Pamukkale Üniversitesi Sosyal Bilimler Enstitüsü Dergisi (2008) 1, 145-164

MÜZİKTE İNSAN SESİ ÜRETİM MODELLERİ VE İŞLEME TEKNOLOJİLERİ ÜZERİNE BİR İNCELEME: TINI ÜRETİM MODELLERİNDEN MÜZİKSEL DOĞRULAMAYA GEÇİŞ

Cihan IŞIKHAN*

Anahtar kelimeler:
Müzik Teknolojisi, Konuşma Makineleri, Ses İşleme, Müzik Doğrulama

Key words:
Music Technology, Speaking Machines, Sound Processing, Music Correction

Özet

Tarihsel süreçte insan sesinin müzikle olan birleşimi, teknolojik açıdan da aynı çizgide bir gelişim süreci izler. İlk olarak konuşma sesi ve hemen ardından şarkı sesi üretim modelleri yaratılır ve bu modeller, müzikte tını üretim kaynağı olarak kullanılabilir bir serüvenin başlangıcını oluşturur. Ancak, kültürel bir birikim olarak ortaya çıkan toplumsal müzik kültürü ve dolayısıyla müzik, insan sesi üretimi ve işleme sırasında teknolojik olanaklarla birleştiğinde kuralcı bir zemine oturma eğilimi gösterir. Önceleri yalnızca tınısal üretim modeli olarak görülen insan sesi aygıtları, teknoloji ilerledikçe detay ses işlemeye doğru yönelir ve özellikle profesyonel ses kayıt endüstrisinde müziksel doğrulama sürecini başlatır. Müziksel doğrulama kısaca, müzik kuramının aynen edime yansıtılma çabası ya da tam tersi, edimin kuralcı bir zemine çekilmesidir. Günümüzde ses işlemeye yönelik teknolojik olanakların ilerlemesiyle birlikte kültürel açıdan hiçbir kaygı olmadığı durumlarda bile üretilen her tür müzikte, müziksel doğrulama, ses kayıt sektöründe bir refleks haline gelmiş ve kendi standardını yaratmıştır. Böyle bir geçiş ve standardın teknolojik sürecini göstermek adına müzikte insan sesi örneğini temel alan bu yazıda, insan sesi üretim modelleri ve işlemenin tarihsel süreçte yapı taşı kabul edilebilecek aygıtlar, teknik ve yöntemler anlatılmıştır. Üretimden analize doğru gelinen tarihsel süreçte konuşma ve şarkı sesi arasındaki teknik farklar belirtilmiş ve tını üretim modellerinden müziksel doğrulamaya geçişin göstergesi, insan sesi özelinde tarihsel bir süreçte araştırılmıştır.

Abstract

An Investigation of Vocal Production Models and Processing Technology in Music: The Transformation from the Production of Timbre Models to Music Correction

The combination of vocal with music follows the same way of technological development in process of historical background. Production models that are used a source of timbre in music are created as a speaking models and singing models, respectively. However, a social

* Öğr. Gör. Dr., Dokuz Eylül Üniversitesi, Güzel Sanatlar Fakültesi Gündoğdu Sokak No:4 Balçova/İzmir. e-posta: cihan.isikhan@deu.edu.tr

music culture appeared as an experimentation of culture, of course music, inclined to full of rules when combining with possibilities of technology for producing and processing of vocal. All of the vocal machines which seem to be only a timbre model previously go forward to analytic sound processing during the time that the technology grows up and have the process of music correction started especially in professional sound recording industry. Music correction, shortly, is an effort to reflect the music theory to performance or vice versa. Nowadays, owing to facilities of technology according with sound processing improve, music correction seems to be a reflection and creates its standard on by itself in every kind of music which is produced that even more there is no preoccupation as a culture. In this paper, basically, it is discussed the process of music technology in an example of vocal production and processing. It is pointed out the technical different between speaking and singing voice and tried to delineated transformation from the production of timbre models to music correction in music technology.

1. GİRİŞ

Dünya'da mühendislik ve/veya sosyal bilimler alanı altında kapsamlı bir çalışma alanı olan müzik teknolojisi, maalesef ülkemizde henüz emekleme aşamasında diyebileceğimiz bir geçmişe sahiptir. Müzik - teknoloji ilişkisinde amaç ve sonuç ister müzik isterse teknoloji hedefli olsun, ülkemizde bu alanda da bilgilendirme kaçınılmaz boyutlardadır. Çünkü teknoloji öyle bir yapıdır ki her alanda kolayca yaşayabilir ve her alandan rahatça beslenebilir (Serhad, 1997). Müzik teknolojisi perspektifinde bilgilendirmek adına, müzik ve teknolojinin iç içe girdiği en önemli konulardan biri de, insan sesinin teknolojiye yansıma süreci ve bu süreçte, günümüz teknolojisinin sunduğu olanaklarla ortaya çıkan sonuçlarıdır.

Teknolojik sonuçlar, kimi zaman arz talep dengesi etrafında kısa bir sürede ama çoğu kez daha hızlıyı ve daha doğru bulma adına uzun bir evrimsel süreçte ortaya çıkar. Bu serüven öyle sonuçlar ortaya çıkarır ki, internet gibi küresel düzeyde sanal bir omurganın topluma olan etkisi gibi bir sonuçla da karşılaşılabilir; üretici - tüketici ilişkisiyle yıllar süren bir haberleşme protokolünün müzikte standartlaşmasına da başlangıç oluşturabilir. Dolayısıyla teknolojisi, her yönüyle müziği ve toplumu etkileyebilecek bir güce sahiptir. Teknolojik boyutuyla müzik teknolojisinde insan sesinin de kendi içinde dinamik bir serüveni vardır ve bu serüven, daha gerilere gidildiğinde müzik teknolojisinin de yapı taşı kabul edilebilecek olan ses kavramını ve onun oluşumunu temel alır.

Ses oluşumunun doğal süreci, ses kaynağı ile başlar ve algılayıcı ile tamamlanır. Bu süreç içinde algılama, doğrudan kulakla ilgilidir ve dolayısıyla bu kavram, genelde biyolojik bir zemine oturtulur. Bu durumun tam tersi olarak kaynak

kavramını niteleyen düzlem, kaynağın yaratımsal sonucu olan titreşimin kendisiyle açıklanabilir. Dolayısıyla titreşimi oluşturan kaynak, canlı-cansız hemen her türlü -şey- olabilir. Fiziksel bir bakış açısıyla bu şekilde tanımlanabilecek bir oluşum, müzik teknolojisi zeminine çekildiğinde, kaynak-algılayıcı ilişkisi altında daha net bir üslup kazanır. Algılayıcı yine biyolojik bir nesne olsa bile kaynak, teknolojik bir bakış açısıyla ve net bir ifadeyle 'titreşimdir'.

Müziği oluşturan hemen her türlü titreşim kaynağının çalgı olarak tanımlanabildiği gibi, bu çalgılardan biri de hiç kuşkusuz insan sesidir. Önceleri salt bir iletişim aracı olarak görülen insan sesinin müzikle olan birleşmesi, teknolojik perspektifte de aynı çizgide bir gelişim süreci izler. Teknolojik süreçte ilk olarak konuşma sesi ve hemen ardından şarkı sesi üretim modelleri yaratılır ve bu modeller, çalgısal olarak kullanılabilecek müziksel, teknolojik ve sosyal bir serüvenin başlangıcını oluşturur.

İnsanda anatomik olarak ses oluşumu, bir başka ifadeyle insan sesi kaynağı, akciğer/diyafram bölgesinden gırtlağa (larynx) itilen hava ile başlar. İletici ortam olan hava, gırtlak bölgesindeki ses tellerini (vocal folds) hareket ettirir ve titreşim oluşur. Titreşim, ağız boşluğu bölgesinden (vocal tract) genliği yükseltilerek (amplification) dışarıya, algılayıcıya iletilir.

Ses telleri, üretimin ilk oluştuğu kaynaktır ve fizyolojide ses kalitesi için anahtar faktör oluşturur (Kim, 2003: 20). Bu bölgede sesin oluşumu, ses tellerinin titreşimiyle elde edilir ve hızlıca tekrarlanan bu titreşim biyolojide fonasyon (phonation) olarak tanımlanır. Titreşimin doğal sonucu olarak bu noktada temel frekans (fundamental frequency) elde edilir. Fonasyon sırasında diyaframdan gelen hava akımı herhangi bir yere çarpmadan ağız yoluyla titreşimi doğrudan algılayıcıya iletebiliyorsa ötümlü (voiced) sesler elde edilir. Bu durumun tam tersi, yani titreşimle oluşmuş hava akımının engellerle karşılaşması yoluyla da ötümsüz (unvoiced) sesler ortaya çıkar. Ötümlü seslerin frekans ve armonik yapısı, ötümsüzlere göre daha düzenli ve analizi nispeten daha kolaydır.

Ağız boşluğu bölgesi dil, diş, burun, dudak vb. organları içinde barındırır ve bu boşluğun şekilsel yapısı insandan insana farklılık gösterdiğinden genlik veya tını gibi değişkenlerin belirleyicisidir. Fonasyon yoluyla oluşan temel frekans üzerine bu bölgede yeni frekanslar eklenir ve tınısal yapı tamamlanır. Bu tınısal yapıya formant (formant), formantı oluşturan enerjisi en yüksek belirleyici frekanslara ise formant frekansı (formant frequency) adı verilir. Örneğin Latin alfabesinde ünlü harfleri belirlemede genelde ilk üç formant frekansı yeterlidir (Turk vd, 2004). Formantı belirleyen en önemli etkenlerden biri de ağız boşluğunda titreşim genliğinin yükseltilmesidir. Bu nedenle ağız boşluğu dil, diş, burun, dudak vb. organlarla birlikte silindirik bir tınlaşım (resonances) kutusuna benzetilebilir (Kim, 2003: 21).

İnsan sesi üretiminin teknolojiye yansıma öyküsü ise somut olarak 18.yy. sonlarında başlar. Teknolojik olarak görülen ilk modeller, konuşmada üretilen ötümlü harflerin mekanik aygıtlarda benzetilmesi olmuştur. İlerleyen yıllarda bu aygıtlar, daha karmaşık ama giderek insan sesine benzeyen sonuçlarıyla daha gerçekçi bir oluşum sağlar. 20.yy. ortalarından itibaren teknolojinin ilerlemesiyle nispeten daha kolay gibi görülen konuşma sesi üretim modelleri, yerini hızlıca müzikte kullanılabilecek şarkı sesi üretim modellerine bırakır.

Şarkı sesi, konuşma sesinin teknolojik olarak bir adım daha ötesinde bir yaygınlaşma gösterir. Bir iletişim aracı olan insanda konuşma sesi, yerini, en basit anlamda eğlence aracı olarak görülebilecek müziğe bırakır ve bu kapsamda konuşma sesleri giderek şarkı seslerine dönüşür. Bunun doğal sonucu olarak da şarkı sesi üretim modellerine yönelik aygıtlar üretilmeye başlanır. Özellikle elektronik müziğin beslenme kaynağını oluşturabilecek düzeyde aygıtlar geliştirilir. Bu aygıtlardan en önemlisi ise hiç kuşkusuz VOCODER'dır. VOCODER, teknolojik süreçte öyle etken bir duruma gelir ki, 2000'li yıllarda kayıt stüdyolarında ve dolayısıyla müzik teknolojisi çalışanları arasında "Cher Efekti" olarak farklı isimle bir standartlaşmanın da atası olur.

Şarkı sesi üretiminde doyum noktasını ise kayıt olanaklarının gelişmesiyle birlikte dijital ortamda örneklenebilen sesler oluşturur. Üretim modellerinin en üst noktası olarak kabul edilebilecek örnekleme yöntemiyle hemen her çeşit insan sesi müzikte rahatça bir model olarak kullanılabilir hale gelir.

Diğer taraftan, bilimsel gelişimini sürdüren müzik teknolojisi, PSOLA ve LPC gibi yöntem ve tekniklerin yaratılmasıyla, Karaoke gibi ses işlemeli eğlence aygıtlarının üretilmesiyle, insan sesinin üretim modelinden giderek onun işlenmesi sürecini hızlandırır. Nihayet 21.yy. başlarından itibaren bilgisayar teknolojisinin baş döndürücü bir hızla gelişmesiyle birlikte Auto-Tune, Melodyn gibi insan sesi üst düzey işleme yazılımları ve dolayısıyla insan sesinin işleme süreci hemen her kullanıcının edinebileceği bir zemine taşınmış olur.

İşte bu nokta, yani insan sesi üretim modelleriyle başlayan ve onun daha detaylı işlenmesiyle doruk noktaya ulaşan teknolojik oluşum, birtakım sosyal sonuçlara yol açar. Artık hemen-hemen her ses kayıt stüdyosundaki sektör çalışanları, müzikte biçim ya da biçem gözetmeksizin ses işleme aygıtlarını kullanır hale gelir. Böylece daha detaylı, daha analitik, edim ile kuram ilişkisinde yok denecek kadar az -hatalı-, kısacası çok daha -doğru- müzikler üretilmeye başlanır. Bu başlangıç, aslında, tınısal üretim modellerinden müziksel doğrulamaya geçişin bir göstergesini oluşturur.

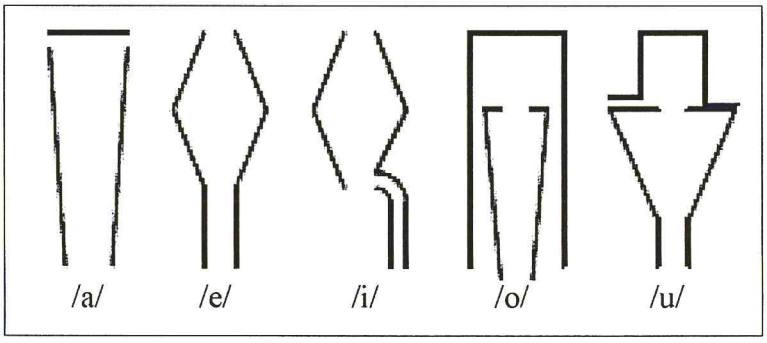
Tınısal üretim çerçevesinde etnomüzikolojik bakış açısıyla müzikte tını üzerine yapılan açıklamaların pek çoğu, tınının ses kaynağı ve konumuna dair en kesin bilgiyi verdiği üzerinde birleşir. Örneğin Cornelia Fales, tüm görüngüler arasında ses kaynağı ve konumunu kesin olarak belirten en önemli değişkenin tını olduğunu söyler. Fales'e göre tınıya nüfuz eden özel bir sıfat yoktur ve tını, değişmeceli anlam içinde veya diğer duyulara yapılan analojiler vasıtasıyla açıklanmalıdır (Fales, 2002: 60). Diğer taraftan Grant Olwage, tınıyı, müzik kültürü çözümlemede en önemli etken faktör olarak kabul eder. Siyah Güney Afrikalıların koroculuk üzerine odaklanan müziksel kültüründe, siyah koro sesinin içinde şekillenen koloni-tarihsel ses kimliğini, sesin tını fonetiği üzerine yaptığı yaklaşımla ispatlamıştır (Olwage, 2004: 210). Bu çerçevede, bir model olarak müzik teknolojisine yansıyan belirleyici tını üretiminin, insan sesi kaynağını da kesin oluşturabilme yeteneğini günümüzde yakaladığını ancak, bunun bir adım ilerisi olan ses işlemenin, belki de müziksel doğrulama olarak nitelenebilecek başka bir görüngüyü ortaya çıkardığını belirtebiliriz.

Müziksel doğrulama kısaca, müzik kuramının aynen edime yansıtılma çabasıdır. Analitik olarak özellikle etnik kökenli birçok müziğin böyle bir kaygısı olmamasına rağmen, çoğunlukla batı müziğinin doğal yönlendirmesi sonucu, bugün hemen-hemen her stüdyo ortamında ses işleme yazılımları yardımıyla tüm müzik türlerinde müziksel doğrulama yapılmaktadır. Dolayısıyla, batı müziği kayıtlarının tamamına yansıyan bu eylem, örneğin halk ya da sanat müziğin de bile görülür. Müzikbilim çerçevesinde böyle bir doğrulamanın kuramsal çalışmaları da süregelen bir araştırma konusudur. Örneğin Martin Clayton, Kuzey Hindistan (Hindustani) sanat müziği tal-kal ile batı müziğindeki ölçü kavramı arasındaki ilişkiyi incelerken, tal-kal müziğinin tartımsal örgüsünü kuramsal ve dolayısıyla kuralcı bir boyuta çekmeye çalışmaktadır (Clayton, 1997: 170).

Müziksel doğrulamanın müzikbilimsel zemini tartışıla dursun, tını üretim modellerinden ses işlemeye geçiş yapan müzik teknolojisi, işleme yöntem ve tekniklerinin gelişmesiyle müziksel doğrulama sürecini bugün hemen-hemen her türlü kayıtta bir standart oluşum haline getirmiştir. Dolayısıyla bu oluşum, kendiliğinden müziksel doğrulamaya geçişin bir göstergesidir. Bu geçişin teknolojik sürecini göstermek adına müzik teknolojisinde insan sesi örneğini temel alan bu yazıda, insan sesi üretim modelleri ve işlemenin tarihsel süreçte yapı taşı kabul edilebilecek aygıtlar, teknik ve yöntemler anlatılmıştır. Üretimden analize doğru gelinen tarihsel süreçte konuşma ve şarkı sesi arasındaki teknik farklar belirtilmiş ve tını üretim modellerinden müziksel doğrulamaya geçişin göstergesi, müzik teknolojisi perspektifiyle tarihsel süreç içinde belirtilmeye çalışılmıştır.

2. KONUŞMA SESİ ÜRETİM MODELLERİ: MEKANİK/ ELEKTRİK AYGITLAR

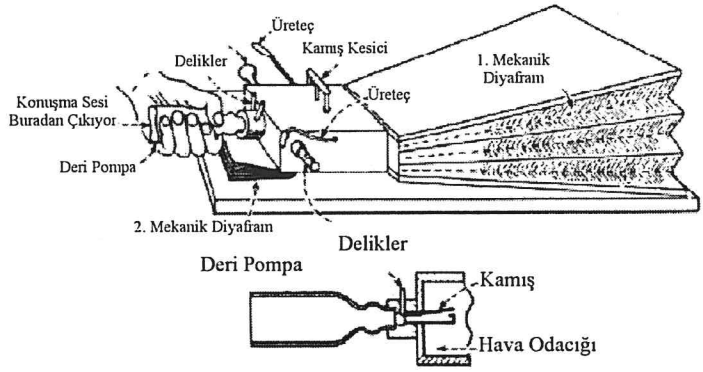
Konuşma seslerinin üretildiği ilk çalışmalar iki yüz yıl öncesine kadar uzanır. 1779 yılında Rus Profesör Christian Kratzenstein, Latin dilinde beş farklı ünlü harfin mekanik üretim modelini açıklar (/a/, /e/, /i/, /o/, /u/) ve ötümlü ses modelini niteler bu harfleri üretebilecek beş adet düzenek oluşturur. Her biri bir tür ağaç üfleme çalgıya benzer bu düzeneklerin hava giriş bölmesine üflenerek beş sesli harf elde edilir.



Şekil 1 Kratzenstein konuşma rezonatörü
(Kaynak: Schroder, 1993: 231)

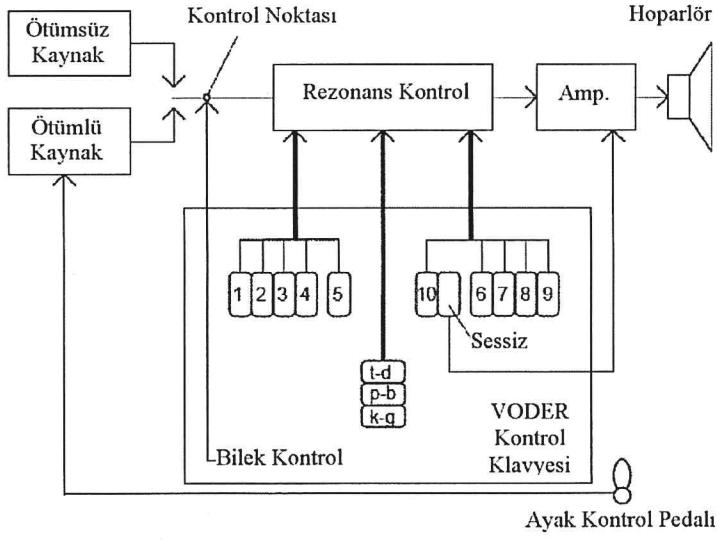
Mekanik konuşma aygıtları 1960'lı yıllara kadar Herman von Helmholz ve Charles Wheatstone gibi bilim adamları tarafından geliştirilmeye devam edildiyse de bu süreçteki en önemli aygıt, Wolfgang von Kempelen'in Viyana'da duyurduğu "Akustik-Mekanik Konuşma Aygıtı" olur (Flanagan, 1973: 50; Shroeder, 1993: 234). Bu mekanik konuşma aygıtı, 1800'lü yılların ortalarında Charles Wheatstone tarafından geliştirilir ve Alexander Graham Bell bu aygıtın şekilde görülen son halini ortaya çıkarır. Oldukça karmaşık yapısıyla ses üretebilen bu cihaz, Latin alfabesindeki tüm sesli harfler dışında birçok sessiz harfi de üretebilecek yeteneğe sahiptir.

Aygıtta tüm üreteç ve delikler kapalı durumdayken el ile hareket ettirilen deri pompa, kamışı harekete geçirerek sesli harfler üretirken, çeşitli üreteç ve delikler açılarak sessiz harflerin oluşumu sağlanır. Harflerin üretimi, aygıtı kullanan kişinin el becerisine bağlıdır. Çünkü sesli harf üretimi sırasında deri pompa ve mekanik diyaframlar yardımıyla sadece kamışa yönlendirilen hava kuvveti, kullanıcının eliyle yaptığı basınçla değişir. Aynı şekilde diğer sessiz harfler için hem bir el delik açmada kullanılır hem de bir taraftan pompaya yapılan basınç diğer el ile kontrol edilir.



Şekil 2 Kratzenstein konuşma rezonatörü
(Kaynak: Schroder, 1993: 231)

Elektriğin devreye girmesiyle sentetik konuşma sesi üretimi sağlayan aygıt yaratma çalışmaları 19.yy. başlarında ivme kazanır. 1922 yılında Stewart ve birkaç yıl sonra Wagner, sesli harfte ilk iki formant frekansını üretebilme yeteneğine sahip cihazlar geliştirirler (Flanagan, 1972: 63). 1932 yılında Japon araştırmacılar Obata ve Teshima, sesli harflerde üçüncü formant frekansını da aygıtlarına eklediklerini duyururlar (Schroeder, 1993: 232). Ancak bu dönem içerisinde hiç kuşku yok ki, 1939 yılında New York Dünya Fuar'ında Homer Dudley tarafından sunulduğunda teknoloji dünyasında büyük bir yankı uyandıran en ünlü elektrikli konuşma sesi üreten aygıt VODER (Voice Operating DEmonstratoR) olur.



Şekil 3 VODER konuşma aygıtının çalışma prensibini gösteriri blok diyagram
(Kaynak: Klatt, 1987: 780)

Bell laboratuarlarında geliştirilen VODER, el ile seçilen mekanik klavyeden gelen ötümlü (voiced) ya da ötümsüz (unvoiced) kaynak sesi, bilek ile kontrol edilen bir mekanizmayla alır ve tınlaşım kontrol ünitesinde elektrikli ayak pedalıyla kontrol edilen daha önceden sabitlenmiş temel frekanslar (fundamental frequencies) üzerinde birleştirir. Böylece aygıtın hoparlör çıkışından sentezlenmiş konuşma sesleri işitilir. Kullanımı oldukça zor olan VODER, çok iyi olmasa da o zamana dek üretilen en yalın konuşma sesi üreteci aygıtı olarak tarihe geçer. Bu tarihten sonra gelişen teknolojiyle birlikte çok farklı yöntemlerle çalışan çeşitli konuşma sesi aygıtları[1] geliştirilse de VODER, ilk olması ve tekniğindeki farklı yöntemle konuşma sesi aygıtı olarak mekanik/elektrikli en önemli aygıttır (Klatt, 1987: 732: Schroder, 1993: 231).

3. KONUŞMA VE ŞARKI SESİ ARASINDAKİ TEMEL FARKLAR

Konuşma sesi ile şarkı sesi üretimi yapısal olarak birbirlerine yakınmış gibi görünse de ses üretimi kapsamında her iki yöntemin önemli farklılıklar göstermesi, zamansal süreçte paralel ama teknolojik süreçte tamamen birbirlerinden bağımsız gelişim süreci izlenmesine yol açar. Örneğin Cook, her iki üretimde en büyük farkın ötümlü (voiced) ve ötümsüz (unvoiced) ses kullanım oranında olduğunu vurgular. Cook'a göre İngilizcede konuşma sesi üretiminde %60 ötümsüz ses kullanılırken, şarkı söylemede bu oran %40'lara düşer (Cook, 1990: 45). Başka bir ifadeyle şarkı söyleme durumunda insan sesi %60 oranında ötümlü ses kullanmaya yönelmektedir Örneğin Bel Canto olarak bilinen klasik dönem şarkı söyleme tekniğinde yorumcular, şarkı seslerinin daha etkili ve daha anlaşılabilir olarak işitilmesi için uzun soluklu sesli harf tınısı üretirler. Konuşma sesi üreteci olarak tasarlanan Kratzestein üreteçlerine baktığımızda da ötümlü seslerin tarihsel süreçte mekanik/elektrik üretim kolaylığı gösterdiğini söyleyebiliriz. Ötümlü ses kullanımının şarkı sesi üretiminde daha fazlalaşması, şarkı sesi üreten aygıtlarının nispeten konuşma sesi üreten aygıtlardan daha kolay tasarlanabileceğini, dahası, sinyal analizi açısından çok daha kolaylık sağlayabileceğini göstermekte.

Her iki ses üretiminde bir diğer önemli fark, temel frekansların vurgulanma genliğiyle ilgilidir. Özellikle batı müziğinde, şarkı sesi üretilirken konuşma sesine göre çok daha yüksek genliklerde temel frekans tınlatılır. Her iki üretimin genliklerinde oluşan bu fark, şarkı sesi üretiminin mekanizmalaşmasını daha kolaylaştırır. Bir diğer fark ise yetisel bilinçtir. Şarkı sesi kullanımı biyolojik açıdan ses teli ve gırtlak hareketi için bir ön hazırlık gerektirirken konuşma sesinde bu etkileşim kendiliğinden rastlamsal bir oluşum sağlar. Bu nedenle teknolojik açıdan konuşma üretiminin kontrolü diğerine göre daha zordur.

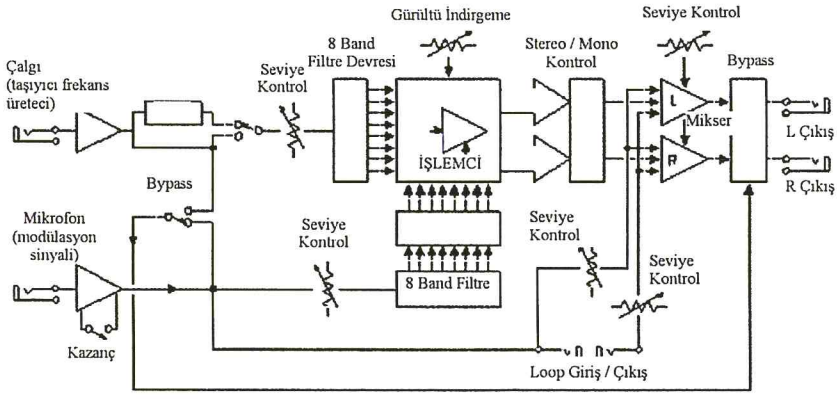
1 Bu aygıtların en bilinenleri arasında, üretim tarihi, üreticisi ve ismine göre: 1951: Franklin Cooper (Pattern Playback Synthesizer), 1953: Walter Lawrence (Parametric Artificial Talker (PAT)), 1952: Gunnar Fant (OVE I (Orator Verbis Electris), 1962: Fant ve Martony (OVE II), 1968: Kungliga Tekniska Högskolan (OVE II ve GLOVE)

Konuşma ve şarkı sesi üretimindeki biyolojik ve teknik ayrımın mekanik/elektrik üretim araçlarına olan etkisi, aynı ayrımın dilbilimsel ve müzikolojik açıdan incelenmesiyle de ortaya çıkar ve aynı oranda teknolojiye etki eder. Stewen Pinker, insan sesinin konuşma ve şarkı adaptasyonunda evrensel olarak daha çok konuşmanın yaşama etki ettiğini, bunun tam tersi olarak şarkı biçiminin, daha geniş bir ifadeyle müziğin, yaşamsal olarak daha dar bir çerçeve oluşturduğunu belirtir (Pinker, 1997: 528) . Patel, algısal olarak her iki işitme yolunun beynin aynı tabakasında oluşması nedeniyle birbirlerinden ayrılmayan iki etkileşim olması gerektiğini düşünürken, konuşma ve şarkı üretim mekanizmasındaki farklılıklar bu tür bir etkileşimde kültürel değişimi farklı ölçülerde ortaya çıkardığını ifade eder (Patel, 1998: 39). Her iki durumda da konuşma ve şarkı üretimi birbirinden ayrılır ve bu ayrım, mekanik/elektrik ses üreteçlerinin de farklı yöntemlerle çalışmasını sağlar.

4. ŞARKI SESİ ÜRETİM MODELLERİ: MEKANİK/ELEKTRİK/ ELEKTRONİK AYGITLAR

Şarkı sesi üretimine yönelik ilk cihazları 19. yy. ortalarında görüyoruz. Bu konuda bilinen en önemli aygıt ise kuşkusuz VODER'ın yaratıcısı Homer Dudley tarafından 1940 yılında üretilen VOCODER (VOice CODER[2])'dır. İlk üretildiği yıllarda VODER ile birlikte anılan bu cihazın çalışma ilkesindeki farklılık, daha sonraki yıllarda teknolojinin gelişmesiyle bu cihaza ticari anlamda çok daha önem kazandırır. Dudley, bu cihazın temel ilkesini VODER'dan farklı olarak bir tür ses sinyali filtreleme ve sentezleme olarak açıklayacaktır ve bu ilke, teknolojinin gelişmesiyle birlikte şarkı sesi üretiminde bir devrim yaratır (Kim, 2003: 24). 1967 yılında VOCODER'ın dijital teknolojiye taşınması (Gold & Rader, 1967: 150), 1978 yılında Fouirer Transform tekniğinin cihaza eklenmesi (Moorer, 1978: 43) ve bunun gibi günümüze dek uzanan teknolojik gelişmelerin VOCODER üretiminde kullanılması, cihazın şarkı sesi üretimindeki devrimsel niteliğini göstermektedir. 1970 yılında Wendy Carlos ve Robert Moog, ticari anlamda ilk VOCODER'ları üreten ve bu cihaza müzikte kullanıma yönelik tam anlamıyla şarkı sesi üretebilen bir yetenek kazandırırlar. Günümüz de bile çalışma sistemi korunarak ama yeni teknolojik gelişmelerin eklenmesiyle üretimine devam edilen VOCODER, şarkı sesi üretiminde kullanılmak üzere tarihte en bilinen mekanik/elektrik aygıttır.

2 VOCODER, bazı kaynaklarda VOice operated reCOrDER ifadesinin kısaltımı olarak belirtilse de, aygıtın orijinal yapısında herhangi bir kayıt özelliği olmadığından gerçek kullanım olarak VOice CODER tercih edilmiştir.



Şekil 4 VOCODER çalışma prensibi
(Kaynak: http://www.paia.com/ProdArticles/vocodwrk.htm)

Amerikan PAiA şirketi tarafından günümüzde üretimine devam edilen ve Şekil 4'de görülen örnek VOCODER devresinde şarkı sesi girişi bir mikrofonla sağlanır ve giriş modülasyon sinyali olarak devreye iletilir. Bu sinyal daha sonra filtre yardımıyla 8 ayrı sabit frekansa bölünerek işlemciye gönderilir. Cihaza bir diğer giriş de tuşlu çalgı girişidir. Bu giriş mikrofon girişiyle aynı özelliğe sahip 8 ayrı frekans filtresine gönderilir ve işlemciye iletilir. İşlemcide her iki giriş sinyali karıştırılır ve mikser yardımıyla çıkışa iletilir. Böylece mikrofonla girilen şarkı sesi, çalgıdan gelen frekanslarla birleştirilir ve çıkışta oldukça düzgün frekanslı sentezlenmiş şarkı sesi tınlatılır.

David Stuff tarafından 20.yy. Alman Elektronik Müzik hareketinin en önemli çalgılarından biri olarak kabul edilen VOCODER[3], Stockhausen'dan Planet Rock'a kadar (özellikle "Africa Bambaataa and Soul Sonic Force" albümü) elektronik müzikte yaygın olarak kullanılmıştır. 1970 yılında Wendy Carlos ve Robert Moog tarafından popüler anlamda kullanılabilir hale getirilen Vocoder Synthesizer 'ın popüler müzikteki ilk örnekleri, Stanley Kubrick'in 1974 yılında çekilen filmi "Otomatik Portakal (A Clock Orange)"da görülür. Filmin içeriğinde bulunan Bethooven'in 9. Sinfonisi'nin bir bölümü ve "Timesteps" adlı parçada kullanılan VOCODER, 1970'li yılların sonlarına doğru pop ve disko türü müziklerde yaygınlaşır. 1977 yılında Giorgio Moroder "From Here to Eternity", birkaç yıl sonra Jeff Lynne "Time" adlı albümlerinde VOCODER kullanacaktır ancak bu cihazın en popüler olduğu çalışmalar Pink Floyd'un "Animals" ve Alan Parsons'un"Alan Parsons Project" albümleri olur.

3 http://www.emusic.com / lists / showlist . html? Lid = 24234587)

1980'li yıllardan sonra müzikte VOCODER kullanımını daha çok özel efektler düzeyinde görülür. Özellikle New Age türü müziklerde efekt ses için yaygın olarak kullanılan VOCODER seslerinin en bilinen örneği 1984 yılında Jean Michel Jarre'ın "Zoolook" ve Mike Oldfield'in "Five Miles Out" albümleridir. Bu cihazın seslerini albümlerinde en çok kullanan grup ise Alman "Kraftwerk" olur. Diğer taraftan Stevie Wonder, Laurie Anderson, The Buggles, Daft Punk gibi pek çok kişi ya da grup, müziklerinde VOCODER kullanan örnekler arasındadır.

2000'li yıllara kadar kimi zaman elektronik müzikte bazen de pop müzik efekti olarak kullanılan VOCODER, popüler anlamda en büyük etkisini 1999 yılında piyasaya sürülen Cher'in "Believe" albümünde yapar. 1999 yılının en çok satan albümü olan "Belive"in hit parçası "Do You Believe in Life After Love"'un tüm vokalinde kullanılan VOCODER piyasada o kadar etkili olur ki, bu etki "Cher Efekti" adıyla yıllar sonra bir tür standart haline gelir. Bu standart, bazen doğrudan VOCODER kullanılarak ama çoğu zaman vokalde perde doğrulama aracı olan Auto-Tune'nun gereğinden fazla (işlenen tınıda VOCODER efekti duyulana kadar) işleme sokulmasıyla halen günümüzde birçok yerli-yabancı pop müzik parçasında kullanılmaktadır.

Albüm piyasaya çıktıktan sonra uzunca bir süre, Cher'in vokalindeki bu etkinin ses işleme ve perde doğrulama yazılımı olarak 1997 sonrasında üretilen Auto-Tune kullanılarak yapıldığı düşünülmüştür. Ancak, albümde kullanılan VOCODER etkisinin nasıl üretildiğine yönelik "Sound on Sound" dergisinde yapılan bir röportajda, albümün yapımcısı Mark Taylor, bu etkinin gerçek anlamda VOCODER ile üretildiğini söyleyecektir[4]. Amacının sadece farklı bir ses efekti aramak olduğunu belirten Taylor, bu etkiyi VOCODER'da yakaladığını ve bunun için ilk olarak 1970'lerin ünlü VOCODER synthesizer'ı olan Korg VC10'u kullandığını belirtir. Böylece uzun bir süre bu etkinin sektör çalışanları tarafından Auto-Tune marifeti gibi görülmesine rağmen ve işin ilginç yanı, günümüzde "Cher Efekt"ini üretmek için yaygın olarak Auto-Tune kullanılmasına rağmen, Cher'in albümünde doğrudan VOCODER kullanıldığı kesindir.

VOCODER'ın şarkı sesi üretimindeki tekniği, örnek bir insan sesini alıp bu sesi çalgıdan gelen ses ile birleştirip sonuçta yeni bir sentezlenmiş şarkı sesi üretimine dayanır. Bu tekniğin dışında tarihsel süreçte farklı teknikler kullanılarak farklı aygıtlar da üretilmiştir. Örneğin Larsson tarafından Stockholm Royal Enstitü'de üretilen MUSSE (Music and Singing Synthesis Equipment), geliştirdiği filtreler yardımıyla herhangi bir kaynaktan gelen sentezlenmiş sinyale formant frekansları ekler ve bas/bariton şarkı sesi üretir. Bu aygıtın gelişmiş dijital bir sürümü 1989 yılında MUSSE DIG adıyla Sundberg tarafından üretilmiştir (Berndtsson & Sundberg 1993: 279). 1970'li yıllarda şarkı sesi için insan gırtlak yapısını örnekleyen "fiziksel model" üretimine yönelik elektronik devreler üretilmeye başlanır.

4 http://www.soundonsound.com/sos/feb99/articles/tracks661.htm

Bu devreler, kaynak sesi üretimi ve filtreleme olarak iki ayrı bölüm içerirler (Titze, 1973: 140). 1980'li yılların başında Rodet tarafından geliştirilen CHANT projesi ile fiziksel modellemeyle örneklenmiş formant frekanslarıyla şarkı sesi üreten aygıtlar gelişimini sürdürür (Rodet vd.,1984: 17). Bu tür modeller, içerdiği tüm teknik ayrıntılarıyla biyolojik açıdan insan sesi üretimini örnek almaktadır ve cihazlar da bu nedenle gırtlağın şekilsel yapısına benzetilmektedir. Örneğin Cook tarafından geliştirilen SPASM, biyolojik olarak insan sesini gösteren bir şekil üzerinde çeşitli parametreler yardımıyla şarkı ve konuşma sesi üreten bir yazılım desteğiyle çalışmaktadır (Cook, 1990).

Teknoloji ilerleyip bu tür aygıtlar dijital devreler üzerinde gelişimini sürdürmeye devam ettikçe; bir başka ifadeyle bilgisayar sektörü aygıtlar üzerinde küresel bir süreç başlattıkça şarkı sesi üretiminde de çok daha farklı teknikler ve farklı üretim modelleri ortaya çıkar. Macon, MIDI verilerinden gelen ezgiyle daha önce örneklenmiş harflerin bilgisayarda birleşimi üzerinde çalışır (Macon vd., 1997). "Sinüsoidal Model" olarak tanımlanan bu teknikle harf ve hecelere dayalı ses bilgisi MIDI verileriyle birleştirilir. Bunun sonucunda, sadece harf ve hece girdisiyle erkek/kadın şarkı sesi elde edilmeye çalışılır. Duh, buna benzer bir yöntemle ancak MIDI yerine doğrudan ses verisi kullanarak söz ile ezgiyi birleştirmeye çalışır (Duh, 2004). Yamaha tarafından geliştirilen Vocaloid[5]; Microsoft tarafından tasarlanan Whistle Music Synthesizer[6], Macintosh tarafından geliştirilen Vocal Writer (Loscos, 2007: 15), Myriad Software'in ürettiği Virtual Singer (Loscos, 2007: 16) gibi yazılımlar, kullanıcı tarafından girilen şarkı sözü ve melodi birleştirme ilkesiyle çalışan ve günümüzde halen geliştirilmeye devam edilen şarkı sesi üretimine yönelik yazılımlar olarak karşımıza çıkar.

Diğer taraftan, örnekleme (sampling) yöntemiyle şarkı ya da konuşma sesi üretimi, 1960'lı yıllarda başlayıp günümüze kadar uzanan bir süreci içermektedir. Bu yöntemde ses, bir mikrofon aracılığıyla dijital ortama veri olarak aktarılır ve oluşturulan bu dijital tını, çeşitli tekniklerle synthesizer'larda çalınabilir hale gelir. Günümüz teknolojisinde kullanımı yaygın olan bu yöntemle, olabildiğince aslına uygun olarak -iyi- örneklenmiş bir ses, elektronik ortamda bağımsız bir çalgı ya da döngü (loop) olarak aslına uygun kullanılabilir ve bu yöntem, konuşma veya şarkı sesi üretiminde dijital teknolojinin sunduğu doruk olanaklardan biridir.

5. SES İŞLEME

Konuşma ve/veya şarkı sesi üretimine yönelik çalışmalar, mekanik ya da teknolojik olanaklardan faydalanarak insansı ses kaynağı yaratmaya yöneliktir. Önceleri konuşmada mekanik üretim modelleri ve sonrasında konuşma veya şarkı sesi

5 http://www.vocaloid.com

6 http://research.microsoft.com/srg/whistmusic/

üretimine yönelik teknolojik ilerlemeler bugün ticari ya da akademik olarak halen sürdürülmekte. İnsan sesi kaynağının mekanik/elektrik/elektronik üretim modeline yönelik bu tür çalışmaların dışında, gelişen dijital teknolojiye koşut olarak bir de konuşma/şarkı sesine yönelik üçüncü bir çalışma alanından bahsetmek mümkün. Müzik teknolojisinde ses işleme (sound processing) olarak tanımlanan bu alandaki hedef ise diğerinin aksine, herhangi bir kaynak üretim modelinden çok, önceden örneklenmiş bir sesin çeşitli tekniklere göre tekrar işlenmesi ve bu işlenen sesin sonucu olarak analizi yapılmış yeni bir ses elde edilmesi olduğunu belirtebiliriz. Bu teknikler arasında silme-kesme-kopyalama-yapıştırma, zaman ekseninde birleştirme, frekans, zaman ya da perde doğrulama, kazanç ayarlama, perde kaydırma, efekt, seviye denetleme gibi yaygın olarak kullanılanlar, önceden örneklenen seslerin işlenip analiz edilerek tekrar üretilmesine olanak sağlar[7].

Ses işleme ve analizi üzerine yapılan çalışmaların gelişimi, dijital teknolojinin gelişimine paralel olarak 1970'li yıllarda başlar. Gelen ses sinyalinin frekansına göre bir sonraki frekansları tahmin etme yöntemiyle çalışan Doğrusal Tahmin Kodlaması (Linear Predictive Coding (LPC)), o yıllarda özellikle konuşma sesi analizi için oldukça önemli bir teknik gelişme olur.

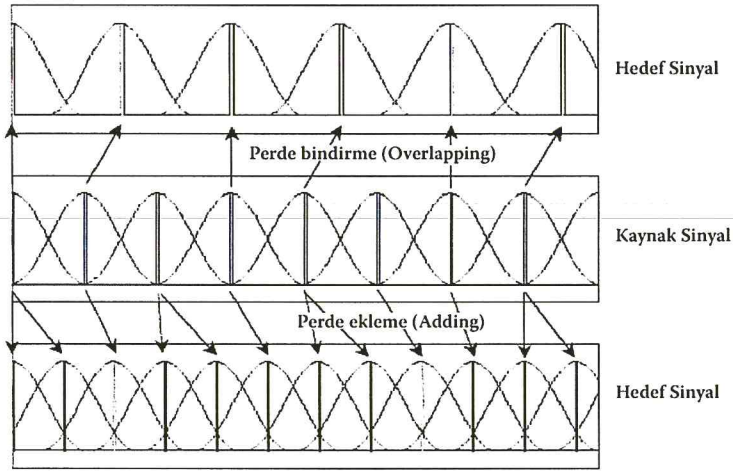
LPC tekniğinin tahmine dayalı frekans analizi, konuşma seslerinde başarı sağlamasına rağmen yeterli gelmez ve bunun yerine, önceden tanımlanmış olası konuşma frekanslarının girildiği bir sayısal kütüphane üretilir. Böylece, üretilen bu kütüphaneden tahminin ötesinde gelen ses sinyali frekansı taranır ve saptanır. Kod Uyarımlı Doğrusal Tahmin (Code-Excited Linear Prediction (CELP)) adı verilen bu yöntemle ses analizi tekniği yeni bir boyut kazanır (Schroeder & Atal, 1984: 938). LPC tekniği ve CELP gibi onun türevleri olan diğer teknikler ses sinyali analizinde temel frekansı tanımlamaya dayalı işlem uygular. Tarihsel süreçte ses sinyali analizi üzerine spektrum analiz teknikleri (Rodet & Depalle, 1992; Goodwin, 1997), gırtlak yapısını örneklemeye dayalı fiziksel bir model olan The Excitation Plus Resonances (EpR) filtreleme (Childers, 1994) gibi birçok analiz tekniği geliştirilse de LPC tekniğinden sonra bu konudaki en büyük gelişme, Pitch Synchronous Overlap-Add (PSOLA) tekniğinin bulunmasıdır.

PSOLA, belirlenen ses sinyalini küçük parçalara ayırarak bu parçalardaki perde temel frekanslarının hedef sinyale göre üst üste ya da yan yana eklenmesi ya da tamamen silinmesi ilkesine dayanır (Valbret, 1991: 346). İşlem üç aşamada gerçekleşir. İlk aşamada sinyal analiz için parçalanır. İkincisinde parçalanan sinyal gruplarında perde temel frekansları seçilir ya da çok küçük değere sahip olanları atılır. Son aşamada geriye kalan perdeler hedef sinyale uyacak biçimde üst üste bindirilerek ya da yan yana eklenerek tüm parçalar yeniden birleştirilir (Kim, 2003: 28). Dikkat edilecek olursa, PSOLA yönteminde mutlaka kaynak ve hedef sinyal bulunur. Kaynak sinyal yöntemin uygulandığı; hedef sinyal ise kaynağın

7 http:// tweakheadz.com/how_to_process_vocal_tracks.htm

benzetilmeye çalışıldığı sinyaldir. Bu sinyaller ise herhangi bir konuşma ya da şarkı sesi olabilir. Diğer bir önemli detay perdelerin analiz eksenidir. PSOLA yönteminde perdeler zaman, frekans vb. eksen (domain) üzerinde analiz edilir. Bu nedenle yöntem, uygulama eksenini ya da temel tekniğini belirtecek şekilde çeşitli adlar almaktadır (Time-Domain PSOLA, Frequency-Domain PSOLA, Linear-Predictive PSOLA vs.).

PSOLA çalışma yöntemini gösterir örnek, Şekil 5'de belirtilmiştir. İlk aşamada analizi yapılarak temel frekansları saptanmış kaynak sinyal perdeleri, ikinci ve üçüncü aşamalarda hedef sinyale benzetilmeye çalışılır. Bunun için uygulanan ve yönteme adını veren teknik, şekilde belirtilen perde bindirme ya da eklemedir (overlapping and adding). Böylece perdelerinde değişiklik yapılan kaynak sinyal hedef sinyale benzetilmeye çalışılır.



Şekil 5 PSOLA yöntemini gösterir işlem.
(Kaynak: www.tc-helicon.com/Files/Whitepapers/Pitch_tshifting.pdf)

1980'li yılların ortalarında konuşma sesleri üzerinde geliştirilen PSOLA, bu tarihten sonra çalışmasındaki farklı teknik nedeniyle önceleri kaydedilen şarkı seslerinde, sonraları 2000'li yıllarda gerçek zamanlı olarak sahnelerde kullanılmaya başlayan bir teknik olarak yaygınlaşmaya başlar (Lemouton vd., 2000).

PSOLA tekniğinin ses işlemeye getirdiği en büyük yenilik, analiz amaçlı sinyal işlemenin ötesinde, perdelere doğrudan müdahale edilerek sinyalin bir başka örnek sinyale benzetime olanak sağlamasıdır. Böyle bir olgunun ortaya çıkardığı sonuçlar, Karaoke olarak bilinen eğlence aygıtlarından profesyonel ses kayıt stüdyolarında vokal ya da çalgı seslerinin düzeltilmesine kadar geniş bir yelpazedeki yazılımların popüler anlamda kullanılmasını sağladı. Tekniğin içinde barındırdığı perde kaydırma (pitch shifting) hareketi, detone kaydedilen vokal ya da çalgı

seslerinin olması gereken perdelere çekilebilmesini; dolayısıyla detonasyon anlamında en bozuk seslerin bile düzeltilerek doğru frekanslarda tınlatılan yaratılar ortaya çıkarılması çalışmalarına kapı açar.

Ses işlemeye yönelik çalışmaların günümüzde geldiği son nokta ise, yukarıda anlatılan neredeyse tüm işlemleri kapsayan karmaşık yazılımların üretilmesi olmuştur. Bu yazılımlar, günümüzde yaygın olarak kullanılan Digidesign Pro-Tools[8], Steinberg Cubase[9], Cakewalk[10], Logic Audio[11] vb. profesyonel ses kayıt yazılımlarıyla bütünleşmiş şekilde eklenti (plug-in) olarak tasarlanırlar ve her biri kendi içinde farklı işleme teknikleriyle kullanıcıya sunulurlar. TC-Helicon tarafından üretilen Voice-Modeler, şarkı sesi üretim modelinin dışında PSOLA yöntemiyle vokal işlemeye yönelik üretilen ilk eklentilerden biridir[12]. Bu tür eklentiler olarak günümüzde kullanılan başlıca örnekler Yamaha'nın ürettiği Vocal Rack[13], Eventide'ın Pro-Tools eklentisi Octavox[14], Voxengo'nun ürettiği Voxformer[15] sayılabilir. Ancak bunların hiçbiri, Anteres tarafından üretilen Auto-Tune[16] ya da Celemony Corp. tarafından geliştirilen Melodyne[17] kadar popüler olamamıştır.

Darwin Grosse tarafından yapılan bir araştırmaya göre, Melodyn ile birlikte dünyanın en fazla satan iki sinyal işleme yazılımından biri olan Auto-Tune, ses kayıtlarında vokal sesleri için geliştirilmiş bir sinyal işleme eklentisi olarak 1997 yılında Anteres Audio Technologies firması tarafından satışa sunulur ve ilerleyen yıllarda vokal kayıtlarındaki perde doğrulama (pitch correction) işlemleri için tüm profesyonel ses kayıt stüdyolarında dünya standardı haline gelir (Grosse, 2007: 84). Temel olarak PSOLA tekniğini kullanan ürün, vokal ya da diğer çalgılardan gelen sinyalleri kaynak sinyal olarak kabul eder ve PSOLA tekniğinin ilk aşaması olarak bu sinyallerin analizini yapar. Böylece perdelerin temel frekansları gruplanır. Ardından ikinci aşama olarak gruplanan bu frekanslar, hedef sinyal olarak kullanıcının seçimine bırakılan aşıt temel frekanslarıyla karşılaştırılır ve son aşamada bu frekanslara üst üste bindirme veya ekleme işlemleriyle kaynak

8 http://www.digidesign.com

9 http://www.steinberg.net/

10 http://www.cakewalk.com/

11 http://www.apple.com/logicpro/

12 http://www.soundonsound.com/sos/feb04/articles/tcvoice.htm

13 http://www.etcetera.co.uk/products/YAM072.shtml

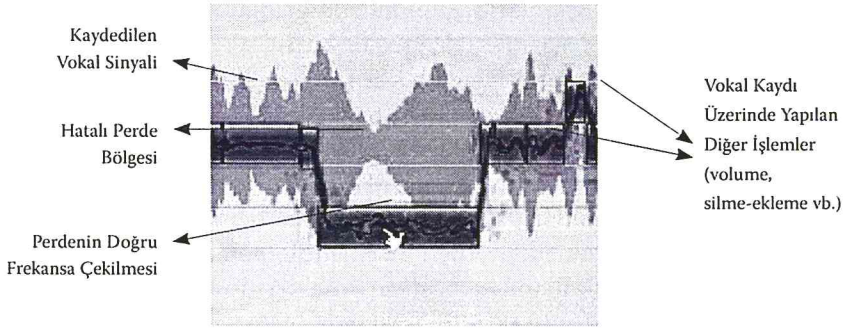
14 http://www.pluginz.com/product/12489?vendor=383

15 http://www.voxengo.com/product/voxformer/

16 http://www.antarestech.com/products/auto-tune5.shtml

17 http://www.celemony.com/cms/index.php?id=358

sinyalin gruplanan temel frekansları bindirilir. Böylece, kaynak sinyaldeki olası farklı frekanslardaki perdeler, aşıt perdelerinin doğru frekanslarına çekilmiş olur ve perde doğrulama işlemi gerçekleşir.



Şekil 6 Ses işleme yazılımlarında perde doğrulama ve diğer bazı işlemlere bir örnek (Kaynak: http://laptopstudio.thunderguy.com/2005/07/page/2/)

Auto-Tune ile temel ses işleme olarak aynı tekniği kullanan Melodyn, Ceremony firması tarafından 2000 yılında piyasaya sürülür. İlk olarak Macintosh bilgisayarlar için üretilen yazılım, kullanım yönünden Auto-Tune'dan farklı ara yüzüyle 2002 yılından itibaren IBM tabanlı PC'ler için de geliştirilmiştir. Yazılımın çalışma ilkesi tıpkı Auto-Tune ile aynı amaca yönelik olsa da (perde doğrulama) ilk sürümlerinde ses yüksekliği, vibrato desteği vb. değişkenleri içinde barındırdığından Auto-Tune ile farklılıklar gösterir. Bu tip farkların dışında Melodyn, Melodyn Bridge adı verilen eklentisiyle desteklenen bağımsız bir yazılım (standalone) olarak bugün halen özellikle vokal işlemede yaygın olarak kullanılmaktadır (Walden, 2007: 44).

6. SONUÇ

Önceleri sadece ses üretiminde kullanılan ancak daha sonraları teknoloji sayesinde detaya inildikçe farklı bir analiz aracı olarak görülen şarkı sesi aygıtlarının müzikte en çok kullanılanı VOCODER'dan sonra hiç kuşkusuz, Cher albümü tartışmasında da belirtildiği gibi, Auto-Tune ya da Melodyn gibi perde doğrulama vokal işlemcileridir. Bu tür işlemciler, bir taraftan gerçek amaçlarına uygun olarak kullanılırken, diğer taraftan vokalde farklı bir tını üretebilen araçlar olarak görülmektedir. Ancak hiç kuşku yok ki teknolojinin ilerlemesiyle bu süreç, sektörün, tınısal üretim modelinden müziksel doğrulamaya geçişinin bir göstergesidir.

Tınısal üretim, tıpkı Cher örneğinde olduğu gibi, müzikte, özellikle vokalde yeni bir tını ortaya çıkarmayı amaçlar. Fales'in de belirttiği gibi bu tını, insan sesi kaynağı ve konumunu kesin olarak belirtir (Fales, 2002: 60). Tınısal üretim

modellerinden VOCODER çalışma prensibinin, tarihsel süreçte şarkı sesi üretim aygıtlarında uzunca bir süre kullanıldığını ve bu durumun ihtiyacı yeterince karşıladığını görüyoruz. VOCODER dışında yazıda belirtilen diğer şarkı sesi üretim araçlarının az da olsa elektronik müzikte; ancak popüler anlamda herhangi bir müzik örneğine rastlanmadığından en fazla teknik/yöntem geliştirme aracı olarak deneysel amaçlarla kullanıldığını vurgulayabiliriz. Bunun en büyük sebebinin, VOCODER'ın gündelik kullanıma indirgenebilecek kadar sağlıklı ve etkili bir üreteç olmasından kaynaklandığı söylenebilir. Özellikle 70'li yıllardan sonra dünyaca ünlü synthesizer üreticisi Robert Moog ve Wendy Carlos'un VOCODER yapımına el atması ve bu hareketin doğal sonucu olarak cihazın endüstriyel müzik sektöründe hızla yaygınlaşması bunun açık bir göstergesidir.

Önceleri doğrudan insan sesinin benzetilmesi amacıyla yapılan aygıt üretimi, sonraları tını sorgulaması nedeniyle farklı bir boyut kazanır. Tınısal sorgulama, beğenideki etken faktörleri içine alan bir sürecin başlangıcıdır. Bu amaca yönelik üretilen aygıtlar da özellikle 1970 sonrası birbirleri arasında yapısal farklılıklar gösterir ancak amaç, tınısal üretim modeline dayanan yeni gereçler üretip bunları çeşitli müzik türlerinde dinleyicinin beğenisine sunmaktır.

1980'li yıllardan sonra dijital teknolojinin hızla gelişmesi ve sektörün mikro düzeyde işlemci üretebilme becerisi, tınısal üretimin kaynağı olan insan sesinin daha detaylı analiz edilebilmesini ve bu analiz sonucunda, orijinaline çok daha benzer tınıları ortaya çıkarabilecek aygıtların üretilmesini sağlar. Bu hızlı süreç, insan sesine sanal ortamda doğrudan müdahale edebilme olanağını mümkün kılar. Böylece üretim modelleri bir taraftan gelişirken, diğer taraftan üretilen seslerin teknik açıdan, üstelik mikro düzeyde doğruluğunun sorgulanması süreci başlar.

Kültürel bir birikim olarak ortaya çıkan toplumsal müzik kültürü ve dolayısıyla müzik, üretimi sırasında teknik olanaklarla birleştiğinde kuralcı bir zemine oturma eğilimi gösterir. İşte bu nokta, müziksel doğrulama sürecinin başlangıcıdır. Böylece özellikle üretim aşamasındaki insan yargısı, önceleri tınısal üretim modeli olarak süre giden bu oluşumu, kendiliğinden, müziksel doğrulama bilinci olarak eyleme geçirir. Kültürel açıdan hiçbir kaygı olmadığı durumlarda bile üretilen her tür müzikte, müziksel doğrulama, stüdyo kayıt uygulamalarında bir standart haline gelir.

Müziksel doğrulamanın en önemli tekniği, üretilen perdelerin kuramsal karşılıklarıyla, yani türdeki aşıta ait perdeler ile eşleştirilmesidir. Müzik teknolojisinde perde doğrulama (pitch correction) olarak tanımlanan bu teknik, doğrulamada en yaygın kullanılan adımdır ve müziksel doğrulamanın bir göstergesidir. Bunun için üretilen sesin çok iyi analiz edilip gerekli perdelerin tam olarak tespit edilmesi gerekir ki günümüz teknolojisinin geldiği noktadaki yöntem ve teknikler, özellikle Auto-Tune veya Melodyn gibi yazılımlar bunu azımsanamayacak ölçüde

mümkün kılar. Örneğin popüler müzik vokal kayıtlarında do perdesi yerine söylenen do'ya çok yakın farklı bir perde, olması gereken do perdesine günümüz teknolojisinin olanaklarıyla rahatça çekilebilir. Bu uygulamanın müzikteki tüm diğer perdeler için de yapılabilineceği düşünülürse, özellikle vokaldeki detonasyon kaygısı kendiliğinden ortadan kalkacaktır. Bu adım sonrasında hatalı her perde doğru zemine çekilir ve müziksel doğrulamanın en önemli adımı tamamlanmış olur. Doğrulamanın diğer adımları ise kesin bir kural olmamakla birlikte genelde tempo ayarlama, silme-kesme-kopyalama-yapıştırma, zaman ekseninde birleştirme, kazanç ayarlama, seviye denetleme vb. ses işlemlerini kapsar. Her bir işlemin amacının kaydedilen sesin doğru bir zemine oturtmak olduğu söylenebilir ve tüm bu işlemler göz önüne alınırsa, ilk başlarda tınısal üretim modeli olarak süregelen müzikteki teknolojik oluşumun müziksel doğrulamaya doğru bir geçiş süreci oluşturduğu açıkça görülebilir.

Müzikbilim ve/veya müzik eğitimi perspektifinde de ele alınabilecek ve müziksel kültürün toplumsal etkilerinin bir göstergesi sayılabilecek müziksel doğrulama süreci, insanda algılamanın zorunlu kıldığı bir ihtiyaç şeklinde de belirtilebilir; teknolojinin kendiliğinden getirdiği bir oluşum olarak da görülebilir. Ancak sonuç ne olursa olsun, bugün hemen her stüdyo ortamında bir standart haline gelen müziksel doğrulama, müzik teknolojisi kapsamında ses üretme ve işleme alanının müziksel ve sosyal sonuçlarını gösteren en önemli görüngülerinden biridir.

KAYNAKÇA

Berndtsson G. – Sundberg, J. (1993). The MUSSE DIG Singing Synthesis. **Proceedings of the Stockholm Music Acoustics Conference / SMAC 1993** (pp. 279–281). Stockholm: School of Computer Science and Communication Press.

Childers, D.G. – C.F. Wong (1994). Measuring and Modeling Vocal Source-Tract Interaction. **IEEE Transactions on Biomedical Engineering**. vol. 41: (pp. 663-671). Washington: University of St.Louis Press.

Clayton, M. (1997). Le metre et le tal dans lamusique de l'Inde du Nord. **Cahiers de Musiques traditionelles,** vol. 10: (pp.169-189). from http://www.open.ac.uk/Arts/music/nclayton/metre.htm; trans. Georges Goormaghtigh.

Cook, P.R. (1990). **Identification of Control Parameters in an Articulatory Vocal Tract Model, with Applications to the Synthesis of Singing**. PhD thesis, Stanford University.

Durmaz, S. (1997). İlişkiler Bağlamında, Müzikoloji ve Müzik Teknolojisi, **1. Müzik Sempozyumu**, Müzik Eğitimi Bölümü, Abant İzzet Baysal Üniversitesi, Bolu

Duh, K. – Kirchhoff, K. (2004). Automatic Learning of Language Model Structure, Proceedings of the 20th International Conference on Computational Linguistics (COLING-2004), Geneva, Switzerland.

Fales, C. (2002). "The Paradox of Timbre", **Ethnomusicology**, Vol. 46, No. 1/Winter, pp: 56-95.

Flanagan, J. (1972). **Speech Analysis, Synthesis, and Perception**, Springer-Verlag, Berlin-Heidelberg-New York, USA

Flanagan, J. – Rabiner, L. (1973). **Speech Synthesis**, Dowden, Hutchinson & Ross, Inc., Pennsylvania, USA

Gold, B. – Rader, C.M. (1967). The Channel Vocoder. **IEEE Transactions on Audio and Electroacoustics**, AU 15(4):148–161

Goodwin, M.M. (1997). **Adaptive Signal Models: Theory, Algorithms, and Audio Applications**. PhD thesis, University of California, Berkeley, USA.

Grosse, D. (2007). Current Tools on Pitch Correction, **Recording Magazine**, May 2007, p. 83-87

Kim, Y.E. (2003). **Singing Voice Analysis/Synthesis**, PhD. Thesis, Program in Media Arts and Sciences, Massachusetts Institute of Technology (MIT).

Klatt, D. (1987). Review of Text-to-Speech Conversion for English. **Journal of the Acoustical Society of America** (JASA) vol. 82 (3), pp.737–793

Larsson, B. (1977). Music and Singing Synthesis Equipment (MUSSE). **Speech Transmission Laboratory Quarterly Progress and Status Report (STL-QPSR)**, (1/1977): 38– 40

Lemouton, S. – Manoury, P. – Schnell, N. – Peeters, G. – Rodet, X. (2000). Synthesizing a Choir in Real-Time Using Pitch Synchronous Overlap Add (PSOLA). **XIII. International Computer Music Conference.** Berkley

Loscos, A. (2007). **Spectral Processing of The Singing Voice**, PhD. Thesis, Department of Information and Communication Technologies, Pompeu Fabra University, Barcelona, Spain

Macon M.W. – Jensen-Link, L. – Oliverio, J. – Clements, M.A. – George, E.B. (1997). Concatenation-Based MIDI-to-Singing Voice Synthesis, 103rd Meeting of the Audio Engineering Society, AES Preprint 459.

Moorer, J.A. (1978). "The Use of The Phase Vocoder in Computer Music Applications", **Journal of the Audio Engineering Society**, 26(1):42–45

Olwage, G. (2004). "The Class and Colour or Tone : An Essay of the Social History of Vocal Timbre", **Ethnomusicology Forum**, Vol 13 [2] : 203-226.

Patel, A.D. (1998). "Syntactic Processing in Language and Music: Different Cognitive Operations, Similar Neural Resources?" **Music Perception**, 16(1): 27- 42.

Pinker, S. (1997). **Evolutionary biology and the evolution of language.** In W. Schopf (Ed.), The origin and evolution of intelligence. Sudbury, MA: Jones & Bartlett. Reprinted in M. Gopnik (Ed.), The inheritance and innateness of grammars. New York: Oxford University Press.

Rodet, X. – Potard, Y. – Barrière, J.-B. (1984). "The CHANT Project: From The Synthesis of The Singing Voice to Synthesis in General", **Computer Music Journal**, 8(3): 15–31

Rodet, X. – Depalle, P. (1992). A New Additive Synthesis Method Using Inverse Fourier Transform and Spectral Envelopes. Proceedings of International Computer Music Conference, San Jose, California, USA.

Schroeder, M.R. – Atal, B.S. (1984). Code-excited linear prediction (CELP): Highquality speech at very low bit rates. **Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing**, pages 937-940.

Schroeder, M. (1993). "A Brief History of Synthetic Speech", **Speech Communication** vol.13, pp. 231-237.

Titze, I.R. (1973). "The Human Vocal Chords: A Mathematical Model", Part I, **Phonetica** 28, 129-170.

Walden, J. (2007). "AutoTune vs Melodyn", **Sound on sound**, March 2007, 43-46.

Turk, O. – Sayli, O. – Ozsoy, A.S. – Arslan, L.M. (2004). "Türkçe'de Ünlülerin Formant Frekans İncelemesi", **18. Ulusal Dilbilim Kurultayı**, Ankara.

Valbret, H. – Moulines, E. – Tubach, J. (1991). "Voice Transformation Using PSOLA Techique". **Journal of Eurospeech,** vol: 91 (1), pp 345-34.