



## Sayısal filtrelerin akustik parametreler, cinsiyet yaş ve duygu durumu üzerindeki etkileri

### The effects of digital filters on acoustic parameters, gender, age and emotion

Turgut ÖZSEVEN<sup>1\*</sup>, Muharrem DÜĞENCİ<sup>2</sup>

<sup>1</sup>Bilgisayar Teknolojileri Bölümü, Turhal Meslek Yüksekokulu, Gaziosmanpaşa Üniversitesi, Tokat, Türkiye.  
turgut.ozseven@gop.edu.tr

<sup>2</sup>Endüstri Mühendisliği Bölümü, Mühendislik Fakültesi, Karabük Üniversitesi, Karabük, Türkiye.  
mdugenci@gmail.com

Geliş Tarihi/Received: 29.11.2015, Kabul Tarihi/Accepted: 26.01.2016  
\* Yazışılan yazar/Corresponding author

doi: 10.5505/pajes.2016.00922  
Araştırma Makalesi/Research Article

#### Öz

Akustik ses analizi makine, insan ve su altı gibi ortamlarda ses sinyalinin özellik çıkarımı yapılarak süreçlerde karar verme amacıyla kullanılmaktadır. Ses analizinde çoğunlukla perde, formantlar, jitter ve shimmer akustik parametreleri kullanılmaktadır. Ortamlardan elde edilen çeşitli ses sinyallerinin tüm frekans bantları her sistem için geçerli olmamaktadır. İnsan seslerindeki duyguların tespiti için düşük frekanslı bileşenler kullanılmaktadır. Ayrıca, ses sinyali gerek ortamdan gerekse de kayıt şeklinden dolayı istenmeyen gürültülere sahip olabilmektedir. Sayısal filtreler yardımıyla ses sinyalinin istenilen frekans bantları elde edilebilme, gürültü giderme, yankı azaltma ve ses kalitesi artırma gibi işlemler gerçekleştirilmektedir. Bu çalışmanın amacı, alçak geçiren, yüksek geçiren ve bant geçiren filtrelerin yaş, cinsiyet ve duygu durumuna bağlı olarak akustik parametrelerdeki değişiminin araştırılmasıdır.

**Anahtar kelimeler:** Jitter, Shimmer, Perde, Sayısal filtre, Duygu

#### 1 Giriş

Akustik ses ile ilgilenen fizik dalı olup psikoakustik ise sesin psikolojik cevaplarıyla ilgilenir. Konuşma akustikleri hem fiziksel hem de psikolojik çalışmaları içerir ve sesin akustik özelliklerini araştırır [1]. Akustik ses analizi, ses bozuklukları değerlendirilirken kullanılan nesnel değerlendirme yöntemlerinden biridir [2]. Akustik analiz sadece insan seslerinde değil makine, araç ve su altı gibi çeşitli ortamlardaki sesler için de kullanılabilir. Bunun için kullanılan temel parametreler perde, jitter ve shimmer'dir.

Perde, ses kıvrımlarının saniyede oluşan açılma-kapanma döngü sayısı olup sesin kalınlık ve inceliğini belirtir. Jitter ise periyotlar arası değişikliği gösteren parametredir. Diğer taraftan Shimmer, genlik pikleri arasındaki periyodik varyasyondur. Formant ses yolundaki rezonanstır. Teorikte sonsuz sayıda format vardır ama pratikte yalnızca ilk 3 formant önemli bilgiler içerir. Formantlar F1, F2 ve F3 gibi formant sayısı ile tanımlanır [1].

Sayısal sinyal işleme (DSP-Digital Signal Processing) bilim ve mühendislik açısından günümüzde güçlü bir teknolojidir ve sinyalin sayısal yöntemlerle çeşitli amaçlar için işlenmesi ile ilgilenir [3]. Sinyaller üzerinde çalışırken sinyalin güçlendirilmesi, bazı frekansların bastırılması veya belirli aralıkta frekansların geçirilmesi gibi işlemlerin yapılmasına ihtiyaç duyulabilir. Filtreler bu şekilde sinyaller üzerinde değişiklikler yapmaya olanak sağlar [4]. Alçak geçiren (LP-low pass) filtreler, sinyalin düşük frekanslarının geçişine izin

#### Abstract

Acoustic sound analysis is used for decision making process performed feature extraction from sound signal in environments such as machine, human and underwater. The pitch, formants, jitter and shimmer of acoustic parameters are mostly used in sound analysis. All frequency bands of various sound signals obtained from environments is not available for every system. The low frequency component is used for emotion detection in the human voice. In addition, sound signal must be have unwanted noise due to both form of records and environment. The desired frequency band of sound signal can be obtained, processes such as noise reduction, echo reduction and sound quality improving are realized with the help of digital filters. The purpose of this study is to investigate the effects of low-pass, high-pass and band-pass filters on acoustics parameters depending on age, gender and emotion.

**Keywords:** Jitter, Shimmer, Pitch, Digital filter, Emotion

verirken yüksek frekanslar atılır. Yüksek geçiren (HP-high pass) filtreler, LP filtrelerin tersine yüksek frekansların geçişine izin verirken düşük frekansların geçişi engellenir. Bant geçiren (BP-band pass) filtreler, belirlenen alt ve üst kesim noktaları arasında kalan frekans bandının geçişine izin verir [5].

Tao, Kang ve Li (2006) akustik parametrelerdeki rastgele değişimi kontrol etmek için beyaz gürültü ve LP filtre kullanmıştır [6].

He ve diğ. (2009) konuşmadan karakteristik özellikleri çıkartmak için iki yeni yaklaşım sunmuşlardır. Kullanılan yaklaşımlardan birisinde 12 log-Gabor filtreleri kullanmıştır. Sonuçta %40-%81 arasında sınıflandırma doğruluğu elde etmişlerdir [7].

MacCallum ve diğ. (2011) yaptıkları çalışmada LP filtrenin kesim frekansının akustik analizdeki etkilerini araştırmış ve en uygun kesim frekansının 300Hz ve üstü olduğu sonucuna ulaşmıştır [8].

Chen ve diğ. (2014) yaptıkları çalışmada 5dB sinyal gürültü oranı seviyesinden düşük 6 farklı gürültü kullanarak sınıflandırma doğruluğunu artırmak için yeni bir özellik kümesi elde etmiştir. Sonuçlar elde edilen özellik kümesinin en iyi sınıflandırma doğruluğunu sağladığını göstermiştir [9].

Siegert ve diğ. (2014) konuşmadaki görünmeyen spektral değişimleri temizlemek ve çevresel gürültüleri elimine etmek için bağıl spektral (RASTA-RelAtive SpecTrAl) filtre ve cepstral

ortalama çıkarım (CMS-Cepstral Mean Subtraction) yöntemlerini kullanmıştır [10].

Ludeña-Choez ve Gallardo-Antolín (2015) yaptıkları çalışmada akustik olay sınıflandırıcı için iki yeni yaklaşım sunmuştur. İlkinde farklı akustik olayların spektral karakteristikleri incelenmiş ve ikincisinde HP filtre tabanlı yeni bir parametre sunulmuştur. Sonuçlar HP filtrelenmiş ses sinyalinin sistem için faydalı olduğu gerçeğini desteklemektedir [11].

Bu çalışmada LP, HP ve BP filtrelerin yaş, cinsiyet ve duyu durumuna bağlı olarak perde, F1, F2, F3, jitter ve shimmer akustik parametrelerinin değişimi araştırılmıştır.

Bu çalışmanın ikinci bölümünde materyal ve yöntem, üçüncü bölümünde uygulama ve dördüncü bölümde sonuçlar verilmiştir.

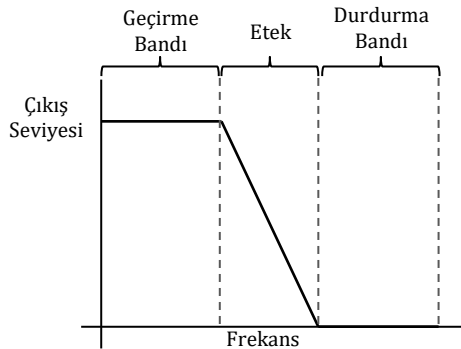
## 2 Materyal ve yöntem

Çalışmada kullanılan konuşma verileri, Berlin Database of Emotional Speech veri tabanından temin edilen 201 adet konuşma kaydından oluşmaktadır [12]. Ses kayıtları 16 kHz örnekleme frekansına sahiptir. Kullanılan veriye ait çeşitli bilgiler Tablo 1'de verilmiştir.

Tablo 1: Veriye ait dağılım.

Demografik Özellik	Adet
Öfke	50
Mutluluk	25
Duygu Durumu	27
İğrenme	21
Anksiyete/Korku	29
Üzüntü	49
Yaş	146
Yaş 31-35	55
Yaş 20-30	92
Cinsiyet	109
Erkek	
Bayan	

Filtreleme, bir sinyalin seçilen süzgeç karakteristiğine göre istenen frekans bileşenlerine ayrılması ve sinyalin genel yapısının geliştirilmesi olarak tanımlanır [13]. Filtre terminolojisinde, sinyalin geçmesine izin verilen sinyal frekans aralığına geçirme bandı (passband), geçmesi engellenen frekans aralığına durdurma bandı (stopband) denir. Geçirme bandı ile durdurma bandı arasında etek yanıtı olarak ifade edilen bir frekans aralığı mevcuttur [14]. Bazı kaynaklarda geçirme bandı ifadesi iletim bandı ve etek ifadesi geçirme bandı olarak geçmektedir [15]. Ayrıca filtrelerde kullanılan kesim frekansı filtrenin sinyal geçişine izin verilen veya durdurulan sınırların belirlenmesini ifade eder. Bir filtreye ait frekans alanı özellikleri Şekil 1'de verilmiştir.

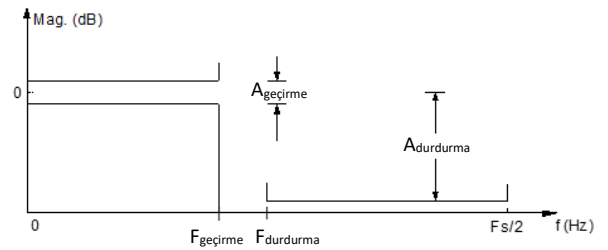


Şekil 1: LP bir filtrenin frekans alanı özellikleri.

Geçirme ve durdurma bandında çıkış seviyesi sabit olmayacak ve kabul edilebilir seviyede dalgalanma içerecektir. Bu dalgalanmaları yok etmek için pencereleme kullanılmaktadır. Pencereleme spektral sızmayı azaltmak için kullanılmaktadır. Konuşma işleme çalışmalarının genelinde Hamming pencereleme kullanıldığı için akustik özelliklerin çıkartımında Hamming pencereleme kullanılmış olup pencere genişliği olarak 30 ms seçilmiştir.

Filtreler frekans yanıtlarına göre LP, HP, BP ve bant durdurucu (BS-band-stop) olmak üzere 4'e ayrılır [14]. LP filtreler konuşmanın içeriği ile çok fazla ilgilenmeyip duygusal içeriği korumaya yönelik çalışmalarda kullanılmıştır [16].

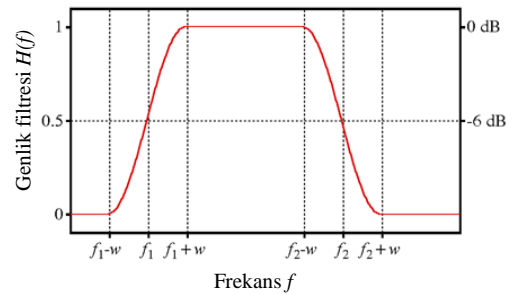
Bir filtrenin genlik cevabı özellikleri aşağıdaki Şekil 2'de gösterildiği gibidir.



$F_{geçirme}$ : Geçirme bandı frekansı,  $F_{durdurma}$ : Durdurma bandı frekansı,  $F_s$ : Örnekleme Frekansı,  $A_{geçirme}$ : Geçirme bandı dalgalanması,  $A_{durdurma}$ : Durdurma bandı zayıflaması.

Şekil 2: LP filtre için genlik cevabı özellikleri.

Konuşma sinyallerine filtrelerin uygulanması ve akustik özelliklerin çıkartımı için Praat yazılımı kullanılmıştır [17]. Filtreleme için Praat üzerinde pass hann band filter kullanılmış olup filtreleme sonrası perde, formant frekanslarının ortalama değerleri, jitter ve yüzde shimmer değerleri hesaplanmıştır. Praat pass hann band filtrenin frekans tepkisi Şekil 3'te verilmiştir.



Şekil 3: Pass hann band filtrenin frekans tepkisi.

Şekilde kullanılan  $f_1$  geçirme bandının alt kenarı,  $f_2$  geçirme bandının üst kenarı ve  $w$  geçirme ve durdurma arasındaki bölgenin genişliğidir.  $f_1$  değeri sıfır alındığında LP,  $f_2$  sıfır alındığında HP filtre elde edilmektedir. Örneğin,  $f_1$  değeri 20Hz ve  $f_2$  değeri 300Hz belirlendiğinde elde edilen filtre BP filtre olup 100Hz-300Hz arasındaki frekansların geçişine izin verecek diğer frekansları bastıracaktır.

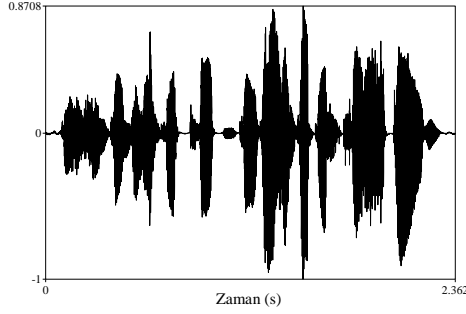
MacCallum ve diğ. (2011) yaptıkları çalışmada LP filtrenin en uygun kesim frekansının 300 Hz ve üstü olduğu sonucuna ulaşmıştır. Perde değeri erişkin erkekler ve kadınlarda sırası ile ortalama 100-150 Hz ve 150-250 Hz arasındadır [8]. Bu bilgiler doğrultusunda çalışmada kullanılan filtrelere ait özellikler Tablo 2'de verilmiştir.

Tablo 2: Kullanılan filtrelere ait özellikler.

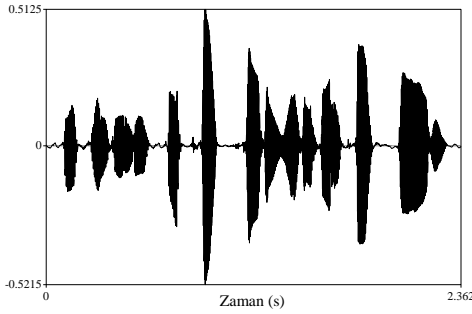
Filtre Türü	$f_1$	$f_2$	$w$
LP	0 Hz	300 Hz	20 Hz
HP	100 Hz	0 Hz	20 Hz
BP	100 Hz	300 Hz	20 Hz

### 3 Uygulama

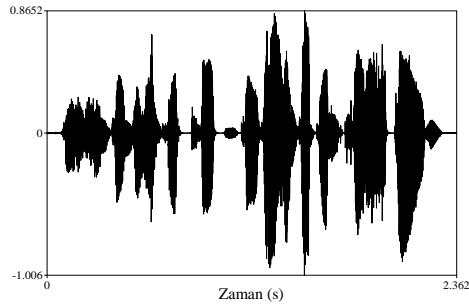
Çalışmada kullanılan 201 adet ses kaydı için sayısal filtrelerin uygulanması Praat yazılımı ile gerçekleştirilmiş olup örnek bir kayıt için orijinal ve filtre uygulanmış sinyal şekilleri Şekil 4, Şekil 5, Şekil 6 ve Şekil 7'de verilmiştir.



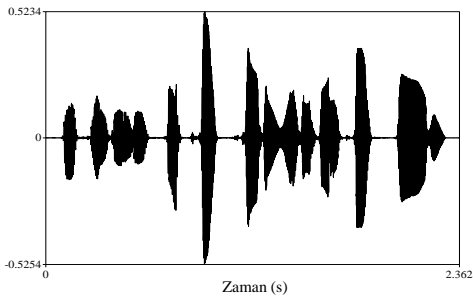
Şekil 4: Örnek bir ses kaydı.



Şekil 5: LP filtre uygulanmış ses kaydı.



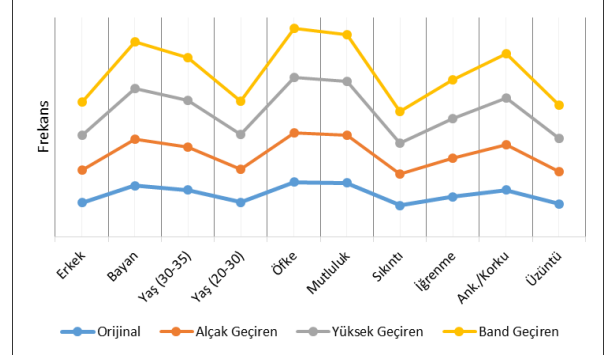
Şekil 6: HP filtre uygulanmış ses kaydı.



Şekil 7: BP filtre uygulanmış ses kaydı.

### 3.1 Filtrelemenin perde değeri üzerindeki etkileri

Üç farklı filtrenin uygulanması sonucu cinsiyet, yaş ve duygu durumuna göre ortalama perde değerinin değişimi Şekil 8'de verilmiştir.

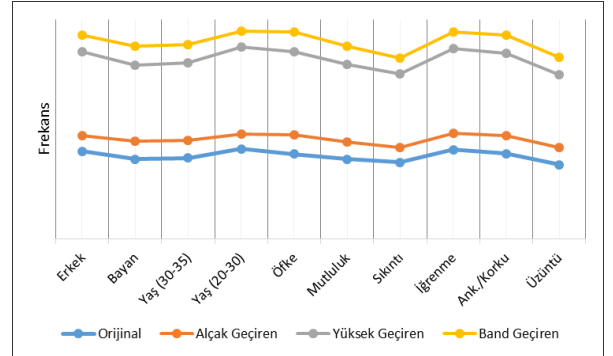


Şekil 8: Filtrelerin ortalama perde üzerindeki etkileri.

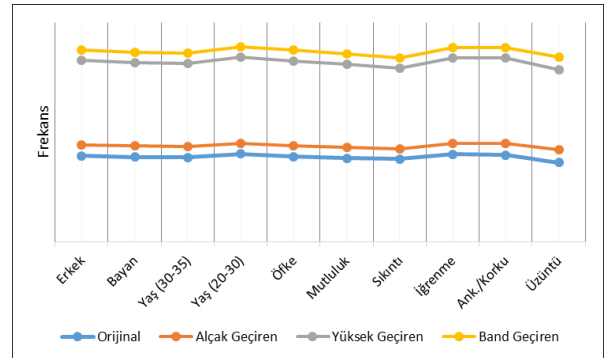
Şekil 8 incelendiğinde bayanlarda LP ve BP ortalama pitch değerini düşürürken erkeklerde filtre önemli bir değişime neden olmamıştır. Üst yaşlarda LP ve BP ortalama perdeyi düşürmüştür. Öfke ve mutluluk duygularında LP ve BP filtreler ortalama perdeyi düşürürken diğer duygularda filtreler ciddi bir değişime neden olmamıştır. Genel anlamda HP filtreler yaş, cinsiyet ve duygu yönünden değişime neden olmamıştır. Ayrıca ortalama perde değeri cinsiyet, yaş ve duygu durumundan etkilenmektedir.

### 3.2 Filtrelemenin formant frekansları üzerindeki etkileri

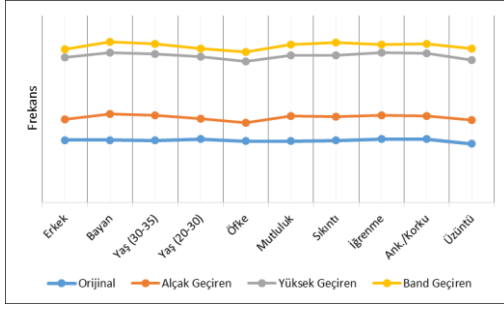
Üç farklı filtrenin uygulanması sonucu cinsiyet, yaş ve duygu durumuna göre formant frekanslarındaki değişimler Şekil 9, Şekil 10 ve Şekil 11'de verilmiştir.



Şekil 9: Filtrelerin F1 üzerindeki etkileri.



Şekil 10: Filtrelerin F2 üzerindeki etkileri.



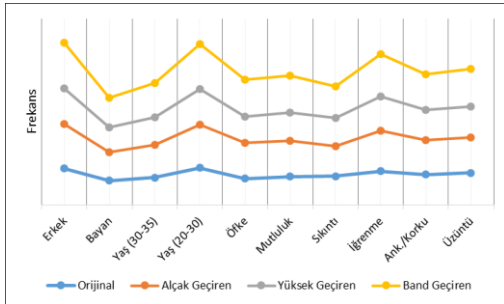
Şekil 11: Filtrelerin F3 üzerindeki etkileri.

Formantlarla ilgili sonuçlar incelendiğinde formantların HP filtreden etkilenmediği gözlemlenmiştir. F1, F2 ve F3 için LP ve BP filtreler formant frekanslarını ciddi şekilde düşürmektedir. F1 ve F2 cinsiyet, yaş ve duygu durumuna göre LP ve BP filtreden aynı oranda etkilenmiştir. F3 ise LP ve BP filtreden farklı oranda etkilenmiştir. Genel olarak F1 tüm özelliklerde değişim gösterirken, F2 ve F3 çok fazla etkilenmez. Ayrıca sıkıntı-üzüntü-mutluluk ve öfke-iğrenme-anksiyete/korku duyguları kendi aralarında yakın F1 değerine sahiptir. Elde edilen sonuçlara göre F1 diğer formantlara göre daha belirleyici olup çalışmaların büyük çoğunluğunda da analiz için F1 kullanılmaktadır.

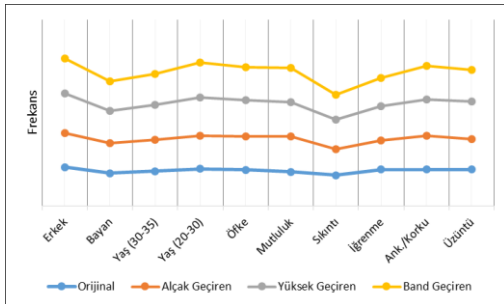
### 3.3 Filtrelemenin Jitter ve Shimmer üzerindeki etkileri

Jitter parametresi üzerinde yapılan analizlerde tüm filtreler ve orijinal sinyal için jitter değeri %1 ve üzerinde çıkmıştır. Jitter'in normal değeri %1'in altında olup yüksek çıkması durumunda ses hastalığı tanısı konulmaktadır [18]. Kullanılan ses kayıtları duygusal durumlar içerdiği için jitter değerinin yüksek çıkması normal olup analiz sonuçlarının doğruluğunda kriter olarak düşünülebilir. Aynı durum shimmer için de geçerli olup shimmer'in normal değeri %3'ün altındadır. Analiz sonuçlarında da shimmer değeri %8 ve üzerinde bulunmuştur.

Şekil 12 ve Şekil 13'te sayısal filtrelerin jitter ve shimmer üzerindeki etkileri verilmiştir.



Şekil 12: Filtrelerin jitter üzerindeki etkileri.



Şekil 13: Filtrelerin shimmer üzerindeki etkileri.

Jitter ve shimmer ile ilgili sonuçlar incelendiğinde jitter'in yaşa ve cinsiyete göre ciddi farklılıklar içerdiği görülmektedir. Aynı durum shimmer için daha düşük orandadır. Jitter duygu durumlarına göre ciddi bir farklılık içermezken shimmer sıkıntı duygusunda diğerlerine göre farklılık içermektedir. Jitter parametresi tüm özellikler için HP filtreden önemli oranda etkilenmezken LP ve BP filtrelerden etkilenmektedir. Aynı durum shimmer için de geçerlidir.

## 4 Sonuçlar

Sonuçlar genel anlamda incelendiğinde HP filtreler yaş, cinsiyet ve duygu durumuna bağlı akustik parametreler üzerinde değişikliğe neden olmamıştır. LP ve BP filtrelerin akustik parametreler üzerinde oluşturduğu değişim aynı oranda olup bu durum LP ve BP filtreler için belirlenen geçiş frekans aralığından kaynaklanmaktadır. Ayrıca LP ve BP filtreler öfke ve mutluluk duygularına bağlı perde değerinde ciddi değişime neden olurken diğer duygularda gözlenen değişim miktarı çok düşüktür. Tüm formant frekansları LP ve BP filtreden etkilenmiştir. Formant frekanslarından öncelikli olarak F1 ve sonrasında F2 duygularda belirleyici olarak kullanılabilir. Bir çok çalışmada da formant frekanslarından sadece F1 analizlere dahil edilmiştir. Jitter ve shimmer konuşmanın gırtlakta oluşumunu temel aldığı için yaş ve cinsiyete göre değişiklikler içermesi de beklenen durumdur.

## 5 Teşekkür

Pamukkale Üniversitesi Mühendislik Bilimleri Dergisine katkı sağlayan tüm yazar ve hakemlere teşekkür ederiz.

## 6 Kaynaklar

- [1] Rezaei N, Abolfazl S. "An introduction to speech sciences (acoustic analysis of speech)". *Iranian Rehabilitation Journal*, 4(4), 5-14, 2006.
- [2] Özbal EA. Septum Devisyonlu Hastaların Septoplast Operasyonu Öncesi ve Sonrası Akustik Ses Analizi ile Değerlendirilmesi. Uzmanlık Tezi, Şişli Etfal Hastanesi, İstanbul, Türkiye, 2008.
- [3] Smith SW. *The Scientist and Engineer's Guide to Digital Signal Processing*. 2<sup>nd</sup> ed. USA, California Technical Publishing, 1999.
- [4] Weeks M. *Digital Signal Processing Using MATLAB and Wavelets*. Infinity Science. Massachusetts, USA, Press LLC, 2007.
- [5] Winder S. *Analog and Digital Filter Design*. 2<sup>nd</sup> ed. USA, Newnes, 2002.
- [6] Tao J, Kang Y, Li A. "Prosody conversion from neutral speech to emotional speech". *Audio, Speech and Language Processing, IEEE Transactions on*, 14(4), 1145-1154, 2006.
- [7] He L, Lech M, Maddage N, Allen N. "Stress and emotion recognition using log-Gabor filter analysis of speech spectrograms". *3<sup>rd</sup> International Conference on Affective Computing and Intelligent Interaction (ACII2009)*, USA, 10-12 September 2009.
- [8] MacCallum JK, Olszewski AE, Zhang Y, Jiang JJ. "Effects of low-pass filtering on acoustic analysis of voice". *Journal of Voice*, 25(1), 15-20, 2011.
- [9] Chen J, Wang Y, Wang D. "A feature study for classification-based speech separation at low signal-to-noise ratios". *IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP)*, 22(12), 1993-2002, 2014.

- [10] Siegert I, Philippou-Hübner D, Hartmann K, Böck R, Wendemuth A. "Investigation of speaker group-dependent modelling for recognition of affective states from speech". *Cognitive Computation*, 6(4), 892-913, 2014.
- [11] Ludeña-Choez J, Gallardo-Antolín A. "Feature extraction based on the high-pass filtering of audio signals for Acoustic Event Classification". *Computer Speech and Language*, 30(1), 32-42, 2015.
- [12] Burkhardt F, Paeschke A, Rolfes M, Sendlmeier WF, Weiss B. "A database of German emotional speech". In *Interspeech*, 25(1), 1517-1520, 2005.
- [13] Ertürk S. *Sayısal İşaret İşleme*. İstanbul, Türkiye, Birsen Yayınevi, 2004.
- [14] Winder S. *Analog and Digital Filter Design*. Woburn, Netherlands, Newness Press, 2002.
- [15] Proakis JG, Manolakis DG. *Digital Signal Processing Principles Algorithms and Applications*. 3<sup>rd</sup> ed. USA, Prentice-Hall, 1996.
- [16] Laukka P, Linnman C, Åhs F, Pissiota A, Örjan F, Faria V. "In a nervous voice: Acoustic analysis and perception of anxiety in social phobics speech". *Journal of Nonverbal Behavior*, 32(4), 195-214, 2008.
- [17] Boersma P. "Praat, a system for doing phonetics by computer". *Glott International*, 5(9-10), 341-345, 2001.
- [18] Sarıca S. Ses Analizinde Kullanılan Akustik Parametreler. Uzmanlık Tezi, Kahramanmaraş Sütçü İmam Üniversitesi Tıp Fakültesi, Kahramanmaraş, Türkiye, 2012.