Türkçe'de Duygu Çözümlemesi

H. Mesut Meral, Hazım K. Ekenel, A. Sumru Özsoy
Boğaziçi Üniversitesi

Özetçe

Bu çalışmada temel sıklık eğrilerinden yararlanarak konuşma sırasında aktarılan dört temel duygunun –kızgınlık, sevinç, üzüntü, duygusuz okuma- çözümlenmesi hedeflenmektedir. Yapılan deneysel çalışmada, ses kaydı için belirlenen dört farklı duyguyla da ifade edilebilecek nitelikteki tümceler, amatör oyuncular tarafından okunmuştur. Alınan ses kayıtlarının temel sıklık eğrileri çıkarılıp, her duygunun temel sıklık eğrilerine yaptığı etki gözlemlenmiştir.

1. Giriş

Konuşma sırasında aktarılan duyguların incelenmesi son yıllarda sesbilim, ruhbilim ve insan makine etkileşimi gibi çalışma alanlarının ilgisini çeken karmaşık bir konudur. Sadece sözcük dizilerinin algılanmasından başka, konuşmada ortaya çıkan ezgi, duygu, perde, vurgu gibi özelliklerin algılanması, iletişim üzerine yapılan bilimsel çalışmalar için son derece önemli hale gelmiş ve son yıllarda bu konular üzerine yapılan çalışmalar artmıştır.

Bu alanlardaki çalışmalar daha çok konuşmadaki duygunun temel sıklık eğrilerine yaptığı etki üzerinde yoğunlaşmakta ve sıradan insanlar, oyuncular tarafından belirli duyguları ifade edecek şekilde okunan tümcelerin ya da filmlerden kesilen duygu içerikli bölümlerin oluşturduğu veri tabanlarının akustik açıdan incelenmesine dayanmaktadır. Bosch (2003), ses sinyalinden duygu tanımlaması yapılabilme olasılığını tartıştığı makalesinde, ses sinyalindeki akustik özelliklerden konuşmacının duygusal durumuna ilişkin bulgular çıkartılmasına odaklanmıştır. Çalışmada, ses sinyalindeki akustik özelliklerin konuşmada aktarılan duyguyu tanımlamada kullanılabileceği sonucuna ulaşılmıştır. İsveççe üzerine yapılan bir çalışmada (Abelin ve Allwood (2000)), insanların, bürünsel örüntülerden değişik duyguları algılamalarındaki tutarlılık, duyguların algılanmasında dinleyicinin sosyal ve kültürel durumunun etkileri ve duygunun tanımlanmasında anlamsal ve akustik özellikler arasındaki ilişki incelenmiştir. Anadili İsveççe olan bir konuşmacının değişik duygularla okuduğu söz öbekleri, anadili İsveççe ve İsveççe'den farklı olan kişilere dinletilmiş, bu kişilerden duyguları tanımlamaları istenmiştir. Sonuçta, kızgınlık, korku, şaşırma, üzüntü gibi duyguların, sevinç, utanma, iğrenme ve egemen olma gibi duygulara oranla daha kolay algılanabildiği, benzer anlamsal içeriğe sahip olan korku, utanma gibi duyguların ses işaretlerinin aynı doğrultuda akustik özelliklere sahip olduğu ortaya çıkmıştır. Mozziconacci ve Hermes (1999), perde düzeyi, perde aralığı gibi ezgi örüntülerinin konuşmada aktarılan ve

dinleyiciler tarafından algılanılan duyguları nasıl etkilediğini araştırmıştır. Üç kişinin, yedi duyguyu (sevinç, sıkıntı, kızgınlık, üzüntü, korku, içerleme ve duygusuz okuma) ifade edecek şekilde okuduğu tümceler ezgileme konusunda uzman iki kişiye dinletilmiş ve bu kişilerden konuşmadaki temel sıklık değişimlerini tanımlamaları istenmiştir. Sonuçlar, tümcelerde ortaya çıkan temel sıklık seviyesi/değer aralığı birleşimleri, ezgi örüntüleri ve dinleyici tanımlamaları açısından incelenmiş ve gerek temel sıklık seviyesi/değer aralığı birleşimleri ile dinleyici tanımlamaları, gerekse ezgileme biçimleri ile dinleyici tanımlamaları arasında büyük oranda etkileşimler olduğu ortaya çıkmıştır. Çıkan sonuçlardan hareketle, tümcenin ezgileme biçiminin duyguyu belirlemede önemli bir etken olduğu belirtilmiştir.

Bu çalışmada, ses sinyalinden yararlanarak, Türkçe'de konuşma sırasında aktarılan duyguların temel sıklık eğrilerine yaptığı etkiler gözlemlenmiştir. İncelemede temel olarak ulaşılmaya çalışılan hedef, Türkçe'de duygusal öğeler ile temel sıklık eğrileri arasında birtakım ilintiler olup olmadığının araştırılması, varsa bu ilintilerin çözümlenmesidir.

2. Yöntem

Bu çalışmada temel sıklık eğrilerinin değişimleri incelenerek, konuşma sırasında aktarılan dört temel duygu –kızgınlık, sevinç, üzüntü, duygusuz okuma- çözümlenmiştir. Göz önüne alınan sıklık değişimleri temel olarak genlikteki ve eğimdeki değişimlerdir. Temel sıklık eğrilerinin incelenme nedeni, daha önceki çalışmaların da gösterdiği gibi, konuşmadaki duygu içeriğinin bu eğrilerdeki değişimler üzerinde belirgin bir etkisi olmasıdır.

Çalışmada kullanılacak veri tabanı dört farklı duyguyla ifade edilebilecek nitelikteki tümcelerin amatör oyuncular tarafından okunması ile hazırlanmıştır. Kayıt için amatör oyuncuların tercih edilme sebebi, diğer diller için yapılan duygu çözümlemesi çalışmalarında da belirtildiği gibi sıradan okuyucuların duygu aktarımı başarılarının düşük olmasıdır. Bu tür çalışmalarda veri tabanı oluşturmanın bir alternatif yolu da filmlerden duygu içerikli bölümlerin kesilmesidir. Ancak bu yöntem çalışmada kullanılmak istenen tümce yapılarının seçilmesinde kısıtlayıcı etkiler göstermektedir. Çalışmada kullanılan veritabanını oluşturan tümceler aşağıdadır.

Dışarıda kar yağıyor.
Sınavdan yetmiş aldım.
Hoca bana yetmiş verdi.
Galatasaray maçı iki sıfır kazandı.
Beni çok şaşırttın.
Telefonum çalıyor.

Kurs yarın bitiyor.
Yarın kar yağacakmış.
Dersler iki hafta ertelendi.
Kapı açık kalmış.
Dersi sadece iki kişi geçemedi.

Kullanılan tümceler sözdizimsel açıdan değişik yapılara sahip, olumlu-bildirici tümcelerdir. Soru tümceleri ve olumsuz tümceler veri tabanının dışında tutulmuştur. Tümceler ad, adıl, ön ad, belirteç, eylem gibi değişik sözcük türlerini içeren, on tanesi etken, biri edilgen yapıya sahip tümcelerdir. Tümcelerin altısı geçissiz, dördü geçişli, biri ise çift geçişli eyleme sahiptir. Tümceler dizim açısından belirteç-özne-yüklem, belirteç-belirtili nesne-yüklem, özne-belirtisiz nesne-belirtili nesne-yüklem, özne-belirtili nesne-belirteç-yüklem, özne-yüklem, özne-belirteç-yüklem, belirtisiz nesne-belirteç-yüklem, belirtili nesne-özne-yüklem sıralamasında basit tümcelerdir.

3. Deneyler

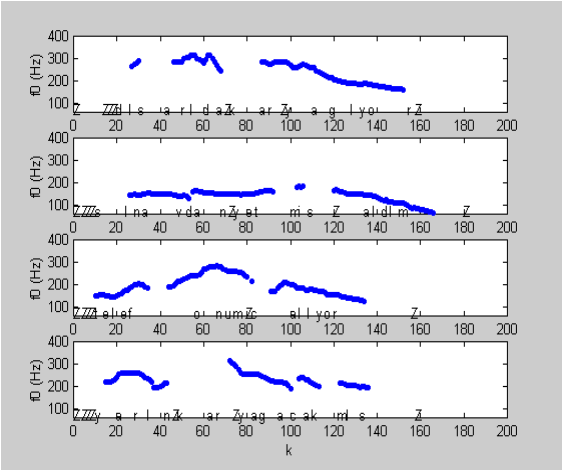
Çalışmada, hazırlanan 11 tümce, 4'ü erkek, 7'si kadın olmak üzere toplam 11 amatör oyuncu tarafından 4 farklı duyguyla okunmuştur. Konuşmacıların tümü üniversite öğrencileridir. Her duyguya ait kayıtlar, duygu aktarımında karışıklığa yol açılmaması için birbirlerinden farklı zamanlarda alınmıştır. Kayıt verenler, tümceleri aynı sırada, bir araştırmacının da bulunduğu ortamda okumuşlardır. Kayıttan önce konuşmacılara okuyacakları tümceler gösterilmiş ve tümceleri birbirinden bağımsız şekilde okumaları istenmiştir. Kayıtlar sonucunda 484 adet ses dosyası içeren bir veri tabanı elde edilmiştir. Kayıtlar tek tek dinlenerek başlardaki ve sonlardaki nefes alıp verme ve diğer etmenlerden kaynaklanan gürültülerden arındırılmıştır. Temizlenen ses dosyaları işlenerek temel sıklık eğrileri çıkartılmış, saklı markov modeli tabanlı bir bölütleme yazılımı ile harfler etiketlenmiştir. Ses kayıtları konuşmacı grubundan farklı kişilere dinletilmiş ve konuşmalardaki duygu aktarımı başarısı algısal açıdan denetlenmiştir.

3.1. Temel Sıklık Eğrisi İncelemeleri

Duyguların temel sıklık eğrilerine yaptığı etkiler, algısal denetleme sonucunda yüksek doğru algılama başarısına (%90 ve üstü) sahip kayıtlardan çıkartılan eğriler üzerinde incelenmiştir. İncelemelerin ana odağı temel sıklık eğrilerinin tümce başı ve tümce sonunda gösterdiği eğilimlerdir. Bu incelemeler sonucunda duyguların temel sıklık eğrilerine yaptığı etkiler şöyle özetlenebilir:

Kızgınlık ve sevinç tümceleriyle duygusuz okunan tümcelerdeki temel sıklık eğrileri, çoğunlukla tümce başında yükselme-düşme, üzüntü duygusuyla okunan tümcelerde ise düşme-yükselme ya da düz devam etme eğilimi göstermektedir.

Temel sıklık eğrilerinin tümce sonunda gösterdiği eğilim kızgınlık tümceleriyle duygusuz okunan tümcelerde düşme, sevinç ve üzüntü duygularıyla okunan tümcelerde yükselmedir. Dört farklı duygunun temel sıklık eğrilerinden hesaplanan çeşitli istatistiksel değerlere göre, en yüksek ortalama temel sıklık değerine sahip olan sevinç duygusu iken en düşük temel sıklık değerine sahip olan duygusuz okumadır. En yüksek değer aralığına ve saçılıma sahip duygu kızgınlıktır. Duygusuz okumanın ve üzüntünün değer aralıkları ve saçılımları diğer duygulara göre düşüktür.



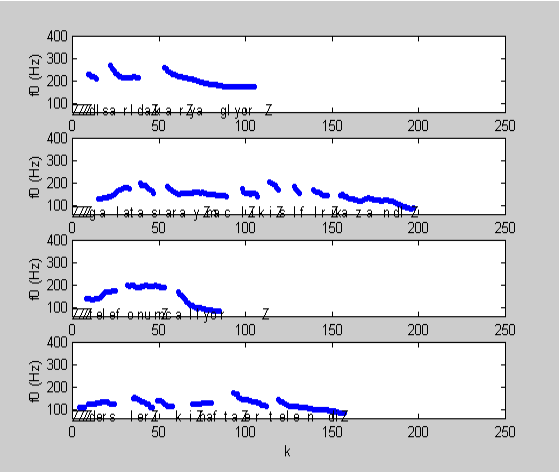
Şekil 1. Kızgınlık içeren kayıtların örnek temel sıklık eğrileri

Kızgınlık duygusuyla okunan tümcelerdeki genel eğilim temel sıklık eğrilerinin tümce sonuna doğru düşmesidir. İncelenen algılanma başarısı yüksek tümcelerin %81'inde tümce sonunda düşme gözlemlenmiştir. Kızgınlık içeren tümcelerin tümce sonunda yükselme oranı %10, düz devam etme oranı ise %9'dur. Kızgınlık duygusuyla okunan tümcelerdeki temel sıklık eğrilerinin tümce başındaki durumlarında kesin bir eğilim olmamakla birlikte; tümce başında yükselme-düşme durumu, düşme-yükselme ve düz devam etme durumlarına oranla daha fazla gözlemlenmiştir. Kızgınlık tümcelerindeki temel sıklık eğrilerinin %48'inde başta yükselme-düşme vardır. Eğrilerinin başta düşme-yükselme içerme oranı ise %29'dur. Kızgınlık tümcelerine ait temel sıklık eğrilerinde gözlemlenen başta düz devam etme oranı ise %23'tür.

Yukarıdaki gözlemler ışığında kızgınlık ile okunan tümcelerin temel sıklık eğrileri çoğunlukla tümce başında yükselme-düşme, tümce sonunda ise düşme eğilimi göstermektedir.

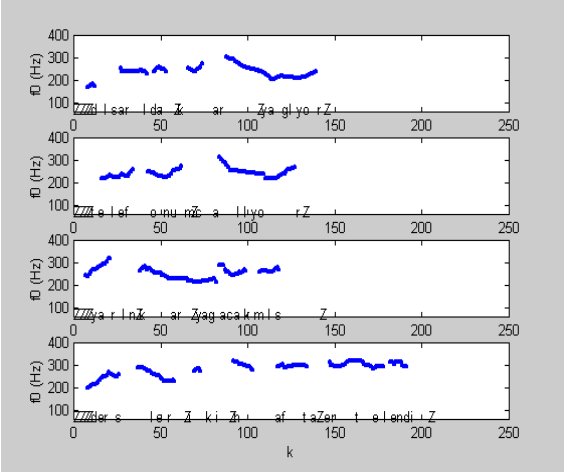
Duygusuz okunuşlarda genel eğilim kızgınlıkta olduğu gibi temel sıklık eğrilerinin sona doğru düşüş göstermesidir. Ancak, sözkonusu tümcelerde sonda yükselme ve düz devam etme durumu da gözlemlenmiştir. Tümcelerdeki temel sıklık eğrilerinin tümce başındaki durumunda ise kızgınlık tümcelerinde olduğu gibi, başta yükselme-düşme durumu ağırlıklı gözlemlenmiştir. Tümcelerin %49'u başta yükselme-düşme içermektedir. Duygusuz okunan tümcelerin %16'sında tümce başında düşme-yükselme, %35'inde ise düz devam etme durumu söz konusudur.

Yukarıdaki bulgulara göre duygusuz okunan tümcelerin temel sıklık eğrileri, kızgınlıkta olduğu gibi, çoğunlukla tümce başında yükselme-düşme, tümce sonunda ise düşme eğilimi göstermektedir. Aradaki fark, saçılım değerlerinin kızgınlıkta duygusuz okumaya göre daha yüksek olmasıdır.



Şekil 2. Duygusuz okuma kayıtlarının örnek temel sıklık eğrileri

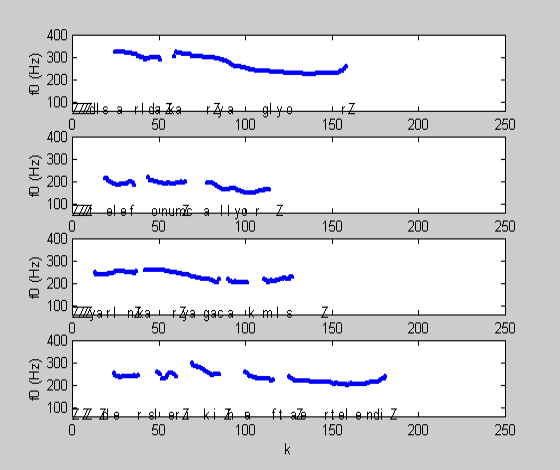
Sevinç duygusuyla okunan tümcelerdeki temel sıklık eğrilerinin tümce sonunda gösterdiği eğilim, kızgınlık duygusuyla ve duygusuz okunan tümcelerdeki temel sıklık eğrilerinin gösterdiği eğilimden farklı olarak tümce sonunda yükselme içermektedir. Doğru algılanma başarıları yüksek sevinç tümcelerindeki temel sıklık eğrileri, tümce sonunda %59 oranında yükselme göstermektedir. Eğrilerin tümce sonundaki düşme oranı %28, düz devam etme oranı ise %13'tür. Eğrilerin tümce başındaki eğilimleri ise kızgınlık tümceleriyle duygusuz okunan tümcelerdeki eğrilerin gösterdiği eğilimlerden daha belirgindir. Sevinçle okunan tümcelerin temel sıklık eğrileri %71 oranında tümce başında yükselme-düşme içermektedir. Temel sıklık eğrilerinin tümce başında düşme-yükselme gösterme oranı %22, düz devam etme oranı ise %7'dir.



Şekil 3. Sevinç içeren kayıtların örnek temel sıklık eğrileri

Özet olarak sevinç duygusu ile okunan tümcelerin temel sıklık eğrileri, kızgınlık duygusunda ve duygusuz okumada olduğu gibi tümce başında yükselme-düşme eğilimi göstermektedir. Tümce sonunda ise temel sıklık eğrileri kızgınlık duygusu ve duygusuz okuma durumunun aksine yükselme eğilimi göstermektedir.

Üzüntü duygusuyla okunan tümcelerdeki temel sıklık eğrilerinin tümce sonunda gösterdiği eğilim, sevinç tümcelerinde olduğu gibi, yükselmedir. Ancak, bu yükselme eğilimi, düşme ve düz devam etme eğilimlerine karşı çok baskın değildir. Eğrilerin %61'i tümce sonunda yükselme, %22'si düşme, %17'si ise düz devam etme eğilimi göstermektedir. Eğrilerin tümce başında gösterdikleri eğilim ise düşme-yükselme ve düz devam etme yönündedir. Üzüntüyle okunan tümcelerin tümce başında düşme-yükselme gösterme oranı %48, düz devam etme oranı ise %39'dur. Eğrilerde %13 oranında ise yükselme-düşme gözlemlenmektedir.



Şekil 4. Üzüntü içeren kayıtların örnek temel sıklık eğrileri

Kısaca, üzüntüyle okunan tümcelerdeki temel sıklık eğrileri tümce başında, genelde, düşme-yükselme ya da düz devam etme eğilimi göstermektedir. Tümce sonunda ise yükselme eğilimi ağırlıklı olarak gözlenmiştir.

	Kızgınlık	**Duygusuz**	**Üzüntü**	**Sevinç**
En yüksek	332	281	321	343
En düşük	180	156	196	208
Değer aralığı	152	125	125	135
Ortalama	251	214	259	274
Ortanca	250	214	259	273
Standart sapma	39,3	33,7	34,5	35,6

Tablo 1. Temel sıklık eğrilerine ait istatistiksel değerler (Hz)

Tablo 1'de dört farklı duygunun temel sıklık eğrilerinden hesaplanan çeşitli istatistiksel değerler görülmektedir. Bu değerlere göre en yüksek ortalama temel sıklık değerine sahip olan sevinç duygusu iken en düşük temel sıklık değerine sahip olan duygusuz okumadır. En yüksek değer aralığına ve saçılıma sahip duygu kızgınlıktır. Duygusuz okumanın ve üzüntünün değer aralıkları ve saçılımları diğer duygulara göre düşüktür.

3.2. Algısal Deneyler

Çalışmanın ikinci aşamasında, elde edilen veri tabanı 18-30 yaşları arasında, eğitimli 10 kişiden oluşan kontrol grubuna dinletilmiş ve kendilerinden dinledikleri duyguları kızgınlık, sevinç, üzüntü ve duygusuz okuma olarak sınıflandırmaları

istenmiştir. Kotrol grubuna dinletilecek kayıtlar üç değişik sıralama ile dizilmiş, böylelikle içerikleri aynı, sıralamaları farklı üç ayrı set elde edilmiştir. Dinleyiciler setleri yarım gün ile bir hafta arasında değişen zaman aralıklarıyla dinlemişlerdir.

Duygu	Kızgınlık	Duygusuz	Üzüntü	Sevinç
Duygusuz	6,40	85,71	6,77	1,12
Kızgınlık	92,37	3,70	1,12	2,81
Üzüntü	3,03	2,92	90,80	3,25
Sevinç	5,80	5,69	6,43	82,08

Tablo 2. Hata matrisi (%)

Tablo 2'de algısal deneyler sonucu ortaya çıkan hata matrisi görülmektedir. Bu matriste, dikey doğrultuda verilen duygular aktarılan duyguları, yatay doğrultuda verilen duygular ise algılanan duyguları göstermektedir. Hata matrisinden de gözlemlenebileceği gibi, kızgınlık duygusunun doğru algılanma başarısı en yüksektir. Bu durum, kızgınlık duygusu ile okunan tümcelerdeki temel sıklık eğrilerinin, diğer duygulara oranla, daha belirgin eğilimler göstermesiyle açıklanabilir. Doğru algılanma başarısı en düşük duygu ise sevinçtir. Duyguların birbirleriyle en fazla karıştırılma durumu, duygusuz okumanın üzüntü olarak algılanmasıdır. Bu karıştırılmadaki ana etken, her iki duygunun da düşük temel sıklık değer aralıklarına ve saçılımlarına sahip olmalarıdır.

Deney Setleri	Hata Oranları
Set 1	14,90
Set 2	12,21
Set 3	9,68
Ortalama	12,26

Tablo 3. Hata oranları (%)

Aktarılan duyguların kestiriminde bir öğrenme süreci olup olmadığını gözlemlemek için, her sette yapılan hatalar ayrı ayrı incelenmiştir. Tablo 3'te yer alan hata oranlarından da anlaşılabileceği gibi, set numarası arttıkça hata değerleri düşmektedir. Bu bulgu gösteriyor ki, duyguların algılanmasında bir öğrenme sürecinden söz etmek doğru olur.

4. Sonuçlar

Bu çalışmada, dört farklı duygunun temel sıklık eğrilerine yaptığı etkiler incelenmiştir. Yapılan incelemeler sonucu,

- Kızgınlık tümceleriyle duygusuz okunan tümcelerin temel sıklık eğrilerinin, çoğunlukla tümce başında yükselme-düşme, tümce sonunda ise düşme eğilimi gösterdiği,

- Duygusuz okunan tümcelerin temel sıklık eğrilerinin, kızgınlıkta olduğu gibi, çoğunlukla tümce başında yükselme-düşme, tümce sonunda ise düşme eğilimi gösterdiği, kızgınlıkla arasındaki farkın, çıkış ve iniş eğimlerinin kızgınlıkta, duygusuz okumaya göre daha yüksek olmasından kaynaklandığı,

- Sevinç duygusu ile okunan tümcelerin temel sıklık eğrilerinin tümce başında yükselme-düşme, tümce sonunda ise yükselme eğilimi gösterdiği,

- Üzüntüyle okunan tümcelerdeki temel sıklık eğrilerinin tümce başında, genelde, düşme-yükselme ya da düz devam etme, tümce sonunda ise ağırlıklı olarak yükselme eğilimi gösterdiği,

- En yüksek ortalama temel sıklık değerine sevinç duygusu sahip iken, duygusuz okumanın en düşük ortalama temel sıklık değerine sahip olduğu,

- En yüksek değer aralığına ve saçılıma sahip duygunun kızgınlık olduğu,

- Duygusuz okumanın ve üzüntünün değer aralıklarının ve saçılımlarının diğer duygulara göre düşük olduğu

gözlemlenmiştir.

Bu incelemelere ek olarak, algısal deneyler yapılmış ve duyguların algılanma başarıları ölçülmüştür. Deney sonuçlarından kızgınlık duygusunun en fazla doğru algılanabilen duygu olduğu, sevincin ise algılanma başarısının en düşük duygu olduğu görülmüştür. Her setteki hataların ayrı ayrı incelenmesi sonucu, duyguların algılanmasında bir öğrenme sürecinin var olduğu ortaya çıkmıştır.

5. Teşekkür

Çalışmaya katkılarından dolayı Boğaziçi Üniversitesi Oyuncuları'na, -Canan Tanır, Deniz Aydın, Didem Karanfil, Erdoğan Çete, Eser Dilsöz, Ozan Uysal, Sezin Gündoğan- Emel Kahraman'a, Işıl Eğrikavuk'a, İlker Öztop'a, Umay Karabaş'a, A. Hakan Büyükbayrak'a, Beste Can Yolcu-Kamali'ye, Duygu Şimşek'e, E. Edip Çiçekli'ye, E. Pınar Cen'e, Erdem Ulukan'a, Gülbin Akgün'e, Görkem Şarkış'a, Murat Erman'a, Onur Sarkan'a, Özge Çağlar'a ve Dilbilim Konuşmaları dinleyicilerine teşekkür ederiz.

6. Kaynakça

Abelin, A ve J. Allwood (2000) "Cross Linguistic Interpretation of Emotional Prosody", *Proceedings of the ISCA Workshop on Speech and Emotion*. http://www.qub.ac.uk/en/isca/proceedings

Bosch, L.T (2003) "Emotions, Speech and the ASR Framework", *Speech Communication,* c. 40, s. 213-215.

Burkhardt, F ve W.F Sendlmeier (2000) "Verification of Acoustical Correlates of Emotional Speech Using Formant-Synthesis", *Proceedings of the ISCA Workshop on Speech and Emotion.* http://www.qub.ac.uk/en/isca/proceedings

Demircan, Ö (1980) "Türkçe Ezgilemeye Giriş", *TDAY-Belleten* s 65-79.

Kornfilt, J (1997) *Turkish*, London: Routledge.

Lee, C.M, S. Narayanan, R. Pierraccini (2001) "Recognition of Negative Emotions from the Speech Signal", *Automatic Speech Recognition and Understanding Workshop*. http://asru01.itc.it/paper/abs/a01cl099.htm

McGilloway, S, R. Cowie, E. Douglas-Cowie, S. Gielen, M. Westerdijk, S. Stroeve (2000) "Approaching Automatic Recognition of Emotion from Voice: A Rough Benchmark", *Proceedings of the ISCA Workshop on Speech and Emotion.* http://www.qub.ac.uk/en/isca/proceedings

Montero, J. M, J. Gutierrez-Arriola, J. Colas, E.Enriquez ve J.M Pardo (1999) "Analysis and Modelling of Emotional Speech in Spanish", *ICPhS'99.* http://lorien.die.upm.es/~macias/doc/pubs/eurosp99/submitted/m058.pdf

Mozziconacci, S.J ve D.J. Hermes (1999) "Role of Intonation Patterns in Conveying Emotion in Speech", *ICPhS'99* .http://www.lpl.univ-aix.fr/projects/sprosig/guests/mozzic/icphs99.pdf

Oudeyer, P.Y (2002) "The Production and Recognition of Emotions in Speech: Features and Algorithms", *International Journal of Human Computer Interaction.* http://www.csl.sony.fr/downloads/papers/2002/oudeyer02b.pdf

Petrushin, V.A (1999) "Emotion in Speech: Recognition and Application to Call Centers", Artificial Neural Networks in Engineering (ANNIE '99) s. 7-10.