5.1. ÜNLÜ FORMANT NORMALİZASYONU: R PROGRAMLAMA DİLİNDE BİR UYGULAMA

Özgür AYDIN & İpek Pınar UZUN
Ankara Üniversitesi, Dilbilim Bölümü
ozguraydin66@gmail.com, pinarbekar@gmail.com

1. Giriş

Formant ya da biçimleyici (*formant*), belirli bir frekansta akustik enerjinin yoğunluğunu ifade etmektedir. Formant frekansı (*formant frequency*), doğrudan ses yoluyla ilişkili olarak konuşma sinyallerinin özelliklerini göstermektedir. Böyle olunca da formant frekans değerleri ünlüye ve konuşura bağlı olarak değişmektedir. Formant frekans değerlerinin konuşurlar arasında değişmesi, ünlülerin formant frekans değerlerine göre ayırt edilmesinde de güçlüklere neden olabilmektedir. Bu nedenle, ünlülerin formant frekans değerlerinin hesaplanmasında, dolayısıyla bir dilde ünlülerin sınıflandırılmasında formant frekans değerlerinin normalizasyonu (*normalization*) önem taşımaktadır.

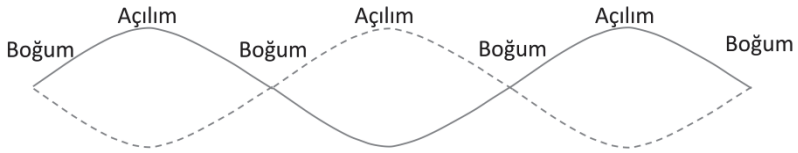
Alanyazında Türkçe ünlülerin formant frekanslarıyla ilgili pek fazla çalışma bulunmamakta (Çiyiltepe, Bekar & Ergenç, 2009; Davutoğlu, 2010; Kılıç, 2003; Korkmaz & Boyacı, 2018; Malkoç, 2011; Türk, Şayli, Özsoy & Arslan, 2004; ayrıca MR ve frekans analizi için bkz. Kılıç & Öğüt, 2004), pek azında normalizasyon işlemlerinin uygulandığı görülmektedir. Bu çalışmada, ünlülerin sınıflandırılmasında formant normalizasyonunun önemi vurgulanarak, normalizasyon işlemlerinin içeriği ve R programlama dilinde (R Core Team, 2014) normalizasyon işlemlerinin nasıl yapılacağı konuları ele alınmaktadır. Çalışmada ayrıca örnek bir uygulama sunularak, ünlülerin sınıflandırılmasında kullanılabilecek normalizasyon işlemlerinin ayırıcı gücü Türkçe üzerinden ortaya konmaktadır.

2. Formant Frekansları

2.1. Temel Frekans (F_0)

Temel frekans (sıklık) (*fundamental frequency/*F_0), ses tellerindeki titreşim sıklığının birim zaman dilimi içindeki devir sayısını gösteren ve Hertz

(Hz) ile ölçülebilir nitelikli akustik bir değişkendir. Cinsiyete dayalı olarak değişmekte olan temel sıklık değeri, genellikle kadınlarda 170-210 Hz, erkeklerde 120-140 Hz ve çocuklarda 220-280 Hz olarak hesaplanmaktadır. Bu frekans değerleri, Kent ve Read (2002)'de belirtildiği gibi, ses tellerinin her bireyde farklı bir anatomi içermesinden ötürü, bireysel farklılıklara dayalı olarak değişebilmektedir. Örneğin oktav atlaması (*octave jump*), enerji yoğunluğundaki (*intensity*) ses şiddetinin artması ya da azalması gibi çeşitli sesbilgisel etkenlere dayalı olarak kimi durumlarda temel sıklık değerleri olağan normun üzerinde ya da altında olabilmektedir.



Şekil 1. Temel sıklık örüntüsü ve ilk altı harmonik değeri

Yukarıda Şekil (1)'de gösterildiği gibi, temel sıklık değerleri harmonik sayısı arttıkça değişim göstermekte ve sıklık alanı artmaktadır. Bu durumda, sesin titreşim değerleri de harmoniklerle uyumlu bir görünüm sergilemektedir. Her bir titreşim alanı, bir boğum *(node*) ve açılım (*antinode*) alanı oluşturarak kesitlere ayrılmaktadır.

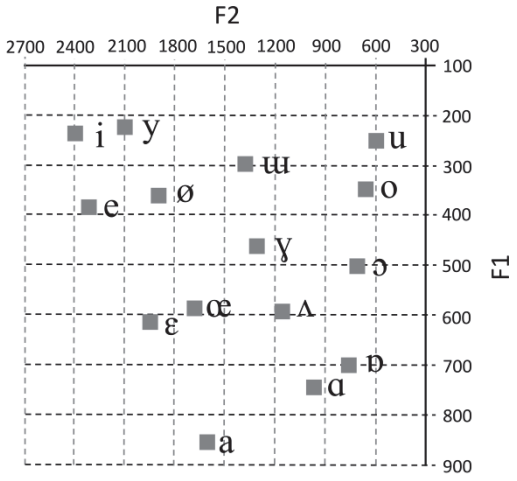
2.2. Ünlü Formant Frekansları

Konuşma seslerinin incelendiği araştırmalarda, formant frekansları değerlerinde ilk beş formantın frekans değerleri hesaplanmaktadır. Cinsiyete dayalı olarak, değişim gösteren formant frekansları erkeklerde 500-5000 Hz, kadınlarda ise 500-5500 Hz aralığında görülmektedir (Kinsler, Frey, Coppens & Sanders, 2000; Kent & Read, 2002). Frekans değerleri, cinsiyete dayalı olarak Tablo (1)'deki gibi görülmektedir:

Tablo 1. Formant frekans değerlerinin cinsiyete dayalı tahmini değişimi

Formant Frekansları	*Tahmini Formant Frekans Aralıkları (Hz)*		
	Kadın	Erkek	Çocuk
Birinci Formant (F_1)	500-1200	500-1000	500-1400
İkinci Formant (F_2)	1200-2200	1000-2000	1400-2400
Üçüncü Formant (F_3)	2200-3200	2000-3000	2400-3400
Dördüncü Formant (F_4)	3200-4200	3000-4000	3400-4400
Beşinci Formant (F_5)	4200-5200	4000-5000	4400-5400

Sesin parçalı birimlerinin incelenmesinde büyük önem taşıyan formant frekansları, titreşim rezonansındaki (*resonance)* değişimlere dayalı olarak farklı sesbilgisel yorumlamalar içermektedir. Örneğin, F_1 formant frekansları dilin ağız boşluğundaki devinimine dayalı olarak arkadil ya da ortadil ünlülerinin belirlenmesinde önemli bir rol oynarken, F_2 ve F_3 formant frekansları öndil ünlülerinin belirlenmesinde etkindir. Aşağıda Şekil (2)'de de sunulduğu gibi, F_1 formant frekansı arkadil ya da ortadil ünlülerinin bulunduğu dörtgenin dikey alanını kapsarken, F_2 formant frekansı ise öndil ünlülerinin ağırlıkta olduğu yatay alanda görülmektedir.



Şekil 2. F_1 ve F_2 formant frekanslarının ünlü dörtgenindeki görünümü

2.3. Ünsüz Formant Frekansları

Formant frekansları, ünlülerde ve kimi durumlarda ünlülere benzer fiziksel örüntüler içeren ünsüzlerde de gözlenmektedir. Genizsil, yan daralmalı ünsüzler ve yarı ünlülerde, pek çok ünsüzden farklı olarak ses tellerinin titreşimi sırasında gürültü ve engel düzeyi daha hafiftir ve ünlülere benzer sesbilgisel özellikler içermektedir. Bu nedenle de formant değerleri üretebilmektedir. Elbette ki, ünsüz formant frekansları, üretim sırasında ünlülerdeki gibi belirgin bir biçim almamaktadır.

Genizsil, yan daralmalı ünsüzler ve yarı ünlüler, üretim aşamasında ünlülerden farklı olarak ses yolunda birden fazla aşama içermektedir. Ünlülerin ses yolunda tek bir tüp içinde üretildiği varsayılırken, ünsüz formantlarında ses yolunda birden fazla işlem içerdiği, özellikle de genizsil ünsüzlerde ses yolunun yanı sıra, geniz yolunun da devreye girdiği varsayılmaktadır.

Genizsil ve yan daralmalı ünsüzlerin üretimi sırasında ses yolu ve geniz yolunun birleştiği yardımcı bir alan olduğu tahmin edilmektedir. Yarı ünlülerde ise, ses yolunun aldığı biçime dayalı olarak formant frekans değerleri belirlenebilmektedir. Ötümlü kapantılı ünsüzlerle ya da ötümlü patlamalı ünsüzlerle birlikte konumlandıklarında ünlülerden farklı örüntüler sergileyen yarı ünlülerde ses yolundaki akustik enerji ünlülere göre değişim göstermektedir. Ancak, ünlüler kadar belirgin olmadıkları için, genizsil ve yan daralmalı ünsüzlere ya da yarı ünlülere ait formantlar belirlenirken, bu ünsüzlerle birlikte konumlanan ünlülerinin de değerleri mutlaka hesaplanmakta ve algoritma hesaplaması yapılırken sesdizimsel etkileşim temel alınmaktadır.

3. Ünlü Formant Normalizasyonu

3.1. Bilgi Türleri ve Ünlü Normalizasyonu

Ünlü normalizasyonu (*vowel normalization*), en genel tanımıyla, akustik veriyi konuşurlar arasında karşılaştırılabilir kılmak için uygulanan bir yöntemdir. Bu yöntem temel olarak, toplumsal, fiziksel ya da doğrudan dilsel ayrımlardan kaynaklanan akustik veriler arasındaki ayrımları belirginleştirmek için kullanılmaktadır. Normalizasyon, cinsiyet farklılıkları, yaş farklılıkları, toplumsal katmanlar arasındaki farklılıklar gibi pek çok farklı bağlamda uygulanabilir. Alanyazında Ladefoged ve Broadbent (1957)'ye dayanarak konuşma sinyallerinin üç tür bilgi yansıttığı ileri sürülmektedir: Sesbirimsel bilgi (*phonemic information*), anatomik/fiziksel bilgi (*anatomical/physical information*), toplumdilbilimsel bilgi (*sociolinguistic information*)[1]. Ladefoged ve Broadbent, günlük konuşmalarımızda, konuşurun söylediği ifadenin kendisinin, yani doğrudan sesbirimsel ayrımların sesbirimsel bilgiyle ilgili olduğunu, ancak konuşurun söylediği ifadelerin ötesinde sesbilimsel olarak artalanıyla ilgili kimi ipuçlarına da sahip olabileceğimizi, bunun da toplumdilbilimsel bilgiler içinde değerlendirilmesi gerektiğini belirtmektedir. Diğer yandan, bireysel farklılıklardan kaynaklı etkenler de anatomik/fiziksel bilgiyi yansıtmaktadır. Aslında, ilk bilgi türü, doğrudan içsel dilsel bilgiyle, yani büyük oranda edinçle ilgiliyken, diğerleri dışsal dilsel bilgilerle yani edimle ilgilidir. Sesbilim incelemelerinde ya da genel olarak ses üzerine farklı disiplinlerin incelemelerinde, söz konusu üç bilgi türü, incelenen konu bakımından farklı değerlere sahiptir. Söz gelimi, iki sesin

[1] Ladefoged ve Broadbent (1957) konuşma sinyallerinin yansıttığı bilgi türlerini sıralarken sesbirimsel bilgiyi "dilbilimsel bilgi", anatomik/fiziksel bilgiyi de "kişisel bilgi" olarak adlandırmıştır.

aynı konuşura ait olup olmadığı sorusu anatomik/fiziksel ya da toplumdilbilimsel bilgiyle ilgiliyken, böyle bir analizde sesbirimsel bilgi veri analizinde gürültü (*noise*) yaratmaktadır. Diğer yandan, söz gelimi, iki konuşur arasındaki cinsiyet, yaş, toplumsal katmanlar gibi özellikler bakımından farklılıkların ünlüler üzerindeki etkisi toplumdilbilimsel ve sesbirimsel bilgilerle ilgilidir, bu defa anatomik/fiziksel bilgi, veri üzerinde gürültüye neden olmaktadır.

3.2. Ünlü Normalizasyonlarının Sınıflandırması

Söz konusu üç tür bilgi türünün ünlülerin formant frekanslarını sistematik olarak etkilediği ileri sürülmüştür (bkz. Peterson & Barney, 1952; Hindle, 1978 ve Labov, 2001). Bu yüzden de geleneksel olarak formant-temelli normalizasyon işlemleri yukarıda sıraladığımız bilgi türlerini temel alarak sınıflanmaktadır. Bu türden sınıflandırmaların ilk örneği Ainsworth (1975)'te görülmektedir. Ainsworth, içsel (*intrinsic*) ve dışsal (*extrinsic*) bilgi türlerine göre bir ayrım yapmaktadır[2]. Ünlü-içsel normalizasyon işlemi, tek bir ünlüyü sınıflandırmak için söz konusu ünlüyü barındıran akustik bilgiyi kullanmaktadır. Bu işlem formant frekansının doğrusal olmayan dönüşümünü (örneğin, log, mel, bark) ya da bununla birlikte formant sıklıklarının birleşimini (örneğin, F_1–F_0) içerebilir. Ünlü-dışsal normalizasyon işlemi için ise konuşurun birden çok ünlü kategorisinin dağılımı gerekmektedir, yani bu işlem konuşurun birden çok ünlüsü arasındaki akustik bilgiyi kullanmaktadır (bkz. Gerstman, 1968; Lobanov, 1971; Nordström & Lindblom, 1975; Nearey, 1978). Ünlü-dışsal normalizasyon işlemleri genellikle ünlülerin sınıflandırılması için kullanılmaktadır[3].

Ünlü-içsel ve ünlü-dışsal kategorilerinin yanı sıra, Ainsworth'un içsel/dışsal ayrımı, formant kategorisine de aktarılmıştır (Rosner & Pickering,

---

[2] Ainsworth'ün içsel ve dışsal bilgi türlerine göre yaptığı ayrımın yanı sıra, Rosner ve Pickering (1994) üçlü bir sınıflandırma önermektedir: kategori-bağımsız, kategori-özgül ve konuşur-özgül. Bu ayrıma göre, konuşur-özgül işlemi, Ainsworth'ün dışsal normalizasyon işlemiyle aynı şeyi ifade etmektedir. İçsel normalizasyon işlemi ise F_0'ın rolü dikkate alınarak kategori-bağımsız ve kategori-özgül olarak ikiye ayrılmıştır.

[3] Bu yazıda sadece akustik ünlü normalizasyon işlemleri ele alınmıştır. Akustik ünlü normalizasyonundan ayrı olarak, algısal (*perceptual*) ünlü normalizasyon işlemleri de bulunmaktadır. Tıpkı, akustik ünlü normalizasyon işlemlerinde olduğu gibi, algısal ünlü normalizasyon işlemlerinde de sesbirimsel, anatomik/fiziksel ve toplumdilbilimsel bilgiler geçerliliğini korumaktadır. Ünlülerin algılanması sırasında, dinleyenlerin ünlüleri doğru olarak sınıflandırabilmesi için akustik sinyallerdeki söz konusu bilgileri kullanmaları gerekmektedir. Algısal ünlü normalizasyon işlemleri konunun algısal yönüyle ilgili farklı yaklaşımlar sunmaktadır (farklı yaklaşımlar için bkz. Traunmüller, 1981; Peterson & Barley, 1952; Syrdal, 1985; Johnson, 1990, Adank, 2003).

1994). Tıpkı tek bir ünlü biriminin (ünlü-içsel) ya da birden çok ünlü biriminin kullanımı (ünlü-dışsal) biçimindeki sınıflandırmada olduğu gibi, formantlarda da içsel/dışsal ayrımına göre dönüştürme işlemi uygulanabilmektedir. Formant-içsel işlemlerde bir formant değerinin normalleştirilmesi, sadece o formanta ilişkin bilginin kullanılmasıyla gerçekleştirilir. Buna karşın, formant-dışsal işlemlerde bir formant değerinin normalleştirilmesinde farklı formantlarla ilgili bilgiler kullanılır. Söz gelimi, F_1 formatının normalleştirilmesinde sadece F_1 değerinin kullanılması formant-içsel işleme örnekken, Syrdal ve Gopal (1986)'nın önerisindeki gibi, F_1 formantının normalleştirilmesinde F_1 ve F_0'ın, F_2 formantının normalleştirilmesinde ise F_2 ve F_3'ün kullanılması, formant-dışsal işlemlere örnek olarak gösterilebilir.

Tablo (2)'de görüldüğü gibi, sırasıyla frekans değerinin log-dönüşümü, mel-dönüşümü, bark-dönüşümü ve Eşdörtgensel Bant Genişliği (*ERB, Equivalent Rectangular Bandwidth*) dönüşümü işlemleri ünlü-içsel/formant-içsel işlemler olarak sınıflandırılmıştır. Buna karşın, Syrdal ve Gopal (1986)'nın bark-fark normalizasyon işlemi ise tek bir ünlüdeki bilgiyi çözümlediği için ünlü-içsel, ancak F_1-farkı için F_1 ve F_0 değerlerini, F_2-farkı için de F_3 ve F_2 değerlerini hesaplamaya kattığı için formant-dışsal özellik sergilemektedir (bkz. Tablo 2). Gerstman (1968)'in dizi normalizasyonu (*range normalization*), Lobanov (1971)'in z-skor dönüşümü (*z-score tranformation*), Watt ve Fabricius (2002)'nin normalizasyonu ve Nearey (1978)'nin Sabit Log-Aralığı Hipotezi (*Constant Log Interval Hyphotesis, $CLIH_{i2}$*) olarak bilinen bireysel log-ortalama işleminde ise her konuşur için birden daha çok ünlüye ilişkin formant değeri gerektiği ve her bir formant ayrı ayrı hesaplandığı için ünlü-dışsal/formant-içsel olarak sınıflandırılmıştır. Nearey (1978)'in paylaşımlı log-ortalama (*shared log-mean)* işlemiyle ($CLIH_{s4}$), Nordström ve Lindblom (1975)'in ses-yolu ölçeği (*vocal-track scaling*) ve Miller (1989)'un formant oranı modeli hem ünlü-dışsal hem de formant-dışsal yöntemlerdir.

Tablo (2)'de ele alınan tüm dönüşüm türleri konuşur-içsel modellerdir, bununla birlikte Labov, Ash ve Boberg (2006)'nın, Nearey (1978)'den uyarladığı model (*Atlas of North American English, ANAE*) konuşur-dışsal bir modeldir. Diğer yandan, Nordström ve Lindblom (1975)'in normalizasyonunda bir gruptaki konuşurların ünlülerine ilişkin tüm formant değerleri bir referans konuşura göre hesaplandığından bu tür normalizasyon işlemi de konuşur-dışsal olarak da ele alınmaktadır (bkz. Flynn, 2011; Adank, 2003:15).

Tablo 2. Ünlü normalizasyon işlemlerinin iki faktöre göre sınıflandırılması

Dönüşüm türleri	Kaynak	Ünlü	Formant
ln / log	–	içsel	içsel
bark	Traunmüller (1990)	içsel	içsel
mel	Stevens ve Volkman (1940)	içsel	içsel
ERB	Moore ve Glasberg (1983)	içsel	içsel
bark-fark	Syrdal ve Gopal (1986)	içsel	dışsal
Gerstman	Gerstman (1968)	dışsal	içsel
Lobanov/z-skor	Lobanov (1971)	dışsal	içsel
Nearley1/$CLIH_{i2}$	Nearey (1978)	dışsal	içsel
Watt ve Fabricius	Watt ve Fabricius (2002)	dışsal	içsel
Nordström ve Lindblom	Nordström ve Lindblom (1975)	dışsal	dışsal
Nearley2/$CLIH_{s4}$	Nearey (1978)	dışsal	dışsal
Miller	Miller (1989)	dışsal	dışsal

4. Ünlü Normalizasyonu İşlemleri ve R Uygulamaları

Ünlü normalizasyon işlemlerinde W. Labov tarafından geliştirilen *PLOTNIK* (https://www.ling. upenn.edu/~wlabov/Plotnik.html) ve *Praat*'a (Boersma ve Weenink, 2020) eklenti olarak çalışan *Akustyk* (https://github.com/Hoetzel2/ akustyk) gibi araçların yanı sıra, R tabanlı olarak Thomas ve Kendall (2007) tarafından geliştirilen *NORM* ve Heeringa ve Van de Velde (2018) tarafından geliştirilen *Visible Vowels* gibi web uygulamaları bulunmaktadır. R tabanlı olarak çalışan her iki uygulama da çeşitli R paketlerinden (sırasıyla *vowels* ve *phonR*) yararlanmaktadır. Bu bölümde, *phonR*, *vowels*, *phonTools* paketleri yardımıyla belli başlı ünlü normalizasyon işlemleri tanıtılmaktadır (bkz. https://github.com/ozguraydin66/vowelnormalization). R'de ünlülerin kullanılan paketler ve normalizasyon türleri Tablo (3)'te görülmektedir.

Tablo 3. Dönüşüm türlerine ilişkin R'de kullanılan kod ve paketler

Dönüşüm türü	phonR	vowels	phonTools
ln /log	normLog()	–	–
bark	normBark()	–	–
mel	normMel()	–	–
ERB	normErb()	–	–
bark-fark	–	norm.bark()	–
Lobanov/z-score	normLobanov()	norm.lobanov()	normalize()
Nearley1/$CLIH_{i2}$	normNearey1()	norm.nearley()	normalize()
Watt ve Fabricius	normWattFabricius()	norm.wattfabricius()	normalize()
Nearley2/$CLIH_{s4}$	normNearey2()	norm.nearley()	normalize()
Labov ANAE	–	norm.labov()	–

4.1. Ünlü-İçsel / Formant-İçsel Ölçekler

Ünlü-içsel/formant-içsel ölçeklerin en bilineni logaritmik (*ln*) ölçektir. Ünlü formant değerlerinin normalizasyonunda logaritmik ölçek kullanımı çok eskilere gitse de sadece ünlü-içsel/formant-içsel ölçeklerde değil (örneğin, mel, ERP), aynı zamanda bazı ünlü-dışsal formant normalizasyonu türlerinde de (bkz. Miller, 1989 ve Nearey, 1978) logaritmik ölçek kullanılmaktadır. Doğal logaritmik ölçek (1)'deki formülle hesaplanmaktadır[4].

$$F_i^{ln} = ln(F_i^{Hz}) \quad (1)$$

R'de log dönüşümleri taban (base v3.6.2) fonksiyonu olarak *log()* ve *log10()* biçiminde kullanılmaktadır (bkz. Liste 1). Liste (1)'deki kodlarda *log()* ya da *log10()* fonksiyonlarıyla *formant.data* veri çerçevesi içindeki F_1ve F_2 değerlerinin doğal ve 10 tabanına göre logaritması hesaplanmaktadır. Burada *with()* fonksiyonuyla log hesaplaması, ilgili veri çerçevesine (formant.data) uygulanmakta ve logaritması hesaplanmış olan F_1 ve F_2 değerleri *log.df* adlı yeni bir veri çerçevesine atanmaktadır. Sonrasında *colnames()* fonksiyonuyla, yeni veri çerçevesindeki sütun adları değiştirilmekte ve son olarak da log-dönüşümünün uyarlandığı *log.df* veri çerçevesiyle *formant.data* veri çerçevesi *cbind()* fonksiyonu kullanılarak birleştirilmektedir.

```
# log-dönüşümü
 log.df = with(formant.data, log(cbind(f1,f2)))
 colnames(log.df) = c("f1.log", "f2.log")
 formant.data = cbind(formant.data, log.df)

# log10-dönüşümü
 log10.df = with(formant.data, log10(cbind(f1,f2)))
 colnames(log10.df) = c("f1.log10", "f2.log10")
 formant.data = cbind(formant.data, log10.df)
```

Liste 1. Log-dönüşüm kodları

Melodi sözcüğünün kısaltmasından gelen mel ölçeği, dinleyenlerin öznel olarak perde (*pitch*) farklılıklarını değerlendirmelerine dayanan deneysel bulgular (Stevens & Volkmann, 1940) çerçevesinde geliştirilmiştir. Farklı mel formülleri önerilmiş olsa da, bu yazıda R'deki *phonR* paketinde (McCloy, 2012, 2016) kullanılan Traunmüller (1990)'ın önerdiği mel formülü en yaygın

[4] Formüllerde F (Formant) kısaltmasının üst iminde dönüşüm türü, alt iminde de formant türü ifade edilmektedir: Birinci formanta mel-dönüşümü için F_1^M, ikinci formanta log-dönüşümü için F_2^{ln}, ikinci formanta ham veri (Hz) için F_2^{Hz}gibi.

olanıdır (bkz.2). Bark ölçeği ise, belirli bir frekansta, bazal zarın aynı bölümünü etkinleştiren frekansların en küçük bant (*band*) düzeyini ifade eden kritik bant kavramına ve 24 birime ya da (Heinrich Georg Barkhausen'un adından türetilen) "bark"a bölümlenen bark ölçeğine dayanmaktadır. Bark ölçeği için de farklı formüller üretilmiş olmasına karşın, *phonR* paketi içinde Traunmüller'in önerdiği bark formülü (bkz.3) kullanılmaktadır.

$$F_i^M = 2595 \times ln\left(1 + \frac{F_i^{Hz}}{700}\right) \quad (2)$$

$$F_i^B = 26.81 \times \frac{F_i^{Hz}}{1960 + F_i} - 0.53 \quad (3)$$

Bu bölümde ünlü-içsel/formant-içsel ölçekler içinde son olarak ele alacağımız Eşdörtgensel Bant Genişliği ölçeği, Moore ve Glasberg (1983) tarafından geliştirilmiştir. ERP ölçeği, temel olarak bark ölçeğine benzese de, iki yönüyle bark ölçeğinden ayrılmaktadır: İlk ve önemli ayrım ERP ölçeğinin logaritmik bir ölçek olması (krş. 3 ve 4), diğeri de ERP'nin daha dar bir bant genişliği içermesidir. ERP ölçeği için *phonR* paketi içinde (4)'teki formül kullanılmaktadır.

$$F_i^E = 21.4 \times ln(0.00437 \times F_i^{Hz} + 1) \quad (4)$$

R'de paketler farklı alanlardan araştırmacıların katkılarıyla oluşturdukları veri setlerini ve fonksiyonları barındıran öğelerdir. R'de mel, bark ve ERP dönüşümlerini hesaplamak için öncelikle *library()* fonksiyonuyla *phonR* paketini yüklemek gerekmektedir. Sonrasında mel, bark ve ERP dönüşümlerini gerçekleştirmek için bu paket içindeki *normMel()*, *normBark()* ve *normErb()* fonksiyonları kullanılmaktadır (bkz. Liste 2).

```
library(phonR)
# mel-dönüşümü
 mel.df = with(formant.data, normMel(cbind(f1, f2)))
 colnames(mel.df) = c("f1.Mel", "f2.Mel")
 formant.data = cbind(formantdata, mel.df)
# bark-dönüşümü
 bark.df <- with(formant.data, normBark(cbind(f1, f2)))
 colnames(bark.df) = c("f1.Bark", "f2.Bark")
 formant.data = cbind(formant.data, bark.df)
# ERB-dönüşümü
 erb.df = with(formant.data, normErb(cbind(f1, f2)))
 colnames(erb.df) = c("f1.ERB", "f2.ERB")
 formant.data = cbind(formantdata, erb.df)
```

Liste 2. Mel, bark ve ERB-dönüşüm kodları

4.2. Ünlü-İçsel / Formant-Dışsal Ölçek

Tablo (3)'te de görüldüğü gibi, ünlü-içsel/formant-dışsal tek ölçek, Syrdal ve Gopal (1986)'nın bark-fark ölçeğidir. Syrdal ve Gopal'ın özgün önerilerinde, öncelikle F_0, F_1, F_2 ve F_3 değerlerine, Peterson ve Barney (1952)'nin formülleriyle bark-dönüşümü uygulanmakta, sonrasında da bu dönüşümlere sırasıyla, $F_1^{S\&G}$ için $F_0^B - F_1^B$ ve $F_2^{S\&G}$ için de $F_3^B - F_2^B$ formülleri uygulanmaktadır. Liste (3)'te Traunmüller (1990)'ın önerdiği bark formülüne göre (bkz.3) bark-fark dönüşümü örneklendirilmektedir. Burada F_0 kullanıldığı için öncelikle F_0'da var olmayan (n/a) veriler temizlendikten sonra, Liste (2)'deki *normBark()* fonksiyonu bu defa F_0 ve F_3'e de uygulanmaktadır. Yukarıda belirttiğimiz çıkartma işlemlerini uygulamak için *dplyr* paketi içindeki *transmute()* fonksiyonu kullanılmaktadır. Bu fonksiyon, sadece veri çerçevesiyle çalıştığı için ve *with()* fonksiyonu veri çerçevesi üretmediği için önce *as.data.frame()* fonksiyonuyla veri seti, veri çerçevesine dönüştürülmektedir.

```
library(phonR)
library(dplyr)

# bark-fark dönüşümü
 formantdata = formantdata[!is.na(formantdata$f0), ]
 df = with(formantdata, normBark(cbind(f0, f1, f2, f3)))
 df = as.data.frame(df)
 df = transmute(df, f1.Bark.dist = f1-f0, f2.Bark.dist = f3-f2)
 formantdata = cbind(formantdata, df)
```

Liste 3. F_0'a dayalı bark-fark dönüşüm kodları

Yukarıdaki kodlarda da görüldüğü gibi, Syrdal ve Gopal (1986)'nın özgün önerilerine koşut olarak $F_1^{S\&G}$'nin hesaplanması için $F_0^B - F_1^B$ işlemi kullanılmaktadır. Bunun nedeni, F_0'ı etkileyen başkaca bir öğe olmadığı durumda, F_0 alçak ünlülere (*low vowels*) nazaran yüksek ünlülerde (*high vowels*) daha büyük bir fark yaratmaktadır, diğer bir deyişle, yüksek ünlülerde F_0 ile F_1 arasındaki fark azalırken, alçak ünlülerde artmaktadır. Kendall ve Thomas (2018), ezgi, ton ve ünsüzlerle ilgili etkenlerin F_0'ı etkilemesi durumunu dikkate alarak *vowels* paketi içinde $F_0^B - F_1^B$ işlemi yerine $F_3^B - F_1^B$ işleminin kullanıldığını belirtmektedir.

Thomas ve Kendall (2007) tarafından geliştirilen *NORM* aracı (http://lingtools.uoregon.edu/norm/norm1.php) ile söz konusu dönüşüm işlemleri gerçekleştirilebileceği gibi, doğrudan NORM aracının kullandığı

vowels paketi ile Liste (4)'teki kodlarla da bark-dönüşümü gerçekleştirilebilir. *Vowels* paketi içinde *norm.data()* fonksiyonu uygun biçimdeki veri setine[5] bark-dönüşümünü doğrudan uygulamaktadır. Liste (4)'te, *norm.data()* fonksiyonunun ardından *rename()* fonksiyonu ile dönüştürme sonucu oluşturulan "Z3-Z2", ve "Z2-Z1" veri çerçevesindeki sütun adları yerine yeni adlar (sırasıyla, "f1.Bark.dist" ve "f2.Bark.dist") atanmakta ve sonrasında ise bark-dönüşümü uygulanan veri çerçevesi (*df*), özgün veri çerçevesiyle (*formantdata*) birleştirilmektedir.

```
library(vowels)
library (dplyr)

# bark-fark dönüşümü
 df = norm.bark(formantdata)
 df = rename(df, f1.Bark.dist = "Z3-Z2", f2.Bark.dist="Z2-Z1")
 formantdata2 = cbind(formantdata, df[,5:6])
```

Liste 4. Vowels paketiyle bark-fark dönüşüm kodları

4.3. Ünlü-Dışsal/ Formant-İçsel Ölçekler

Yukarıda da belirtiğimiz gibi, ünlü-dışsal ölçekler konuşurun birden çok ünlüsü arasındaki akustik bilgiyi kullanmaktadır. Bunlardan en bilineni Lobanov (1971)'in z-skor ölçeğidir. Lobanov'un ölçeğinde (bkz. 5), formant değerleri, her bir konuşurun F_1 ve F_2 değerinden ortalamanın (μ_i) farkının alınması ve elde edilen farkın standart sapmaya (σ_i) bölünmesiyle hesaplanmaktadır.

$$F_i^{lobonov} = \frac{F_i^{Hz} - \mu_i}{\sigma_i} \tag{5}$$

Liste (5)'te R'de *phonR* paketi içindeki *normLobanov()* fonksiyonu ile Lobanov (1971)'in z-skor ölçeğinin nasıl hesaplandığı örneklendirilmektedir. *normLobanov()* fonksiyonunda ayrıca eklenebilecek *group* faktörü (vektörü) ile birlikte normalizasyonu gerçekleştirilecek satırları göstermektedir. Liste (5)'te konuşurlara göre konuşur-içsel bir hesaplama da örneklendirilmektedir. Ayrıca, Liste (5)'te normalizasyon işlemi F_1 ve F_2'ye uygulandık-

[5] *Vowels* paketi (benzer olarak NORM aracı) belirli bir sırada veri seti gerektirmektedir. Bunun için veri setinin başlıklarının "speaker","vowel", "context", "F_1", "F_2", "F_3" biçiminde düzenlenmesi gerekmektedir. Diğer bir deyişle, F_1, F_2 ve F_3 değerlerinin 4-6 sütunlarında yer alması gerekmektedir.

tan sonra, Liste (1) ve (2)'de açıklandığı gibi, yeni veri çerçevesindeki sütun adları değiştirilmekte ve veri çerçeveleri birleştirilmektedir.

```
library(phonR)
library(dplyr)

# Lobanov dönüşümü
 lab.df = with(formant.data, normLobanov(cbind(f1,f2), group=Speaker))
 colnames(lab.df) = c("f1.Lobanov", "f2.Lobanov")
 formant.data = cbind(formantdata, lab.df)
```

Liste 5. phonR paketiyle Lobanov dönüşüm kodları

NORM aracının (Thomas ve Kendall, 2007) kullandığı *vowels* paketiyle de Liste (6)'daki kodlar yardımıyla da Lobanov (1971)'in z-skor hesaplaması yapılabilmektedir. *Vowels* paketi içinde *norm.lobanov()* fonksiyonu uygun biçimdeki veri setine bark-dönüşümünü doğrudan uygulamaktadır. Liste (6)'da, *norm.lobanov()* fonksiyonunun ardından *rename()* fonksiyonu ile dönüştürme sonucu oluşturulan veri çerçevesindeki "F*1", ve "F*2" sütun adları yerine yeni adlar atanmakta ve sonrasında ise yeni veri çerçevesi (*df*), özgün veri çerçevesiyle (*formantdata*) birleştirilmektedir. Burada yeni veri çerçevesiyle birleştirme işleminde Liste (4)'tekinin tersine 4–5. sütunlar kullanılmaktadır.

```
library(vowels)
library(dplyr)

# Lobanov dönüşümü
 df = norm.lobanov(formantdata, f1.all.mean=NA, f2.all.mean=NA)
 df = rename(df, f1.Lobanov = "F*1", f2.Lobanov="F*2")
 formant.data = cbind(formantdata, df[,4:5])
```

Liste 6. vowels paketiyle Lobanov dönüşüm kodları

Burada ünlü-dışsal/formant-içsel ölçekler içinde ele alacağımız bir başka ölçek Nearey (1978)'in Tekil Log-Ortalaması Yöntemi (*Single Log-Mean Method*, Adank, Smith & van Hout, 2004) ya da Bireysel Formant Ortalaması Yöntemi (*Individual Formant Mean Method*, Clopper, 2009) olarak adlandırılan ve alanyazında $CLIH_{i2}$ kısaltmasıyla (*Constant Log Interval Hypothesis*) gösterilen ölçektir. Burada *i2* dizini, iki ayrı log-ortalamasının (yani F_1, F_2); söz gelimi *i4* dizini, dört ayrı log-ortalamasının (yani F_0, F_1, F_2 ve F_3) ayrı ayrı hesaplanma durumunu temsil etmektedir. (7)'de de görüldüğü gibi, bu ölçekte normalizasyon her bir konuşur için tüm ünlülerin log-

dönüşümlü formant frekans değerlerinin ortalamasının (μ) log-dönüşümlü formant değerinden çıkarılmasıyla hesaplanmaktadır.

$$F_i^{CLIHi2} = ln(F_i^{Hz}) - \mu_{ln(F_i^{Hz})} \quad (6)$$

Liste (7)'de *phonR* paketi içindeki *normNearley1()* fonksiyonu ile $CLIH_{i2}$ ölçeğinin nasıl hesaplandığı görülmektedir. Bu paket içindeki *normNearley1()* fonksiyonu *normLogmean()* fonksiyonu ile aynı anlama gelmektedir. Diğer taraftan, *vowels* paketi yardımıyla da $CLIH_{i2}$ ölçeği hesaplanabilmektedir. Bu paketle normalizasyon işlemi için *norm.nearey()* fonksiyonu yukarıda belirtiğimiz uygun biçimdeki veri setine uygulanmaktadır. Bu paket içinde *use.f3* öğesine TRUE değeri yüklenmesi log-ortalamasına F_1, F_2 dışında, F_3 değerinin de eklenmesi anlamına gelmektedir.

```
library(phonR)
library(dplyr)

# CLIHi2 dönüşümü
 N1.df= with(formant.data, normNearley1(cbind(f1,f2), group=Speaker))
 colnames(N1.df) = c("f1.Nearley1", "f2.Nearley1")
 formant.data = cbind(formantdata, N1.df)
```

Liste 7. PhonR paketiyle $CLIH_{i2}$ dönüşüm kodları

```
library(vowels)
library(dplyr)

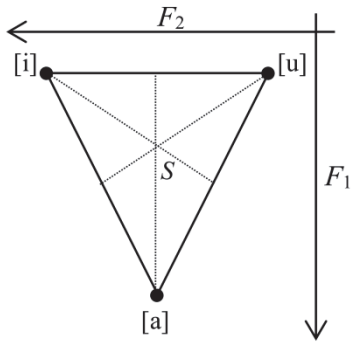
# CLIHi2 dönüşümü
 df = norm.nearey(formantdata, formant.int=TRUE, use.f3=FALSE)
 df = rename(df, f1.nearey1 = "F*1", f2.nearey1="F*2")
 formant.data = cbind(formantdata, df[,4:5])
```

Liste 8. Vowels paketiyle $CLIH_{i2}$ dönüşüm kodları

Ünlü-dışsal/formant-içsel ölçekler içinde son tanıtacağımız ölçek Watt ve Fabricius (2002)'nin ünlü boşluğunun merkezine göre hesaplama yapan ölçeğidir. Bu ölçekte, konuşurun ünlü boşluğu, [i], [a] ve [u′] ünlülerinden oluşan tepe noktaları bulunan bir üçgen olarak ele alınmaktadır, bu noktalar en yüksek ve en düşük F_1 ve F_2 değerlerini temsil etmektedir (bkz. Şekil 3). Bu üçgenin merkezi konumundaki *S* noktası, [i], [a] ve [u′] ünlülerinin büyük ortalamasıyla bulunmaktadır (bkz. 7), bundan sonra formant değerleri *S* noktasına göre hesaplanmaktadır (bkz. 8).

$$S(F_i^{Hz}) = \frac{F_i^{Hz}[i] + F_i^{Hz}[a] + F_i^{Hz}[u']}{3} \quad (7)$$

$$F_i^{w\&f} = \frac{F_i^{Hz}}{\mathrm{S}(F_i^{Hz})} \quad (8)$$



Şekil 3. Watt ve Fabricius dönüşümünde merkez belirleme

Liste (9)'da *phonR* paketi yardımıyla Watt ve Fabricius dönüşümü gösterilmektedir. Bu paket içinde *normWattFabricius()* fonksiyonuyla yapılan işlemde [i] ünlüsünün formant değeri ortalamasını belirlemek yerine sol üst köşedeki ünlünün formant değeri ortalaması alınmaktadır. Böyle olunca da katılımcıların tümünde maksimum F_1 ve F_2 değerlerine sahip ünlüler örtüşmezse *normWattFabricius()* fonksiyonu hata vermektedir. Bu durumda Liste (10)'daki gibi, *phonTools* paketi içindeki *normalize()* fonksiyonuyla köşelerin elle ayarlanması gerekmektedir. Bu paket içinde normalizasyon işlemlerinde *normalize()* fonksiyonu içinde *method* öğesine çeşitli değerler yüklenmesi gerekmektedir: Lobanov dönüşümü için *lobanov*, Nearey1/CLIH$_{i2}$ dönüşümü için *neareyI*, Nearey2/CLIH$_{s4}$ dönüşümü için *neareyE* ve Watt ve Fabricius dönüşümü için *wandf*.

```
library(phonR)
library(dplyr)

# W&F dönüşümü
 df = with(formantdata, normWattFabricius(cbind(f1, f2),
 group = speaker, vowel = vowel))
 colnames(df) = c("f1.WF", "f2.WF")
 formantdata = cbind(formantdata, df)
```

Liste 9. PhonR paketiyle Watt ve Fabricius dönüşüm kodları

```
library(phonTools)
library(dplyr)

# W&F dönüşümü
 df = with(formantdata, normalize(cbind(f1, f2), Speaker, Vowel,
 method = 'wandf', corners = c("i", "u")))
 colnames(df) = c("f1.WF", "f2.WF", "Speaker_n", "Vowel_n")
 formantdata = cbind(formantdata, df)
```

Liste 10. PhonTools paketiyle Watt ve Fabricius dönüşüm kodları

Watt ve Fabricius dönüşümüne ilişkin alanyazında farklı hesaplama türleri bulunmaktadır (bkz. Flynn, 2011). Bunlardan biri, Fabricius, Watt ve Johnson (2009)'un önerisidir, buna göre merkezin x-ekseni hesaplanmaktadır. (9)'da görüldüğü gibi, Watt ve Johnson'un bu ikinci versiyonunda F_2[a] dışarıda bırakılmakta; S merkezi, F_2 ekseninde [i] ve [u]'ya eşit uzaklıkta yer almaktadır. Kısacası merkezi hesaplamada F_1 için (7)'deki formül kullanılırken, F_2 için (9)'daki formül geçerlidir.

$$S(F_2^{Hz}) = \frac{F_2^{Hz}[i] + F_2^{Hz}[u']}{2} \quad (9)$$

Watt ve Fabricius dönüşümüne ilişkin bu ikinci versiyonun hesaplanması *phonTools* ya da *phonR* paketleriyle gerçekleştirilememektedir, ancak *vowels* paketindeki *mod.WF* öğesine FALSE değerinin yüklenmesi durumunda Watt ve Fabricius (2002)'deki dönüşüm, TRUE değerinin yüklenmesi durumunda ise Fabricius, Watt ve Johnson (2009)'daki dönüşüm hesaplanmaktadır (bkz. Liste 11).

```
library(vowels)
library(dplyr)

# W&F dönüşümü
 df = norm.wattfabricius(formantdata, norm.means=FALSE, mod.WF=TRUE)
 df = rename(df, f1.WF = "F*1", f2.WF="F*2")
 formant.data = cbind(formantdata, df[,4:5])
```

Liste 11. Vowels paketiyle Watt ve Fabricius dönüşüm kodları

4.4. Ünlü-Dışsal/ Formant-Dışsal Ölçek

Ünlü-dışsal/formant-dışsal ölçeklerin en bilineni ölçek Nearey (1978)'in Paylaşımlı Log-Ortalaması Yöntemi (*Shared Log-Mean Method*, Adank ve

diğ., 2004) ya da Büyük-Ortalama Yöntemi (*Grand-Mean Method*, Clopper, 2009) olarak adlandırılan ve alanyazında $CLIH_{s4}$ kısaltmasıyla gösterilen ölçektir. Bu yöntemin Nearey1/$CLIH_{i2}$ dönüşümünden farkı her formant frekansı (F_0, F_1, F_2, F_3) değeri için log ortalamalarının büyük ortalamasının alınmasıdır. Alanyazında Nearey2/$CLIH_{s4}$ olarak adlandırılan bu dönüşümün formülü (10)'da sunulmuştur.

$$F_i^{CHIHs4} = \frac{ln\left(F_i^{Hz}\right) - \left(\mu_{\ln\left(F_0^{Hz}\right)} + \mu_{\ln\left(F_1^{Hz}\right)} + \mu_{\ln\left(F_2^{Hz}\right)} + \mu_{ln\left(F_3^{Hz}\right)}\right)}{4} \qquad (10)$$

Nearey2/$CLIH_{s4}$ dönüşümü için R'de *phonR* paketi içinde *normNearey2()* fonksiyonu kullanılabilmektedir (bkz. Liste 12). Görüldüğü gibi, kodlarda F_0, F_1, F_2, F_3 değerlerinin tümü kullanılmaktadır. *Vowels* paketi içinde Nearey2/$CLIH_{s4}$ dönüşümünü uygulamak için Liste (8)'deki *formant.int* öğesine Nearey1/$CLIH_{i2}$ dönüşümünde olduğu gibi TRUE değil, FALSE değeri yüklenmelidir.

```
library(phonR)
library(dplyr)

# CLIHs4 dönüşümü
 df = with(formantdata, normNearey2(cbind(f0, f1, f2, f3),
                          group = Speaker))
 colnames(df) = c("f0.N2", "f1.N2", "f2.N2", "f3.N2")
 formantdata = cbind(formantdata, df)
```

Liste 12. phonR paketiyle Nearey2/$CLIH_{s4}$ dönüşüm kodları

5. Uygulama: Türkçede Formant Normalizasyonu

5.1. Yöntem

Ankara Üniversitesi Beyin Araştırmaları Uygulama ve Araştırma Merkezinde bulunan Dilbilim Laboratuvarında yürütülmüş olan bu çalışmanın örneklemi, anadili Türkçe olan 9 kadın (ort= 20.15, SD= 0.42) ve 11 erkek (ort= 21.34 SD=1.2) olmak üzere toplam 20 katılımcıdan oluşmaktadır. Dış sesten büyük oranda yalıtılmış olan Dilbilim Laboratuvarında *Shure* marka dinamik mikrofon kullanılarak toplanan sözlü dil verileri, erişime açık *Audacity Ses Kayıt Programında* (Crowder, 2015) kaydedilmiş ve *Praat Ses Analizi Programında* (Boersma & Weenink, 2020) analize hazır hale getirilmiştir. Tüm katılımcılar, üniversitede öğrenim görmekte olan ya da üniversiteden mezun bireylerden oluşmaktadır. Katılımcıların, tanı konmuş

herhangi bir işitme bozukluğu, konuşma bozukluğu, nörolojik ya da nöropsikiyatrik bir geçmişi ve aile öyküsü bulunmamaktadır.

Sözlü üretim deneyleri, arkaplanı gri renkten oluşan bir bilgisayar ekranında katılımcılara sunulmuştur. Temel deneylere başlamadan önce tüm katılımcılara yaklaşık 10 adet en küçük çift sunularak, alıştırma deneyi uygulanmıştır. Deneylere alışma süreci tamamlandıktan sonra, temel deneylere geçilmiştir. Bu aşamada, katılımcılara gri arkaplanda tek tek sunulan her bir en küçük çifti en az üçer defa ve çiftler arasında kısa aralıklar bırakarak sesletmeleri istenmiştir. Sözlü üretim deneyleri her bir katılımcı için yaklaşık olarak beşer dakika sürmüştür.

Praat Ses Analizi Yazılımında (Boersma & Weenink, 2020) analiz edilen sözlü üretim verileri, çalışmanın izleyen bölümlerinde ayrıntılı olarak sunulmakta olan formant sentezleme yöntemleri ve algoritmaları eşliğinde R programlama diliyle (R Core Team, 2014) sesbilgisel gözlem tekniği kullanılarak istatistiksel olarak analiz edilmiştir. İstatistiksel analiz öncesinde, tüm sözlü dil verileri, Praat 6.1 ses analizi yazılımında etiketleme (*textgrid*) yöntemi kullanılarak, akustik sesbilgisel analize hazır hale getirilmiştir. Etiketleme yöntemi aşamasında öncelikli olarak tüm sözlü dil verileri temizlendikten sonra, Elvira-Garcia (2013) tarafından geliştirilen *blank_TextGrid_creation* komut dizisi kullanılarak, her bir sözlü dil verisi ayrı ayrı etiketlenmektedir. Bu aşamadan sonra, yine Elvira-Garcia (2017) tarafından geliştirilen *extract_vowel_data* komut dizisi aracılığıyla, sözlü dil verilerine ilişkin formant frekansları, sesletim süreleri ve enerji yoğunluğuna ilişkin değerler elde edilmiştir. Bir sonraki bölümde, elde edilen verilerin normalizasyonu ve istatistiği sunulmaktadır.

5.2. Bulgular

Tablo (4)'te her bir ünlünün ortalama F_1, F_2 ve F_3 formant sıklığı değerleri sunulmuştur. Formant sıklığı ile ilgili Türkçede daha önce yapılmış çalışmalarla (Türk ve diğ., 2005; Davutoğlu, 2010; Malkoç, 2011) bu çalışmada elde edilen bulgular karşılaştırıldığında, F_1 formant değerine ilişkin olarak iki yönlü ANOVA bulguları çalışmalar arasında belirgin bir farklılığın olmadığını göstermektedir: [kadın: $F(3,21)= 3.02$, $p=0.06$[6]; erkek: $F(3,21)= 1.32$, $p=0.293$] Benzer biçimde, F_2 formant değerlerine ilişkin olarak da çalışmalar arasında (Türk ve diğ., 2005; Davutoğlu, 2010; Malkoç, 2011)

[6] Kadın katılımcılarla ilgili post hoc karşılaştırmalar (Tukey HSD) bu çalışmanın bulgularıyla Davutoğlu (2010) arasındaki farkın −106.5 olduğunu (p=0.06) göstermektedir.

farklılık görülmemektedir: [kadın: $F(3,21)= 1.33$, $p=0.290$; erkek: $F(3,21)= 1.59$, $p=0.223$]. F_3 formant değerine ilişkin olarak bu çalışmanın bulgularıyla diğer iki çalışmaların bulguları (Türk ve diğ., 2005; Malkoç, 2011) karşılaştırıldığında, erkek katılımcıların bulguları arasında fark olmadığı [$F(2,14)=1.83$ $p=0.196$], ancak kadın katılımcıların bulguları arasında belirgin bir fark olduğu [$F(2,14)=27.50$, $p<0.001$] görülmektedir. Post hoc karşılaştırmalar (Tukey HSD) bu çalışmadaki bulgularla diğer çalışmaların bulguları arasında $p<0.05$ düzeyinde farklılığın olduğunu, diğer iki çalışma arasında ise $p<0.001$ düzeyinde farklılığın olduğunu göstermektedir.

Tablo 4. Ünlülere ilişkin ortalama formant frekans ve SS değerleri

	F_1^{Hz}		F_2^{Hz}		F_3^{Hz}	
Kadın						
[ɑ]	885.56	(87.37)	1407.00	(88.73)	2758.78	(160.99)
[ɛ]	679.00	(87.93)	2016.44	(427.15)	2824.22	(214.21)
[ɯ]	535.22	(50.98)	1683.33	(82.20)	2749.89	(126.43)
[i]	426.78	(47.71)	2435.56	(115.97)	3111.33	(196.97)
[o]	657.78	(70.06)	1066.11	(114.95)	2769.00	(215.09)
[œ]	661.00	(63.23)	1541.22	(122.89)	2549.33	(55.32)
[ʊ]	482.11	(32.56)	1102.11	(79.88)	2783.56	(123.83)
[ʏ]	448.00	(36.26)	1875.67	(100.87)	2496.33	(106.77)
Erkek						
[ɑ]	687.73	(40.66)	1209.18	(89.08)	2625.36	(156.44)
[ɛ]	511.36	(32.11)	1834.64	(81.48)	2557.91	(197.79)
[ɯ]	442.73	(46.13)	1524.55	(95.03)	2443.09	(148.66)
[i]	340.55	(34.09)	1999.18	(121.54)	2722.82	(106.52)
[o]	540.36	(48.68)	912.27	(88.66)	2630.36	(176.72)
[œ]	516.73	(35.32)	1335.64	(76.87)	2322.09	(172.02)
[ʊ]	370.73	(74.21)	985.00	(84.26)	2451.64	(155.99)
[ʏ]	328.09	(46.98)	1630.82	(80.41)	2199.82	(133.17)

Buradaki analizlerde, ünlü-içsel/formant-içsel normalizasyon işlemi olarak *phonR* paketi yardımıyla, log, bark, mel ve ERP dönüşümleri (bkz. sırasıyla Liste 1 ve 2) uygulanmıştır. Ünlü-içsel/formant-dışsal normalizasyon işlemlerinden olan bark-fark dönüşümünün hesaplanmasında ezgi, ton ve ünsüzlerle ilgili etkenlerin F_0'ı etkilemesinin önüne geçmek için (Kendall & Thomas, 2018), *vowels* paketi kullanılmıştır (bkz. Liste 4). Diğer yandan, ünlü-dışsal/formant-içsel normalizasyonu için *phonR* paketi yardımıyla sırasıyla Lobanov, Nearey1/$CLIH_{i2}$ ve Watt ve Fabricius dönüşümleri (bkz. Liste 5, 7 ve 9-11) uygulanmıştır. Son olarak, ünlü-dışsal/formant-dışsal

işlemi olarak *phonR* paketi yardımıyla Nearey1/CLIH$_{s4}$ dönüşümü (bkz. Liste 12) uygulanmıştır.

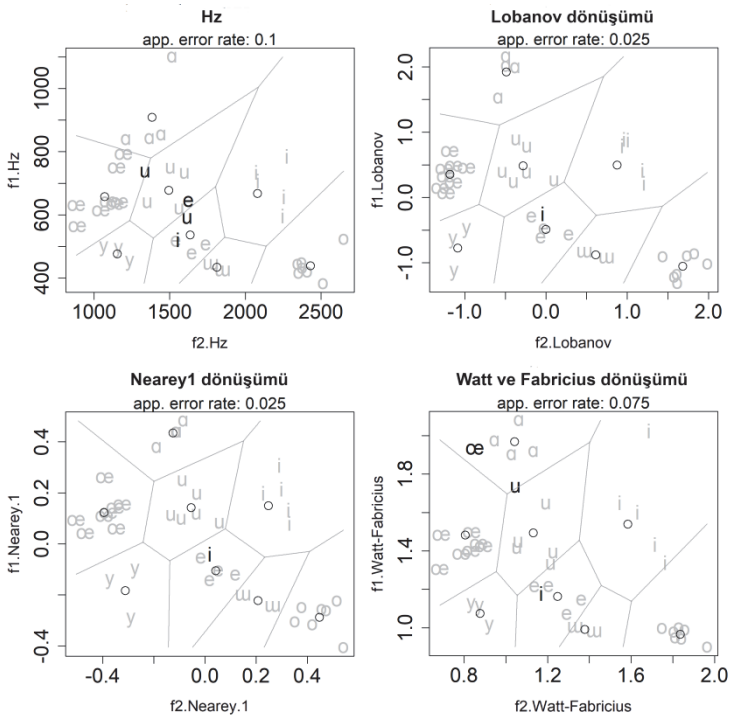
Ünlülerin formant analizinde kullanılacak normalizasyon yöntemlerinin hangisinin sınıflandırmada daha iyi sonuç verdiğini belirlemek için belirli özelliklere sahip birimleri, doğadaki gerçek sınıflarına en doğru şekilde atamayı amaçlayan çok değişkenli istatistiksel bir yöntem olan azaltılmış dereceli (reduced-*rank*) Doğrusal Ayırıcı Analiz (*LDA, Linear Discriminant Analysis*) ve Karesel Ayırıcı Analiz (*QDA, Quadratic Discriminant Analysis*) kullanılmıştır. Bu çalışmadaki LDA ve QDA analizlerinde R dilinde (R Core Team, 2014) *MASS* paketi (Venables & Ripley, 2002) içinde yer alan *lda()* ve *qda()* fonksiyonları kullanılmıştır, ayrıca görselleştirmede *klaR* paketi (Weihs, Ligges, Luebke & Raabe, 2005) içinde *partimat()* fonsiyonu kullanılmıştır. Tablo (5)'te dokuz ayrı normalizasyon işlemine dayalı hesaplamalardan elde edilen bulguların LDA ve QDA doğruluk değerleri görülmektedir. Tablo (5)'te görüldüğü gibi, gerek LDA'da gerekse QDA'da Nearey2/CLIH$_{i4}$ dönüşümü dışındaki hiçbir dönüşüm modelinde referans değeri koşulundaki (Hz) doğruluk ortalamasının (sırasıyla %77 ve %75) altında bir değer görülmemektedir. En yüksek değerler de birden çok ünlü kategorisinin dağılımının gerektiği, diğer bir deyişle konuşurun birden çok ünlüsü arasındaki akustik bilginin kullanıldığı ünlü-dışsal/formant-içsel dönüşüm türlerinde (Lobanov, Nearey1/CLIH$_{i2}$ ve Watt ve Fabricius) görülmektedir. Bu nedenle burada bu üç dönüşüm işleminin sonuçlarına odaklanılacaktır.

Tablo 5. LDA doğruluk ortalamaları

Dönüşüm türleri	LDA	QDA
Hz	77	75
log	80	75
Bark	80	75
Mel	80	75
ERB	80	75
Bark-fark	85	85
Lobanov	97	95
Nearey1/CLIH$_{i2}$	95	95
Watt ve Fabricius	95	92
Nearey2/CLIH$_{s4}$	70	70

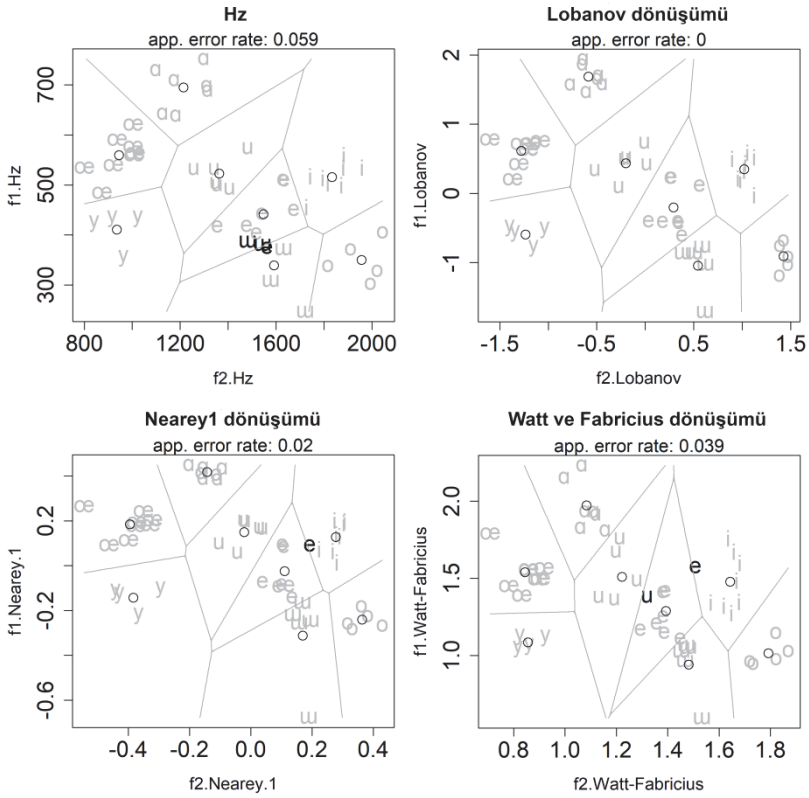
LDA analizine bağlı olarak Lobanov, Nearey1/CLIH$_{i2}$ ve Watt ve Fabricius dönüşüm işlemlerinin sonuçları sonsal olasılıklar (*posterior probabilities*) kullanılarak yorumlanacaktır. Kadın katılımcılarda, referans

değeri koşulunda (Hz) [o] ünlüsünün sonsal olasılığı %100'dür, buna karşın [ɛ] ve [ʊ] ünlülerinde sırasıyla %68 ve %69'dur. Diğer ünlülerde oran %82'nin üzerindendir. Bu görünümüyle referans değeri koşulunda büyük oranda belirsizlik olduğu görülmektedir (bkz. Şekil 4). Nearey1/$CLIH_{i2}$ dönüşümünde ise [ɯ] ve [i] ünlülerine ilişkin değerler %83'tür, diğer değerler ise %94'ün üzerindedir. Watt ve Fabricius dönüşümünde ise [ɯ] ve [ɛ] ünlülerine ilişkin sırasıyla %67 ve %68 sonsal olasılık gözlenirken [i] ve [ʊ] ünlülerinde sırasıyla, %77 ve %71 sonsal olasılık görülmektedir. Genel olarak tüm dönüşüm işlemlerinin sonuçları [o] ünlüsünün en yüksek sonsal olasılık sergilediğini göstermektedir, Watt ve Fabricius'ta da bu oran %100'dür. Lobanov dönüşümü belirsizliğin en az görüldüğü dönüşüm işlemidir. Bu işlemde [i] ünlüsü %83 oranında en düşük sonsal olasılık gösterirken diğer ünlüler %96'nın üzerinde bir olasılık göstermektedir (bkz. Şekil 4). Hata oranlarını incelediğimizde Lobanov ve Nearey1/$CLIH_{i2}$ dönüşümünde dönüşümlerinin %2.5, Watt ve Fabricius dönüşümünün %7.5, temel (*baseline*) koşulunun (Hz) ise %10, hata oranına sahip olduğu görülmektedir.



Şekil 4. Kadın katılımcılarda LDA sonuçları

Erkek katılımcılarda, her bir ünlünün ayrı ayrı sonsal olasılık dağılımlarını incelediğimizde, referans değeri koşulunda (Hz), [ɯ] ve [ɛ] ünlülerinin sırasıyla %63, %78 sonsal olasılık değerlerine sahip oldukları (bkz. Şekil 5), diğer ünlülerin %95 üzerinde sonsal olasılık sergiledikleri gözlenmektedir. Nearey1/CLIH$_{i2}$ dönüşümünde ise [ɯ] ve [ɛ] ünlülerine ilişkin değerlerin sırasıyla %86 ve %88'e ulaştığı gözlenmektedir. Watt ve Fabricius dönüşümünde [ɯ], [ɛ] ve [ʊ] ünlülerine ilişkin sırasıyla %85, %80 ve %83 sonsal olasılık gözlenirken diğer ünlülerde bu oran %87'nin üzerindedir. Buna karşın Şekil (5)'te de açıkça gözlenebildiği gibi, Lobanov dönüşümünde belirsizlik gözlenmemekte, tüm ünlüler %97'nin üzerinde sonsal olasılık değerleri sergilemektedir (bkz. Şekil 5). Hata oranları da bu sonuçları doğrulamaktadır. Lobanov dönüşümünde hata gözlenmezken, referans değeri koşulunda (Hz), %6, Nearey1/CLIH$_{i2}$ dönüşümünde %2, Watt ve Fabricius dönüşümünde ise %4 hata oranı görülmektedir.



Şekil 5. Erkek katılımcılarda LDA sonuçları

6. Sonuç

Bu çalışmada akustik veriyi konuşurlar arasında karşılaştırılabilir kılmak için uygulanan ünlü normalizasyon işlemleri genel hatlarıyla tanıtılarak R programlama dilinde (R Core Team, 2014) toplam üç pakette (*phonR*, *vowels*, *phonTools*) ünlü normalizasyonunun nasıl yapılacağı örnekleriyle sunulmuştur. Alanyazında pek çok normalizasyon yöntemi bulunmakla birlikte bu yazıda R programlama dilinde ele aldığımız paketlerin izin verdiği normalizasyon yöntemleri tanıtılmıştır.

Bu çalışmada, R'de normalizasyon işlemini örneklendirmek amacıyla, Türkçedeki sekiz ünlüye ([ɑ], [ɛ], [ɯ], [i], [o], [œ], [ʊ], [ʏ]) dokuz farklı normalizasyon işlemi uygulanmıştır. Uygulanan normalizasyon yöntemlerinin hangisinin sınıflandırmada daha belirgin sonuç verdiğini belirlemek için LDA ve QDA kullanılmıştır. Analizler, ünlü-dışsal/formant-içsel yöntemlerin Türkçede ünlüleri ayırma konusunda daha belirgin sonuçlar verdiğini göstermektedir.

Ünlü-dışsal/formant-içsel üç dönüşüm işlemine (Lobanov, Nearey1/$CLIH_{i2}$, Watt ve Fabricius) ve referans koşulu olarak da Hz'ye ilişkin LDA ve QDA bulguları betimsel olarak karşılaştırıldığında en belirgin ayırıcı dönüşüm işleminin Lobanov olduğu görülmektedir. Daha sonra da Nearey1/$CLIH_{i2}$ ve Watt ve Fabricius gelmektedir. Yine de üç dönüşüm yöntemi arasında çok belirgin bir farklılığın olduğunu söylemek güçtür.

LDA Lobanov > Nearey1/$CLIH_{i2}$ ≥ Watt ve Fabricius > Hz

QDA Lobanov ≥ Nearey1/$CLIH_{i2}$ > Watt ve Fabricius > Hz

Bulgular Fabricius, Watt ve Johnson (2009)'un bulgularıyla benzerlik göstermektedir. Fabricius ve diğerleri de çalışmalarında Lobanov yönteminin diğerlerine göre daha başarılı olduğunu ileri sürmekte, ancak farklı analizlerde Watt ve Fabricius ile benzer ayırıcı nitelikler taşıdığını ortaya koymaktadır. Tüm bu bulgular, ünlülerin sınıflandırılmasında akustik veriler arasındaki ayrımları belirginleştirmek amacıyla Hz yerine ünlü-dışsal/formant-içsel üç dönüşüm işlemlerinin kullanılmasının daha belirgin sonuçlar sunacağını göstermektedir.

Kaynakça

Adank, P. (2003). *Vowel normalization: A perceptual-acoustic study of Dutch vowels*. [Yayımlanmamış Doktora Tezi]. University of Nijmegen.

Adank, P., Smits, R., & van Hout, R. (2004). A comparison of vowel normalization procedures for language variation research. *Journal of the Acoustical Society of America* 116, 3099-3107.

Ainsworth, W. A. (1975). Intrinsic and extrinsic factors in vowel judgements. İçinde G. Fant, & M. A. A. Tatham (Yay.), *Auditory analysis and perception of speech* (ss. 103-113). Academic Press.

Boersma, P., & Weenink, D. (2020). *Praat: Doing phonetics by computer [bilgisayar yazılımı]*. Version 6.1.09.

Clopper, C. G. (2009). Computational methods for normalising acoustic vowel data for talker differences. *Language and Linguistic Compass 3*(6), 1430-1442.

Çiyiltepe, M., Bekar, İ. P., & Ergenç, İ. (2009). Changing formant values: Synthesis of four regions of Turkey. İçinde S. Ay, Ö. Aydın, İ. Ergenç, S. Gökmen, S. İşsever, & D. Peçenek (Yay.), *Essays on Turkish linguistics: Proceedings of the 14th international conference on Turkish linguistics, Turcologica series, 79* (ss. 3-9). Harrassowitz Verlag.

Davutoğlu, A. Y. (2010). *Standart Türkçedeki ünlülerin akustik analizi ve fonetik altyapı*. [Yayımlanmamış Doktora Tezi]. İstanbul Üniversitesi.

Elvira-Garcia, W. (2013). Blank-TextGrid-creation. v.1. [praat script].

Elvira-Garcia, W. (2017). vowelFormants. v.1. [praat script].

Fabricius, A. H., Watt, D., & Johnson, D. E. (2009). A comparison of three speaker-intrinsic vowel formant frequency normalization algorithms for sociophonetics. *Language Variation and Change 21*(3), 413-435.

Flynn, N. (2011). Comparing vowel formant normalisation procedures. *York Papers in Linguistics Series 2*(11), 1-28.

Gerstman, L. (1968). Classification of self-normalized vowels. *IEEE Transactions of Audio Electroacoustics AU-16*, 78-80.

Heeringa, W., & van de Velde, H. (2018). *Visible vowels: A tool for the visualization of vowel variation*. Proceedings CLARIN annual conference. 8-10 Ekim 2018, İtalya. CLARIN ERIC.

Hindle, D. (1978). Approaches to formant normalization in the study of natural speech. İçinde D. Sankoff (Yay.), *Linguistic variation, models and methods* (ss. 161-171). Academic Press.

Johnson, K. (1990). The role of perceived speaker identity in F0 normalization of vowels. *Journal of the Acoustical Society of America* 88, 642-654.

Kendall, T., & Thomas, E. R. (2018). Package 'vowels': Vowel manipulation, normalization, and plotting. Version 1.2-2. (https://cran.r-project.org/web/packages/vowels/index.html).

Kent, R. D., & Read, C. (2002). *The acoustic analysis of speech*. Singular/Thomson Learning Yayınları. (2. Baskı).

Kılıç, M. A. (2003). *Türkiye Türkçesi'ndeki ünlülerin sesbilgisel özellikleri*. Proceedings of the 10th international conference in Turkish linguistics (ss. 3-18).

Kılıç, M. A., & Öğüt, F. (2004). A high unrounded vowel in Turkish: Is it a central or back vowel? *Speech Communication, 43*(1-2), 143-154.

Kinsler, L. E., Frey, A. R., Coppens A. B., & Sanders, J. V. (2000). *Fundamentals of acoustics.* John Wiley.

Korkmaz, Y., & Boyacı, A. (2018). Classification of Turkish vowels based on formant frequencies. *2018 International conference on artificial ıntelligence and data processing (IDAP)* (ss. 1-4), IEEE.

Labov, W. (2001). *Principles of linguistic change (2): Social factors.* Blackwell Publishers.

Labov, W., Ash, S., & Boberg, C. (2006). The atlas of north American English: Phonology, phonetics, and sound change. A multimedia reference tool. de Gruyter Mouton.

Ladeforged, P., & Broadbend, D. E. (1957). Information conveyed by vowels. *Journal of the Acoustical Society of America* 29, 98-104.

Lobanov, B. M. (1971). Classification of Russian vowels spoken by different speakers. *Journals of the Acoustical Society of America* 49, 606-608.

Malkoç, E. (2011). Türkçe ünlü formant frekans değerleri ve bu değerlere dayalı ünlü dörtgeni. *Dil Dergisi* 146, 71-85.

McCloy, D. R. (2012). Vowel normalization and plotting with the phonR package. *Technical Reports of the UW Linguistic Phonetics Laboratory* 1, 1-8.

McCloy, D. R. (2016). phonR: tools for phoneticians and phonologists. R package, 1-0.7 (https://cran.r-project.org/web/packages/phonR/index.html).

Miller, J. D. (1989). Auditory-perceptual interpretation of the vowel. *Journal of the Acoustical Society of America, 85*(5), 2114-2134.

Moore, B. C., & Glasberg, B. R. (1983). Suggested formulae for calculating autory-filter bandwidths and excitation patterns. *Journal of the Acoustical Society of America* 74, 750-753.

Nearey, T. M. (1978). *Phonetic feature systems for vowels*. [Yayınlanmış Doktora Tezi, Indiana University]. Indiana University Linguistics Club.

Nordström, P. E., & Lindblom, B. (1975). *A normalization procedure for vowel formant data*. Proceedings of the 8th international congress of phonetic sciences. Leeds.

Peterson, G. E., & Barney, H. L. (1952). Control methods used in the study of the vowels, J. *Acoustical Society of America* 24, 175-184.

R Core Team. (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing. Austria. URL http://www.R-project.org/.

Rosner, B.S., & Pickering, J. B. (1994). *Vowel perceptation and production*. Oxford University Press.

Stevens, S. S., & Volkmann, J. (1940). The relation of pitch to frequency: A revised scale. *American Journal of Psychology* 53, 329-353.

Syrdal, A. K. (1985). Aspects of a model of the auditory representation od American English vowels. *Speech Communication* 4, 121-135.

Syrdal, A. K., & Gobal, H. S. (1986). A perceptual model of vowel recognition based on the auditory representation of American English vowels. *Journal of the Acoustical Sociaety of America* 79, 1086-1100.

Thomas, E. R., & Kendall, T. (2007). NORM: The Vowel Normalization and Plotting Suite. Erişim için Online Resource. http:ncslaap.lib.ncsu.edutoolsnorm, accessed:171108.

Traunmüller, H. (1981). Perceptual dimensions of openness in vowels. *Journal of the Acoustical Society od America* 69, 1465-1475.

Traunmüller, H. (1990). Analitic expressions fort he tonotopic frequency scale. *Journal of the Acoustical Society of America* 88, 97-100.

Türk, O., Şayli, Ö., Özsoy A. S., & Arslan L. M. (2004). *Türkçede ünlülerin formant analizi*. 18. Ulusal dilbilim kurultayı. Ankara Üniversitesi.

Venables, W. N., & Ripley, B. D. (2002). Modern Applied Statistics with S. (4. Baskı). ISBN 0-387-95457-0, Erişim için http://www.stats.vox.ac.uk/pub/ MASS4. Springer.

Watt, D. J. L., & Fabricius, A. H. (2002). Evaluation of a technique for improving the mapping of multiple speakers' vowel spaces in the F1-F2 plane. *Leeds Working Papers in Linguistics and Phonetics* 9, 159-173.

Weihs, C., Ligges, U., Luebke, K., & Raabe, N. (2005). klaR analyzing German business cycles. İçinde D. Baier, R. Decker, & L. Schmidt-Thieme (Yay.), Data analysis and decision support (ss. 335-343). Springer.

Kısaltmalar

F_0	Temel Sıklık
ERB	Eşdörtgensel Bant Genişliği (*Equivalent Rectangular Bandwidth*)
F_1	Birinci Formant Frekansı
F_2	İkinci Formant Frekansı
F_3	Üçüncü Formant Frekansı
F_4	Dördüncü Formant Frekansı
F_5	Beşinci Formant Frekansı
LDA	Doğrusal Ayırıcı Analiz (*Linear Discriminant Analysis*)
QDA	Karesel Ayırıcı Analiz (*Quadratic Discriminant Analysis*)

Bölümde Kullanılan Terimler Listesi

Açılım	Antinode
Alçak Ünlüler	Low Vowels
Algısal	Perceptual
Anatomik/Fiziksel Bilgi	Anatomical/ Physical Information
Azaltılmış Dereceli	Reduced-Rank
Bant	Band
Biçimleyici	Formant
Bireysel Formant Ortalaması Yöntemi	Individual Formant Mean Method
Boğum	Node
Büyük-Ortalama Yöntemi	Grand-Mean Method
Dışsal	Extrinsic
Dizi Normalizasyonu	Range Normalization
Doğrusal Ayırıcı Analiz	Lda, Linear Discriminant Analysis

Enerji Yoğunluğu	Intensity
Eşdörtgensel Bant Genişliği	Equivalent Rectangular Bandwidth
Etiketleme	Textgrid
F_0	Fundamental Frequency
Formant	Formant
Formant Frekansı	Formant Frequency
Gürültü	Noise
İçsel	Intrinsic
Karesel Ayırıcı Analiz	Qda, Quadratic Discriminant Analysis
Normalizasyon	Normalization
Oktav Atlaması	Octave Jump
Paylaşımlı Log-Ortalaması Yöntemi	Shared Log-Mean Method
Perde	Pitch
Rezonans	Resonance
Sabit Log-Aralığı Hipotezi	Constant Log Interval Hyphotesis, $Clih_{i2}$
Sesbirimsel Bilgi	Phonemic Information
Ses-Yolu Ölçeği	Vocal-Track Scaling
Sonsal Olasılıklar	Posterior Probabilities
Temel	Baseline
Temel Sıklık	Fundamental Frequency
Toplumdilbilimsel Bilgi	Sociolinguistic Information
Ünlü Normalizasyonu	Vowel Normalization
Yüksek Ünlüler	High Vowels
Z-Skor Dönüşümü	Z-Score Transformation