

MAHREC: Mobil Tabanlı Harf Çıkış Bozukluklarının İyileştirilmesi

Seyfullah Uysal, Emre Yılmaz, Süleyman Eken, Ahmet Sayar

Bilgisayar Mühendisliği, Kocaeli Üniversitesi,
41380 İzmit, Türkiye
seyfullah.ysl@gmail.com, ylmazemre95@gmail.com
{suleyman.eken, ahmet.sayar}@kocaeli.edu.tr

Özet. Fonolojik gerilik veya bozukluğun tespit edilmesi ve düzeltilmesi bireylerin kişilik üzerlerindeki negatif etkisinin kalkması ve sosyal yaşama entegrasyon açısından önemlidir. Telaffuz bozukluklarında konuşmacılar, normal olarak konuşabilmesine rağmen bazı ses veya ses gruplarını yanlış telaffuz etmekte. Bu çalışmada mobil tabanlı olarak izole olmuş bazı seslerin (r, s-ş, z) telaffuzlarındaki bozukluk seviyesi tespit edilerek kişinin bozukluk seviyesine (başlangıç, orta, ileri) göre bir takım dinleme egzersizleri önerilmiştir. İlgili aşamalar sonunda ses tekrar alınarak gerçekte olması gerekene göre kalitesi değerlendirilerek iyileşme varsa üst seviyeye çıkması (başlangıçtan orta seviyeye gibi) sağlanacaktır. Böylelikle telaffuz bozukluğu yaşayan bireylerin gelişiminin sağlanması hedeflenmiştir.

Anahtar kelimeler: Artikulasyon bozukluğu, r sesi bozukluğu, s ve z sesi bozukluğu

MAHREC: Mobile Based Improvement of Letter Pronunciation Disorders

Abstract. Detection and correction of phonological retardation is important for the development of the negative effect of individuals on personality and integration into social life. Speakers in pronunciation disorders often misrepresent some sounds or groups of sounds, although they can normally speak. In this study, we implement a mobile application to determine the level of impairment of some isolated sounds (r, s-ş, z) and a number of listening exercises are proposed according to the level of the person's impairment (beginning, middle, advanced). At the end of the relevant stages, the sound will be taken again and the quality will be assessed according to the actual necessity, and if it is improved, it will be raised to the upper level (from the beginning to the middle level). Thus, it is aimed to provide the development of individuals with impaired pronunciation.

Keywords: Articulation disorders, rotasizm, sigmatizm

1 Giriş

Artikülasyon (telaffuz etme), nefesin gırtlaktan çıktıktan sonra yutak, ağız ve burundan oluşan üçüncü küme organlarında (dil, diş, damak, dudak) konuşma dilimizin geleneksel seslerine dönüşüp biçimlenmesidir. Artikülasyon bozukluğu ise bireyin yaşına ve konuşma gelişimi dönemine uygun olarak beklenen şekilde ana dilinin bağımsız ya da bileşik seslerini doğru ve anlaşılır biçimde çıkaramaması ve birbirine gereği gibi uylayamaması şeklindedir. Konuşma üretiminde birbirleri ile ilgili dört süreç vardır [1]. Bunlar;

- Seslenim (Fonasyon): Vokal kasların titreşimi ile seslerin üretilmesidir.
- Solunum (Respirasyon): Konuşma için motive edici gücü sağlar.
- Yankılama (Rezonasyon): Kişinin sesinin niteliğini etkileyen konuşma sesinin farkında önemli rol oynayan bir süreçtir.
- Söyleme-Ekleme (Artikülasyon): Bireysel konuşma seslerinin dil, çene, dudaklar ve yumuşak damağın yardımıyla ses tonu ve nefes akışının çıkarılmasında üretilen ve bu yolla sesin yankılanmasını da içeren bir süreçtir.

Konuşma seslerinin etkilenmesine göre her artikülasyon bozukluğuna özel bir terim verilmiştir: a-Rotasizm (r sesi bozukluğu), b-Sigmatizm (s ve z sesi bozukluğu), c-Gamatizm (g sesi bozukluğu), d-Kapasizm (k sesi bozukluğu). Artikülasyon bozukluğu çocuklarda dört değişik türde görülebilmektedir. Sesin düşürülmesi veya atlanması, (omissions) bir sözcüğü oluşturan seslerin tümü çıkarılmadan sözcüğün söylenmeye çalışılması durumunda ortaya çıkmaktadır (hayır-ayır gibi). Ses eklenmesi (additions) sözcükte olmayan başka seslerin eklenmesi durumunda ortaya çıkmaktadır (Recep-irecep). Sesin değiştirilmesi (substitutions), sözcük içinde çıkarılması güç gelen bir sesin çıkarılması kolay gelen bir sesle değiştirilmesi durumunda ortaya çıkar. Değişiklikler bazen sözcüğün başındaki seste, bazen de ortasındaki seslerde olur. Bazendeğişiklikler, sözcük içindeki seslerin yerleri değiştirilerek de yapılabilir (Süleyman-Sümeylan, yüzük-yüsük gibi). Sesin bozulması (distortions), bu üç duruma uymayan durumlar da olabilir. Burada sözcük oluşturulurken esas çıkarılması gereken ses, olduğundan başka ses çıkarılarak konuşulur. Bu da konuşmayı engelli hale getirir (Karagöz-Kaxgöz-Kağagöz gibi). Biz bu çalışmada sesin değiştirilmesi durumuyla ilgileneceğiz. Konuşmanın anlaşılmasının zorluğu artikülasyon bozukluğunun derecesi ile ilgilidir.

Yukarıda bahsi geçen bozukların birtakım farklı nedenleri olabilir: Yapısal (organik) nedenler konuşma organlarındaki özellikle üçüncü küme organlarındaki (dudak, diş, damak, dil vs.) bir ya da birkaç organik bozukluktan ötürü artikülasyon bozukluğu olabilir. İşitme engelli ya da işitme duyarlılığındaki yetersizlik artikülasyonu olumsuz yönde etkilemektedir. Evde konuşulan dil, çocuğun konuşma şevkinin kırılması, konuşmanın engellenmesi, konuşmayı pekiştirmeye olanak vermeyen ortam artikülasyon bozukluğuna neden olan etmenlerdir. Bazı durumlarda konuşma bozuklukları duygusal çatışmaya bağlı olarak gelişebilir. Sonuç olarak ne

sebeple olursa olsun artikülasyon engeli türleri (düşürme, ekleme, değiştirme, bozma) erken yaşlarda saptanmalı ve artikülasyon derecesine göre önlemler alınmalıdır [2].

Bu güne kadar konuşmacı ve ses tanıma üzerine pek çok çalışma yapılmış ve günümüzde artık çok başarılı sonuçlar elde edilmektedir. Son yıllarda Türkçe sesler üzerine yapılan çalışmalarda belirgin bir artış görülmektedir. Yörüklü ve Koçal [4] Türkçe Sesli harflerin SOM (Self Organizing Maps- Kendi Kendini Düzenleyen Haritalar) modellenmesi yaptıktan sonra konuşmacı ve sesli harf sınıflandırmaları incelenmiştir. Çalışmada, ses ve konuşmacı tanıma işlemlerinin genelde periyodikimsi bir yapı sergileyen sesli harfler üzerinde çalışılarak yapıldığı düşünülerek sesli harfleri baz almışlardır. Inge ve arkadaşları [5] da rotasizm üzerine bir çalışma yapmışlardır. Mel-süzgeci (Mel-cepstrum) katsayısını özellik çıkarma metodu, kNN'ni de sınıflandırıcı olarak kullanmışlardır.

Bu çalışmanın geri kalanı şu şekilde organize edilmiştir. İkinci bölümde rotasizm ve sigmatizmin derecesinin saptanması ve iyileştirilmesine yönelik geliştirilen mimari açıklanmıştır. Üçüncü bölümde geliştirilen mobil uygulamaya ait arayüzler verilmiştir. Son kısımda ise gerçekleştirilen birtakım testler ve analizler verilip sonuçlar değerlendirilecek ve gelecek çalışmalardan bahsedilmiştir.

2 Harf Çıkış Bozuklarını Saptama ve İyileştirme Mimarisi

Bazı seslere ait artikülasyon bozukluklarının tespiti ve işlemi için önerilen sistemin bileşenlerine ilişkin detaylı açıklama ve önerilen sistemde kullanılacak materyaller takip eden alt bölümlerde verilmiştir.

2.1 Konuşma Örneklerinden Oluşan Veri Setinin Oluşturulması

Konuşma seslerinin etkilenmesine göre her artikülasyon bozukluğuna (r, s-ş, z) sahip erkek ve kadınlardan örnekler alınmıştır. Bu örnekler ilgili harflerin ve bu harfleri içeren kelimelerin (izole edilmiş) en az 10'ar kere tekrar edilmesinden oluşmaktadır. Aynı harf ve kelimeleri sağlıklı şekilde telaffuz eden erkek ve kadınlara ait örnekler de veri setinde yer almaktadır. Artikülasyon bozukluğunun ilerleme seviyelerine göre aynı kişilerden aynı harf ve kelimelerle setin genişletilmesi de düşünülmektedir. Veri seti, eğitim ve test için kullanılmaktadır.

2.2 Öznitelik Çıkarımı

Ses tanımda en önemli şey, konuşma sinyalinden özellikleri çıkaran özellik çıkarımıdır. Özellik özütleme, girdi verisini özellik kümesine dönüştüren ve öznitelik çıkarma denilen bir işlemdir. Öznitelik çıkarımı sırasında bir hoparlörün ayırt edici özelliğini korurken giriş vektörünün boyutunu küçültür. Günümüzde en çok

kullanılan özellik cepstral katsayısıdır. Yaygın kullanılan iki tip cepstral katsayısı Linear predictive cepstral coefficient (LPCC) ve Mel frequency cepstral coefficient (MFCC) dir. Aşağıda özellik çıkarımı ile ilgili alt adımlar verilmiştir.

2.2.1. Ön İşlemler

Konuşma kaynağı olarak akıllı telefonlardan yararlanılacaktır. Konuşma tanıma uygulamalarında ses işareti frekans dönüşümüne uğramadan önce bazı ön işlemlerden geçirilmektedir. Uygulanan ön işlemler aşağıda belirtilmiştir.

- Normalizasyon: Ses veri değerlerinin belirli bir aralığa çekilmesi
- Ön vurgulama (Preemphasis): Yüksek frekanslı bölgelerin güçlendirilmesi
- Dither: Giriş işaretinde büyük miktarda sıfır içermesi durumunda azaltma

2.2.2 Çerçeve Üretimi ve Pencereleme

Telaffuz bozukluğu olan ilgili harfin tespitinde çerçeve üretimi yapılması gerekmektedir. Ses verisinin tamamı değil kısa harflere karşılık düşecek küçük parçaları ile işletilir. Bu nedenle ön işlemeden geçirilmiş konuşma işareti “çerçeve” adı verilen sabit uzunluklu parçalara bölünür. (Pencereleme işlemi ile) örtüşmüş çerçeveler üretilir. Yani ses işareti üzerinde 25ms uzunluğundaki bir pencerenin belirli bir süre mesafesince kaydırılması ile çerçeveler üretilir.

2.2.3 Dönüşüm, Filtreleme ve Özelliklerin Belirlenmesi

Ayrık Fourier dönüşümü ile zaman düzlemindeki işaretlerin frekans düzlemindeki karşılıkları elde edilebilir. Her çerçeve için bu işlemin yapılması gerektiğinden bu dönüşümü hızlı şekilde hesaplayan Hızlı Fourier Dönüşümü (FFT) kullanılmıştır. Mel süzgeci ile FFT sonucunda ortaya çıkan enerji değerlerinin farklı frekans bantlarındaki toplam değerleri hesaplanır [6]. Bu şekilde konuşma sinyali parçasını daha az sayıda parametre ile ifade edilmiş olur. Daha sonra Mel Filtre Bankası analizi, Kepstrumların elde edilmesi ve kepsral ağırlıklandırmadan sonra Karşılıklı ilinti [7], işlemeyle özilinti vektörü [8] hesaplanır. Sonuç olarak, her çerçeve için 8 LPC ve MFCC öznitelik değerleri elde edilir. Bu aşamalardan sonra ilgili artikülasyon bozukluğunu saptamak için özellik vektörü elde edilmiştir.

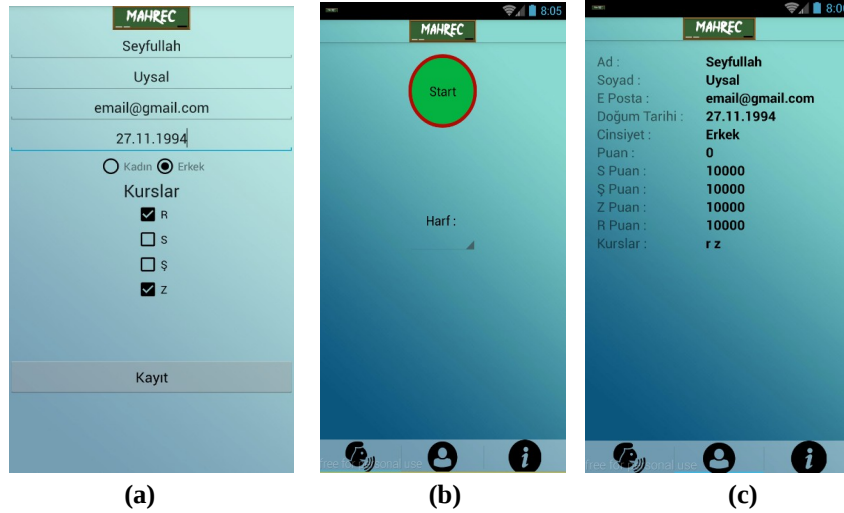
2.3 Sınıflandırma

Sınıflandırma yöntemi olarak gizli markov modeli (GMM) temelli bir sınıflandırıcı kullanılmıştır. GMM, ses sinyallerinin istatistiksel olarak modelleyen bir metottur.

SMM, en başarılı konuşma tanıma metodlarından biri olmuştur. Çünkü, GMM, ses sinyallerini çok uygun bir şekilde matematiksel olarak karakterize etme özelliğe sahiptir. Mobil platformda ses tanıma işlemlerini gerçekleştirmek için pocketsphinx-android kütüphanesinden yararlanılmıştır.

3 Kullanıcı Arayüzleri

Kullanıcı sisteme ilk kayıt esnasında birtakım kişisel bilgilerini ve zayıf olduğu karakterleri/kursları seçer (Şekil 1a). Daha sonraki oturumlarda alıştırma sekmesine gelerek kayıt esnasında seçilen karakterlerle alıştırma yapabilir. Bunun için alıştırma yapılması istenilen harfi seçtikten sonra “start” butonuna basarak ses kaydını başlatır ve alıştırma harfini söyler (Şekil 1b). Yapılan alıştırmalara göre kullanıcı puanı hesaplanır ve bu puanlar göz önüne alınarak belirli aralıklarla alıştırma yapılması için kullanıcılara bildirim gönderilir. Puanlar ve diğer bilgiler profil sekmesinden görüntülenebilir (Şekil 1c).



Şekil 1 Mobil uygulama arayüz görüntüleri

4 Sonuç ve Gelecekteki Çalışmalar

Geliştirilen sistemde r,s,z harfleri için bay ve bayanlardan çeşitli seviyelerde (düşük,orta ve yüksek) örnekler alınmış ve bu örnekler ile uygulama eğitilmiştir. Test aşamasında ise rotasizm ve sigmatizmlili kişiler üzerinde test edilmiş ve sonuçları Tablo 1'de gösterildiği gibidir. Elde edilen sonuçlara göre kesinlik ve recall değerleri sırası ile %68.18, %65.20'dır.

Tablo 1. Karışıklık Matrisi

	Pozitif Sınıflandırma	Negatif Sınıflandırma
Pozitif Örnekler	15	8
Negatif Örnekler	7	10

Artikülasyon bozukluğuna sahip bireyleri topluma kazandırabilmek, eğitim hayatına daha iyi adaptasyon sağlayabilmeleri için mobil tabanlı olarak kullanıcıların sesleri alınarak izole olmuş bazı seslerin (r, s-ş, z) telaffuzlarındaki bozukluk seviyesi tespit edilmiş ve iyileştirilmesi sağlanmıştır. İleriki çalışmalarda daha fazla kişiden örnek alınarak veriseti genişletilecek ve YSA, SVM gibi farklı sınıflandırıcılar ile performans karşılaştırması yapılacaktır.

Teşekkür

Bu çalışma, TÜBİTAK tarafından 1919B011602169 nolu proje ile desteklenmektedir. Desteklerinden dolayı TÜBİTAK'a teşekkür ederiz.

Kaynaklar

1. Konuşma bozuklukları, <http://www.kekemelikegitimi.net/konusma-bozukluklari/artikulasyon-harf-soyleyememe> (Erişim Tarihi, 15 Mart 2017)
2. Söyleme kusurları, <http://www.kendinigelistir.com/soyleme-kusurlari/> (Erişim Tarihi, 15 Mart 2017)
3. A.O. Özcan and A.F. Özcan: “Türk Çocuklarının Ses Gelişim Özellikleri ve İlk Okuma Yazma Öğrenme”, İstanbul Gelişim Üniversitesi Sosyal Bilimler Dergisi, 1(2): 67-86, 2014.
4. E. Yörüklü and O.H. Koçal: “Kendi Kendini Düzenleyen Haritalar Yöntemiyle Türkçe Sesli Harflerin Sınıflandırılması Ve Tanınması”, Uludağ Üniversitesi Mühendislik-Mimarlık Fakültesi Dergisi, 17(1), 2012.
5. G. Inge, O. Grigore, V. Velican: “Impaired Speech Recognition: Case Study on Recognition of Initial ‘r’ Consonant in Rhotacism Affected Pronunciations”, in Proc. of 6th Conference on Speech Technology and Human-Computer Dialogue, pp. 1-6, 2011.
6. S.K. Gaikwad, B.W. Gawali, P. Yannawar: “A Review on Speech Recognition Technique”, International Journal of Computer Applications, 10(3): 16-24, 2010.

7. J.G. Proakis, D.G.Manolakis, Digital Signal Processing: Principles and Application, Prentice-Hall, Upper Saddle River, NJ, 1996.
8. L. Rabiner, B.H. Juang, Fundamentals of Speech Recognition, Prentice-Hall, Englewood Cliffs, NJ, 1993.