

Konuşma Terapisine Yönelik Otomatik Konuşma Tanıma Yöntemleri

Oytun Türk

Levent M. Arslan

Elektrik-Elektronik Müh. Bölümü, Boğaziçi Üniversitesi, Bebek, İstanbul
Sestek A.Ş., ARI-1 Teknopark Binası, İTÜ Ayazağa Kampüsü, Maslak, İstanbul
oytun@sestek.com.tr arslanle@boun.edu.tr

Özetçe

Konuşma terapisinde amaç, konuşma bozukluklarının terapi yoluyla tedavisidir. Bu çalışmada Türkçe konuşma terapisinde kullanılabilecek konuşma tanıma yöntemlerinin incelenmesi amaçlanmaktadır. Yalıtılmış Türkçe fonemlerin ve sözcüklerin tanınması için sürekli dağılımlı Saklı Markov Modelleri kullanılmıştır. Sözcük tanımada birbirine çok yakın sözcüklerin tanınması problemi üzerinde durulmuştur. Deneyleerde kullanılmak üzere Türkçe bir veri tabanı tasarlanmış ve çeşitli konuşmacılardan toplanmıştır. İlk aşamada konuşmacı bağımsız eğitilen modeller ile terapide kullanılan fonemler için %84.9, sözcükler için %94.2'lik doğru tanıma oranları elde edilmiştir. Konuşmacı bağımlı eğitimle Türkçe konuşma terapisinde kullanılabilecek sözcükler için %97.2'lik orana ulaşılmıştır. Sıklıkla birbirine karıştırılan Türkçe fonemler içeren yakın sözcük çiftleri arasında %88.0'lik tanıma oranı elde edilmiştir. Bu çalışmada incelenen konuşma tanıma yöntemlerinin kullanıldığı, eğitilebilir ve hastaya uyarlanabilir bir Türkçe konuşma terapisi yazılımının geliştirilmesine başlanmıştır.

Abstract

Speech therapy focuses on methods for the treatment of speech and language disorders. In this study, speech recognition methods are investigated for computer assisted speech therapy in Turkish. Continuous-mixture Hidden Markov Models are employed for isolated phoneme and isolated word recognition tasks. Special care is taken for the recognition of confusable words. A Turkish database is designed and collected from native speakers for the evaluations. Initial experiments indicate 84.9% correct recognition rate for isolated phonemes and 94.2% for isolated words when the system is tested in speaker-independent mode. A correct recognition rate of 97.2% is achieved with speaker-dependent training for a list of Turkish words that can be used in speech therapy. The recognition rate between word pairs that contain confusable Turkish phonemes is 88.0%. A software tool for

speech therapy in Turkish is being developed that employs the speech recognition methods investigated in this study.

1. Giriş

Konuşma terapisinde amaç, doğru seslendirmenin, vurgu ve ritim kontrolünün öğretilmesi, konuşmadaki hata ve bozuklukların giderilmesidir. Bu amaçla terapist hastaya çeşitli testler uygulayarak konuşma bozukluğu sorunlarını tespit etmekte, uygun tedavi yöntemleriyle sorunları gidermeye çalışmaktadır. Konuşma terapisinden yararlanan kişiler arasında işitme engelliler, afazi hastaları, yeni bir dil öğrenenler ve çeşitli fonemleri/sözcükleri seslendirmekte zorluk çeken çocuklar sayılabilir.

Konuşma terapisinde teşhis ve tedavi aşamalarında otomatik konuşma tanıma yöntemlerinin kullanımı yaygınlaşmaya başlamıştır [1], [2], [3]. Artikülasyon eğitimi için tasarlanan yazılım araçlarında konuşma tanıma kullanıcısıyla bilgisayar arasında birebir etkileşim kurulmasını sağlamaktadır [4]. Telaffuz hatalarının otomatik olarak belirlenmesi ve doğru telaffuzun öğretilmesi için fonem tanıma [5] ve ses perdesi eğrilerinin analizi [6] gibi teknikler uygulanmaktadır.

Terapide konuşma tanıma yöntemlerinin kullanımı, terapist ve hastaya çeşitli yararlar sağlamaktadır:

- Teşhis sürecinin hızlandırılması
- Hastanın terapi boyunca kaydettiği ilerlemenin izlenmesi ve değerlendirmelerde nesnel ölçütlerin kullanılması
- Tedavi amaçlı egzersizlerin tasarlanıp uygulanabileceği kullanımı kolay ve eğitilebilir yazılım araçlarının geliştirilip terapide kullanılması
- Hastanın çeşitli egzersiz ve testleri kendi başına uygulama imkanının sağlanması

2. Amaç

Bu çalışma, Türkçe konuşma terapisinde kullanılacak konuşma tanıma yöntemlerinin incelenmesini amaçlamaktadır. Başlıca iki problem üzerinde durulmuştur: yalıtılmış Türkçe fonem tanıma ve yalıtılmış Türkçe sözcük tanıma. Tanıma için sürekli dağılımlı Saklı Markov Modelleri (SMM) kullanılmıştır [7]. Bölüm 3, konuşma terapisinde kullanılması amaçlanan konuşma tanıma yöntemlerini açıklamaktadır. Terapide

kullanılabilecek üç konuşma tanıma görevi belirlenmiş ve bu görevler için SMM tabanlı konuşma tanıma yöntemleri kullanılmıştır. Deneyler için uygun Türkçe veri tabanları tasarlanmış ve çeşitli konuşmacılardan toplanmıştır. Deney sonuçları ayrıntılı olarak incelenmiştir. Bölüm 4’te bulgular özetlenmiştir. Çalışma, sonuçların tartışıldığı Bölüm 5 ile sona ermektedir.

3. Yöntem

Konuşma terapisinde konuşmadaki sorunların belirlenmesi için çeşitli deneyler uygulanmaktadır. Bu amaçla hastaya önceden hazırlanmış çeşitli sözcükler ve cümleler okutulmakta ve terapist tarafından ayrıntılı olarak incelenmek üzere kaydedilmektedir. Bu ses kayıtlarından belirlenen sorunlara yönelik terapi egzersizleri uygulanmakta ve hastanın kaydettiği ilerleme terapist tarafından sürekli olarak izlenmektedir.

Sık karşılaşılan konuşma bozukluklarından biri çeşitli fonemlerin birbiriyle karıştırılmasıdır. Tablo 1’de konuşma bozukluklarında sıklıkla birbiri yerine kullanılan Türkçe fonemler verilmiştir. Bu tür sorunların giderilmesi için terapist hastaya uygulamalı olarak ilgili fonemlerin seslendirilişini göstermektedir. Daha sonra hasta, ilgili fonem çiftlerinin geçtiği sözcük ve cümleler okumaktadır. Terapist bu sözcük ve cümleleri dinleyerek hastanın kaydettiği ilerlemeyi izlemekte ve gerektiğinde müdahale etmektedir.

Fonem Türü	Fonem Çiftleri	Örnekler
Diş-Dudak Ünsüzleri	/f/-/v/	de <u>f</u> a-de <u>v</u> a
Dil Ucu-Dişeti Ünsüzleri	/s/-/z/ /t/-/d/	ka <u>s</u> -ka <u>z</u> ka <u>tı</u> -ka <u>dı</u>
Dişeti-Damak Ünsüzleri	/ş/-/j/ /c/-/ç/	be <u>ş</u> -be <u>j</u> <u>c</u> am- <u>ç</u> am
Çift Dudak Ünsüzleri	/p/-/b/	<u>p</u> ut - <u>b</u> ut
Art Damak Ünsüzleri	/k/-/g/	<u>k</u> ar- <u>g</u> ar
Dişeti Ünsüzü ve Sert (Ön) Damak Ünsüzü	/r/-/y/	bi <u>r</u> -bi <u>y</u>

Tablo 1: Konuşma bozukluklarında sıklıkla birbiriyle karıştırılan Türkçe fonemler.

Bu çalışmada çeşitli fonemlerin birbirine karıştırılmasından ve doğru seslendirilememesinden kaynaklanan sorunların teşhisi ve tedavisinde kullanılabilecek

otomatik konuşma tanıma yöntemleri incelenecektir. Otomatik konuşma tanıma, konuşma işaretinin makineler tarafından tanınması olarak tanımlanabilir. Programlama kolaylığı, hızı ve güvenilirliği nedeniyle bilgisayarlar konuşma tanıma yöntemlerinin gerçekleşmesinde sıklıkla kullanılmaktadır. Daha fazla hız gerektiren uygulamalar için özel elektronik yongaların tasarlanması da mümkündür.

Kullanılan konuşma tanıma sisteminin akış şeması Şekil 1’de gösterilmiştir. Sistem üç görev için kullanılmıştır:

- Türkçe yalıtılmış fonemlerin tanınması
- Belirli bir Türkçe sözcük listesinden yalıtılmış sözcük tanıma
- Birbirine çok yakın (başta, ortada ya da sondaki tek fonemin farklı olduğu) sözcüklerin tanınması



Şekil 1: Kullanılan konuşma tanıma sistemi.

Konuşma tanıma için sürekli dağılımlı SMM kullanılmıştır. SMM ile konuşma işareti istatistiksel olarak modellenmekte ve farklı konuşma işaretlerine karşılık gelen modeller arasında tanıma yapılabilmektedir. Yüksek tanıma performansları nedeniyle SMM tabanlı yöntemler konuşma tanımada en sık kullanılan yöntemlerdir.

Eğitim ve tanıma öncesi konuşma işaretlerinin başlangıç ve bitiş anlarının sezimi, enerji eşikleme tabanlı bir yöntemle gerçekleştirilmiştir. Akustik öznitelikler Mel-Frekanslı Kepstrum Katsayıları (MFKK), enerji ve ötümlülük olasılığı ile bunlara karşılık gelen fark ve ivme parametreleridir. MFKK ve enerji parametreleri ile konuşma işaretinin kısa süreli spektrumu (20-30 ms.) modellenmektedir. Ötümlülük olasılığı, konuşma işaretinin periyodikliği ile ilgili bir öznitelik olup ötümlü seslerde yüksek (1.0’a yakın), ötümsüz seslerde düşük (0.0’a yakın) değerler almaktadır. Fark ve ivme parametreleri MFKK, enerji ve ötümlülük olasılığı özniteliklerinin konuşma işareti boyunca değişimlerinin hesaplanmasıyla

elde edilmektedir. Böylece konuşma işaretinin zamanla değişen özellikleri de modellenebilmektedir.

Her konuşma tanıma görevi için geçerleme kümesi üzerinde en yüksek tanıma oranını veren parametreler kullanılmıştır. Her Markov durumundaki akustik öznitelik vektörlerinin olasılık yoğunluk işlevi, eğitimde kullanılan konuşmacı sayısına bağlı sayıda Gauss dağılım bileşeni içeren bir Gauss Karışım Modeli (GKM)'yle modellenmiştir. Eğitim Baum-Welch algoritmasıyla gerçekleştirilmiş, tanıma için Forward algoritması kullanılmıştır. Farklı konuşma tanıma görevleri için kullanılan parametreler ve diğer ayrıntılar ilgili bölümlerde verilmiştir.

3.1. Yalıtılmış Fonem Tanıma

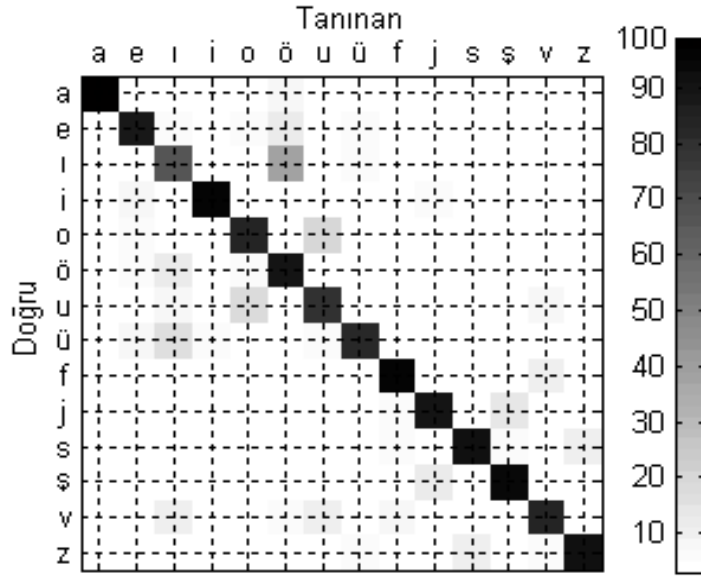
Konuşma terapisinde fonemlerin doğru seslendirilmesi için gerçekleştirilen uygulamalardan biri fonemlerin uzun bir şekilde hastaya söylenmesidir. Böylece hastanın ilgili fonem için doğru dil ve dudak hareketlerini öğrenmesi sağlanmaktadır. Bu amaçla kullanılan fonemlerin bir kısmı Tablo 2'de verilmiştir.

Konuşma tanıma sisteminin eğitimi için 12 konuşmacıdan (dört bayan, sekiz bay) Tablo 2'de listelenen 14 fonem kaydedilmiştir. Örnekleme sıklığı 16 KHz'dir. Konuşmacılardan her fonemi ortalama iki saniye boyunca olabildiğince sabit ses seviyesinde ve sabit vurguyla söylemeleri istenmiştir. Gerektiğinde kayıt öncesi örnekler dinletilmiştir. Her fonem toplam beş kez kaydedilmiştir. Kayıt listesi hazırlanırken ardışık kayıtlarda aynı fonemin yer almamasına dikkat edilmiştir.

Ünlüler	/a/, /e/, /ı/, /i/, /o/, /ö/, /u/, /ü/
Ünsüzler	/f/, /j/, /s/, /ş/, /v/, /z/

Tablo 2: Fonem listesi.

En iyi SMM parametrelerinin belirlenmesi için iki konuşmacı (bir bayan, bir bay) geçerleme kümesi olarak ayrılarak diğer 10 konuşmacıyla farklı parametrelerle eğitim gerçekleştirilmiştir. Geçerleme kümesinde en az hatayı veren parametre kümesi belirlenmiştir. Sonuçlara göre her fonem tek durumlu bir SMM ile modellenmiş ve dört Gauss dağılım bileşeni kullanılmıştır. Fonemler olabildiğince sabit vurguyla söylendiğinden fark ve ivme parametrelerinin kullanılması tanıma oranlarını azaltmaktadır.



Şekil 2: Fonem tanıma oranları (%).

Fonem tanıma oranının belirlenmesi için çapraz-geçerleme yöntemiyle 12 kez eğitim ve tanıma gerçekleştirilmiştir. Her eğitim ve tanıma aşamasında farklı bir konuşmacı test için ayrılıp kalan 11 konuşmacıyla fonem modelleri eğitilmiştir. Test için ayrılan konuşmacı eğitim kümesinde kullanılmadığı için sonuçlar sistemin konuşmacı bağımsız performansını göstermektedir. Yalıtılmış fonem tanıma oranı, 12 çapraz-geçerleme adımının ortalaması alınarak %84.9 olarak hesaplanmıştır. Çapraz geçerleme tanıma oranlarının standart sapması 6.8'dir.

Şekil 2'de her fonem için doğru ve yanlış tanıma oranları ayrı ayrı gösterilmiştir. İlgili fonem için köşegen üzerindeki değer doğru tanıma oranına karşılık gelmektedir. En düşük oranla tanınan üç fonem /ı/ (%63.3), /u/ (%77.5) ve /ü/ (%79.6)'dır. Köşegen üzerinde olmayan değerler ilgili fonem yerine hangi fonemin tanındığını göstermektedir. Birbirine en çok karışan üç fonem çifti sırasıyla /ı/-/ö/, /o/-/u/ ve /u/-/o/ şeklindedir. /ı/ %34.7 olasılıkla /ö/ olarak, /o/ %16.3 olasılıkla /u/ olarak, /u/ ise %14.3 olasılıkla /o/ olarak tanınmıştır.

3.2. Yalıtılmış Sözcük Tanıma

Konuşma terapisinde kullanılan bir başka teşhis yöntemi çeşitli sözcüklerin hastaya söylenilerek yanlış seslendirilen fonemlerin tespit edilmesidir. Tedavi için yine bu fonemlerin geçtiği sözcük listeleri kullanılmaktadır.

Bu bölümde terapide kullanılan çeşitli sözcüklerin otomatik olarak tanınması için deneyler gerçekleştirilmiştir. Deneylerde terapide kullanılabilecek 126 Türkçe sözcükten oluşan bir listeden tanıma gerçekleştirilmiştir. Bu sözcükler 11 konuşmacıdan (dört bayan, yedi bay) birer kez kaydedilmiştir. Konuşmacıların beşinden (iki bayan, üç bay) ayrı bir zamanda ikinci kez kayıt alınmıştır. İkinci kez alınan kayıtlar konuşmacı bağımlı eğitim ile tanıma oranının belirlenmesi için test amaçlı kullanılmıştır.

En iyi SMM parametreleri Bölüm 2.1’de açıklanan geçerleme yöntemiyle belirlenmiştir. Sonuçlara göre her Markov durumu için iki Gauss karışımı kullanılmış, sözcükteki her harf için modele bir durum eklenmiştir. Ayrıca konuşma başlangıcı ve sonu için fazladan iki durum daha kullanılmıştır. Örneğin beş harfli “şapka” sözcüğünün SMM modeli $5+2=7$ durumdan oluşmaktadır. Akustik özneliklere fark parametreleri de eklenmiştir.

Yalıtılmış sözcük tanıma oranı iki farklı durum için belirlenmiştir:

- Konuşmacı bağımsız şekilde eğitim ve test
- Konuşmacı bağımlı şekilde eğitim ve test

Konuşmacı bağımsız eğitim ve test için Bölüm 2.1’de açıklanan çapraz geçerleme yöntemi kullanılmıştır. Her çapraz geçerleme adımında 10 konuşmacı eğitim, bir konuşmacı test için kullanılmıştır. 11 çapraz geçerleme adımının ortalaması olarak %94.2’lik doğru tanıma oranı elde edilmiştir. Tanıma oranlarının standart sapması 7.0’dır.

Konuşmacı bağımlı eğitim ve test için beş konuşmacıdan ayrı bir zamanda toplanan kayıtlar kullanılmıştır. Sistem, önce 11 konuşmacının kayıtlarıyla eğitilmiş, sonra eğitim kümesinde bulunan beş konuşmacının test kayıtlarıyla denenmiştir. Sistemin konuşmacı bağımlı şekilde kullanılmasıyla hata oranı ilk duruma göre %51.7 azalmış ve tanıma oranı %97.2’ye çıkmıştır. Sonuçlar Tablo 3’te özetlenmiştir.

Eğitim ve Test Şekli	Tanıma Oranı
Konuşmacı bağımsız	%94.2
Konuşmacı bağımlı	%97.2

Tablo 3: Konuşma terapisinde kullanılan 126 yalıtılmış Türkçe sözcük için sözcük tanıma oranları.

3.3. Birbirine Çok Yakın Sözcüklerin Tanınması

Bu bölümde SMM tabanlı konuşma tanıma sistemiyle sözcük içindeki fonem hatalarının belirlenmesi için deneyler yapılmıştır. Tablo 1’de verilen fonem çiftleri göz önünde bulundurularak birbirine çok yakın sözcük çiftleri oluşturulmuştur. Her fonem çifti için ilk fonemin başta, ortada ve sonda bulunduğu üç sözcük seçilmiştir (Örneğin /f/-/v/ fonem çifti için “fıdan”, “defter” ve “çarşaf”). İlk fonem yerine ikinci fonemin kullanımıyla sözcüklerin bozuk halleri türetilmiştir. Mümkün olduğunca bozuk sözcüğün de anlamlı olmasına dikkat edilerek bozuk sözcüklerin okunmaları kolaylaştırılmıştır. Sekiz fonem çifti için Tablo 4’te gösterilen 48 sözcük seçilmiştir. Bu sözcükler dokuz konuşmacıdan (dört bayan, beş bay) beşer kez kaydedilmiştir.

Fonem Çiftleri	/f/-/v/	/s/-/z/	/ş/-/j/
Başta	fıdan-vidan	sarı-zarı	şale-jale
Ortada	defter-devter	aşker-azker	ajan-aşan
Sonda	çarşaf-çarşav	kaş-kaz	beş-bej
Fonem Çiftleri	/c/-/ç/	/p/-/b/	/t/-/d/
Başta	çeket-çeket	pasta-basta	tabak-dabak
Ortada	açı-aççı	kapı-kabı	katı-kadı
Sonda	avuç-avuç	dolap-dolab	yakut-yakud
Fonem Çiftleri	/k/-/g/	/r/-/y/	
Başta	kez-gez	raket-yaket	
Ortada	basket-basget	çorap-çoyap	
Sonda	gözlük-gözlüg	bir-biy	

Tablo 4: Birbirine çok yakın sözcükler listesi.

En iyi SMM parametreleri Bölüm 2.1’e benzer şekilde belirlenmiştir. Sonuçlara uygun olarak sözcükteki her harf için bir Markov durumu ve her durum için iki Gauss karışımı kullanılmıştır. Tanıma, her sözcük çifti için eğitilen iki modelin log-olabilirliklerinin hesaplanması ve en yüksek log-olabilirlik veren modelin seçilmesiyle gerçekleştirilmiştir.

İlk aşamada sistemin önceki bölümlere benzer şekilde konuşmacı bağımsız tanıma performansı ölçülmüştür. Ortalama tanıma oranı %80.3’tür. Tanıma oranı terapide kullanılacak bir yöntem için oldukça düşük olduğundan konuşmacı bağımlı bir yöntem

tasarlanmıştır. Böylece her hasta için kolay bir şekilde uyarlanabilecek, tanıma oranı yüksek bir yöntem elde edilmiştir. Konuşmacılardan toplanan kayıtlar eğitim ve deney kümesi olarak ikiye ayrılmıştır. Her sözcük için üç kayıt eğitimde, iki kayıt deneylerde kullanılmıştır. Her konuşmacı için eğitim gerçekleştirilerek deney kümesi üzerinde tanıma oranı belirlenmiş ve ortalama tanıma oranı tüm konuşmacıların ortalaması alınarak hesaplanmıştır.

Tablo 5’te tanıma oranları fonem çiftleri için ayrı ayrı gösterilmiştir. Her fonem çiftinde ilgili fonemlerin başta, ortada ve sonda olduğu durumlardaki tanıma oranları tablonun son sütununda hesaplanmıştır. Ayrıca farklı fonem çiftlerinin sözcüğün farklı yerlerinde bulunmaları durumları için ayrı ayrı ortalama doğru tanıma oranları tablonun son satırında verilmiştir.

Fonem Çiftleri	Başta	Ortada	Sonda	Ortalama
/p/-/b/	97.5	75.0	82.5	85.0
/v/-/f/	87.5	97.5	82.5	89.2
/s/-/z/	87.5	87.5	82.5	85.8
/c/-/ç/	67.5	95.0	87.5	83.3
/t/-/d/	97.5	95.0	100.0	97.5
/k/-/g/	72.5	75.0	77.5	75.0
/j/-/ş/	92.5	85.0	92.5	90.0
/r/-/y/	95.0	100.0	100.0	98.3
Ortalama	87.2	88.8	88.1	88.0

Tablo 5: Birbirine çok yakın sözcüklerin tanınma oranları (Konuşmacı bağımlı)

Tablo 5’e göre konuşma tanımada yanlış kullanımları en az hataya yol açan fonem çifti %98.3’lük doğru tanınma oranıyla /r/-/y/’dir. Bu çifti %97.5’lik oranla /t/-/d/ izlemektedir. Konuşma tanıma için en fazla sorun yaratan ve sistemin yanlış karar vermesine neden olan fonem çiftleri /k/-/g/ (%75.0) ve /c/-/ç/ (%83.3) olarak belirlenmiştir. En düşük tanıma oranları /c/-/ç/ ve /k/-/g/ çiftleri sözcük başında olduklarında elde edilmiştir (%67.5 ve %72.5). Baştaki fonemin yanlış söylendiği durumlarda ortalama tanıma oranı %87.2, ortadaki yanlış söylendiğinde %88.8, sondakinde ise %88.1’dir.

4. Bulgular

Farklı testler için elde edilen sonuçlar Tablo 6’da özetlenmiştir. Bazı fonemlerin tanınmasında diğerlerine göre daha düşük tanıma oranları elde edilmiştir. En düşük oranın

elde edildiđi /ı/-/ö/ çifti insanlar tarafından da en zor ayırt edilen fonemlerdir. Sesliler için formant üçgeni incelendiđinde de bu iki fonemin birbirine çok yakın oldukları görölmektedir. Konuşma tanımada birbirine karıştırılan fonemlerin insanlar tarafından da birbirinden daha zor ayırt edilebilen fonemler olması, fonem tanıma hatalarının terapi açısından ciddi bir sorun oluşturmayacağını göstermektedir.

Görev	Veritabanı	Konuşmacı Sayısı	Eđitim ve Test Şekli	Tanım Oranı
Fonem Tanıma	14 fonem	12	Konuşmacı Bağımsız	%84.9
Sözcük Tanıma	126 sözcük	11	Konuşmacı Bağımsız	%94.2
Sözcük Tanıma	126 sözcük	11	Konuşmacı Bağımlı	%97.2
Sözcük Tanıma	24 sözcük çifti (*)	10	Konuşmacı Bağımlı	%88.0

Tablo 6: Konuşma terapisine yönelik konuşma tanıma yöntemleri için tanıma oranları.

(*) Birbirine çok yakın sözcükler.

Birbirine çok yakın sözcüklerin tanınması konuşma tanıma sistemlerinde çözümü en zor problemlerden biridir. Konuşma terapisinde hastanın sesine kolayca uyarlanabilecek bir sistemin kullanılması önemlidir. Bu amaçla Bölüm 2’de ayrıntılı olarak açıklanan konuşma tanıma yöntemlerini içeren bir terapi yazılımı geliştirilmiştir. Yazılım iki kısımdan oluşmaktadır. Terapist arayüzünde hastadan alınan fonem, sözcük ya da cümle kayıtları SMM eğitimi için kullanılmaktadır. Ayrıca hastaya uygulanacak test tasarlanmaktadır. Hastanın kullanacağı arayüzde test uygulanmakta ve sonuçlar verilmektedir. Böylece terapistin olmadığı durumlarda da hasta kendi kendini değerlendirme imkanına sahip olmaktadır.

Konuşma terapisinde kullanılan sözcükler, küçük sözcük dađarcıklı SMM tabanlı konuşma tanımayla yüksek oranda tanınabilmektedir. Elde edilen tanıma oranları konuşma tanımanın çeşitli terapi egzersizlerinde kullanılabileceğini göstermektedir.

5. Tartışma

Bu çalışmada konuşma terapisine yönelik konuşma tanıma yöntemleri incelenmiştir. SMM tabanlı yöntemler kullanılarak fonem tanımda %84.9, terapide kullanılan bir sözcük listesi üzerinde %94.2 ve birbirine çok yakın sözcüklerde %88.0 doğru tanıma oranı elde edilmiştir. Bu oranlar, konuşma tanınmanın terapide yardımcı bir araç olarak kullanılabilmesini göstermektedir. Bu amaçla terapistler ve hastalarca kullanılacak bir yazılım geliştirilmektedir. Bu yazılıma üç-boyutlu yüz senteziyle doğru seslendirmeyi öğretecek görsel öğelerin de eklenmesi planlanmaktadır.

Doğru seslendirmede konuşma işaretinin fonetik içeriği kadar uzun dönemli akustik öznelikler de önem taşımaktadır. Örneğin ses perdesi eğrisinin, enerji değerlerinin ve fonem sürelerinin referans cümlelerle karşılaştırılması ve nicel puanlama yapılması terapide kullanılacak yöntemler olup halen üzerinde araştırma yaptığımız konulardır.

6. Kaynakça

- [1] Russel, M., Brown, C., Skilling, A., Series, R., Wallace, J., Bonham, B. ve Barker, P., “Applications of Automatic Speech Recognition to Speech and Language Development in Young Children”, *Proceedings of ICSLP 1996*, Philadelphia, PA, ABD, 1996, sf. 176-179.
- [2] Neumeyer, L. F. H., Weintraub, M. ve Price, P., “Pronunciation Scoring of Foreign Language Student Speech”, *Proceedings of ICSLP 1996*, Philadelphia, PA, ABD, 1996, sf. 1457-1460.
- [3] Neumeyer, L. F. H., Digalakis, V. ve Weintraub, M., “Automatic scoring of pronunciation quality”, *Speech Communication*, 30 (2-3), sf. 83-93, 2000.
- [4] Bunnell, H. T., Yarrington, D. M., ve Polikoff, J. B., “STAR: Articulation training for young children”, *Proceedings of ICSLP 2000*, c. 4, sf. 85-88.
- [5] Witt, S. M. ve Young, S. J., “Phone-level pronunciation scoring and assessment for interactive language learning”, *Speech Communication*, 30 (2-3), sf. 95-108, 2000.
- [6] Chanwoo K. C. ve Sung, W., “Implementation of an intonational quality assessment system”, *Proceedings of ICSLP 2002*, Denver, Colorado, ABD, c. 2, sf. 1225-1228.

- [7] Rabiner, L. R. ve Juang, B.-H., *Fundamentals of Speech Recognition*, Prentice-Hall, Inc., New Jersey, NJ, 1993.
- [8] Girgin, C., M., *Türkçe konuşan doğal işitsel sözel yöntemle eğitim gören işitme engelli kız çocukların konuşma anlaşılabilirliği ile süre ve perde özellikleri ilişkisi*, Anadolu Üniversitesi Yayınları, No. 1167., Eğitim Fakültesi Yayınları, No. 63., 1999.