileşeni gözlenmiştir. Bu da aşinalığın beyin ön bölgelerini de ilgilendirdiğini ortaya koymuştur [75]. Daha sonrasında Schweinberger ve diğerlerinin 2011 yılında yürüttüğü bir başka kişi sesi aşinalığı çalışmasında aşına seslerde beyin sağ ön temporal bölgesinde etkinlik gözlenmiştir. Araştırmacılar yapılan ERP çalışmalarında Beauchemin ve diğerlerinin sonuçlarına yakın sonuçlara ulaşmışlardır [76].

Graux ve diğerlerinin 2013 yılında yaptıkları çalışmada araştırmacılar kişilere kendi seslerini ve başka kişilere ait sesleri dinletmişlerdir. Alınan EEG kayıtlarından elde edilen ERP'ler sonucunda beyinde kişilerin kendi seslerine bariz bir şekilde tepki gösterdikleri ortaya koyulmuştur. Buna göre orta çizgi üzerindeki bölgelerde kişilerde kendi sesleri için daha düşük genlikte P3a bileşeni gözlemlenirken, başka kişilere ait seslerde genlik daha yüksek genlikte gözlenmiştir [77].

Graux ve diğerleri 2015 yılında gerçekleştirdikleri bir diğer çalışmalarında ise kişilere kendi sesleri, tanıdıkları kişilere ait sesler ve tanımadıkları kişilere ait sesler olmak üzere üç farklı uyarım dinletmişlerdir. Kişilerin kendi sesleri ve tanıdıkları sesler arasındaki bağlantının araştırıldığı bu çalışmada ERP verileri kişilerin kendi sesleri ve tanıdık sesler arasında benzer örüntüler olduğunu göstermiştir. Bu benzerlik örüntü olarak ortaya çıksa da



kişilerin kendi seslerinde oluşan P3a bileşenlerinin genlik olarak daha düşük olduğu da gözlenmiştir [78].

Holeckova 2006 yılında yaptığı çalışmasında kişilere tanıdıkları ve tanımadıkları kişilerin ismini söyledikleri ve vokal olmayan sesler olarak üç farklı uyarım sunmuştur. Kişiler tanıdıkları kişiler tarafından söylenen kendi isimlerini duyduklarında orta çizgi üzerinde P3a bileşenine rastlanmıştır [79].

Klinikte karşılaşılan kişi seslerini birbirinden ayıramama hastalığı olarak bilinen Fonagnozi bu çalışmalarda da araştırma konusu olmuştur. von Kriegstein ve diğerlerinin yürüttüğü çalışmada, araştırmacılar bu hastaların kişi tanıma süresince beyinlerini incelemiştir. Yapılan çalışma sonucunda kişi ses tanıma sürecinin sadece ses tanıma ile ilgili bölgelerde sürdürülmediği aynı zamanda yüz tanıma bölgelerinin de bu süreçte etkin olduğunu ortaya koymuştur [80].

Gonzalez ve diğerleri aşinalığın sinirsel tabanını araştırmak için yola çıktıkları çalışmalarında kişilere sadece yüz, sadece ses ve yüz ve ses olarak üç farklı deneyde tanıma ve tanıyamama durumunu incelemiştir. ERP verilerinde görülmüştür ki 250 ms sonra ortaya çıkan aşinalığı bağlı N250 bileşeni sağ temporal bölgelerde en yüksek genliktedir [81].

Sun ve diğerlerinin 2012 yılında yaptıkları çalışmalarında kişilere yine tanıdıkları ve tanımadıkları sesler olarak iki farklı uyarım dinletilmiştir. Araştırmacılar bu çalışmalarında beyin orta, parietal ve temporal bölgelerinde pozitif takip eden N400f bileşenine rastlamışlardır [82].

Zaske ve diğerlerinin 2014 yılında yaptıkları çalışmaları kişi seslerinin öğrenilmesindeki süreçlere ışık tutmuştur. Araştırmacılar kişilere daha önce tanıdıkları kişilere ve yeni tanıdıkları kişilere ait sesler olmak üzere iki farklı uyarım dinletmişlerdir. Bu çalışma sonucunda kişilerin seslerini yeni tanıdıkları seslerde Beta salınımları gözlenmiştir [83].

Aşinalık kişilerin gerçekleştirdikleri eylemlerdeki performanslarını da etkilemektedir. Bu konuyla ilgili yapılan bir çalışmada Thammasan ve diğerleri duygu tanıma ile ilgili bir çalışmada bu etkiyi incelemiştir. Çalışmada kişilerden dinledikleri müzik parçalarının duygusal olarak sınıflandırmalarını istemiş ve bu deneyi kişilerin tanıdıkları müzik parçaları ve tanımadıkları müzik parçaları olmak üzere iki farklı şekilde gerçekleştirmiştir. Daha sonra EEG verileri üzerinden sınıflandırma yapan araştırmacılar tanıdık müziklerde yapılan sınıflandırmanın daha başarılı olduğunu ortaya koymuşlardır [84].

Aşinalığın belirlenmesi literatürde hala muhtaç kalmış konular arasındadır. Kişilerin aşinalığının tespit edilmesinin uygulamada çok önemli katkılar sunacağından yola çıkılarak bu çalışmada aşinalığın tespiti amaçlanmıştır. Bu konuyla ilgili az sayıdaki çalışmalardan biri olarak Ernest ve diğerleri yüz aşinalığının belirlenmesini amaçladıkları çalışmalarını sunmuşlardır. Araştırmacılar kişilere tanıdıkları ve tanımadıkları kişi yüzlerini göstererek aşinalığının tespiti üzerine çalışmışlardır. Yapılan çalışma sonucunda Teta bandında ortalama %70,71 başarımla yüz aşinalığını belirleyebilmişlerdir [85]. Benzer bir başka çalışma olarak Smitha ve diğerleri tanıdık ve tanımadık seslere ait aşinalığı belirlemek için çalışma yürütmüşlerdir. Kişilere dinletilen tanıdık ve tanımadık sesleri Teta bandı üzerinde ortalama %72,2 başarımla belirleyebilmişlerdir [86].

Literatüre göz atıldığında Aşinalığın belirlenmesi ve aşinalığı bağlı salınımların ortaya koyulması hala tam olarak sağlanmış değildir. Yapılan çalışmalar daha çok ERP tabanlı olmakla beraber fMRI ve PET çalışmaları bunları takip etmektedir. Çalışmalar arasında aşinalığın bağlı olduğu salınımlara rastlanamamaktadır. Bellek ile ilgili çalışmalardan yola çıkılarak aşinalığının Teta bandına bağlı olması gerektiği düşünülmektedir. Fakat bunun konuşmacı aşinalığının belirlenmesinde etkin ve gerekli olduğunu ortaya koyan çalışmalar sayıca çok değildir. Bunun yanında aşinalığın belirlenmesinin adli suçlara muhtemel katkısı düşünüldüğünde bu tip çalışmalar bir ihtiyaca dönüşmüştür.

Bu çalışmada eksik görülen bu konular üzerine yoğunlaşmıştır. Kişi sesi aşinalığı üzerine yapılan çalışmalarda uyarımların seçilmesinde takip edilen yöntemlerin geliştirilmesi ve daha güvenilir sonuçlar elde edilmesi amaçlanmıştır. Konuşma işaretlerinin yapısı ve kişilerin bunlar içinde çıkarabilecekleri bağıl başka bilgiler bu çalışmaların güvenilirliğini azaltmaktadır. Bu tip bilgisel girişimlerin ortadan kaldırılması amacıyla bu çalışmada uyarımların belirlenmesinde konuşmacı tanıma alanından esinlenilmiş bir yöntem akışı takip edilmiştir.

Bunun yanında aşinalığı bağlı salınımların belirlenmesinde de her alt banda eşit şekilde yaklaşarak içlerinden en etkin olanının belirlenmesi amaçlanmıştır. Gerek yapılan çalışmaların azlığı gerekse yapılan sınıflandırma başarımlarının düşük olması her alt banda eşit yaklaşılması gerektiği gerçeğini ortaya koymaktadır. Bu çalışma da olabildiğince gerçekleştirilebilir ve uyarımların belirli bir mantığa uygun olarak seçildiği ayrıca her alt banda eşit şekilde yaklaşan bir çalışma olarak sunulmuştur.

## 2. YAPILAN ÇALIŞMALAR, BULGULAR VE İRDELEME

### 2.1. Ön Çalışma: Bilgisayar Tabanlı Otomatik Konuşmacı Tanıma

#### 2.1.1. Materyal

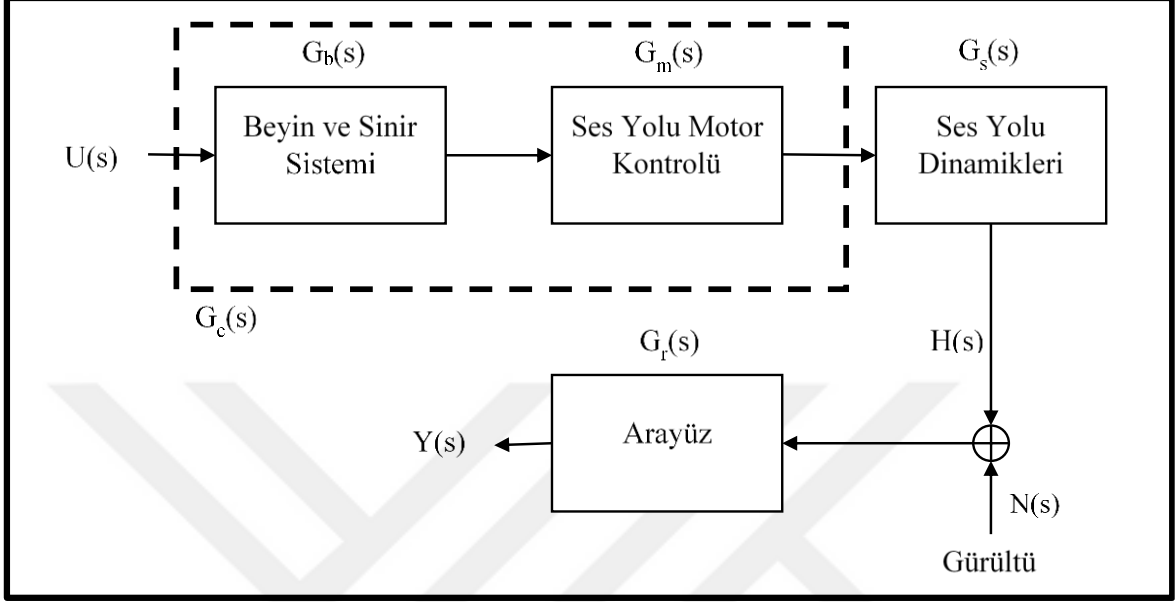
Bu çalışmada “LibriSpeech ASR corpus” isimli veri kümesine ait “dev-clean” ve “test-clean” isimli alt kümeler birleştirilerek kullanılmıştır [87]. Her bir alt kümede 20’si kadın, 20’si erkek olmak üzere 40’ar adet konuşmacı kaydı bulunmaktadır. Bu alt kümelerin birleşiminden ortaya çıkarılan veri kümesinde ise 40 erkek, 40 kadın olmak üzere toplam 80 kişiye ait konuşmacı kaydına ulaşılmıştır. Veri kümesi içerisinde her bir konuşmacıya ait parça parça sunulmuş toplam 8’er dakikalık kayıtlar bulunmaktadır. Kayıtlar farklı kişilere okutulmuş farklı metinlerden oluşmaktadır. Bu yüzden bu çalışma metinden bağımsızdır. Veri kümesi bu çalışma için yeniden düzenlenmiştir. Farklı uzunluktaki farklı kayıtlar kişi bilgisine sadık kalınarak rastgele arka arkaya eklenmiştir. Daha sonra elde edilen kişilere ait 8’er dakikalık kayıtlar 5’er saniyelik parçalar halinde ayrılmış ve karıştırılmıştır. Karıştırma işleminden sonra kişi başı 96 adet olan kayıtlar içinden rastgele 80 adeti çalışmada kullanılmak üzere ayrılmıştır. Konuşma işaretleri 16 kHz ile örneklenmiş olup veri kümesi itibariyle minimum gürültü seviyelerindedir.

#### 2.1.2. Yöntem

İnsan konuşma sistemi bir kontrol sistemi olarak ele alınırsa, konuşma işareti nin ses yolu karakteristiğine bağlı olduğu görülür. Şekil 2.1’de insan konuşma sisteminin kontrol sistemi olarak sunulduğu akış diyagramı gösterilmektedir.

Burada  $U(s)$  konuşma sisteminin başlangıcı olan beyinde sürecin başlamasından sorumlu  $U(t)$  işaretinin Laplace Dönüşümünü temsil etmektedir. Beyin ve sinir sisteminin Transfer fonksiyonu olan  $G_b(s)$  ve ses yolu motor kontrolüne ait transfer fonksiyonu olan  $G_m(s)$ ,  $G_c(s)$  olarak birleştirilebilir. Ses yolu karakteristiğinin transfer fonksiyonu ise  $G_s(s)$  olarak gösterilmiştir. Sistemin çıkışı olan  $H(s)$ , üretilen ses işaretlerini temsil eder. Bu ses işaretlerinin taşınmasında işarete karışan gürültü  $N(s)$  ile ifade edilmiştir. Kayıt cihazlarının veya kişisel kulak yolları karakteristiğinin etkisinin gösteren arayüzler  $G_r(s)$  olarak ifade

edilmiştir. Son olarak üzerinde çalışılmak üzere elde edilen işaret  $Y(s)$  olarak temsil edilmiştir.



Şekil 2.1. Bir kontrol sistemi olarak konuşma akışı

Bu kontrol sistemi üzerinde  $U(s)$  girişi ile  $Y(s)$  çıkışı arasında matematiksel ifade türetildiğinde;

$$H(s) = G_s(s)G_c(s)U(s) \quad (2.1)$$

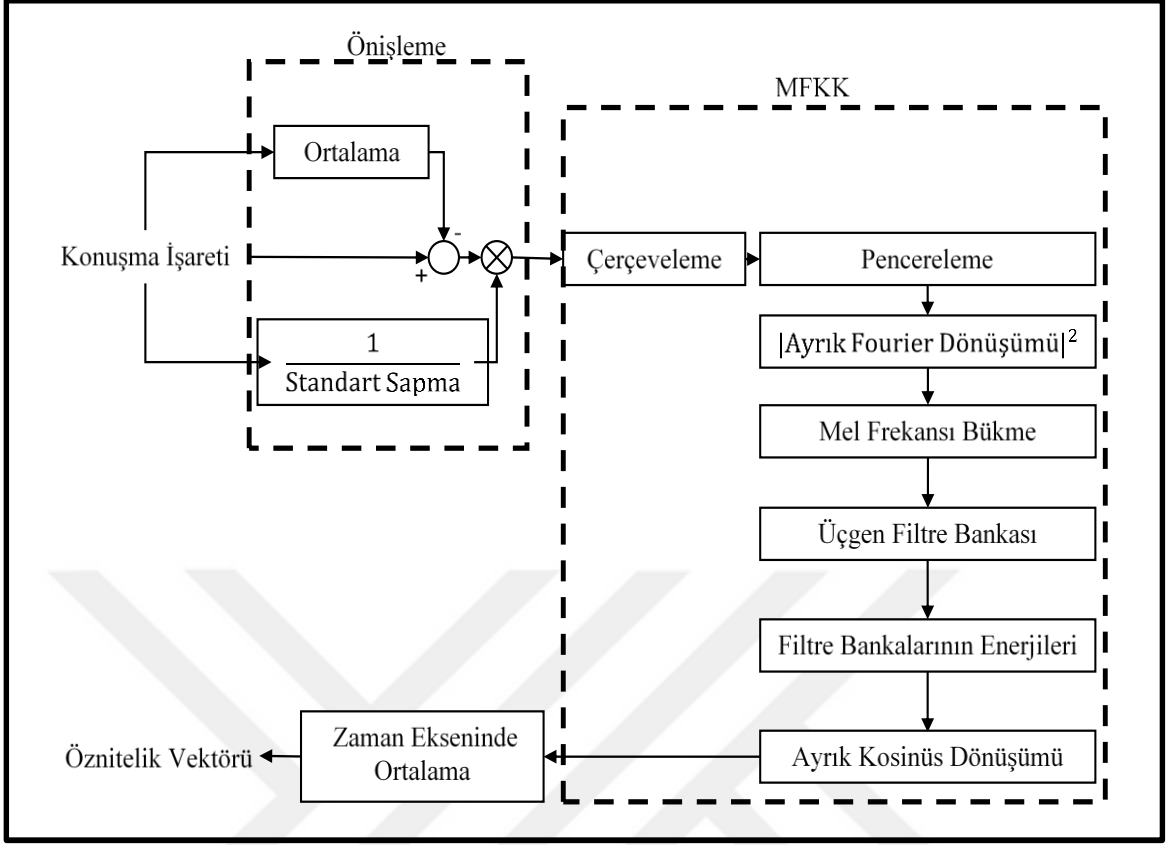
$$Y(s) = G_r(s)[H(s) + N(s)] \quad (2.2)$$

elde edilir. Bu iki denklem birleştirildiğinde aşağıdaki denklem elde edilir:

$$Y(s) = G_r(s)[G_s(s)G_c(s)U(s) + N(s)] \quad (2.3)$$

Burada amaç, gözlenen  $Y(s)$  üzerinden sistemin tanınmasını sağlamaktır. Konuşma tanıma uygulamalarında bu sistem üzerinden  $G_c$  kestirilmeye çalışılır. Burada  $G_s$  konuşma tanıma uygulamaların istenmeyen bir gürültüdür. Bunun tam tersi olarak metinden bağımsız konuşmacı tanıma uygulamaları içinse  $G_s$  kestirilmesi gereken bileşen olurken  $G_c$  istenmeyen bir gürültüdür.

Bu çalışmada konuşmacı tanımda doğru model oluşturma ve  $G_s$ 'nin en iyi temsil edilmesi amacıyla Şekil 2.2'de gösterilen yöntem akışı kullanılmıştır. Elde edilen  $Y(s)$  işaretleri bir ön işleme aşamasında geçirildikten sonra, Mel Frekansı Kepstrum Katsayıları hesaplanmış ve bu katsayılar üzerinde sınıflandırma yapılmıştır. Yapılan işlemlere ait detaylı anlatım bu bölümde sunulacaktır.



Şekil 2.2. Konuşma işaretinin önişleme ve MFKK hesaplanmasına ait yöntem akışı

### 2.1.2.1. Önişleme

Konuşmacı tanımada sınıflandırma performansını etkileyen en önemli etkenlerden biri kayıt esnasında kişinin yüksek sesle konuşması, kayıt cihazının kişi ile arasındaki mesafe, kayıt esnasında kayıt cihazının ses seviyesi ve kayda karışmış DC frekans bileşenleridir. Hesaplanan özniteliklerin sınıflandırıcılarda doğru bir sonuç gösterebilmesi için normalleştirme işleminin yapılması gerekir. Bu çalışmada ortalama ve standart sapma normalleştirmeleri yapılmıştır. Eşitlik 1’de bu işleme ait matematiksel ifade verilmiştir.

$$x_{Nor} = \frac{x - \bar{x}}{std(x)} \quad (2.4)$$

Burada  $x$  bir denemeye ait işaret olmak üzere  $\bar{x}$ ,  $std(x)$  ve  $x_{Nor}$  sırasıyla denemenin ortalama değeri, denemeye ait standart sapma değeri ve normalleştirilmiş denemeyi ifade eder.

### 2.1.2.2. Mel Frekans Kepstrum Katsayıları

#### 2.1.2.2.1. Çevçeveleme ve Pencereleme

Konuşma işaretlerinin zamanda sürekli değişmesi ve kişi bilgisinin bu değişen konuşma işaretindeki ses yolu karakteristiğinde saklı olması konuşma işaretlerinin parçalar halinde incelenmesini gerektirir. Parçalanmış konuşmalardan elde edilen kesitler Ayırık Fourier dönüşümünün sonsuz uzunluk etkisi nedeniyle bir pencere fonksiyonu ile çarpılır. Farklı pencere uzunlukları çok farklı etki göstermese de bir pencere fonksiyonunun kullanılması gerekli görülmektedir [1]. En çok kullanılan pencere fonksiyonlarında bazılarının ait matematiksel ifadeler şu şekildedir:

Hamming Penceresi:

$$w(n) = 0.54 - 0.46 \cos\left(\frac{2\pi n}{N-1}\right) \quad (2.5)$$

Hann Penceresi:

$$w(n) = 0.5 \left(1 - \cos\left(\frac{2\pi n}{N-1}\right)\right) \quad (2.6)$$

Üçgen Pencere:

$$w(n) = 1 - \left|\frac{2n - N + 1}{N - 1}\right| \quad (2.7)$$

Gauss Penceresi:

$$w(n) = e^{-\frac{1}{2} \left(\frac{n - \frac{N-1}{2}}{\sigma \frac{N-1}{2}}\right)^2}, \quad \sigma \leq \frac{1}{2} \quad (2.8)$$

Bu matematiksel ifadelerde  $w(n)$  pencere fonksiyonunu,  $N$  pencere uzunluğunu,  $n$  pencere eleman indisini ve  $\sigma$  ise standart sapmayı ifade etmektedir.

#### 2.1.2.2.2. Ayırık Fourier Dönüşümü

Fourier serisi bir fonksiyonun kosinüsler ve sinüslerin toplamı halindeki açılımıdır. Bu sayede gerekli katsayıların hesaplanması dahilinde işaretler bu formata dönüştürülebilir. Fourier serileri açılım ifadesi:

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos(n\omega t) + b_n \sin(n\omega t)) \quad (2.9)$$

Fourier serisi hesaplamaları armonik analiz olarak bilinir ve seçilmiş herhangi bir fonksiyonun bir dizi basit terimlere ayrılarak, ayrık terimler olarak çözülmesi ve yeniden birleştirilip orijinal problemin çözümü için oldukça kullanışlı bir yoldur. Böylelikle problem istenilen ya da pratik olan bir yaklaşıklık ile çözülebilir.

Burada;

$\omega$ : Temel Frekans

$$a_0 = \frac{2}{T} \int_0^T f(t) dt \quad (2.10)$$

$$a_n = \frac{2}{T} \int_0^T f(t) \cos(n\omega t) dt \quad (2.11)$$

$$b_n = \frac{2}{T} \int_0^T f(t) \sin(n\omega t) dt \quad (2.12)$$

olarak hesaplanır.

Burada integral limitleri olarak  $\int_{-T/2}^{T/2} \dots$  kullanılabilir. Euler eşitliğinden;

$$e^{\pm jx} = \cos(x) \pm j\sin(x) \quad (2.13)$$

Burada  $\cos(x)$  ve  $\sin(x)$  çekilirse;

$$\cos(x) = \frac{e^{jx} + e^{-jx}}{2} \quad (2.14a)$$

$$\sin(x) = \frac{e^{jx} - e^{-jx}}{j2} \quad (2.14b)$$

Bu ifadeler aşağıdaki ifadede yerine yazılacak olursa;

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos(n\omega t) + b_n \sin(n\omega t)) \quad (2.15)$$

$$= \frac{a_0}{2} + \sum_{n=1}^{\infty} \left( a_n \left[ \frac{e^{jn\omega t} + e^{-jn\omega t}}{2} \right] - jb_n \left[ \frac{e^{jn\omega t} - e^{-jn\omega t}}{2} \right] \right) \quad (2.16)$$

$$= \frac{a_0}{2} + \sum_{n=1}^{\infty} \left( e^{jn\omega t} \left[ \frac{a_n - jb_n}{2} \right] + e^{-jn\omega t} \left[ \frac{a_n + jb_n}{2} \right] \right) \quad (2.17)$$

$$= \frac{a_0}{2} + \sum_{n=1}^{\infty} e^{jn\omega t} \left( \frac{a_n - jb_n}{2} \right) + \sum_{n=1}^{\infty} e^{-jn\omega t} \left( \frac{a_n + jb_n}{2} \right) \quad (2.18)$$

Burada;

$$c_0 = \frac{a_0}{2} \quad (2.19a)$$

$$c_n = \left( \frac{a_j - jb_n}{2} \right) \quad (2.19b)$$

$$c_{-n} = \left( \frac{a_n + jb_n}{2} \right) \quad (2.19c)$$

olduđuna gore;

$$f(t) = \sum_{-\infty}^{\infty} c_n e^{jn\omega t} \quad (2.20)$$

Denklem 2.19b zerinden  $c_n$  katsayısı hesaplandıđında;

$$= \frac{1}{2} \cdot \frac{2}{T} \int_0^T f(t) \cos(n\omega t) dt - \frac{j}{2} \cdot \frac{2}{T} \int_0^T f(t) \sin(n\omega t) dt \quad (2.21)$$

$$= \frac{1}{T} \int_0^T f(t) [\cos(n\omega t) - j \sin(n\omega t)] dt \quad (2.22)$$

$$= \frac{1}{T} \int_0^T f(t) e^{-jn\omega t} dt \quad (2.23)$$

Tm bu hesaplamalar sonucunda Ayırık Fourier Donşm ifadesi benzer olarak:

$$f(k) = \sum_{n=0}^{N-1} f(n) e^{\frac{-j2\pi kn}{N}} \quad (2.24)$$

olacaktır. Ters Ayırık Fourier Donşm ise:

$$f(n) = \frac{1}{N} \sum_{k=0}^{N-1} f(k) e^{\frac{j2\pi kn}{N}} \quad (2.25)$$

olarak hesaplanır.

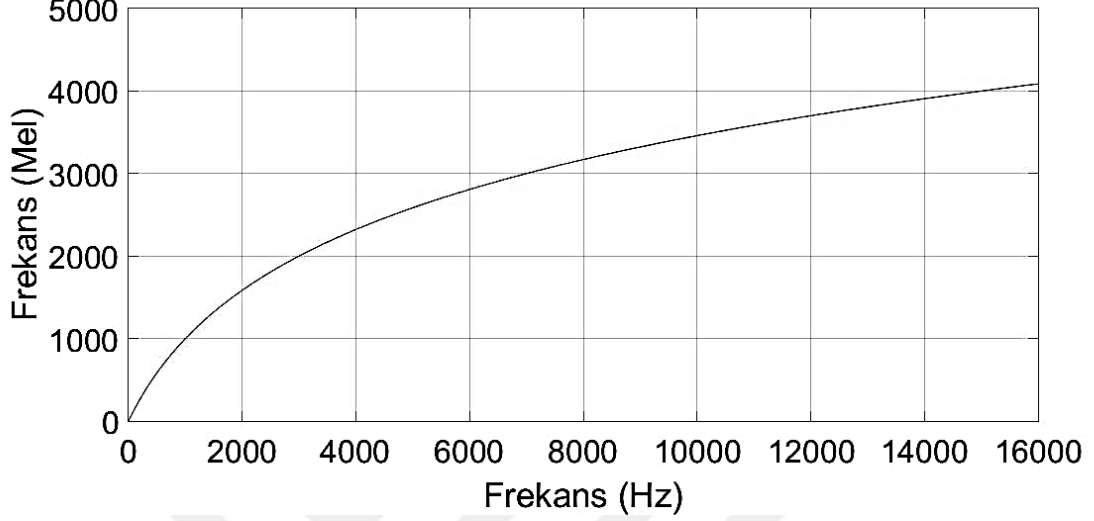
### 2.1.2.2.3. Mel Frekansı Bkme

İnsan iřitme sistemi 1 kHz' e kadar lineer bir algıya sahipken 1 Khz ve stnde logaritmik bir algıya sahiptir. Bu nedenle psikoakustik alanından esinlenerek oluřturulmuř perdenin tespitine dayalı Mel frekans skalası sunulmuřtur [88]. Denklem (2.26a) ve (2.26b)'de frekanstan Melodi (Mel) Frekansının hesaplandıđı formller verilmiřtir.

$$f_{mel} = \frac{1000}{\ln\left(1 + \frac{1000}{700}\right)} \ln\left(1 + \frac{f_{Hz}}{700}\right) \quad (2.26a)$$



$$f_{mel} = \frac{1000}{\log(2)} \log\left(1 + \frac{f_{Hz}}{1000}\right) \quad (2.26b)$$



Şekil 2.3. Mel-Frekans değişimi

#### 2.1.2.2.4. Üçgen Filtre Bankaları ve Enerjileri

MFKK genellikle üçgen filtre bankaları kullanılarak hesaplanır. Filtrelerin her biri üçgen pencere şeklindedir. Bu üçgen filtre bankaları frekans ekseninde Mel frekansı skalasında eşit aralıklarda olacak şekilde yerleştirilir. Bu filtreler şu şekilde ifade edilir:

$$H_m[k] = \begin{cases} 0, & k < f[m-1] \text{ ve } k \geq f[m+1] \\ \frac{k - f[m-1]}{f[m] - f[m-1]}, & f[m-1] < k < f[m] \\ \frac{f[m+1] - k}{f[m+1] - f[m]}, & f[m] \leq k \leq f[m+1] \end{cases} \quad (2.27)$$

Burada  $H_m[k]$  Üçgen filtre bankasının ifadesini  $f[m]$  m'inci filtrenin fonksiyonunu  $k$  ise eleman indisini temsil etmektedir. Şekil 2.3'te 10 adet üçgen filtre bankasının frekans uzayında yerleşimi gösterilmiştir. Bu filtre bankalarına ait logaritmik enerji ifadeleri denklem 2.28'deki gibi hesaplanır.

$$S[m] = \sum_{k=0}^{N-1} |x[k]|^2 H_m[k], \quad m = 0, 1, 2, \dots, M-1 \quad (2.28)$$

### 2.1.2.2.5. Ayrık Kosinüs Dönüşümü

MFKK'nın hesaplanmasında ilk zamanlar Ters Ayrık Fourier dönüşümü kullanılmış olsa da işaretin enerji katsayılarına artık bu aşamada genellikle Ayrık Kosinüs Dönüşümü uygulanmaktadır. Ayrık Kosinüs Dönüşümü ifadesi aşağıda verilmiştir.

$$x[n] = \sum_{m=0}^{M-1} S[m] \cos \left[ \left( m + \frac{1}{2} \right) \frac{\pi n}{M} \right], \quad n = 0, 1, 2, \dots, M - 1 \quad (2.29)$$

### 2.1.2.3. Sınıflandırma: Doğrusal Destek Vektör Makineleri

Vapnik [89] tarafından 1960'ların sonunda ortaya atılan Destek Vektör Makineleri karmaşık sınıflandırma problemlerine kısa ve basit çözüm sunma açısından muadilleri arasında ön plana çıkmaktadır. Temel ve basit kurallara dayanması daha hızlı ve güvenilirliği yüksek sonuçlar sunmasına olanak sağlar. Uygulama da “Görüntü Tanıma”, “Ses Tanıma”, “Yüz Tanıma” gibi birçok uygulamada sınıflandırıcı olarak tercih edilmektedir [90].

Doğrusal Destek Vektör Makineleri en büyük sınırlı sınıflandırma olarak bilinir. Sadece doğrusal olarak ayrılabilen veriler için geçerli olsa da temel Destek Vektör Makineleri mantığının temelini oluşturur. Bu yöntemde verileri en iyi şekilde ayıran ve maksimum sınırlı sınıflandırmayı yapabilen hiperdüzlemin bulunması amaçlanır. İki sınıflı uygulama olarak sunulan yöntemde bir sınıf  $\{1\}$  diğer sınıf ise  $\{-1\}$  etiketi ile temsil edilir. Bu etiketlere karşılık gelen veriler  $x^-$  ve  $x^+$  olarak isimlendirilecek olursa;

$$\langle wx^+ \rangle + b = +1 \quad (2.30a)$$

$$\langle wx^- \rangle + b = -1 \quad (2.30b)$$

olacaktır. Sınıflandırıcının sınırı  $\gamma$  olmak üzere;

$$\gamma = \frac{1}{2} \left( \left\langle \frac{w}{\|w\|_2} x^+ \right\rangle - \left\langle \frac{w}{\|w\|_2} x^- \right\rangle \right) \quad (2.31)$$

$$= \frac{1}{2\|w\|_2} (\langle wx^+ \rangle - \langle wx^- \rangle) \quad (2.32)$$

$$= \frac{1}{\|w\|_2} \quad (2.33)$$

ifadesi olarak hesaplanır. Bu denklemlerle verilen ifadelerin çözümü zor olacağından bu problemi uygun bir ikili forma çevirmek iyi olacaktır.

$$L(w, b, a) = \frac{1}{2} \langle w \cdot w \rangle - \sum_{i=1}^n a_i [y_i (\langle wx_i \rangle + b) - 1] \quad (2.34)$$

Burada  $a_i \geq 0$  değerli Langrange çarpanlarıdır. Yukarıdaki eşitlikte  $w$  ve  $b$  'ye göre türev alınırsa,

$$\frac{\partial L(w, b, a)}{\partial w} = w - \sum_{i=1}^n y_i a_i x_i = 0 \quad (2.35)$$

$$\frac{\partial L(w, b, a)}{\partial b} = \sum_{i=1}^n y_i a_i = 0 \quad (2.36)$$

Bu iki ifade birleştirilirse;

$$w = \sum_{i=1}^n y_i a_i x_i \quad (2.37)$$

$$\sum_{i=1}^n y_i a_i = 0 \quad (2.38)$$

ifadeleri elde edilir.

Hesaplanan tüm değerler birleştirilirse;

$$L(w, b, a) = \frac{1}{2} \langle w \cdot w \rangle - \sum_{i=1}^n a_i [y_i (\langle wx_i \rangle + b) - 1] \quad (2.39)$$

$$= \frac{1}{2} \sum_{i,j=1}^n y_i y_j a_i a_j \langle x_i x_j \rangle - \sum_{i,j=1}^n y_i y_j a_i a_j \langle x_i x_j \rangle + \sum_{i=1}^n a_i \quad (2.40)$$

$$= \sum_{i=1}^n a_i - \frac{1}{2} \sum_{i,j=1}^n y_i y_j a_i a_j \langle x_i x_j \rangle \quad (2.41)$$

ifadesi elde edilir.

$S = ((x_1 y_1), \dots, (x_n y_n))$  şeklinde verilen doğrusal olarak ayrılabilen bir eğitim kümesinde  $a^+$  parametrelerinin ikinci dereceden bir optimizasyon sorununu giderdiğini düşünürsek,

$$W(a) = \sum_{i=1}^n a_i - \frac{1}{2} \sum_{i,j=1}^n y_i y_j a_i a_j \langle x_i x_j \rangle \quad (2.42)$$

$$= \sum_{i=1}^n y_i a_i = 0, \quad a_i \geq 0, \quad i = 1, 2, 3, \dots, n \quad (2.43)$$

Ağırlık vektörü  $w^* = \sum_{i=1}^n y_i a_i^* x_i$  en büyük sınırlı hiperdüzlemin geometrik sınırı  $\gamma = 1/\|w^*\|_2$  şeklinde hesaplanır.

$b$  değeri ikili problemde bulunmaktadır. Bu sebeple birincil sabitler kullanılarak  $b^*$  değeri bulunmalıdır.

$$b^* = - \frac{\text{Maksimum}_{y_i=-1}(\langle w^* x_i \rangle) + \text{Minimum}_{y_i=1}(\langle w^* x_i \rangle)}{2} \quad (2.44)$$

Karush-Kuhn-Tucker şartları kullanılarak,

$$a_i^* [y_i (\langle w^* x_i \rangle + b^*) - 1] = 0, \quad i = 1, 2, \dots, n \quad (2.45)$$

ifadesi bulunur.

Hiperdüzleme en yakın noktalar, yani Destek Vektörleri için  $a_i^*$  değerleri 0 değildir. Diğer noktalar için tüm  $a_i^*$  değerleri 0'dır. Bu sebeple ağırlık vektörü  $w$  için yalnızca Destek Vektörlerinin  $a_i^*$  değerleri geçerlidir. Bu açıklamalar sonunda optimal hiperdüzlem şöyle verilir:

$$f(x, a, b^*) = \sum_{i=1}^n y_i a_i^* \langle x_i x \rangle + b^* \quad (2.46)$$

$$= \sum_{i \in DV} y_i a_i^* \langle x_i x \rangle + b^* \quad (2.47)$$

$j \in DV$  için,

$$y_j f(x_j, a^*, b^*) = y_j \left( \sum_{i \in DV} y_i a_i^* \langle x_i x \rangle + b^* \right) = 1 \quad (2.48)$$

$$\langle w^*, w^* \rangle = \sum_{i,j=1}^n y_i y_j a_i^* a_j^* \langle x_i x_j \rangle \quad (2.49)$$

$$= \sum_{j \in DV} a_j^* y_j \sum_{i \in DV} y_i a_i^* \langle x_i x_j \rangle \quad (2.50)$$

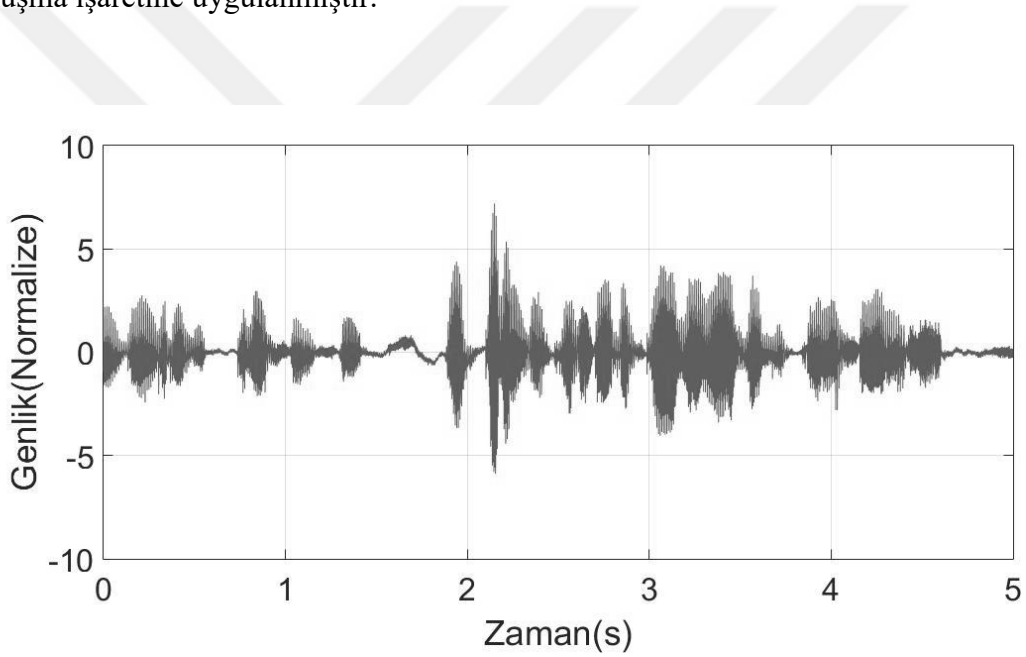
$$= \sum_{j \in DV} a_j^* (1 - y_j b^*) \quad (2.51)$$

$$= \sum_{i \in DV} a_i^* \quad (2.52)$$

### 2.1.3. Öznitelik Çıkarma ve Sınıflandırma

#### 2.1.3.1. Uygulanan Adımlar

Bu çalışmada 2.1.2. Yöntem kısmında anlatılan adımlar Şekil 2.4'te gösterilen ham konuşma işaretine uygulanmıştır.

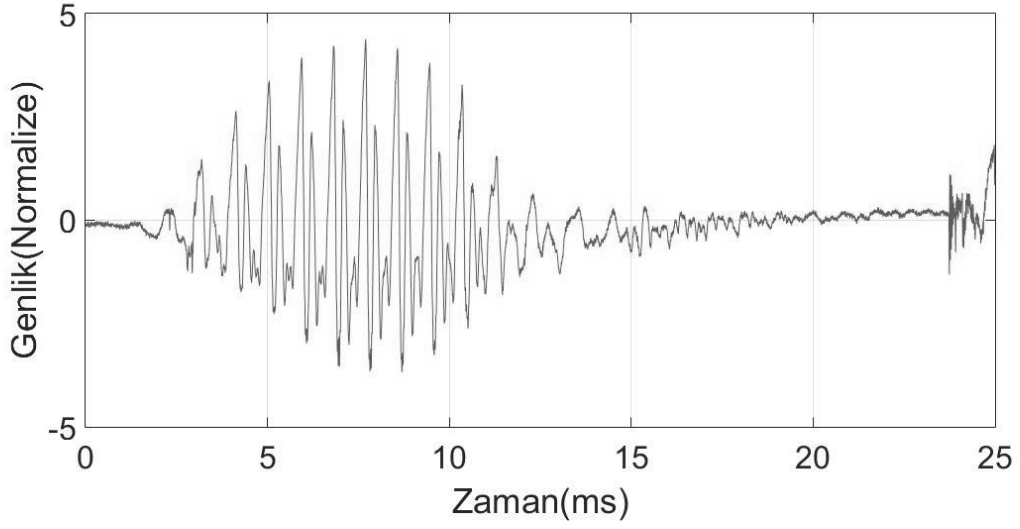


Şekil 2.4. Ham konuşma işareti

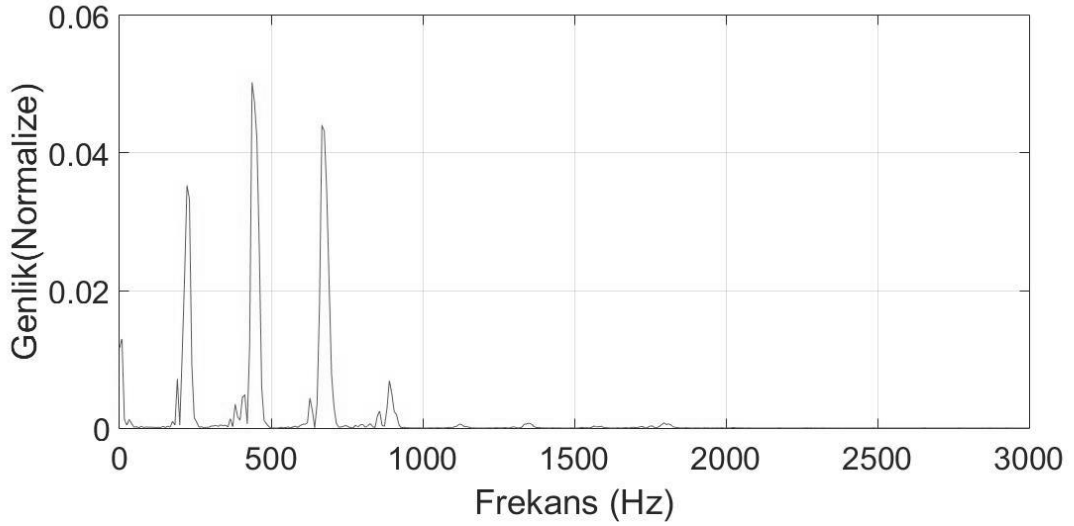
Beş saniye uzunluğunda olan ham veri 20 milisaniyelik parçalara ayrıldığında parçalardan biri Şekil 2.5'te gösterilmiştir.

Daha sonra bu parça Ayırık Fourier Dönüşümüne tabii tutulmuş, parçaya ait Ayırık Fourier Dönüşümü Şekil 2.6'da verilmiştir.

Mel Filtre Bankalarına ayrılan Ayırık Fourier Dönüşümünün karelerinde, her bir filtre bankası için hesaplanan enerjiler Şekil 2.7'de gösterilmiştir.



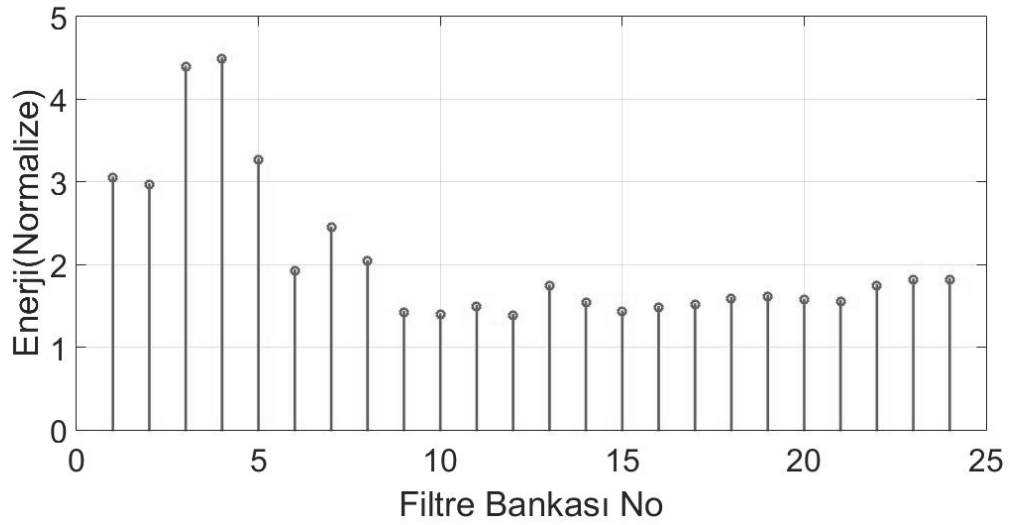
Şekil 2.5. Ham konuşma işaretinden alınmış işaret parçası



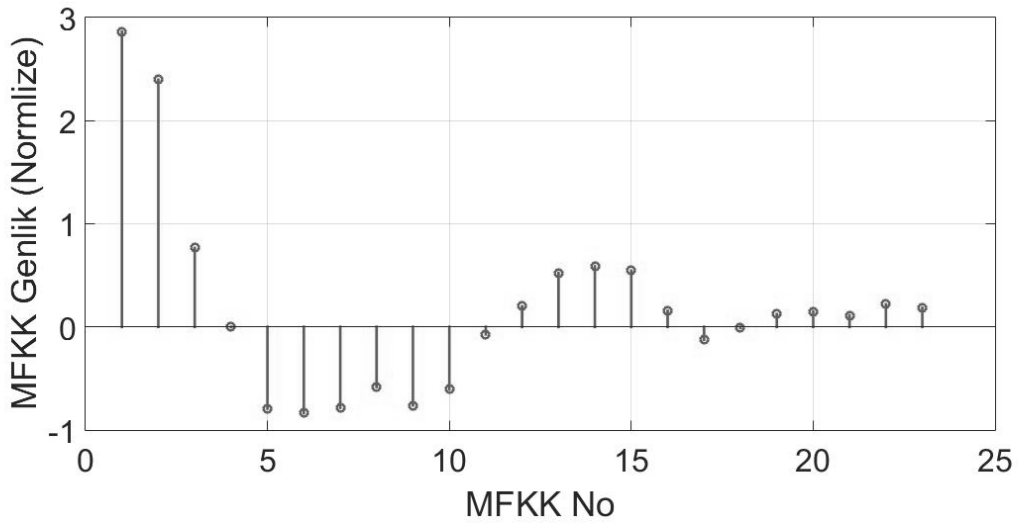
Şekil 2.6. İşaret parçasının Ayrık Fourier Dönüşümü karesi

Bu enerji ifadelerine Ayrık Kosinüs Dönüşümü uygulandığında ise hesaplanan Mel Frekanslı Kepstrum Katsayıları Şekil 2.9’de gösterilmiştir. Son olarak elde edilen bu MFKK zaman ekseninde ortalamaları alınarak öznelik vektörlerine ulaşılmıştır.

Sınıflandırma konuşmacı tanımada doğru modelin ortaya çıkarılması ve yüksek başarı ile kişilerin konuşmalarının birbirinde ayrılmasında en önemli etkenlerden biridir [91]. Oluşturulan model test işaretlerinin doğru sınıflandırılmasında doğrudan etkilidir.



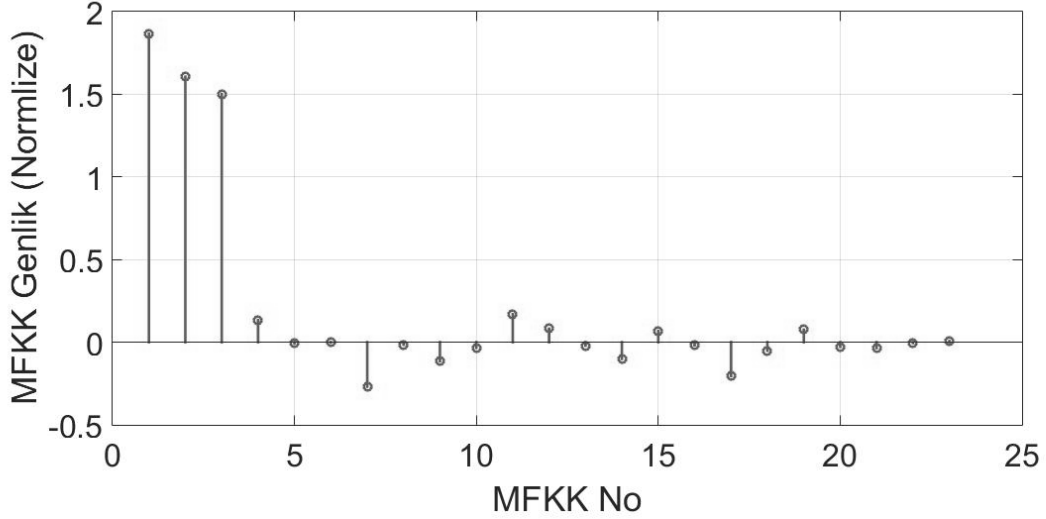
Şekil 2.7. Melodi filtre bankalarının enerjileri



Şekil 2.8. Bir parça işarete ait MFKK

Bu çalışmada, hesaplanan öznelik vektörleri her kişiye ait deneme sayısı eşit olacak şekilde eğitim ve test öznelik vektörleri olarak ikiye ayrılmıştır. Eğitim öznelik vektörleri sınıflandırıcılar üzerinden çapraz doğrulama ile eğitilerek başarının en iyi olduğu sınıflandırıcı değişkenleri tespit edilmiştir. Yüksek eğitim başarısı gösteren sınıflandırıcı değişkenleri kullanılarak sınıflandırma modelleri elde edilmiştir. Bu modeller test öznelik vektörlerinin sınıflandırılmasında kullanılmış ve modellerin belirlediği kimlikler tespit

edilmiştir. Modelin belirlediği kimlikler ile test öznelik vektörlerinin asıl kimlikleri ile karşılaştırılarak “Sınıflandırma Doğruluğu” hesaplanmıştır.



Şekil 2.9. Bir denemeye ait MFKK

Bu çalışmada sınıflandırıcı olarak, “Destek Vektör Makineleri (DVM)” ve “Doğrusal Ayırıcı Analizi (DAA)” kullanılmıştır. Ayrıca sınıflandırıcılara ait doğru değişkenlerin hesaplanması amacıyla eğitim aşamasında çapraz doğrulamaya başvurulmuştur. Eğitim aşamasında çapraz doğrulama sonucu en yüksek doğruluk gösteren sınıflandırma değişkenleri model oluşturulurken seçilmiştir. Yapılan hesaplamalar MATLAB programı ile gerçekleştirilmiştir [92].

Tablo 2.1. Otomatik konuşmacı tanıma eğitim ve test sınıflandırma sonuçları

| MFKK MİKTARI |                      | 5     | 10    | 15    | 20    |
|--------------|----------------------|-------|-------|-------|-------|
| <b>DAA</b>   | Eğitim Doğruluğu (%) | 55,84 | 93,78 | 99,28 | 99,88 |
|              | Test Doğruluğu (%)   | 51,53 | 88,13 | 95,91 | 98,06 |
| <b>DVM</b>   | Eğitim Doğruluğu (%) | 57,13 | 90,97 | 96,47 | 97,53 |
|              | Test Doğruluğu (%)   | 49,09 | 81,72 | 88,56 | 90,97 |



Bu çalışmada kullanılan 80 farklı kişiye ait eğitim ve test kümesi için 40'ar adet olmak üzere 5'er saniyelik kişi başı toplam 80, veri kümesi için toplam 3200 adet test ve 3200 adet eğitim kümesi olmak üzere 6400 farklı işaret ile çalışılmıştır. DAA ve DVM olmak üzere 2 farklı sınıflandırıcılara ait sonuçlar bu başlık altında sunulmaktadır. Öznitelik olarak MFKK kullanılmıştır. Adım adım takip edilen sonuçlarda MFKK sayısı beşinci katsayıdan başlamak üzere beşer artırılarak sonuçlar incelenmiştir. Katsayı miktarının n adet olduğu düşünülürse, n adet katsayı birinci katsayıdan n'inci katsayıya kadar olan bütün katsayıların öznitelik olarak kullanıldığı anlamındadır. Eğitim ve test aşamalarına ait "Sınıflandırma Doğruluğu" değerlerine Tablo 2.1'den ulaşılabilir.

### 2.1.3.2. Hata Analizi

Bu bölümde bu ön çalışmada yapılan sınıflandırmada hatalı sınıflandırılan verilerin analizi yapılmıştır. Murry ve Singh [7] davranışsal çalışmalarında konuşmacı eşleştirme aşamasında dinleyicilerin yaptıkları tercihlerin konuşmacıların cinsiyetlerine bağlı olduğunu ortaya koymuştu. Bu çalışmada da yapılan hatalarda denemelerin asıl cinsiyetlerinin ve yapılan hata sonucu tahmin edilen etiketin cinsiyeti hakkında irdeleme yapılmıştır.

Tablo 2.2. Yapılan hataların cinsiyete bağlı yüzdeler dağılımı

|            |  | <b>MFKK MİKTARI</b> |           |           |           |
|------------|--|---------------------|-----------|-----------|-----------|
|            |  | <b>5</b>            | <b>10</b> | <b>15</b> | <b>20</b> |
| <b>DAA</b> | Bayan Etiketinin Bayana Verildiği Hata | 35%                 | 36%       | 31%       | 31%       |
|            | Erkek Etiketinin Erkeğe Verildiği Hata | 31%                 | 30%       | 37%       | 48%       |
|            | Erkek Etiketinin Bayana Verildiği Hata | 19%                 | 19%       | 20%       | 13%       |
|            | Bayan Etiketinin Erkeğe Verildiği Hata | 15%                 | 15%       | 12%       | 8%        |
| <b>DVM</b> | Bayan Etiketinin Bayana Verildiği Hata | 34%                 | 35%       | 34%       | 36%       |
|            | Erkek Etiketinin Erkeğe Verildiği Hata | 32%                 | 33%       | 36%       | 40%       |
|            | Erkek Etiketinin Bayana Verildiği Hata | 18%                 | 16%       | 16%       | 12%       |
|            | Bayan Etiketinin Erkeğe Verildiği Hata | 16%                 | 15%       | 14%       | 12%       |

Asıl cinsiyeti bayan olan konuşmacılar için yapılan hatalı eşleşmenin yine bir bayana yapıldığı durum DAA için ortalama %33,25, DVM için ortalama %34,75 olarak hesaplanmıştır. Yine asıl cinsiyeti erkek olan konuşmacılar için yapılan hatalı eşleşmenin yine bir erkeğe yapıldığı durum DAA için ortalama %36,50, DVM için ortalama %35,25 olarak hesaplanmıştır.

Asıl cinsiyeti bayan olan konuşmacılar için yapılan hatalı eşleşmenin bir erkeğe yapıldığı durum DAA için ortalama %12,5, DVM için ortalama %14,25 olarak hesaplanmıştır. Benzer olarak asıl cinsiyeti erkek olan konuşmacılar için yapılan hatalı eşleşmenin bir bayana yapıldığı durum DAA için ortalama %17,75, DVM için ortalama %15,50 olarak hesaplanmıştır.

Tüm bu sonuçlara bir bütün olarak bakıldığında yapılan bir hatanın bir cinsiyetteki konuşmacıyı yine aynı cinsiyetteki konuşmacı olarak seçilmesi durumu DAA için ortalama %69,75, DVM için ortalama %70 olarak hesaplanmıştır. Yapılan bir hatanın asıl cinsiyetten karşı cinsiyete olduğu durum ise DAA için ortalama %30,25, DVM için ise ortalama %30 olduğu görülmektedir.

Bu sonuçlar açıkça göstermektedir ki yapılan hatalar yine aynı cinsiyet içinde seçimler yapılmasına daha çok olanak sağlamaktadır. Yani bir sistem sistemin yaptığı hatada konuşmacının asıl cinsiyeti ile tahmin edilen hatalı konuşmacıların cinsiyetleri birbiri ile aynı olma eğilimindedirler. Bu sonuçlar Murry ve Singh'in [7] davranışsal çalışmaları ile paralellik göstermektedir.

## **2.2. Ana Çalışma: Konuşmacı Aşinalığının EEG Tabanlı Belirlenmesi**

### **2.2.1. Materyal**

Bu çalışmada her biri sağ elini kullanan 22 ila 27 yaşları arasında toplam altı kişiden EEG kaydı yapılmıştır. Kişiler rahat olabilecekleri sessiz bir deney ortamında deney senaryosunu gerçekleştirmişlerdir. Kişilere rahatsız etmeyecek ses seviyesinde, ses seviyeleri her bir denek için eşit olacak şekilde hem erkek hem bayan sesleri dinletilmiştir.

### 2.2.1.1. Uyarım

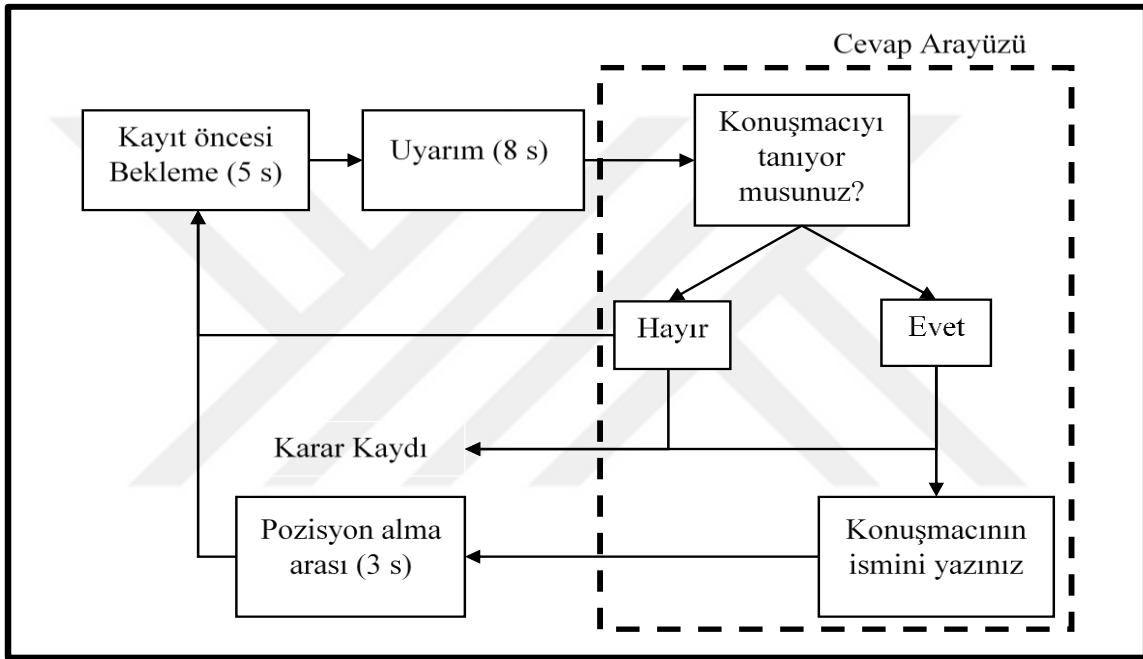
Bu çalışmada 16 erkek, 16 bayan olmak üzere, toplam 32 farklı yetişkin konuşmacıya ait konuşma kayıtları dinletilmiştir. Hem erkek hem bayan konuşmacılar kendi içlerinde yarısı dinleyicilerin kolayca tanıyabilecekleri politika, sinema, tiyatro ve spor dünyasından tanınmış ünlü kişilerden seçilmiş, diğer yarısı ise dinleyicilerin tanımadığı rastgele kişilerden seçilmiştir. Her bir uyarım sekiz saniye uzunluğunda olup 44100 Hz örnekleme frekansında saklı tutulmuştur. Dinletilen kayıtlar her bir kişi için aynı ses seviyesinde olmak kaydıyla maksimum 95 dB ses seviyesinde çıkış verebilen kulaklıktan en yüksek ses seviyesinde dinletilmiştir. Seçilen konuşma kayıtlarının gürültüden uzak, konuşmacının sesinden farklı müzik veya dış seslerden arınmış kayıtlar olmasına özellikle dikkat edilmiştir. Kişilere konuşmacıların kimlikleri hakkında herhangi bir bilgi kesinlikle verilmemiştir. Kayıtlarda konuşmacıların kimliğini açığa vuran sözlerin bulunmaması veya kişinin kimliğini ön plana çıkaran söz gruplarının geçmemesine özen gösterilmiştir. Deneyde konuşmacıların tahminde bulunmasını zorlaştırmak amacı ile konuşmacı ses renkleri birbirine benzer seçilmiştir. Bayan sesleri kendi arasında, erkek sesleri ise kendi aralarında yarı yarıya ünlü ve tanınmayan kişiler olacak şekilde ikişer grup olarak ayrılmıştır. Dinleyicilerin seçim yaparken bayan ve erkek seslerini kolayca ayırabilmeleri göz önünde bulundurularak deney senaryosunun zorluğunu artırmak amacıyla bayan ve erkek sesleri ayrı ayrı gruplar haline getirilmiştir. İki erkek seslerinden ikisi ise bayan seslerinden oluşmak üzere 8 konuşmacıya ait ses kaydı bulunan dört farklı veri seti elde edilmiştir. Konuşmacıların kayıtlarında bölgesel ağızdan uzak temiz ve akıcı Türkçe konuşulan kısımlar kullanılmıştır. Konuşmacıların hiçbiri kayıtlarında bir başka kişiyi taklit etmemektedir. Konuşmacıların anlık duygusal durumlarının etkili olmaması amacıyla kayıtlarda konuşmacıların aşırı duygusal içerikli kayıtlarından uzak durulmuştur. Her veri setinde 8 kişiye ait kayıtlardan kişi başı iki farklı konuşma kaydı bulunmaktadır. Her bir veri seti toplam 16 farklı konuşmacı kaydından meydana gelmektedir. Deney sonunda elde edilen verilerde kişi başı toplam 64 farklı EEG kaydı yapılmıştır.

### 2.2.1.2. Deney Senaryosu

Deney senaryosu dört farklı veri seti ve bir de deneğin dinletilen kişileri isimlerinden tanıyıp tanımadıklarını anlamak amacıyla oluşturulan bir deneyle birlikte 5 farklı deneyden

oluşmaktadır. Kullanılan konuşmacı veri seti itibariyle birbirinden farklı, deney senaryosu olarak birbirinin aynısı olan ilk dört deneye ait senaryo şu şekildedir:

Senaryo 5 saniyelik uyarım öncesi bekleme zamanı ile başlamaktadır. Bekleme zamanının bitmesi ile birlikte rastgele sıralanmış konuşmacı kayıtlarından biri kişiye dinletilmektedir. Uzunluğu 8 saniye olan kaydın kişi tarafından sonuna kadar dinlenilmesi istenmiştir. Kişi, konuşmacının kimliğini kayıt bitmeden önce belirlese bile konuşmanın sonuna kadar dinlemesi gerektiği bildirilmiştir.



Şekil 2.10. EEG kayıt aşaması deney senaryosu

Konuşmacı kaydının sona ermesi akabinde program otomatik olarak kişiye konuşmacıyı tanıyıp tanımadığını bildirmesini isteyen bir arayüz sunmuştur. Arayüz basit olarak kişinin konuşmacıyı tanıyıp tanımadığını yazılı olarak sormuştur ve cevap vermesi için fare üzerinden “Evet” veya “Hayır” seçeneklerini seçmesi istemiştir. Kişinin konuşmacıyı tanımaması ve cevap olarak “Hayır” seçeneğini seçmesi durumunda bir sonraki konuşmacıya geçmiştir. Her bir konuşmacı kaydı arasında 5 saniyelik bekleme aralığı bulunmaktadır. Kişinin konuşmacıyı tanıması ve cevap olarak “Evet” seçeneğini seçmesi durumunda ise yeni bir arayüz kişiye konuşmacının kimliğini yazması için gösterilmiştir. Kişi tanıdığı konuşmacının ismini bu arayüzde klavye üzerinde yazmıştır. Kişinin konuşmacının sesini tanımasına rağmen ismini hatırlamaması durumunda ise kişiler

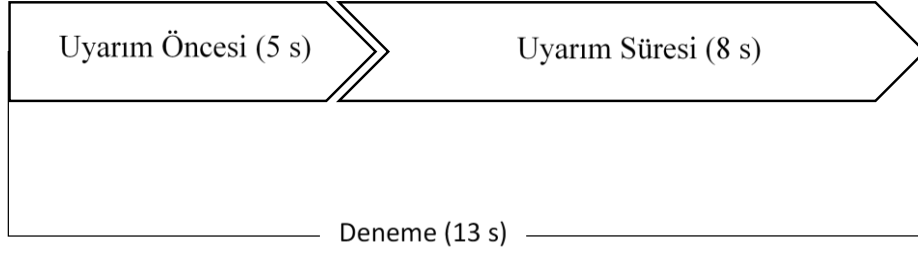
konuşmacının kimliğini belirten bir özelliğini yazmak durumundadırlar. Bu durum bilgilendirmeleri deney öncesi kişilere bildirilmiştir. Kişinin konuşmacının ismini klavye üzerinden yazmasının akabinde ise uygun pozisyona geçmesi için kişiye 3 saniyelik bir ara tanınmıştır. Bu aranın ardından bir sonraki konuşmacı kaydından önceki bekleme süresi başlatılmıştır.

Deney senaryosu tüm konuşma kayıtları bitene kadar bu döngüde devam etmiştir. Tüm bu kişi davranışları EEG kaydı altında takip edilmiştir. Bu senaryo her bir kişiye 4 seans olarak uygulanmıştır. Beşinci deneyde ise kişilerin konuşmacıları gerçekte tanıyıp tanımadıklarının anlaşılabilmesi için konuşmacıların isimleri söylenmiştir ve kişiler bu kişileri tanıyıp tanımadıklarını bildirmişlerdir. Her bir deney arasında kaydı alınan kişi daha fazla süre talep etmediği sürece 2 ila 3 dakikalık ara verilir. Toplam 5 deney oluşan bir deney seansı 75 ila 90 dakika arası sürmektedir. Deneye ilişkin senaryo akışı Şekil 2.10'da ve elde edilen her bir denemenin sürelerine göre içeriği Şekil 2.11'de gösterilmiştir.

### 2.2.1.3. EEG Verisi

EEG verileri Micromed Brain amplifikatörü ve Micromed Brainspy yazılımı üzerinde toplam 21 elektrottan kaydedilmiştir. Uluslararası 10-20 sisteminde Fp1, Fpz, Fp2, F3, F7, Fz, F4, F8, C3, Cz, C4, T3, T4, P3, P7, Pz, P4, P8, O1, Oz ve O2 elektrotlarından alınan EEG verileri 256 Hz örnekleme frekansında ve 16 bitlik kuantalama seviyelerinde sayısallaştırılmıştır. Kayıt esnasında referans olarak sağ mastoid noktası alınmış ve daha sonra sayısal ortamda ortalama referansa çevrilmiştir. Her bir elektrot için kanal empedans değeri 20 k $\Omega$  altında tutulmuştur. Şebekeye bağlı gürültülerden kaçınmak adına 50 Hz'te dar bantlı, band durduran Notch filtre kullanılmıştır. Kayıt aşamasında EEG verileri 0.008 Hz ve 100 Hz aralığında bant geçiren filtreye tabii tutulmuştur.

Deney esnasında rastgele sıra ile dinletilen konuşmacı kayıtlarının başlangıç ve bitiş anları MATLAB programı üzerinden kaydedilmiştir. Deney senaryosunun kişiye sunulduğu bilgisayar ve EEG kaydı için kullanılan bilgisayar olarak iki farklı bilgisayar kullanılmıştır. Bu bilgisayarlar arasındaki senkronizasyon, sol tuşları birleştirilerek senkron olarak başlatılabilen iki fareden sağlanmış. Birleştirilen iki farenin birinde yapılan işlem eş zamanlı olarak diğerinde de gerçekleşecek şekilde donanımsal birleşme sağlanmış ve iki bilgisayar da bu sayede aynı anda başlatılmış ve durdurulmuştur.



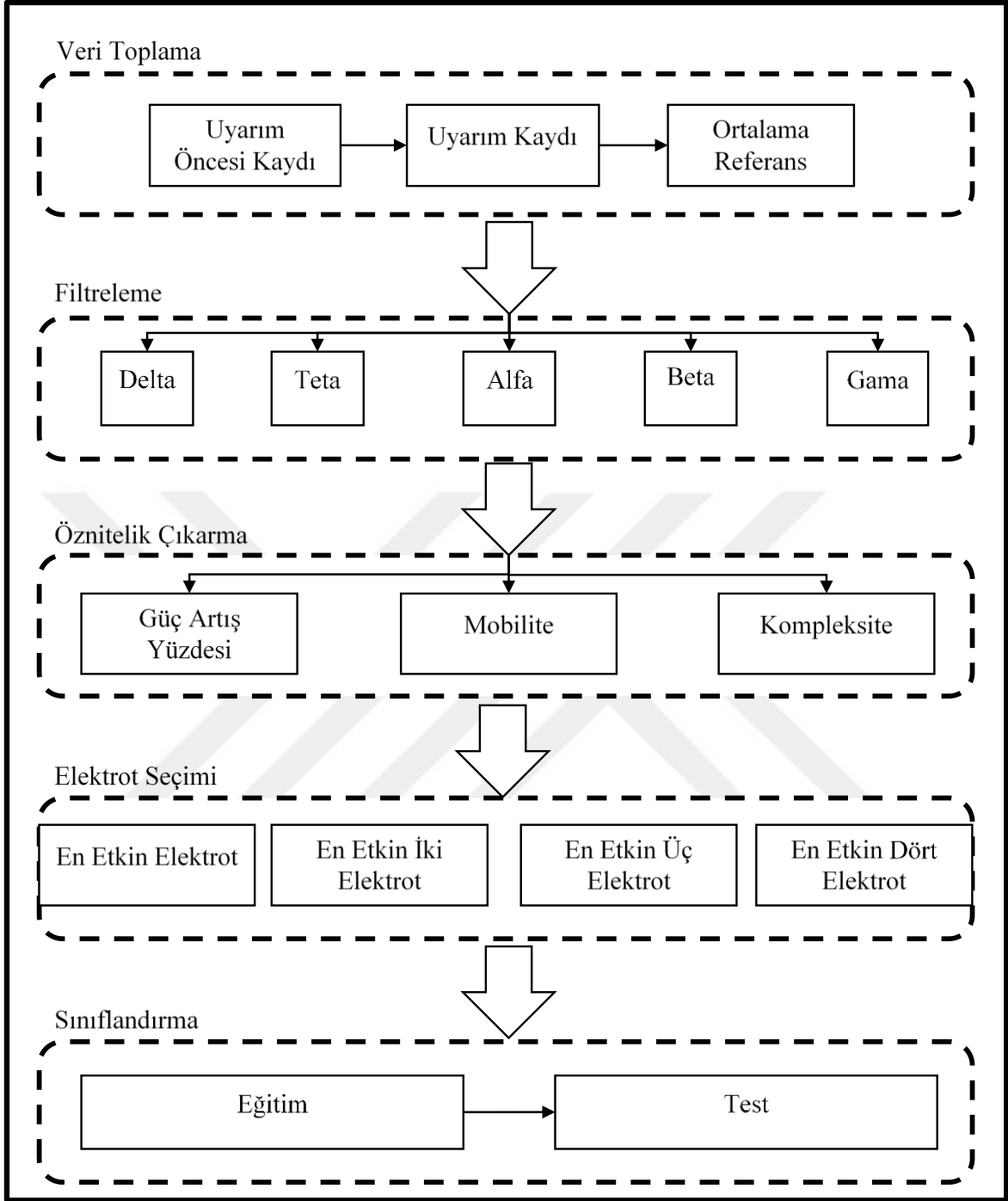
Şekil 2.11. Deneme verisi içeriği

Her bir deney için alınan kayıtlar ilgili programlarda kaydedilen başlangıç ve bitiş zamanlarına göre çevrimdışı olarak parçalara ayrılmış ve denemeler elde edilmiştir. Denemelere ait olan etiketler yine kişinin senaryo süresinde ilgili zamanda verdiği cevaplar yardımıyla birleştirilmiştir. Elde edilen denemeler 5 saniyelik uyarım öncesi ve 8 saniyelik uyarım süresi olarak 13 saniyelik parçalar ve bu parçalara ait etiketle olarak saklı tutulmuştur. Bir denemeye ait içeriğin bilgi ve sırası Şekil 2.11’de gösterilmiştir.

### 2.2.2. Yöntem

Bu çalışmada kullanılan EEG verilerinin kayıt aşaması ve sonrasında üzerinde gerçekleştirilen çalışmalar Şekil 2.12’de gösterilmiştir. Deney anında Uyarım kısmında anlatılan konuşmacı verileri, Deney Senaryosu kısmında anlatılan senaryo üzerinden kişilere sunulmuş ver EEG veri toplama kısmında detaylı anlatılan şekliyle önışlemlerden geçirilmiş, sayısallaştırılmış ve ortalama referansa çevrilmiştir.

Elde edilen denemeler her bir elektrot için filtrelerden geçirilmiştir. Filtre sonunda elde edilen 5 farklı bant içinde “Güç Artış Yüzdesi”, “Mobilite” ve “Kompleksite” olmak üzere 3’er öznitelik çıkarılmıştır. Bu öznitelikler üzerinden sınıflandırmanın en yüksek olduğu elektrotlar ve bantlar, kişiler için hesaplanmıştır. Kullanılan yöntemle ait akış Şekil 2.12’de gösterilmiştir.



Şekil 2.12. EEG verisi yöntem akışı

### 2.2.2.1. Filtreleme

Bu çalışmada ham EEG verisini ilgili beş alt ana banda ayırmak amacıyla filtreleme yapılmıştır. “Delta”, “Teta”, “Alfa”, “Beta” ve “Gama” alt bantlarının elde edilmesi için kullanılan filtrelerin alt ve üst kesim frekansları sırasıyla “0.1-4 Hz”, “4-8 Hz”, “8-12 Hz”, “12-30 Hz” ve “30-50 Hz” olarak seçilmiştir. Butterworth filtresi kullanılan çalışmada filtre derecesi 2 olarak seçilmiştir. Her elektrot noktası için ayrıca yapılan bu işlem sonucunda 21 elektrot için toplam 105 farklı alt bant elde edilmiştir. Bu bantlar her biri kendi içinde olmak üzere 5 farklı gruba ayrılmıştır. Öznitelik çıkarma ve sınıflandırma işlemleri her bant için ayrı ayrı yapılmıştır.

### 2.2.2.2. Öznitelik Çıkarma

Öznitelik çıkarma aşaması bir sınıflandırma uygulamasında doğruluğun, güvenilirliğin ve etkinliğin yüksek olması konusunda doğrudan etkilidir. Doğru özniteliklerin kullanılmış olması ve özniteliklerin işaretlerin genliğinden veya uzunluğundan etkilenmiyor olması etkin bir sınıflandırma açısından oldukça önemli bir değişkendir. Bu çalışmada da tüm bu özellikler dikkate alınarak normalleştirilmiş öznitelikler kullanılmış ve verilerin etiketlerden bağımsız olan değişkenliklerin ortadan kaldırılması amaçlanmıştır. Kullanılan öznitelikler “Güç Yüzde Artışı”, “Mobilite” ve “Kompleksite” olarak seçilmiştir. Bu öznitelikler bütün alt bantlarda her bir elektrot için aynı şekilde hesaplanmıştır.

#### 2.2.2.2.1. Hjorth Değişkenleri

Hjorth değişkenleri zaman-frekans öznitelikleri arasında sık kullanılan öznitelik yöntemlerindedir. Kullanılan çalışmalar açısından farklılık gösterse de yüksek başarı sağlaması ve normalleştirilmiş özniteliklerin arasında olması bu çalışmada seçilmelerinin en önemli nedenlerindedir. Hjorth değişkenleri “Aktivite”, “Mobilite” ve “Kompleksite” olmak üzere üç değişkenden oluşur.

Aktivite işaretlerin varyansını belirtir. Aktivite değişkeninin hesaplanmasında kullanılan matematiksel ifade;

$$Aktivite(x) = var(x) = \sigma^2 \quad (2.53a)$$



$$= \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2 \quad (2.53b)$$

olarak verilmiştir. Mobilite ise ortalama frekansı temsil eder ve bir işaretin türevinin varyansının işaretin kendi varyansına oranının karekökü olarak hesaplanır. Bu değişkeninin hesaplanmasında kullanılan matematiksel ifade;

$$Mobilite(x) = \frac{Aktivite\left(\frac{dx}{dt}\right)}{Aktivite(x)} \quad (2.54a)$$

$$= \sqrt{\frac{var\left(\frac{dx}{dt}\right)}{var(x)}} \quad (2.54b)$$

olarak verilmiştir. Kompleksite bir işaretin sinüzoidal bir işarete benzerliğinin ölçüsüdür ve bir işaretin türevinin mobilitesinin işaretin kendi mobilitesine oranı olarak hesaplanır. Bu değişkenin hesaplanmasında kullanılan matematiksel ifade;

$$Kompleksite(x) = \frac{Mobilite\left(\frac{dx}{dt}\right)}{Mobilite(x)} \quad (2.55a)$$

$$= \frac{\sqrt{\frac{var\left(\frac{d^2x}{dt^2}\right)}{var(x)}}}{\sqrt{\frac{var\left(\frac{dx}{dt}\right)}{var(x)}}} \quad (2.55b)$$

$$= \frac{\sqrt{var\left(\frac{d^2x}{dt^2}\right) \cdot var\left(\frac{dx}{dt}\right)}}{var\left(\frac{dx}{dt}\right)} \quad (2.55c)$$

olarak verilmiştir.

#### 2.2.2.2.2. Güç Yüzde Artışı

Güç yüzde artışı işaretlerin referans bir parçasının gücüne oranla etkin bölgenin gücündeki yüzdesel değişim ölçüsüdür. Öznitelik olarak kullanılması durumunda normalleştirilmiş olmamasından dolayı işaretin genlik ve uzunluğuna bağlı olan güç değeri, işaretin referans alınan bir parçası üzerinden yüzdesel olarak hesaplanması nedeniyle

normalleştirilmiş duruma getirilmiştir. Bu çalışmada kullanılan güç yüzde artışının hesaplandığı ifade;

$$P_{yüzde} = \frac{P_{uyarım\ süresi} - P_{uyarım\ öncesi}}{P_{deneme}} \times 100 \quad (2.56)$$

olarak verilmiştir.

### 2.2.2.3. Elektrot Seçimi

EEG verilerinin sınıflandırılmasında doğru elektrot bölgelerinin kullanılmış olması sınıflandırma başarı açısından doğrudan etkilidir. Kafatası üzerinde senaryo ile ilişkili bölgelerin sınıflararası farklılığın en yüksek olduğu yer olması beklenir. Farklı kişilerde farklı bölgeler bu bağlamda farklılık gösterebilir. Kişi beyinlerinin kendine has yapısı nedeniyle bir senaryoya ait sınıflararası farklılıkların en fazla olduğu bölgeler, nitelik ve nicelik açısından da farklılık gösterebilir. Bir kişinin sınıflandırma başarısının en yüksek olduğu elektrot yer ve sayısı bir başkası ile aynı olmayabilir.

Bunun yanında kişiler arası farklılıklardan biri de ilgili salınım türlerinin kişilerarası değişebileceğidir. Bir olaya ilişkin yapılan bir çalışmada bir kişide bir alt bant aktivasyonu gözlemlenirken bir başka kişide bu alt banttan farklı bir alt bant aktivasyonu gözlemlenebilir. Bu yüzden doğru alt bant, elektrot pozisyonları ve sayılarının belirlenmesi açısından bir çalışma yapılması gerekir.

Bu çalışmada da en iyi, en iyi iki elektrot, en iyi üç elektrot ve en iyi dört elektrot eşleşmeleri yapılmıştır. Her bir elektrot ilgili eşleşme aşamasında muhtemel tüm olasılıklarda birbirleriyle eşlenmiş ve eğitilmiştir. Eğitim sonucu elde edilen eğitim doğrulukları içinde en yüksek sırasıyla “20”, “50”, “100”, “100” eğitim sonucu test aşamasına sokulmuştur. Test sonuçları içinde en yüksek başarıyı gösteren elektrot veya elektrot grupları o kişi için ilgili elektrot sayısı ve alt bandında en yüksek sonuç olarak verilmiştir. Bu işleme her bir elektrot katılmış ve eşleşmelerde eşit şekilde ele alınmışlardır.

### 2.2.2.4. Sınıflandırma: Doğrusal Ayırıcı Analizi

Ayırıcı analizi R.A. Fisher tarafından, 1936 yılında, çok değişkenli bir sınıflandırma metodu olarak geliştirilmiştir [93]. Ayırıcı analiz, bağlı değişkenin sürekli değil, koşullu olması dışında regresyon analizi ile oldukça benzerdir. Ayırıcı analizde amaç; ayırık

gözlemlerin hangi sınıfa ait olduğunu, belirlenmiş kestirim değişkenleri ile saptayabilmektir. DAA genellikle, en iyi ayrık gözlem gruplarına ait kestirim değişkenlerinin doğrusal kombinasyonlarını bulmaya çalışır. İşte bu kombinasyonlara ayrık fonksiyonlar denir.

Her birinin ortalama vektörü  $\mu_k$  ( $k = 1, 2, \dots, K$ ) ve kovaryans matrisi  $\Sigma$  olan çok değişkenli ve normal dağılımlı  $K$  farklı grup olduğu varsayalım. Gerçek ortalama vektör ve kovaryans matrisi neredeyse her zaman denklemin bilinmeyenleridir, dolayısı ile bu parametreleri bulabilmek için en olası kestirimler kullanılır.

DAA'nın amacı  $x_i$  den  $k$ . gruba kadar tüm gözlemleri en az grup çeşitliliği ile sınıflandırabilmektir.

Örnek verecek olursak:

$$k = \operatorname{argmin}_k (x_i - \mu_k)^T \Sigma^{-1} (x_i - \mu_k) \quad (2.57)$$

Çok değişkenli normal varsayımlar göz önüne alındığında, bu denklem ile olabirliği en yüksek gözlem grubunu elde edilebilir. Genel anlamda, her gruptaki gözlem sayısı oranı kullanılarak önsel olasılık kestirilebilir. Örnekleme gerekirse;  $\pi_1 + \dots + \pi_k = 1$  için  $k$  grubunun orantısı  $\pi_k = \frac{n_k}{n}$  olsun. Bu durumda olabirliği maksimuma çıkarmak yerine sonsal olasılık maksimuma çıkartılırsa; belirli bir gruba ait gözlem:

$$k = \operatorname{argmax}_k \left[ -\frac{1}{2} (x_i - \mu_k)^T \Sigma^{-1} (x_i - \mu_k) + \log \pi_k \right] \quad (2.58)$$

İndirgenmiş haliyle  $k$  DAA fonksiyonu:

$$d_k(x) = x^T \Sigma^{-1} \mu_k - \frac{1}{2} \mu_k^T \Sigma^{-1} \mu_k + \log \pi_k \quad (2.59)$$

İkili durumlarda, ayırıcı fonksiyonlar aşağıdaki gibi kurulur:

$$d_1(x) = x^T \Sigma^{-1} \mu_1 - \frac{1}{2} \mu_1^T \Sigma^{-1} \mu_1 + \log \pi_1 \quad (2.60a)$$

$$d_2(x) = x^T \Sigma^{-1} \mu_2 - \frac{1}{2} \mu_2^T \Sigma^{-1} \mu_2 + \log \pi_2 \quad (2.60b)$$

Eğer  $d_1(x) > d_2(x)$  ise,  $x$  gözlemi grup bir, tersi durumda ise grup iki olarak alınır. Bu iki ayırıcı fonksiyon ayrıca aşağıdaki gibi de yazılabilir,

$$d(x) = d_1(x) - d_2(x) \quad (2.61a)$$

$$= x^T \Sigma^{-1} (\mu_1 - \mu_2) - \frac{1}{2} (\mu_1 + \mu_2)^T \Sigma^{-1} (\mu_1 - \mu_2) + \log \frac{\pi_1}{\pi_2} \quad (2.61b)$$

Eğer  $d(x) > 0$  ise,  $x$  gözlemi grup bir, tersi durumda ise grup iki olarak alınır. Denklemdaki (2.61b) son iki kısım, ayırıcı fonksiyon katsayıları  $D = \Sigma^{-1}(\mu_1 - \mu_2)$  ile belirtilmiş olan veri setinde verilmiş sabitlerdir. Katsayılar değişkenlerin fonksiyon üzerindeki ortak katkısını, dolayısı ile her bir katsayının diğerlerine kıyasla ne kadar etkili olduğunu gösterir. Normalleştirilmiş katsayılar  $D^* = \text{diag}(\Sigma)D$ , her bir katsayının ilişkili değişkenlerin standart sapmaları ile çarpılması ile bulunur. Normalleştirilmiş katsayı vektörü, ölçeklerin büyük oranda değişim gösterdiği durumlarda, ayırıcı fonksiyonun bağlı dağılımlı her bir değişkeni ile ilgili daha sağlam bilgiler verir.

Ayırıcı fonksiyonlar temel olarak iki varsayıma dayanır; her bir gruptaki çoklu-normalite ve gruplara arası kovaryansların düzgün dağılımı. Eğer çok fazla kategorik kestirici değişken varsa, bu iki varsayım genelde göz ardı edilir ve bu durum, model ve kestirim kalitesini kötü yönde etkileyebilir. Ayrık analizin bir diğer sınırlaması ise grupların ortalama vektörlerinin fark edilebilir ve her bir gruptaki gözlem sayısının değişken boyutundan daha büyük olma zorunluluğudur. Eğer ortalama vektörler yeterince farklı değilse, DAA ile yüksek oranda sınıflandırma yapabilmek mümkün olmayabilir. Eğer ki bazı gruplarda gözlemler sınırlı ise, DAA kullanılmadan önce aşamalı metod ile önemli değişkenlerin seçilmesi gerekebilir.

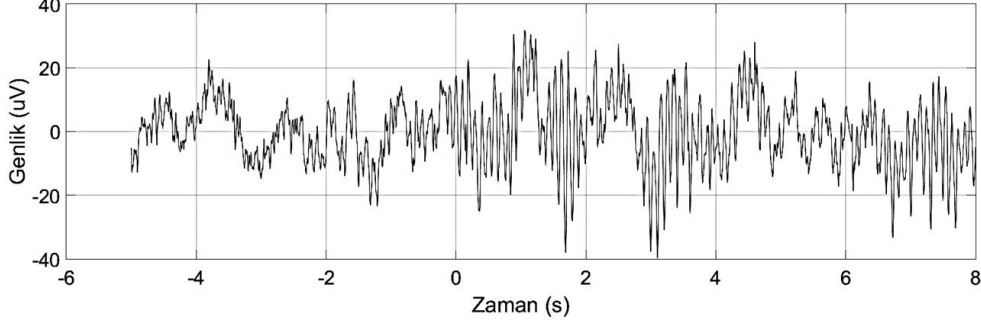
Eğer bir sınıflandırma yapılıyorsa, sınıflandırma oranı önceden hesaplanmış olmalıdır. Doğru sınıflandırma olasılığını ölçen metotlardan biriside  $k$ -katlama çapraz-doğrulama metodudur [94].  $K$ -katlama çapraz doğrulama metodunda orijinal örnek  $k$  tane alt örneğe ayrılır. Her defasında, seçilen bir alt örnek, diğer  $k-1$  alt örnek ile hazırlanan modelin doğrulaması için ayrı tutulur ve bu işlem  $k$  defa tekrarlanır. Böylece her bir alt örnek bir defa doğrulama için kullanılmış olur. Elde edilen sonuçlar tek bir sınıflandırma oranı elde edebilmek için kullanılır. Özelleştirilmiş bir  $k$ -katlama metodunun adı da  $k$ 'nın orijinal veri setindeki gözlem sayısına eşit olduğu "birini-dışarıda-bırak" çapraz doğrulama metodudur.

### 2.2.3. Öznitelik Çıkarma ve Sınıflandırma

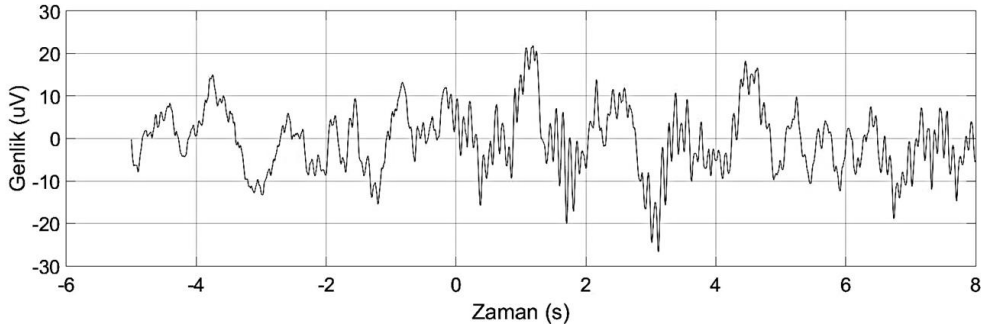
#### 2.2.3.1. Uygulanan Adımlar

Bu çalışmada ortak referansa çevrilmiş olan bütün elektrotlarda aynı işlem yapılmıştır. Elektrotlarda her işaret beş ana alt banda filtrelenmiştir. Bu filtreleme işleminin yapıldığı

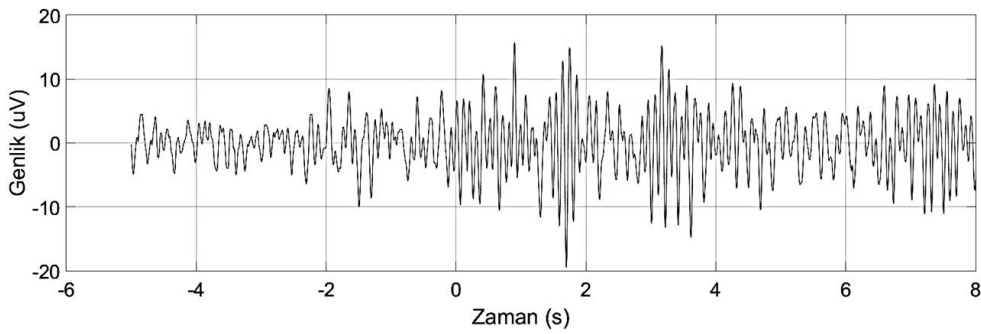
örnek ham EEG verisine Şekil 2.12'den filtreleme sonrasında elde edilen alt bantlara ait işaretlere ise Şekil 2.13, Şekil 2.14, Şekil 2.15, Şekil 2.16 ve Şekil 2.17'den ulaşılabilir.



Şekil 2.13. Ham EEG işareti

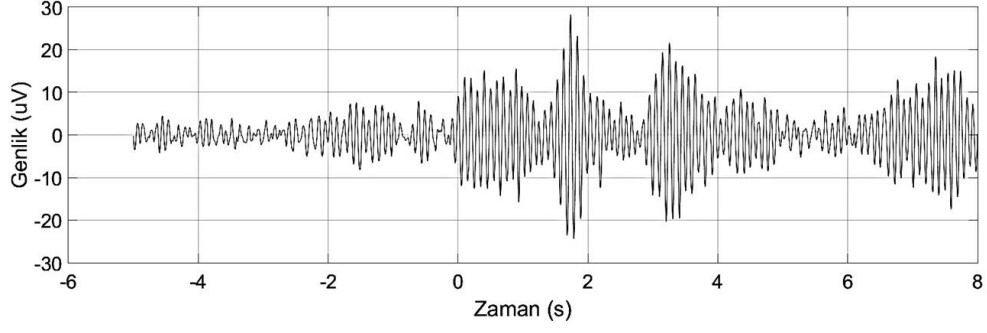


Şekil 2.14. Delta bandında elde edilen işaret

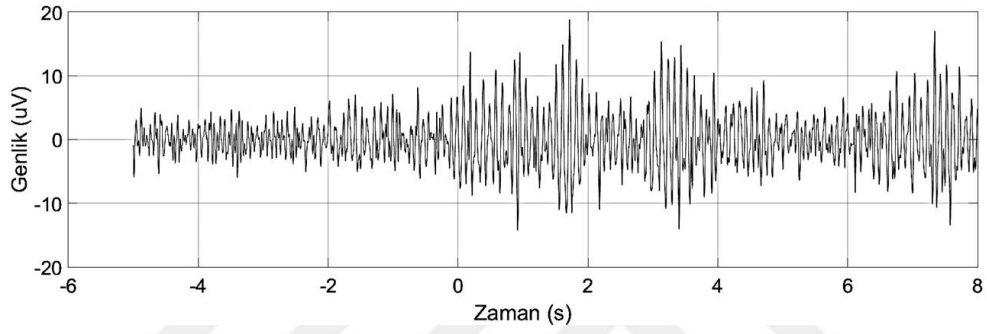


Şekil 2.15. Teta bandında elde edilen işaret

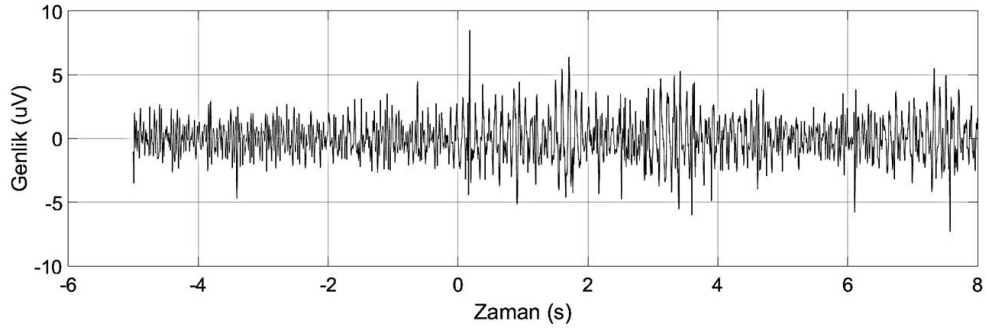
Filtrelenmiş olan işaretler alt bantlarda ayrı ayrı muhafaza edilmiş ve öznitelik çıkarılmıştır. Bütün elektrotlar için öznitelik olarak mobilite, kompleksite ve güç yüzde artışı kullanılmıştır. Bu çalışmada kullanılan elektrot isimleri bu kısımda itibaren üç farklı vektörden oluşan öznitelik vektörlerini temsil etmektedir.



Şekil 2.16. Alfa bandında elde edilen işaret



Şekil 2.17. Beta bandında elde edilen işaret



Şekil 2.18. Gama bandında elde edilen işaret

### 2.2.3.2. Davranışsal Bulgular

Uyarım kısmında belirtildiği üzere bu çalışmada, kişiler uyarımlarda kullanılan konuşmacıların kimliği hakkında hiçbir bilgiye sahip değillerdir. Başlangıçta konuşmacı kimliğinden yoksun olarak başlatılan senaryo sunucunda kişilerin verdiği cevaplar ele

alınmıştır. Erkek ve bayan grupları için davranışsal olarak verdikleri “Tanıma” ve “Tanıyamama” durumları için elde edilen verilere Tablo 2.3’ten ulaşılabilir.

Deney sonunda kişilere konuşmacıların kimlikleri sunulmuş ve gerçekte konuşmacıları tanıyıp tanımadıklarını belirtmeleri istenmiştir. Deney sonrasında yapılan bu işlem sonucunda kişiler gerçekte tanıdıkları kişileri seçmişlerdir. Bu sonuçlardan yararlanarak oluşturulan erkek ve bayan konuşmacıların gerçekte tanınıp tanımadığını gösteren bulgular Tablo 2.3.’te verilmiştir.

Tablo 2.3. Davranışsal bulgular ve sınıf dağılımları

|               | Deney Performansı |          |          |          | Deney Sonrası Bildirilen Cevaplar |           |          |           |
|---------------|-------------------|----------|----------|----------|-----------------------------------|-----------|----------|-----------|
|               | Bayanlar          |          | Erkekler |          | Bayanlar                          |           | Erkekler |           |
|               | Tanıdı            | Tanımadı | Tanıdı   | Tanımadı | Tanıyor                           | Tanımiyor | Tanıyor  | Tanımiyor |
| <b>Kişi 1</b> | 12                | 20       | 8        | 24       | 12                                | 20        | 16       | 16        |
| <b>Kişi 2</b> | 14                | 18       | 17       | 15       | 20                                | 12        | 18       | 14        |
| <b>Kişi 3</b> | 15                | 17       | 16       | 16       | 18                                | 14        | 18       | 14        |
| <b>Kişi 4</b> | 15                | 17       | 13       | 19       | 18                                | 14        | 18       | 14        |
| <b>Kişi 5</b> | 15                | 17       | 12       | 20       | 16                                | 16        | 16       | 16        |
| <b>Kişi 6</b> | 15                | 17       | 14       | 18       | 20                                | 12        | 16       | 16        |

Tablo 2.3’ün yorumlanmasında ortaya çıkan ilk sonuç kişiler aslında tanıdıklarını bildirdikleri bazı konuşmacıları kimlik bilgisinden yoksun olduklarında konuşmacı seslerinden tanıyamamıştır. Ayrıca bir konuşmacı için iki farklı konuşmada bir konuşmada kimlik bilgisini çıkarabilirken bir diğerinde konuşmacıyı tanıyamama durumu söz konusudur. Bunun yanında bazı kişiler de aslında tanımadıklarını bildirdikleri bazı konuşmacıları hatalı olarak tanıdıkları yönünde seçim yapmışlardır. Bu bağlamda deney performansı olarak kişilerin de kişisel performansının etkili olabileceği öngörülmüştür. Kişilerin konuşmacıları tanımadıkları halde bir başkasına benzeterek tanıdıklarını ifade etmeleri, tanıdıkları halde dinletilen konuşma içinde kimlik bilgisini doğru çıkaramama durumu ve kişilerin EEG kaydı süresi içinde değil de kayıt sonrası cevap verme sırasında tanıyıp tanınamasına karar vermesi durumları bu çalışmanın en büyük sorunlarından. Sınıflandırma sonuçlarının değerlendirilmesi bu gibi etkenlerin de göz önünde bulundurulması doğru bir analiz için hayati önem taşımaktadır.

### 2.2.3.3. Sınıflandırma ve Bulgular

Bu bölümde sınıflandırma sonucunda her bir alt bant için kişilere göre farklı elektrot sayılarında en iyi sonuçlar araştırılmıştır. Elektrot sayısı bir, iki, üç ve dört olması durumunda her bir kişi en iyi sınıflandırma başarısını farklı elektrot grupları ve farklı alt bantlarda göstermiştir. Yapılan sınıflandırma işleminin sonucunda kişilerin elektrot sayısı ve alt bant türleri için sınıflandırma başarısına Tablo 2.4 ve Tablo 2.7 arasında kalan tablolarda ulaşılabilir.

Tablo 2.4. En etkin elektrot için bant ve kişiye bağlı Eğitim ve test başarıları

| Alt Bant     | Sınıflandırma Aşaması | Kişi 1 | Kişi 2 | Kişi 3 | Kişi 4 | Kişi 5 | Kişi 6 |
|--------------|-----------------------|--------|--------|--------|--------|--------|--------|
| <b>Delta</b> | Eğitim (%)            | 76,56  | 65,63  | 65,63  | 73,44  | 67,19  | 68,75  |
|              | Test (%)              | 73,44  | 64,06  | 64,06  | 70,31  | 64,06  | 67,19  |
| <b>Teta</b>  | Eğitim (%)            | 76,56  | 70,31  | 65,63  | 67,19  | 67,19  | 65,63  |
|              | Test (%)              | 76,56  | 64,06  | 59,38  | 62,50  | 62,50  | 64,06  |
| <b>Alfa</b>  | Eğitim (%)            | 79,69  | 60,94  | 70,31  | 62,50  | 70,31  | 65,63  |
|              | Test (%)              | 78,13  | 57,81  | 70,31  | 59,38  | 65,63  | 65,63  |
| <b>Beta</b>  | Eğitim (%)            | 78,13  | 68,75  | 60,94  | 59,38  | 73,44  | 68,75  |
|              | Test (%)              | 76,56  | 64,06  | 60,94  | 57,81  | 70,31  | 67,19  |
| <b>Gama</b>  | Eğitim (%)            | 78,13  | 67,19  | 65,63  | 62,50  | 70,31  | 64,06  |
|              | Test (%)              | 76,56  | 62,50  | 60,94  | 60,94  | 67,19  | 64,06  |



Tablo 2.5. En etkin iki elektrot için bant ve kişiye bağı Eğitim ve test başarıları

| <b>Alt Bant</b> | <b>Sınıflandırma Aşaması</b> | <b>Kişi 1</b> | <b>Kişi 2</b> | <b>Kişi 3</b> | <b>Kişi 4</b> | <b>Kişi 5</b> | <b>Kişi 6</b> |
|-----------------|------------------------------|---------------|---------------|---------------|---------------|---------------|---------------|
| <b>Delta</b>    | Eğitim (%)                   | 84,38         | 73,44         | 75,00         | 76,56         | 71,88         | 73,44         |
|                 | Test (%)                     | 81,25         | 68,75         | 71,88         | 67,19         | 67,19         | 68,75         |
| <b>Teta</b>     | Eğitim (%)                   | 90,63         | 73,44         | 71,88         | 70,31         | 73,44         | 73,44         |
|                 | Test (%)                     | 82,81         | 65,63         | 70,31         | 68,75         | 70,31         | 68,75         |
| <b>Alfa</b>     | Eğitim (%)                   | 87,50         | 78,13         | 76,56         | 75,00         | 76,56         | 71,88         |
|                 | Test (%)                     | 85,94         | 71,88         | 73,44         | 70,31         | 71,88         | 68,75         |
| <b>Beta</b>     | Eğitim (%)                   | 81,25         | 73,44         | 79,69         | 71,88         | 76,56         | 70,31         |
|                 | Test (%)                     | 79,69         | 71,88         | 70,31         | 68,75         | 70,31         | 67,19         |
| <b>Gama</b>     | Eğitim (%)                   | 82,81         | 71,88         | 78,13         | 75,00         | 79,69         | 73,44         |
|                 | Test (%)                     | 76,56         | 65,63         | 70,31         | 65,63         | 73,44         | 67,19         |

Tablo 2.6. En etkin üç elektrot için bant ve kişiye bağlı eğitim ve test başarıları

| <b>Alt Bant</b> | <b>Sınıflandırma Aşaması</b> | <b>Kişi 1</b> | <b>Kişi 2</b> | <b>Kişi 3</b> | <b>Kişi 4</b> | <b>Kişi 5</b> | <b>Kişi 6</b> |
|-----------------|------------------------------|---------------|---------------|---------------|---------------|---------------|---------------|
| <b>Delta</b>    | Eğitim (%)                   | 84,38         | 76,56         | 79,69         | 81,25         | 76,56         | 78,13         |
|                 | Test (%)                     | 79,69         | 70,31         | 73,44         | 73,44         | 70,31         | 70,31         |
| <b>Teta</b>     | Eğitim (%)                   | 84,38         | 75,00         | 73,44         | 78,13         | 81,25         | 76,56         |
|                 | Test (%)                     | 82,81         | 70,31         | 70,31         | 73,44         | 76,56         | 70,31         |
| <b>Alfa</b>     | Eğitim (%)                   | 89,06         | 82,81         | 78,13         | 79,69         | 82,81         | 81,25         |
|                 | Test (%)                     | 82,81         | 73,44         | 73,44         | 71,88         | 76,56         | 73,44         |
| <b>Beta</b>     | Eğitim (%)                   | 84,38         | 73,44         | 81,25         | 79,69         | 81,25         | 79,69         |
|                 | Test (%)                     | 79,69         | 70,31         | 76,56         | 71,88         | 78,13         | 71,88         |
| <b>Gama</b>     | Eğitim (%)                   | 85,94         | 76,56         | 79,69         | 79,69         | 82,81         | 76,56         |
|                 | Test (%)                     | 79,69         | 65,63         | 75,00         | 71,88         | 78,13         | 65,63         |

Tablo 2.7. En etkin dört elektrot için bant ve kişiye bağlı Eğitim ve test başarıları

| <b>Alt Bant</b> | <b>Sınıflandırma Aşaması</b> | <b>Kişi 1</b> | <b>Kişi 2</b> | <b>Kişi 3</b> | <b>Kişi 4</b> | <b>Kişi 5</b> | <b>Kişi 6</b> |
|-----------------|------------------------------|---------------|---------------|---------------|---------------|---------------|---------------|
| <b>Delta</b>    | Eğitim (%)                   | 90,63         | 79,69         | 84,38         | 81,25         | 82,81         | 84,38         |
|                 | Test (%)                     | 84,38         | 71,88         | 78,13         | 76,56         | 71,88         | 70,31         |
| <b>Teta</b>     | Eğitim (%)                   | 90,63         | 81,25         | 79,69         | 81,25         | 87,50         | 79,69         |
|                 | Test (%)                     | 84,38         | 70,31         | 73,44         | 70,31         | 76,56         | 68,75         |
| <b>Alfa</b>     | Eğitim (%)                   | 89,06         | 82,81         | 85,94         | 85,94         | 84,38         | 84,38         |
|                 | Test (%)                     | 82,81         | 75,00         | 76,56         | 81,25         | 81,25         | 73,44         |
| <b>Beta</b>     | Eğitim (%)                   | 87,50         | 76,56         | 84,38         | 79,69         | 85,94         | 85,94         |
|                 | Test (%)                     | 81,25         | 68,75         | 79,69         | 73,44         | 81,25         | 76,56         |
| <b>Gama</b>     | Eğitim (%)                   | 87,50         | 78,13         | 85,94         | 85,94         | 84,38         | 79,69         |
|                 | Test (%)                     | 82,81         | 68,75         | 82,81         | 73,44         | 81,25         | 67,19         |

### 3. SONUÇLAR VE TARTIŞMA

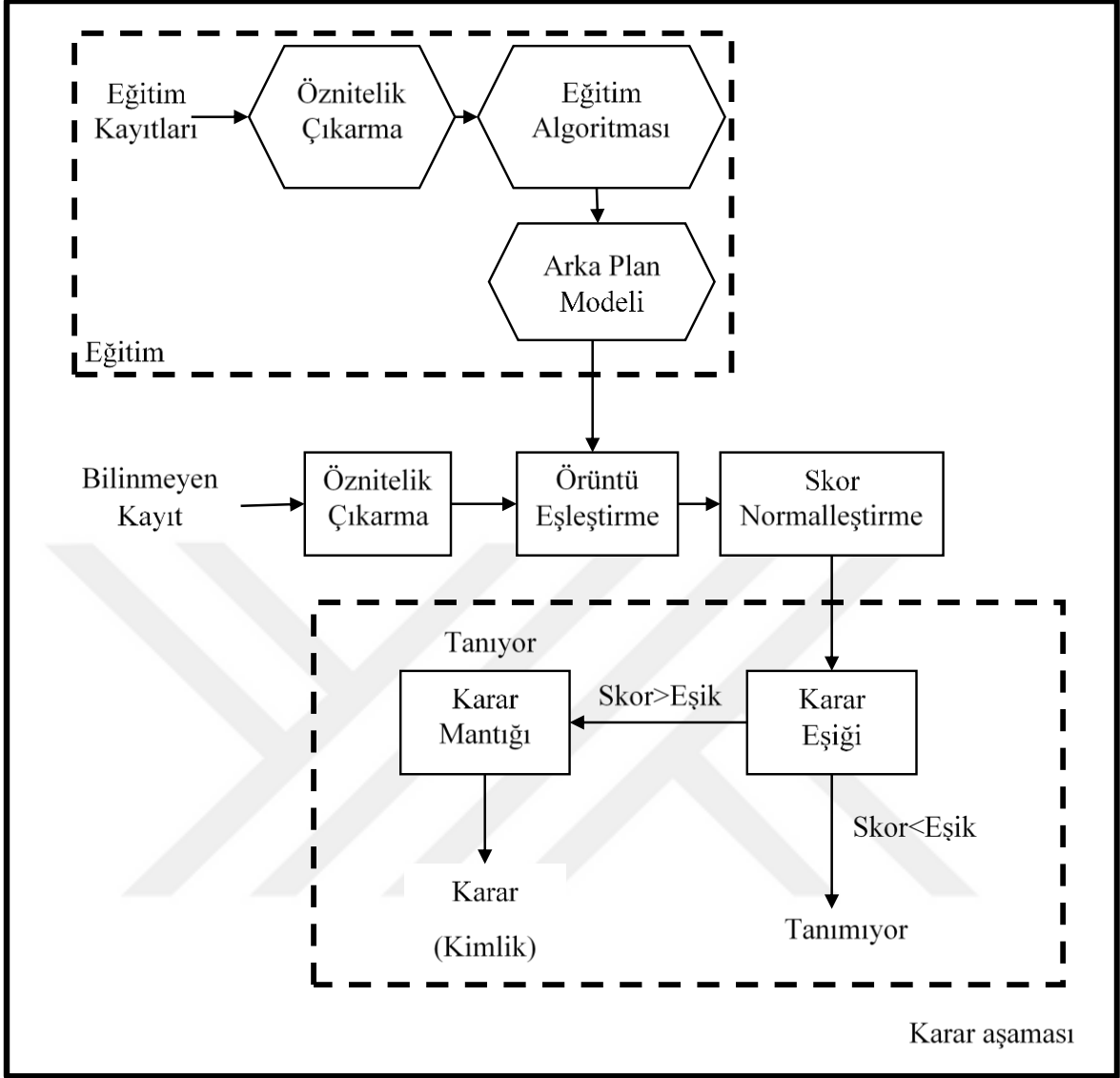
Yapılan çalışmalar ve bulgular kısmında sunulan veriler incelendiğinde sonuçlar iki farklı koldan yorumlanabilir. Bunlar davranışsal bulgular üzerinden beyin konuşmacı tanıma ile otomatik konuşmacı tanıma aşamalarının karşılaştırılması ve EEG verilerinin kişi, elektrot ve alt bantların başarımlar üzerinden karşılaştırılmasıdır.

#### 3.1. Beyin Konuşmacı Tanıma ve Otomatik Konuşmacı Tanıma

Bu çalışmada uygulanan deney senaryosu bir otomatik konuşmacı tanıma uygulamasına benzetilecek olsa kişinin metinden bağımsız, açık küme, konuşmacı teşhis etme uygulaması yaptığı görülebilir. Dinletilen konuşmaların içeriğine bakıldığında konuşmacıların arasında aksan farklılıklarının olmaması, konuşmacıların cinsiyetlerine göre ayrılmış ve farklı cinsiyetteki konuşmacıların farklı zamanlarda sunulmuş olması, konuşmacıların her birinin yetişkin kişiler olması ve konuşmacıların belirgin bir duygusal içeriği göstermiyor olması kişilerin yüksek seviyeli öznitelikler üzerinden belirleme yapabilmesini zorlaştırmaktadır.

Tablo 2.3'te verilen davranışsal bulgular incelendiğinde, kişilerin başlangıçta konuşmacı kimliklerini bilmediği durumda yapılan deneyler ile deney sonrası bildirdiği cevaplar arasında farklılıklar görülmektedir. Buna göre kişiler daha sonra tanıdıklarını bildirdikleri kişileri başlangıçta tanıyamamışlardır. Hatta konuşmacı başına iki kayıt olmasına rağmen bazı kişiler bu konuşmacıları bir kaydında tanıyabilmişken bir diğer kaydında konuşmacıyı tanıyamamışlardır. Bazı kişiler ise deney sonrası tanımadığını bildirdikleri konuşmacıları deney senaryosunda tanıdıklarını söylemiş ve farklı bir kimlik atamışlardır. Yani konuşmacıyı tanıdığı başka birine benzetmişlerdir. Bu da beyin konuşmacı tanıma sürecinde hatanın varlığını ortaya koymuştur.

Deney süresince bayan ve erkek seslerinin ayrı ayrı dinletilmiş olması yapılan hataların cinsiyete bağlı olup olmadığının tespitini olanaksız kılmış olsa da Murry ve Singh'in çalışmasında ortaya koyduğu yapılan hatalarda kimsenin farklı cinsiyetten birini hatalı olarak seçmemesi bu açığı kapatmıştır.



Şekil 3.1. Beyin konuşmacı tanıma akışı

Beyin konuşmacı tanıma süreci bir akış halinde Şekil 1.1 üzerinden yeniden uyarlanacak olursa Şekil 3.1'deki akış ortaya koyulabilir. Sunulan bu akış ile beyin işitsel hafızası, dahilinde konuşmacıların modellerini bulundurduğu beyin arka plan modellerini saklar. Bir uyarım geldiğinde ve konuşma işaretinin içinden konuşmacı kimlik bilgisinin çıkarılması gerektiği durumlarda beyin bu arka plan modeline danışır ve teşhis yapar. Kişilerin özel olarak tecrübe ve hafızasında saklı tuttuğu beyin arka plan modeli alınan konuşmanın içindeki kimliği içeriğindeki ile karşılaştırır. Her bir arka plan model elemanı ile yapılan karşılaştırma sonucunda bilinmeyen konuşmacının her bir model ile benzerliği bir skor olarak bulunur. Bu skorlar arasında en yüksek olan konuşmacı ile bilinmeyen

konuşmacı eşleştirilir ve teşhis etme sonuçlanır. En yüksek skor kişiye bağlı bir eşiğin altında ise bu bilinmeyen konuşmacının model içinde biri olmadığı daha önce karşılaşılmayan bir kişi olduğu belirlenir.

Davranışsal bulgulara göz atıldığında kişilerin gerçekte tanıdıklarını bildirdikleri kişileri deney süresinde tanıyamamış olması beyin arka plan modelinin otomatik konuşmacı tanıma sistemlerindeki gibi statik değil dinamik yapıya sahip olduğu savunulabilir. Hafızanın zamanla zayıflaması veya çok kullanılmayan verilerin beyin arkaplan modelinde öncelik sırasında geriye düşmesi aslında tanınan kişilerin tanınmamasına neden olmuştur. Buna göre beyin en yakın zamanda duymuş olduğu veya uzun süre birlikteliklerle modellerini oluşturduğu konuşmacıları bir teşhis etme sürecinde karşılaştırılacak ve skor tahmini yapılacaklar arasında en üst sıralara koyarken, kullanımı azalan konuşmacı modellerini ise alt sıralara koyar. Skor hesaplanması yapılırken ise bu sıra üzerinden devam eder ve tüm hafızayı taramadan sadece en muhtemel olduğunu düşündüğü üst sıraları kullanır.

### **3.2. Konuşmacı Aşinalığı**

Ana çalışma olarak sunulan konuşmacı aşinalığının EEG tabanlı belirlenmesinde Yöntem incelendiğinden olabilecek tüm eşleşmeler denenmiş ve alt bantlara sadık kalınarak sınıflandırma yapılmıştır. Kişiler arasında elektrot bölgeleri ve alt bant bakımından en yüksek başarıların elde edildiği değişkenler farklılık göstermiştir. Kişiler için sınıflandırma test başarısının en yüksek olduğu alt bant ve elektrot bölgeleri Tablo 3.12’de verilmiştir. Buna göre 6 kişilik denek grubunda 4 kişi alfa bandında yüksek başarı göstermiştir. Diğer deneklerin de alfa bandı başarımları incelendiğinde aslında kabul edilebilir sonuçlar elde edilmiştir. Kişi 3’e ait dört elektrotlu alfa bandı başarımları %76,56 ve kişi 6’ya ait dört elektrotlu alfa bandı başarımları %73,44 olarak göze çarpmaktadır. Fakat en yüksek başarımları sırasıyla Gama ve Beta alt bantlarında vermiş olmaları nedeniyle Tablo 3.1’e bu şekilde eklenmiştir. Eğer bir genelleme yapılmak istenirse, alfa bandının konuşmacı aşinalığının belirlenmesi için en uygun bant olacağı söylenebilir.

Bunun yanında elektrot bölgelerine bakıldığında kişiler için birbirine yakın bölgeler elde edilmiştir. Ne kadar sağ elini kullandıkları bilinse de deneklerin beyinlerinin hangi bölgesinin dominant olduğu bilinmemektedir. Konuşma, Dil ve İşitmeye Dair Fizyolojik Süreçler başlığında ulaşılabileceği üzere dilsel özellikler, müziksel özellikler konuşmacı

tanımının beyinde işlendiği yerler kişinin beyinin dominant bölgesine göre farklılık göstermektedir. Bu savdan yola çıkılarak kişiler arasında nispeten simetrik sonuçlara ulaşılmıştır.

Tablo 3.1. En yüksek başarılar ve elektrot bölgeleri

| <b>Kişiler</b>                              | <b>Kişi 1</b>       | <b>Kişi 2</b>        | <b>Kişi 3</b>        | <b>Kişi 4</b>        | <b>Kişi 5</b>        | <b>Kişi 6</b>        |
|---|---------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| <b>Elektrot Sayısı</b>                      | 2                   | 4                    | 4                    | 4                    | 4                    | 4                    |
| <b>Alt Bant</b>                             | Alfa                | Alfa                 | Gama                 | Alfa                 | Alfa                 | Beta                 |
| <b>Eğitim Doğruluğu (%)</b>                 | 87,50               | 82,81                | 85,94                | 85,94                | 84,38                | 85,94                |
| <b>Test Doğruluğu (%)</b>                   | 85,94               | 75,00                | 82,81                | 81,25                | 81,25                | 76,56                |
| <b>Seçilen Elektrotlar</b>                  | Fpz                 | F8                   | F3                   | F8                   | T3                   | F3                   |
|   | T3                  | Cz                   | Fz                   | C3                   | C4                   | C4                   |
|   |                     | T5                   | P4                   | T6                   | T6                   | T4                   |
|   |                     | O2                   | O2                   | O2                   | Oz                   | T5                   |
| <b>Elektrotlara Karşılık Gelen Bölgeler</b> | Orta Çizgi          | Sağ Broca bölgesi    | Sol Broca Bölgesi    | Sağ Broca Bölgesi    | Sol İşitsel Korteks  | Sol Broca Bölgesi    |
|   | Sol İşitsel Korteks | Orta Çizgi           | Orta Çizgi           | Sol İşitsel Korteks  | Sağ İşitsel Korteks  | Sağ İşitsel Korteks  |
|   |                     | Sol Wernicke Bölgesi | Sağ Wernicke Bölgesi | Sağ Wernicke Bölgesi | Sağ Wernicke Bölgesi | Sağ İşitsel Korteks  |
|   |                     | Görsel Korteks       | Görsel Korteks       | Görsel Korteks       | Görsel Korteks       | Sol Wernicke Bölgesi |

Tablo 3.1’de ayrıca elektrot bölgelerine en yakın ilgili bölgeler de gösterilmektedir. Buna göre elektrotlar konuşma ve dil ile ilişkilendiren bölgelerde toplanmıştır. Kişi bazlı olarak elektrot isimleri farklılık gösterse de her bir kişinin kendine has bir kafa yapısı olması ve kullanılan elastik başlığın genel bir kafa yapısına göre tasarlanmış olması bu farklılıkların hoş görülebilir olduğunu göstermektedir.

Öte yandan konuşmacı aşinalığının EEG tabanlı belirlenmesinde bu uygulamanın zorluklarından ve başarımının artırılmasında etkili olan faktörlerden bahsetmek gerekir.

Davranışsal bulgulara göz atıldığında kişilerin deney süresince hata yaptığı gerçeği söz konusudur. Bu uygulamada kullanılan etiketlerin kesinliğinin sorgulanmasına neden olan bu gerçek sınıflandırma başarısını düşüren en büyük etkenlerden biridir. Kişilerin gerçekte tanıdığı kişileri tanıyamamasının beyinde nasıl bir etki yapacağı henüz bilinmemektedir. Bu yüzden hatalı olarak eşleştirilen konuşmacılara ait EEG verilerinin burada sınıflandırma başarısının artmasına engel olduğu görülmektedir.

Bir başka açıdan bakıldığında kişilerin cevap verirken emin olup olmama durumudur. Kişi aslında tanıdığı birinin konuşmasını dinlerken şüphelendiği kişinin gerçekten konuşmacı olması ve emin olamama durumu nedeniyle tanımadığını belirtmesi etiketlerin içinde tanınmış kişinin tanınamış olarak işlenmesine neden olmuştur. Bu da sınıflandırma başarısını artırmada önemli bir engeldir.

Ayrıca kişinin tanıyıp tanımama kararını verdiği anın EEG kayıtlarının aktif olmadığı bölgelerde vermiş olma ihtimali de önemli bir engel olarak göze çarpmaktadır. Kişi senaryo süresince alınan EEG kaydında kararını verememiş ama EEG kaydı bittikten sonra cevap aşamasında kararını vermiş olabilir. Bu da alınan işaretin içinde önemli olan karar anının olmaması anlamına gelmektedir.

Son olarak kişinin kararını erkenden vermesi durumu da söz konusudur. Konuşma kaydı başladıktan kısa bir süre sonra kişi konuşmacıyı tanımış olabilir. Tanıma anından sonra kişi deney senaryosu ile ilgili yapacağı başka bir işlem olmadığından dolayı odağı konuşmacıdan kaymış olabilir. Bu da deney senaryosundan ayrı zihinsel aktivitelere sebep olacağından EEG verisinin içinde önemli bilgi olan karar anının etkisini azaltacaktır.



#### 4. ÖNERİLER

Bu tez çalışmasında konuşmacı aşinalığının EEG tabanlı belirlenmesi ilgili başlıklarda verilen yöntemlerle araştırılmıştır. Dinleyiciler deney öncesi konuşmacıların kimliği hakkında bilgilendirilmemiştir. Takip edilen deney senaryosunda dinleyicilerin başlangıçta konuşmacı kimlikleri hakkında bilgi sahibi olması durumunda sonuçlar elde edilebilir ve bu çalışma ile karşılaştırılabilir.

Bu çalışma dahilinde kadın ve erkek sesleri ayrı ayrı gruplar halinde sunulmuş olsa da başka bir çalışma için kadın ve erkek seslerinin karışık olarak sunulduğu senaryolar hazırlanabilir.

Senaryoda dinleyicilerden konuşma kayıtlarını sonuna kadar dinlemeleri istenmiş ve bitiminde cevap vermeleri beklenmiştir. Başka bir çalışma için dinleyicilerin istedikleri zamanda cevap verebilecekleri bir senaryo oluşturulabilir ve sınıflandırma doğruluğuna etkisi incelenebilir.

Elektrot pozisyonları ve alt bant türleri bu çalışma için kişiler arası farklılık göstermiş olsa da her kişi için kabul edilebilir ve yüksek sınıflandırma doğruluğu gösteren elektrot pozisyonları ve alt bant türü belirlenebilir. Böylece genel kitleye de uyumluluk sağlayabilecek bir kümeleme yöntemi geliştirilerek aşinalık tespiti yapılabilir.

Bu tez çalışması konuşmacı aşinalığı üzerine yoğunlaşmış olsa da anadil ve yabancı dil aşinalığı, erkek ve kadın denekler üzerinde yapılabilecek cinsiyet tanıma uygulaması, konuşmacıların yaşının belirlenmesi, konuşmacıların duygusal durumunun belirlenmesi gibi çalışmalar da yapılabilir.

## 5. KAYNAKLAR

1. T. Kinnunen and H. Li, An overview of text-independent speaker recognition: From features to supervectors, *Speech Commun.*, 52,1(2010)12–40.
2. B. S. Kisilevsky et al., Effects of experience on fetal voice recognition, *Psychol. Sci.*, 14,3(2003) 220–224.
3. M. Beauchemin et al., Mother and stranger: An electrophysiological study of voice processing in newborns, *Cereb. Cortex*, 21,8(2011) 1705–1711.
4. P. Belin and M. H. Grosbras, Before Speech: Cerebral Voice Processing in Infants, *Neuron*, 65,6(2010)733–735.
5. R. Winston, *The human mind and how to make the most of it*, London: Transworld/BBC, 2003.
6. N. J. Lass, K. R. Hughes, M. D. Bowyer, L. T. Waters, and V. T. Bourne, Speaker sex identification from voiced, whispered, and filtered isolated vowels., *J. Acoust. Soc. Am.*, 59,3(1976)675–678.
7. T. Murry and S. Singh, Multidimensional analysis of male and female voices., *J. Acoust. Soc. Am.*, 68,5(1980)1294–1300.
8. C. Clopper and D. Pisoni, Some Acoustic Cues for the Perceptual categorization of American English regional dialects, *J. Phon.*, 32,1(2004)111–140.
9. D. R. Preston, *Handbook of perceptual dialectology*, vol. 1. John Benjamins Publishing, 1999.
10. D. R. Preston, Folk dialectology, *Am. dialect Res.*, pp. 333–377, 1993.
11. R. E. Remez, S. C. Wissig, D. F. Ferro, K. Liberman, and C. Landau, A search for listener differences in the perception of talker identity, *J. Acoust. Soc. Am.*, 116,4(2004)2544.
12. D. Abercombie, *Elements of general phonetics*. Aldine Pub. Company, 1967.
13. L. Cerrato, M. Falcone, and A. Paoloni, Subjective age estimation of telephonic voices, *Speech Commun.*, 31,2(2000)107–112.
14. D. Read and F. I. M. Craik, Earwitness identification: Some influences on voice recognition., *J. Exp. Psychol. Appl.*, 1,1(1995)6–18.
15. K. R. Scherer, Vocal communication of emotion: A review of research paradigms, *Speech Communication*, 40,1(2003)227–256.

16. M. Schröder, *Speech and Emotion Research: An overview of research frameworks and a dimensional approach to emotional speech synthesis*. Institut für Phonetik, 2004.
17. A. D. Yarmey, A. L. Yarmey, M. J. Yarmey, and L. Parliament, *Commonsense Beliefs and the Identification of Familiar Voices*, *Appl. Cogn. Psychol.*, 15,3(2001)283–299.
18. S. Cook and J. Wilding, *Earwitness testimony .2. Voices, faces and context*, *Appl. Cogn. Psychol.*, 11,6(1997)527–541.
19. J. J. Wolf, *Efficient Acoustic Parameters for Speaker Recognition*, *J. Acoust. Soc. Am.*, 51,6(1972)2044.
20. H. Jonathan and S. Cassidy, *Techniques in speech acoustics*. Dordrecht, The Netherlands: Kluwer Academic Publishers, 1999.
21. A. V Oppenheim, *Discrete-time signal processing*. Pearson Education India, 1999.
22. K. Paliwal and L. Alsteris, *Usefulness of phase spectrum in human speech perception.*, *Interspeech*, (2003)2117–2120.
23. R. M. Hegde, H. A. Murthy, and G. V. R. Rao, *Application of the modified group delay function to speaker identification and discrimination*, 2004 *IEEE Int. Conf. Acoust. Speech, Signal Process.*, 1,1(2004)517.
24. G. R. Doddington, *Speaker recognition based on idiolectal differences between speakers*, in *Eurospeech 2001. Proceedings of the 7th European Conference on Speech Communication and Technology, 2nd Interspeech Event, 2001*, 2521–2524.
25. H. S. M. Beigi, S. H. Maes, U. V Chaudhari, and J. S. Sorensen, *IBM model-based and frame-by-frame speaker recognition*, *Speak. Recognit. its Commer. Forensic Appl. Avignon, Fr.*, 4(1998).
26. S. Maes and H. Beigi, *Open sesame! Speech, password or key to secure your door?*, *Comput. Vision—ACCV’98*, (1997)531–541.
27. H. E. Cetingul, Y. Yemez, E. Erzin, and A. M. Tekalp, *Discriminative lip-motion features for biometric speaker identification*, in *Image Processing, 2004. ICIP’04. 2004 International Conference on*, 3(2004)2023–2026.
28. a. Kanak, E. Erzin, Y. Yemez, and a. M. Tekalp, *Joint audio-video processing for biometric speaker identification*, 2003 *Int. Conf. Multimed. Expo. ICME ’03. Proc. (Cat. No.03TH8698)*, 3(2003)1–4.
29. R. Miller, *Nature of the vocal cord wave*, *J. Acoust. Soc. Am.*, 203,1940(1959) 858–865.

30. J. L. Flanagan, *Speech analysis synthesis and perception*, Springer Science & Business Media, 3(2013).
31. D. a. Reynolds, T. F. Quatieri, and R. B. Dunn, *Speaker Verification Using Adapted Gaussian Mixture Models*, *Digit. Signal Process.*, 10,1(2000)19–41.
32. C. Chau, C. Lai, and B. Shi, *Feature vs. model based vocal tract length normalization for a speech recognition-based interactive toy*, *Act. Media Technol.*, (2001)134–143.
33. G. Peterson and H. L. Barney, *Control Methods Used in a Study of the Vowels*, *Jarnal Acoust. Soc. Am.*, 24,2(1952)175–184.
34. R. Winston and D. Hill, *Human Mind*. Bantam Press, Ltd., 2004.
35. J. H. Breasted, *The Edwin Smith Papyrus*. New-York Historical Society, 1922.
36. H. Beigi, *Fundamentals of speaker recognition*. Springer Science & Business Media, 2011.
37. J. C. Marshall and G. R. Fink, *Cerebral localization, then and now*, in *NeuroImage*, 20,1(2003).
38. B. Parker, *Good vibrations: the physics of music*. JHU Press, 2010.
39. J. L. Flanagan, K. Ishizaka, and K. L. Shipley, *Synthesis of speech from a dynamic model of the vocal cords and vocal tract*, *Bell Syst. Tech. J.*, 54,3(1975)485–506.
40. B. R. Schofield and D. L. Coomes, *Pathways from auditory cortex to the cochlear nucleus in guinea pigs*, *Hear. Res.*, 216(2006)81–89.
41. D. J. Levitin, *This is your brain on music: The science of a human obsession*. Penguin, 2006.
42. D. Eagleman, *Incognito (Enhanced Edition): The Secret Lives of the Brain*. Knopf, 2011.
43. F. Gougoux, P. Belin, P. Voss, F. Lepore, M. Lassonde, and R. J. Zatorre, *Voice perception in blind persons: a functional magnetic resonance imaging study*, *Neuropsychologia*, 47,13(2009)2967–2974.
44. H. Berger, *Über das elektrenkephalogramm des menschen*, *Eur. Arch. Psychiatry Clin. Neurosci.*, 87,1(1929)527–570.
45. G. Dietsch, *Fourier-analyse von elektrencephalogrammen des menschen*, *Pflügers Arch. Eur. J. Physiol.*, 230,1(1932)106–112.
46. F. A. Gibbs, W. G. Lennox, and E. L. GIBBS, *Cerebral blood flow preceding and accompanying epileptic seizures in man*, *Arch. Neurol. Psychiatry*, 32,2(1934)257–272.

47. A. L. Loomis, E. N. Harvey, and G. Hobart, Potential rhythms of the cerebral cortex during sleep., *Science* (80-. ),, 1935.
48. W. G. Walter, The location of cerebral tumours by electro-encephalography, *Lancet*, 228,5893(1936)305–308.
49. H. H. Jasper, The ten twenty electrode system of the international federation, *Electroencephalogr. Clin. Neurophysiol.*, 10(1958)371–375.
50. A. J. F. Griffiths, *An introduction to genetic analysis*. Macmillan, 2005.
51. D. A. Pizzagalli, Electroencephalography and high-density electrophysiological source localization, *Handb. Psychophysiol.*, 3(2007)56–84.
52. P. C. Gilmore and R. P. Brenner, Correlation of EEG, computerized tomography, and clinical findings: study of 100 patients with focal delta activity, *Arch. Neurol.*, 38,6(1981)371–372.
53. A. Fernández-Bouzas et al., Sources of abnormal EEG activity in the presence of brain lesions, *Clin. Electroencephalogr.*, 30,2(1999)46–52.
54. E. R. John, H. Ahn, L. Prichep, M. Trepetin, D. Brown, and H. Kaye, Developmental equations for the electroencephalogram, *Science* (80-. ),, 210,4475(1980)1255–1258.
55. D. L. Schacter, EEG theta waves and psychological phenomena: A review and analysis, *Biol. Psychol.*, 5,1(1977)47–82.
56. E. Niedermeyer and F. H. L. da Silva, *Electroencephalography: basic principles, clinical applications, and related fields*. Lippincott Williams & Wilkins, 2005.
57. V. N. Murthy and E. E. Fetz, Coherent 25-to 35-Hz oscillations in the sensorimotor cortex of awake behaving monkeys., *Proc. Natl. Acad. Sci.*, 89,12(1992)5670–5674.
58. E. Rodriguez, N. George, J.-P. Lachaux, J. Martinerie, B. Renault, and F. J. Varela, Perception's shadow: long-distance synchronization of human brain activity, *Nature*, 397,6718(1999)430–433.
59. W. H. R. Miltner, C. Braun, M. Arnold, H. Witte, and E. Taub, Coherence of gamma-band EEG activity as a basis for associative learning, *Nature*, 397,6718(1999)434–436.
60. S. Uchida et al., Suppression of gamma activity in the human medial temporal lobe by sevoflurane anesthesia, *Neuroreport*, 11,1(2000)39–42.
61. H. Cross, Fisch & Spehlmann's *Eeg Primer: Basic principles of digital and analog Eeg* 3rd Edition., *Epilepsia*, 42,2(2001)287.
62. S. E. P. Nomenclature, American Electroencephalographic Society Guidelines for, *J. Clin. Neurophysiol.*, 8,2(1991)200–202.

63. R. Oostenveld and P. Praamstra, The five percent electrode system for high-resolution EEG and ERP measurements, *Clin. Neurophysiol.*, 112,4(2001)713–719.
64. S. J. Luck, *An Introduction to Event-Related Potentials and their Neural Origins* (Chapter 1). Cambridge: MIT Press, 2005.
65. M. G. Woldorff, S. A. Hackley, and S. A. Hillyard, The effects of channel-selective attention on the mismatch negativity wave elicited by deviant tones, *Psychophysiology*, 28,1(1991)30–42.
66. P. Belin, S. Fecteau, and C. Bedard, Thinking the voice: neural correlates of voice perception, *Trends Cogn. Sci.*, 8,3(2004)129–135.
67. D. R. Van Lancker and G. J. Canter, Impairment of voice and face recognition in patients with hemispheric damage, *Brain Cogn.*, 1,2(1982)185–195.
68. F. Neuner and S. R. Schweinberger, Neuropsychological impairments in the recognition of faces, voices, and personal names, *Brain Cogn.*, 44,3(2000)342–366.
69. D. Van Lancker and J. Kreiman, Voice discrimination and recognition are separate abilities, *Neuropsychologia*, 25,5(1987)829–834.
70. I. Peretz et al., Functional dissociations following bilateral lesions of auditory cortex, *Brain*, 117,6(1994)1283–1301.
71. P. Belin, R. J. Zatorre, P. Lafaille, P. Ahad, and B. Pike, Voice-selective areas in human auditory cortex, *Nature*, 403,6767(2000)309–312.
72. S. E. Barrett, M. D. Rugg, and D. I. Perrett, Event-related potentials and the matching of familiar and unfamiliar faces, *Neuropsychologia*, 26,1(1988)105–117.
73. S. R. Schweinberger, Human brain potential correlates of voice priming and voice recognition, *Neuropsychologia*, 39,9(2001)921–936.
74. S. R. Schweinberger, V. Huddy, and A. M. Burton, N250r: a face-selective brain response to stimulus repetitions, *Neuroreport*, 15,9(2004)1501–1505.
75. M. Beauchemin et al., Electrophysiological markers of voice familiarity, *Eur. J. Neurosci.*, 23,11(2006)3081–3086.
76. S. R. Schweinberger, N. Kloth, and D. M. C. Robertson, Hearing facial identities: brain correlates of face–voice integration in person identification, *Cortex*, 47,9(2011)1026–1037.
77. J. Graux, M. Gomot, S. Roux, F. Bonnet-Brilhault, V. Camus, and N. Bruneau, My voice or yours? An electrophysiological study, *Brain Topogr.*, 26,1(2013)72–82.

78. J. Graux, M. Gomot, S. Roux, F. Bonnet-Brilhault, and N. Bruneau, Is my voice just a familiar voice? An electrophysiological study, *Soc. Cogn. Affect. Neurosci.*, 10,1(2015)101–105.
79. I. Holeckova, C. Fischer, M.-H. Giard, C. Delpuech, and D. Morlet, Brain responses to a subject's own name uttered by a familiar voice, *Brain Res.*, 1082,1(2006)142–152.
80. K. Von Kriegstein, A. Kleinschmidt, and A.-L. Giraud, Voice recognition and cross-modal responses to familiar speakers' voices in prosopagnosia, *Cereb. Cortex*, 16,9(2006)1314–1322.
81. I. Q. González, M. A. B. León, P. Belin, Y. Martínez-Quintana, L. G. García, and M. S. Castillo, Person identification through faces and voices: An ERP study, *Brain Res.*, 1407(2001)13–26.
82. D. Sun, C. C. H. Chan, and T. M. C. Lee, Identification and classification of facial familiarity in directed lying: an ERP study, *PLoS One*, 7,2(2012)31250.
83. R. Zäske, G. Volberg, G. Kovács, and S. R. Schweinberger, Electrophysiological correlates of voice learning and recognition, *J. Neurosci.*, 34,33(2014)10821–10831.
84. N. Thammasan, K. Moriyama, K. Fukui, and M. Numao, Investigation of Familiarity Effects in Music-Emotion Recognition Based on EEG, in *International Conference on Brain Informatics and Health*, 2015, pp. 242–251.
85. Z. H. E. Tan, K. G. Smitha, and A. P. Vinod, Detection of Familiar and Unfamiliar Images Using EEG-Based Brain-Computer Interface, in *Systems, Man, and Cybernetics (SMC), 2015 IEEE International Conference on*, 2015, pp. 3152–3157.
86. K. G. Smitha, A. P. Vinod, and K. Mahesh, Voice familiarity detection using EEG-based Brain-Computer Interface, in *Systems, Man, and Cybernetics (SMC), 2016 IEEE International Conference on*, 2016, pp. 1626–1631.
87. V. Panayotov, G. Chen, D. Povey, and S. Khudanpur, Librispeech: an ASR corpus based on public domain audio books, in *Acoustics, Speech and Signal Processing (ICASSP), 2015 IEEE International Conference on*, 2015, pp. 5206–5210.
88. E. Zwicker, Subdivision of the audible frequency range into critical bands (Frequenzgruppen), *J. Acoust. Soc. Am.*, 33,2(1961)248, 1961.
89. V. N. Vapnik and V. Vapnik, *Statistical learning theory*, vol. 1. Wiley New York, 1998.
90. C. J. C. Burges, A tutorial on support vector machines for pattern recognition, *Data Min. Knowl. Discov.*, 2,2(1998)121–167.
91. H. S. Jayanna and S. R. M. Prasanna, Analysis, feature extraction, modeling and testing techniques for speaker recognition, *IETE Tech. Rev.*, 26,3(2009)181–190.

92. U. G. Matlab, The mathworks, Inc., Natick, MA, vol. 1992, 1760.
93. R. A. Fisher, The use of multiple measurements in taxonomic problems, *Ann. Eugen.*, 7,2(1936)179–188.
94. S. Geisser, A predictive approach to the random effect model, *Biometrika*, pp. 101–107, 1974.





## ÖZGEÇMİŞ

1991 yılında İstanbul'da doğdu. İlk ve orta öğrenimini Hasan Ali Yücel İlköğretim Okulu, lise öğrenimini Kuleli Askeri Lisesi'nde tamamladı. 2010 yılında Karadeniz Teknik Üniversitesi Mühendislik Fakültesi Elektrik-Elektronik Mühendisliği Bölümü'nde lisans programına başladı ve 2014 yılında bu bölümden mezun oldu. Aynı yıl Karadeniz Teknik Üniversitesi Fen Bilimleri Enstitüsü Elektrik-Elektronik Mühendisliği Anabilim Dalı'nda yüksek lisans öğrenimine başladı. 2015 yılı bahar döneminden itibaren Elektrik-Elektronik Mühendisliği bölümünde araştırma görevlisi olarak çalışmaktadır. Yabancı dil olarak iyi seviyede İngilizce bilmektedir.