



POLİTEKNİK DERGİSİ

JOURNAL of POLYTECHNIC

ISSN: 1302-0900 (PRINT), ISSN: 2147-9429 (ONLINE)

URL: <http://dergipark.gov.tr/politeknik>

Yerleşik ve gömülü uygulamalarda kontrol işlemleri ve PC’de yazı yazmak için kullanabilen düşük maliyetli genel amaçlı bir konuşma tanılama sistemi

A low cost general purpose speech recognition system used to control processes in the embedded and stationary system and write texts on the PC

Yazar(lar) (Author(s)): Mustafa BURUNKAYA¹, Melek DİJLE²

ORCID¹: 0000-0002-3971-0590

ORCID²: 0000-0002-0900-7968

Bu makaleye şu şekilde atıfta bulunabilirsiniz(To cite to this article): Burunkaya M. ve Dijle M., “Yerleşik ve gömülü uygulamalarda kontrol işlemleri ve pc’de yazı yazmak için kullanabilen düşük maliyetli genel amaçlı bir konuşma tanılama sistemi”, *Politeknik Dergisi*, 21(2): 477-488, (2018).

Erişim linki (To link to this article): <http://dergipark.gov.tr/politeknik/archive>

DOI: 10.2339/politeknik.389641

Yerleşik ve Gömülü Uygulamalarda Kontrol İşlemleri ve PC’de Yazı Yazmak İçin Kullanabilen Düşük Maliyetli Genel Amaçlı Bir Konuşma Tanılama Sistemi

Araştırma Makalesi / Research Article

Mustafa BURUNKAYA^{1*}, Melek DİJLE²

¹Teknoloji Fakültesi, Elektrik-Elektronik Mühendisliği Bölümü, Gazi Üniversitesi, Türkiye

²Düzce Borsa İstanbul Mesleki ve Teknik Anadolu Lisesi, Merkez Düzce, Türkiye

(Geliş/Received : 11.12.2016 ; Kabul/Accepted : 04.01.2017)

ÖZ

Bu çalışmada yeni bir düşük maliyetli, mikrodenetleyici kullanılan, konuşmacı bağımlı ses tanıma sistemi tasarlanarak yüksek doğrulukla konuşma tanımayı etkileyen parametreler araştırılmış ve genel amaçlı olarak kullanılacak bir sesli kontrol sistemi elde edilmiştir. Gerçekleştirilen sistemde konuşma ile yazı yazabilmek için veriler bir bilgisayara (PC) aktarılmıştır. Bu amaçlarla sistem engelli insanlar veya sağlıklı insanlar tarafından kullanılabilir. Konuşma tanımlama devresi kullanılarak, alfabeyle ait harfler ve kontrol işlemleri için önceden seçilen bazı kelimeler eğitilebilen sisteme kayıt edilmiştir. Sistemde mikrofon tarafından algılanan sesler karşılığı olan sayısal sinyallere çevrilir. Tanınan seslere ait veriler birer mikrodenetleyici kullanılarak karşılaştırıldıktan sonra, RS232 ve PS/2 iletişimi kullanılarak PC’ye aktarılır. Sistemin esnek yapısı ve gerektiğinde PC den bağımsız olarak da kullanılabilmesi kullanışlılığı artırmaktadır. Sistem gürültüsüz ortamda ve farklı gürültü seviyelerinde, harfler, eşsesli ve eşsesli olmayan kelimeler ile test edilmiştir. Elde edilen veriler SPSS paket programı kullanılarak istatistiksel analize tabi tutulmuştur. Deneyler sırasında gürültü ölçümleri için Rion NL-21 cihazı kullanılmıştır.

Anahtar Kelimeler: Konuşmacı bağımlı, yazı yazma, RS232, PS/2, kontrol.

A Low Cost General Purpose Speech Recognition System Used to Control Processes in The Embedded and Stationary System and Write Texts on The PC

ABSTRACT

In this study, a low cost, microcontroller based speaker dependent speech recognition system was designed and parameters affecting speech recognition with high accuracy were investigated and a speech control system which can be used as a general purpose was obtained. In the realized system, the relevant data is transferred to a computer (PC) in order to write text. For these purposes, the system can be used by people with disabilities or healthy people. Using the speech recognition circuit, the letters of the alphabet and some selected words for control operations are recorded in advance in this system which can be trained. Speech voices detected by the microphone in the system are converted into digital signals corresponding to them. Once the data for the recognized voices are compared using a microcontroller, they are transferred to the PC using RS232 and PS2 communication protocol. The flexible structure of the system and the ability to be used stand-alone without PC when necessary also improves its usability. The system has been tested in noiseless environment and with different noise levels, with letters, homophones and non-homophones words. The obtained data were analyzed statistically by using SPSS package program. During the experiments, Rion NL-21 instrument was used for noise measurements.

Keywords: Speaker dependent, texting letter, RS232, PS/2, control

1. GİRİŞ (INTRODUCTION)

Otomatik konuşma tanıma, bir bilgisayarın bir akustik konuşma sinyalini metne eşleştirmesi işlemi olarak tanımlanabilir [1]. Ses tanıyıcı sistemler 1960’lı yılların başlarında ortaya çıkmaya başlamıştır. Bu konuda IBM firması, “Shoebbox” adı verilen öncü ses tanıma sistemlerini geliştirmiştir. 16 kelimeyi tanıyabilen bu sistemde

elektromekanik devre elemanlarına yer verilmiştir [2]. Başlangıçta düşük performanslı sayılabilecek olan bu gibi çalışmalar zamanla hızla gelişmişlerdir. Günümüzde de bu gelişimleri devam etmekte olup, uygulama alanları da gittikçe yaygınlaşmaktadır.

Ses tanıma sistemleri sağlık, askeri, otomasyon, robotik, iletişim, eğlence, biyometrik uygulamalar, müzik vb. çok çeşitli alanlarda ve amaçlarla kullanılabilir. Bu sistemler felçli, görme ve zihinsel engelli, konuşma sorunları olan vb. insanların iletişimleri, tedavi ve terapileri için veya

*Mustafa Burunkaya (Corresponding Author)
e-posta : bmustafa@gazi.edu.tr

ihtiyaç duydukları çeşitli cihazları kontrol edebilmelerine olanak sağlamak için de kullanılabilir [3-15].

Bu amaçlar için kullanılacak yüksek performans gerektiren sistemlerde ses tanıma işlemleri daha çok bir PC ve bu amaçla geliştirilen yazılımlar yolu ile yapılmaktadır. Bu sistemlerde PC uyumlu ses kartları gerekli olup, yazılımlar işletim sisteminin arkasında sürekli olarak çalışmaktadır. Bu tür sistemlerde tanılama işlemleri için karmaşıklık, maliyet, PC gereksinimi, sürekli çalışan yazılımlar vb. bazı önemli olabilen dezavantajlar olarak değerlendirilebilir [16,17].

Bu tür yazılımların geliştirilmesi açısından konuşma tanıma tipleri, konuşmacı bağımlı ve konuşmacı bağımsız yazılımlar olmak üzere gruplandırılabilir. Konuşmacı bağımlı yazılımlar genellikle yazılı hale getirilmek üzere sözcükleri yüksek sesle söyleme işlemleri (dictation) için kullanılmaktadır. Konuşmacıya bağımlı yazılım, ses tanıma benzer şekilde, bilgisayarda tek bir kişinin sesinin benzersiz özelliklerini öğrenerek çalışır. Yeni kullanıcılar için, önce yazılımla konuşularak eğitilmeleri gereklidir. Geliştirilmesi kolay, ucuz ve daha doğru tanıma yapabilir [18]. Fakat bunlar konuşmacı uyumlu (adaptif) olarak da sınıflandırılabilen sistemler veya bağımsız sistemler kadar esnek değildir.

Konuşmacı bağımsız sistemler ise kişiye bağlı olmadığından eğitilmesi gerekli değildir. Daha çok telefon uygulamalarında vb. kullanılmaktadır. Esnek olmalarına karşın, pahalı ve doğruluğu konuşmacıya bağımlı sistemlerden daha düşüktür [1,18].

Bazı çalışmalarda, konuşmacı bağımsız sistemlerde konuşma motoruna hangi seslerin matematiksel olarak benzediğini hesaplayan bir veri tabanı kullanılmaktadır [18]. Konuşmacı uyumlu sistemler ise operasyonlarını yeni konuşmacıların özelliklerine uyarlamayabilmektedirler [1].

Konuşma tanıma sisteminin kelime dağarcığı, karmaşıklığı vb. işleme gereksinimlerini ve sistemin doğruluğunu etkiler. Bazı uygulamalarda yalnızca birkaç kelimenin tanınması yeterli olabilirken, örneğin dikte makineleri gibi bazıları çok daha fazlasını gerektirebilir. Kelime tanıma sayısına göre konuşma tanıyıcılar şu şekilde gruplandırılabilir:

- Küçük kelime - onlarca kelime
- Orta kelime - yüzlerce kelime
- Büyük kelime - binlerce kelime
- Çok büyük kelime - on binlerce kelime.

Konuşma tanıma sistemleri kelimeler arasındaki boşluk açısından izole edilmiş ve sürekli konuşma sistemi olarak da gruplandırılabilir.

İzoleli sistemde kelimeler arasında duraklamalar olup, ayrı ayrı söylenen kelimeler tanınabilir. Bu sistemde bir sözcüğün telaffuzu diğerlerini etkilememe eğiliminde olduğundan, sözcüklerin ortaya çıkışı daha tutarlı olup, tanınması daha kolaydır. Sürekli konuşma sistemi, sözcükler arasında duraklarla ayrılmayan konuşmaların tanınmasında kullanılır. Bu sistemin, konuşmadaki çeşitli etkiler, ortak oluşum, her bir ses biriminin üretiminin

çevresindeki fonemlerden etkilenmesi, sözcüklerin başlangıç ve bitişi, önceki ve sonraki sözcüklerden, konuşma hızından etkilenmesi vb. sebeplerle tanıma doğruluğu düşüktür [1].

Bu çalışmada bütün bunlara göre yerleşik ve gömülü uygulamalarda daha ucuz ve doğru tanıma yapabileceği göz önüne alınarak konuşmacı bağımlı bir tanıma ve kontrol sistemi tasarlanmış ve gerçekleştirilmiştir [18]. Sistemin görme ve bazı ilaveler ile duyma, konuşma ve bazı diğer engellilerin yaşamlarını kolaylaştırabileceği düşünülmektedir. Bilindiği gibi görme engelliler Braille Alfabeti olarak adlandırılan kabartma noktalar ile karakterize edilen bir yazı kullanılmaktadırlar [19]. Bu durum bazı açılardan çözümler sunsa da, bu şekilde yazı yazmanın bazı zorlukları ve yazılmış yazının diğer insanların anlayabilmeleri yönünden kullanışlı olmadığı düşünülebilir. Konuşma ve ses tanıma yöntemleri kullanılarak bu tür bazı sorunların üstesinden gelinmesi mümkün olabilir [3,4,20]. Gerçekleştirilen sistem kullanılarak onlar için de PC ortamında yazı ve çeşitli metinlerin konuşma ile yazdırılması hem mümkün hem de daha kolay ve hızlı olabilir.

Bu çalışmada gerçekleştirilen sistem bazı değişiklikler ile endüstri, biyomedikal sistemler, güvenlik, robotik, günlük hayattaki cihaz vb. alanlarda ses ile kontrol işlemleri için de kullanılabilir. Böylece insan rahatı ve konforunun artırabileceği, pratiklik sağlanabileceği ve verim artışı sağlanabileceği düşünülmektedir [6,20,21]. Sistem modüler ve yazılım yolu ile kontrol edilebilen esnek bir yapıya sahip olup, gerektiğinde tasarımcıların ihtiyacına göre farklı amaçlarla birçok uygulamalarda kullanılabilir.

Geliştirilen sistemde önceden hafıza birimine kaydedilen sesler kullanılarak, daha sonraki konuşmalar mikrofonla algılanarak tanınmakta ve bu sesin karşılığı olan sayısal elektrik sinyal elde edilmektedir. Elde edilen sinyaller bilgisayara farklı yöntemler ile gönderilebilir [22]. Bu çalışmada kullanımı en yaygın yöntemlerden olan RS232 ve PS/2 iletişimi yöntemleri kullanılarak veri iletişimi yapılmıştır. Bu çalışmada veri iletişimde her iki yöntemde de ayrı birer mikrodenetleyici (MCU: Microcontroller) kullanılmaktadır. RS232 metodunda önce bir RS232 dönüştürücüsü kullanılarak MCU ile PC arasındaki lojik voltaj seviyeleri uyumlu hale getirilmiştir. Görüntüleme işlemleri için PC’de geliştirilebilecek olan bir kullanıcı ara yüzü veya işletim sisteminde bulunan Hyper Terminal kullanılabilir. Ayrıca ilgili port, baudrate hızı, kaç bit gönderileceği, verilerin hangi pin ile gönderileceği ve alınacağı belirlenmelidir. PS/2 metodunda, MCU donanımı ve yazılımı ile veriler PS/2 girişinden gönderilmektedir. Mikrofon ile algılanan seslere ait veriler önceden açılan bir text dosyasına yazdırılabilir [22-25]. Bu çalışmanın devam eden kısımlarında önce Material ve metod verilmekte, akabinde gerçekleştirilen sistem açıklanmaktadır. Daha sonra sırası ile yapılan testler, ulaşılan sonuçlar ve kaynaklara yer verilmektedir.

2. MATERYAL METOD (MATERIAL AND METHOD)

2.1. Ses Tanılama (Speech Recognition)

Giriş bölümünde de ifade edildiği gibi, yüksek performanslı ses tanıma işlemleri çoğunlukla bir PC ve bu amaçla geliştirilen yazılımlar yolu ile yapılmaktadır. Bu sistemlerde ilave olarak PC uyumlu ses kartları da gerekli olabilmekte ve yazılımlar işletim sisteminin arkasında sürekli çalışmaktadır. Bu tür sistemlerde minyatürizasyon, ağırlık, maliyet ve tanılama işlemleri için ise karmaşıklık vb. durumları göz önüne alındığında PC gereksinimi ve sürekli çalışan yazılımlar bir dezavantaj olarak değerlendirilebilir [16,17].

Taşınabilir ve gömülü uygulamalarda düşük maliyet, doğru tanıma vb. tür sorunlara çözüm getirebilmek için konuşmacı bağımlı tanıma işlemi gerçekleştirebilen sistemler, yazılımlar ve donanımsal ve yazılımsal çözümler sunan ses sentezleyici entegre (IC: Integrated Circuit) devreler vb. çalışmaları yapılmaktadır. Şekil 1' de verilen HM2007 ses sentezleyici IC kullanılarak geliştirilecek devreler ile 0.96 s'lik 40 kelime ya da 1.92 s'lik 20 kelime tanınabilir. Hafıza için devrede 8k x 8 Statik RAM kullanılmaktadır. Bu IC' nin iki çalışma modu vardır: manuel mod ve CPU modu. CPU modu ile bir ana bilgisayar kontrolünde çalışması da sağlanabilir. Burada verilen devrede manuel modda çalışma gerçekleştirilmektedir. İşlemci ile iletişim kurmak için tasarımda tuş takımı ve 2 digit sayısal göstergeye yer verilmiştir. Böylece hem ses kaydı ve testi yapılabilir ve hem de silinebilir veya hata kodları görülebilir.

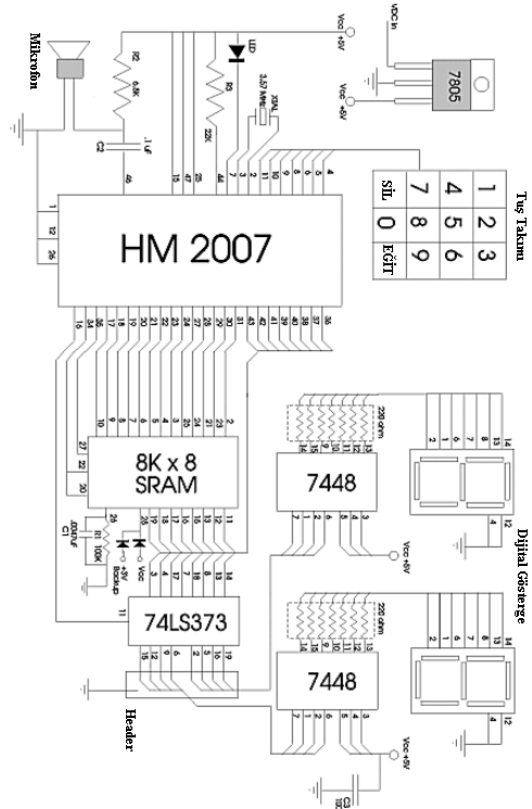
Bu devreye ilk enerji verildiğinde statik RAM hafıza kontrol edilir. Ekranda "00" görülür ve kırmızı LED ışık verir. Bu devrenin yukarıdaki işlemler için hazır olduğunu gösterir. Örneğin bir ses kaydı için önce TRN (Train: Eğitim) butonuna basılır. Sonra 1 ile 40 arasında ses kaydının yapılacağı bir dizin numarası girilir. Eğer LED yanıp sönerse kayıt başarılıdır. Tekrar TRN butonuna basılarak işlem bitirilir. Arzu edilen sayıda ses kayıt edildikten sonra test yapılabilir. Ses mikrofona konuşulduğunda, sayısal ekranda kayda ait dizin adresi görüntülenir [17,18,26,27].

2.2. Seri ve Paralel İletişim (Serial and Parallel Communication)

MCU ve diğer cihazlar ile PC arasında veri transferi için seri ve paralel iletişim yöntemleri kullanılabilir. Paralel veri iletişimde kullanılan hat sayısı fazla ve maliyeti daha yüksektir. Seri veri iletişimde ise hız düşük olmasına karşın, sadece 2 hat kullanılır.

Seri haberleşme protokolü senkron ve asenkron iletişim olmak üzere ikiye ayrılır:

Senkron iletişim: Senkron iletişimde bir saat işareti kullanılarak senkronize iletişim yapılır.



Şekil 1. HM2007 IC işlemci kullanılan ses tanıma devresi (Speech recognition circuit used HM 2007 IC processor)

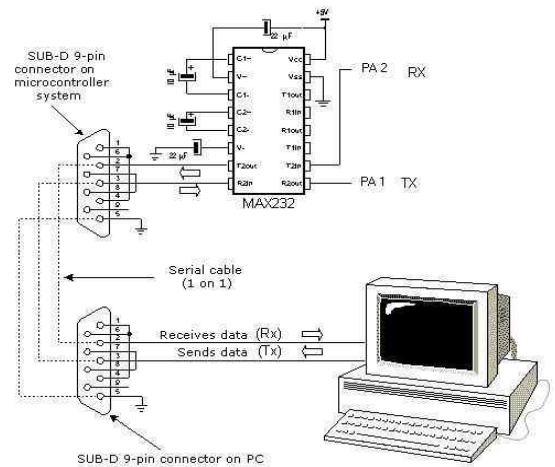
Asenkron İletişim: Asenkron iletişimde, saat işareti yerine başlangıç, bitiş bitleri kullanılır. Senkron iletişime göre daha yavaştır. Seri iletişimin 3 metodu vardır:

Full-duplex seri iletişim: Bu yöntemde her iki taraf da alıcı ve verici olarak çalışır.

Half-duplex seri iletişim: Bu yöntemde de her iki taraf alıcı ve verici olarak çalışabilir.

Simplex seri iletişim: Bu yöntemde bir taraf verici, diğer taraf ise alıcıdır.

2.3. RS232 Sürücü ve MAX232 Entegre Devresi (RS232 Driver and MAX232 Integrated Circuit)



Şekil 2. RS232 seri veri iletişimde kullanılan MAX232 IC bağlantısı (MAX232 IC connection in the RS232 circuit using serial communication)

RS232 seri asenkron veri iletişimi için geliştirilen bir standarttır. Veri hatlarında tipik gerilim seviyeleri +12 V ve -12V'dur. MCU'lar ise TTL Lojik 1 (+5V) veya Lojik 0 (0V) seviyelerinde çalışırlar. Bu değerleri PC ve denetleyici arasında her iki yönde de uygun seviyeye getirmek için MAX232 IC kullanılabilir (Şekil 2). 25 pinli DB25 veya 9 pinli DB-9 konnektörü kullanılabilir [22,23,28,29].

2.4. Hyper Terminal (Hyper Terminal)

Hyper Terminal alınan verileri görüntülemek için kullanılabilen bir ara yüzdür. PC'de (Start>Programs>Accessories>Communications>Hyper terminal) yolu ile ulaşılabilir. Veriler PC'ye seri veya paralel olarak iletebilir. RS232'nin bağlanacağı port ve seri iletişim hızının belirtilmesi, yazılımda bit sayısı ve baud rate hızının ayarlanması gereklidir [25,28].

2.5. PS/2 İletişimi ve PC Klavye (PS/2

Communication and PC Keyboard)

Klavyelerin en çok kullanılan bağlantı şekilleri DIN ve PS/2' dir. Klavye haberleşmesinde her bir byte veri için 11 bit gönderilir. 1 bit start biti, 8 bit data biti, 1 bit parity biti ve stop bitidir. PS/2 portu mouse ve klavye bağlantısı için üretilen, 6 pinli konektörden oluşan, düşük hızlı bir seri porttur. Herhangi bir tuşa basıldığında klavyeden PC'ye 8 bitlik özel tuş verileri data ve clock hatları kullanılarak yani senkron seri iletişim protokolü ile gönderilir. Bir tuşa basıldığında 1 Byte'lık, aynı tuş bırakıldığında 2 Byte'lık bir Hex kodu gönderilir. Böylece PC tarafından basılan veya bırakılan tuş belirlenir. Örneğin klavyede "A" tuşuna basıldığında "1C" hex kodu, basma işlemi sona erdiğinde önce "F0" sonra da "1C" hex kodları üretilir [24,30].

3. GERÇEKLEŞTİRİLEN SİSTEM (REALIZED SYSTEM)

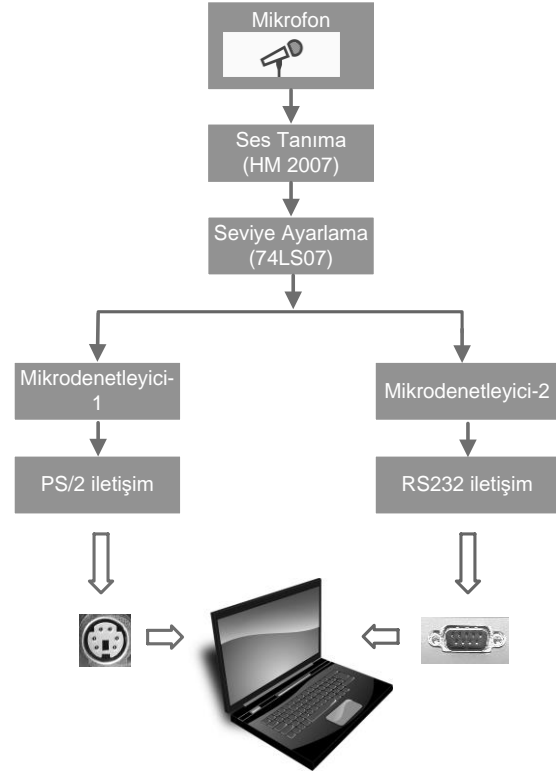
Bu çalışmada ses tanılama işlemi yapılarak, bir sesli kontrol sistemi ve PC' de yazı yazılabilen bir sistem tasarlanmış ve gerçekleştirilmiştir. Sistemin temel blok yapısı ve genel görünümü Şekil 3 ve Şekil 4' de görülmektedir. Tasarıma ait donanım, yazılım, bunların açıklamaları ve testleri izleyen bölümlerde verilmiştir.

Gerçekleştirilen sistem temel olarak altı donanım birimi ve ilgili yazılımlardan oluşur. Bunlar şöyle sıralanabilir:

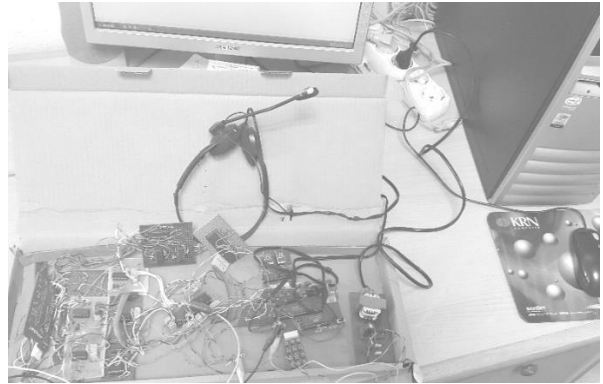
- Ses tanılama birimi,
- Seviye düzenleme ve izolasyon birimi [26],
- MCU birimi,
- RS232 veri iletişim birimi,
- PS/2 veri iletişim birimi.

Bu sistemde mikrofon ile algılanan sesler HM2007 IC ile işlenir ve veriler harici bir 8K x 8 SRAM IC'de daha önceden kaydedilip saklanan seslere ait veriler ile karşılaştırılarak eşleşen ses bulunur ve ses tanılama işlemi gerçekleştirilir. Daha sonra voltaj seviyesi ayarlanarak veriler iki ayrı MCU'ya uygulanır. Burada verilerin karşılaştırması yapılarak ilgili karakter kodu elde edilir. Bu sistem yazılım yolu ile kontrol edilen bir birim olduğundan, sunulan esneklik sayesinde gerektiğinde ses ile kontrol amaçları içinde kullanılabilir. Bu MCU'lar ile tanılama işlemi yapılmış olan sese ait

veriler PC' ye biri RS232 diğeri ise PS/2 iki iletişim yöntemi kullanılarak aktarılır.



Şekil 3. Sistemin temel blok yapısı (Basic block structure of the system)



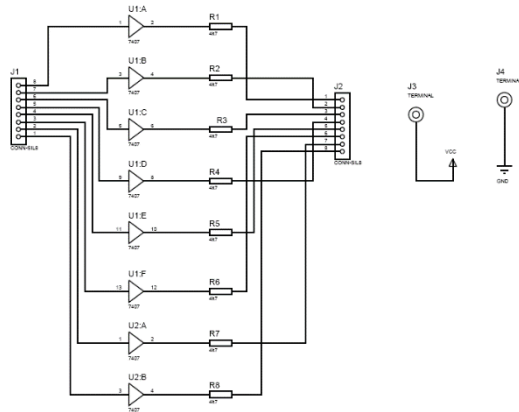
Şekil 4. Sistemin genel görünümü (Overview of the system)

3.1. Voltaj Seviye Düzenleme ve İzolasyon Birimi (Level Regulation of Voltage and Isolation Circuit)

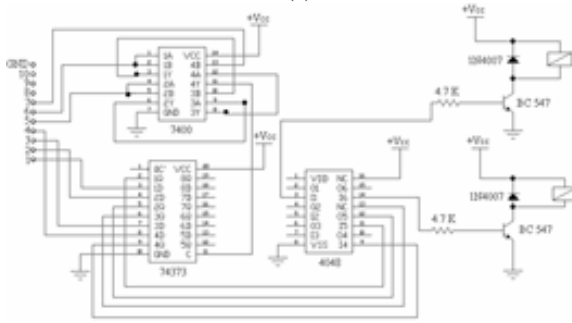
Ses tanıma devresinde yüksek çıkış 2.5 V, düşük çıkış 0.5 V civarında ölçülmüştür. Ses tanılama işleminden sonra elde edilen sayısal işaretin MCU' ya uygulanmadan önce Lojik Seviyesini TTL uyumlu hale getirmek için 74LS07 buffer IC kullanılarak, sinyal ihtiyaç duyulan TTL H(5V) ve L(0V) seviyelerine elektriksel bir izolasyon sağlanarak çekilmiştir (Şekil 5-a) [26, 28].

Ses tanıma devresi kullanılarak elektrikli cihazların kontrolü Şekil 5-b' de verilen sürücü devre ile gerçekleştirilir. Bu devrenin girişi ses tanıma devresinin sayısal gösterge için de kullanılan 10 pinli çıkışına bağlanmaktadır. Ses tanındığında, ilgili çıkış yüksek seviyeye çekilir. Bu şekilde 10 farklı yük sürülebilir.

Sistemde kelimenin tanınması ile rölenin enerjilenmesi arasındaki süre yaklaşık 0.25 s'dir [26].



(a)



(b)

Şekil 5. (a) Voltaj seviye düzenleme ve izolasyon birimi (b) Sürücü devre (a) Voltage regulation and isolation circuit (b) Driver circuit

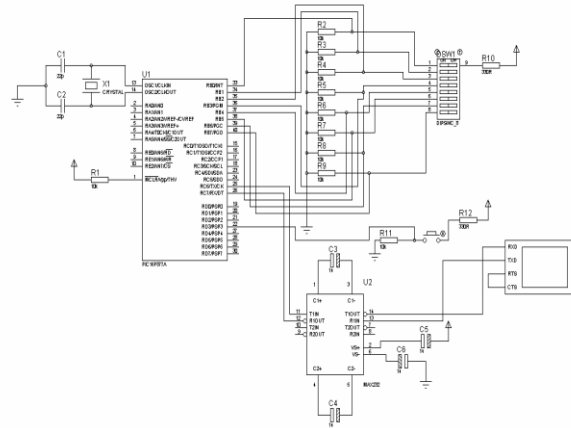
3.2. PC ile İletişim (Communication with the PC)

PC ile MCU kullanılan kontrol devresi arasında veri iletişimi iki yöntem kullanılarak gerçekleştirilmiştir. Bunlar RS232 ve PS/2 yöntemleridir:

3.2.1. RS232 sürücü ve MCU kullanılan kontrol devresi (RS232 driver and control circuit used MCU)

Birinci yöntem asenkron veri iletişiminin kullanıldığı RS232 standardıdır. Bu yöntemde -12V ve +12V "Lojik 1", "Lojik 0" değerine karşılık gelir. Bu voltajlar ile TTL seviye uyumunu her iki yönde de sağlayabilmek için MAX232 IC kullanılmıştır [22,23,28,29]. MAX232 ve MCU arasında sadece veri gönderme işlemi yapıldığı için tek yönlü iletişim (Simplex Seri İletişim) kullanılmıştır. RS 232 iletişimini gerçekleştirmek için kullanılan donanım ve yazılımın akış şeması Şekil 6 ve Şekil 7'de verilmiştir. Devrede görülen 8x1 DIP Switch ses tanımlama devresinin çıkışındaki sayısal işareti sembolize eder ve bir PIC16F877A denetleyicinin girişine uygulanmıştır [29, 31]. Gerçekleştirilen sistemde ses tanımlama işleminin nasıl gerçekleştiği şöyle açıklanabilir: Örneğin "ALI" sesleri için ses tanımlama devresine önceden kayıtlar yapılır. "A" sesi şayet 01. adrese kayıt edilmiş ise, mikrofon "A" sesini algıladığında ses tanıma devresi çıkışında 01 in 8 Bit'lik

karşılığı olan sinyal elde edilir ve Seven Segment Display' da "01" yazar.



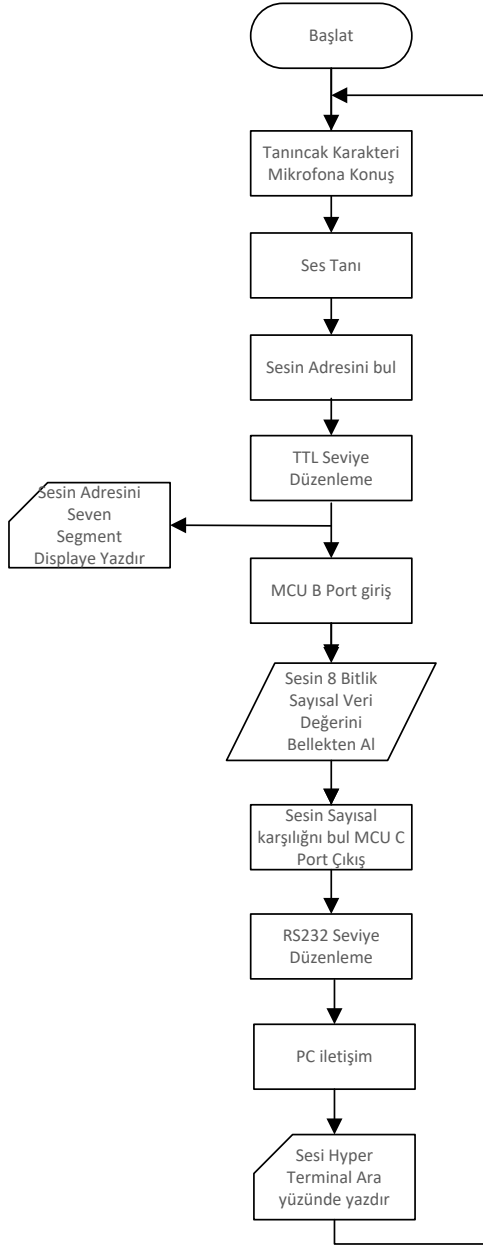
Şekil 6. RS232 sürücü ve MCU kullanılan kontrol devresi (RS232 driver and control circuit used MCU)

Bu veri aynı zamanda denetleyicinin B portuna giriş olarak uygulanır ve kayıtlı olan diğer verilerle karşılaştırılarak "A" harfinin karşılığı olan veri C portundan alınır, MAX232 IC ile gerilim seviyesi düzenlenerek RS232 portundan PC'ye gönderilir. PC işletim sisteminde bulunan Hyper terminalde ise gönderilen "A" harfi gözlemlenir. Daha sonra aynı işlemler "L" ve "I" sesleri içinde yapılır. Devrede kullanılan buton açıkken hiçbir işlem yapılmaz. "A" harfi yazdırılmak istendiğinde buton basılı iken konuşma yapılır. Eğer buton sürekli basılı kalırsa "A" yazmaya sürekli devam eder.

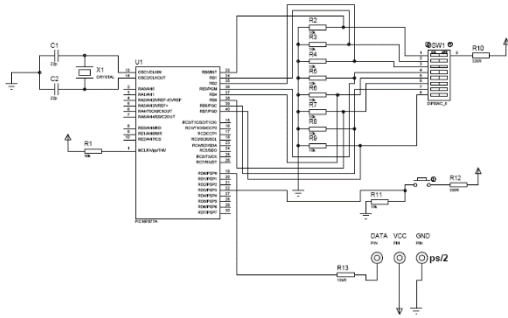
3.2.2. PS/2 klavye sürücü ve MCU kullanılan kontrol devresi (PS/2 keyboard driver and control circuit used MCU)

Gerçekleştirilen sistemde ses tanıma devresi ile PC arasında veri iletişimi için kullanılan ikinci yöntem PS/2 yöntemidir. RS232 veri iletişimine benzer biçimde yine HM2007 ses tanıma devresi çıkışından alınan veriler "Seviye Düzenleme ve İzolasyon Birimi" devresi kullanılarak TTL uyumu sağlanmış ve daha sonra 2. MCU kontrol birimine verilmiştir (Şekil 8). Burada farklı olarak, daha önce MCU'ya kaydedilen PC ASCII kodları [28] ile, gelen ses tanıma verisi Şekil 9'da akış şeması verilen yazılım ile karşılaştırılmaktadır. Örneğin mikrofondan "A" algılandığında MCU'ya "A" verisinin kodları gelir. MCU ile gelen bu kod "A" harfinin ASCII kodu olan Dec "97" ile karşılaştırılarak PS/2 girişine gönderilir.

Daha önce Bölüm 2.5'te açıklanan her bir klavye tuş kodu, burada verilen PIC MCU kullanılan kontrol birimi ile tekrar ölçülerek kontrol edilmiş ve kontrol işlemleri bunlara göre yapılmıştır. Gerçekleştirilen sistemde PIC MCU devresi ile 10 kHz değerinde bir clock sinyali üretilmiş ve veriler, bu sinyalin düşen kenarlarında gönderilmiştir.

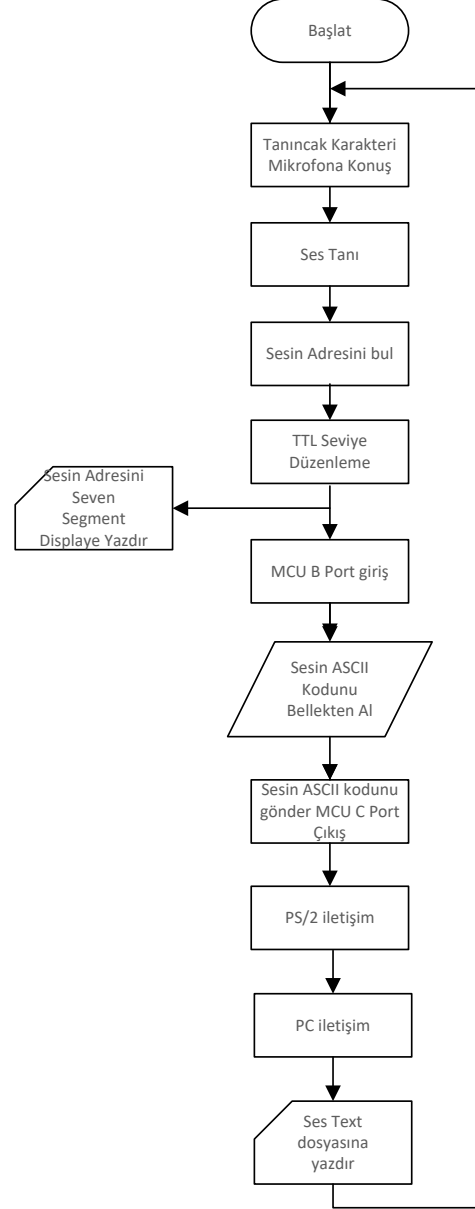


Şekil 7. Gerçekleştirilen sistemde ses tanıma ve RS232 seri iletişimin akış şeması (The flow chart of speech recognition and RS232 serial communication).



Şekil 8. PS/2 klavye sürücü ve MCU kullanılan kontrol devresi (PS/2 keyboard driver and control circuit used MCU)

Gerçekleştirilen devrede klavye tuşlarının 275mA'den fazla akım çekmesi nedeni ile, 7805 IC kullanılan ilave bir regüleli güç kaynağı tasarlanmış ve gerilim önerilen +4.5V ile +5.5V aralığında tutulmuştur. Klavyenin data ve clock uçları açık kollektör (open collector) yapısı nedeni ile 4k7 değerinde Pull-Up dirençleriyle Lojik 1 (+5V) seviyesine çekilmiştir [24].



Şekil 9. Gerçekleştirilen sistemde ses tanıma ve PS/2 iletişim akış şeması (Flow chart of speech recognition and PS/2 communication)

4. TESTLER (TESTS)

Gerçekleştirilen sistemde temel olarak iki grup altında testler yapılmıştır. Birincisi bu çalışmada yazı yazma işlemi harf tanıma temelli olduğu için Harf Tanıma Temelli Testler, diğeri ise Kelime Tanıma Temelli Testlerdir. Testlerde örneklem gruplarında 3 bayan 6 kişi yer almıştır. Testler değişik gürültü değerlerinde, % 38 bağıl neme sahip, 28 °C sıcaklıkta ve

ortam basıncındaki 5x4x3 m boyutlara sahip bir ortamda yapılmıştır [32]. Elde edilen veriler SPSS paket programı (SPSS Inc., Chicago, IL, USA) kullanılarak istatistiksel analize tabi tutulmuştur. Gürültü ölçümleri için 1/3 oktav bant filtresine sahip, Rion NL-21 cihazı kullanılmış [33] ve alınan sonuçlar aşağıda verilmiştir:

4.1. Harf Tanıma Temelli Testler (Character Recognition Based Testing)

Harf tanıma temelli testler yapılırken Çizelge 1' de verilen Türkçe konuşma seslerinin özellikleri göz önüne alınarak elde edilen sonuçlar yorumlanmıştır.

Çizelge 1. Türkçe konuşma sesleri (Turkish speech letters)

Ünlüler	Geniş	Dar	Geniş	Dar
Kalın	a	ı	o	u
İnce	e	i	ö	ü
Ünsüzler	Sürtünme	Patlamalı	Burunsal	Kaygan
Nefesli	c,j,v,z	b,d,g	m,n	ğ,l,r,y
Nefessiz	ç,f,h,s,ş	t,k,p	-	-

Ses tanıma sisteminde mikrofon aracılığı ile algılanan harflere ait sesler, bunların tanınma yüzdeleri ve grafikleri sırası ile Çizelge 2 ve Şekil 10' da verilmiştir. Burada elde edilen sonuçlara göre, ünlü sesler, ses tellerinin titreşmesi ile üretilen ve ses yolunda bir engelle uğramayan sesler olduğu için tanınma oranlarının daha yüksek olduğu düşünülmektedir. En düşük %95, en yüksek %100 ve ortalama %99.375 tanınma oranı elde edilmiştir. Ünsüz seslerin ise ses yolunda karşılaştıkları engeller nedeni ile tanınma oranlarının düştüğü değerlendirilmektedir (%93.4). Ünlü ve ünsüz harfler için ortalama %95.2 doğrulukla tanıma gerçekleştirilmiştir.

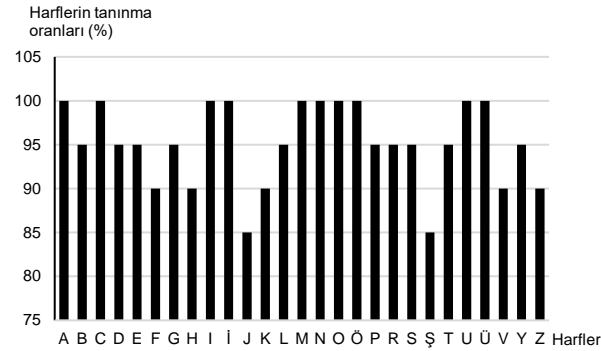
Çizelge 2. Harflere ait sesler ve bunların tanınma yüzdeleri (Speech recognition of the letters and their percentages)

Harf	A	B	C	D	E	F	G	H	I	İ
Tanınma Oranı(%)	100	95	100	95	95	90	95	90	100	100
Harf	J	K	L	M	N	O	Ö	P	R	S
Tanınma Oranı(%)	85	90	95	100	100	100	100	95	95	95
Harf	Ş	T	U	Ü	V	Y	Z			
Tanınma Oranı(%)	85	95	100	100	90	95	90			

Yine bu sonuçlara göre düzenli sesbirimi geçişleri olmayan, gürültü faktörünün mevcut olduğu patlamalı ünsüzler ve sürtünmeli ünsüzler için tanınma oranlarının nispeten düşük bulunduğu, ağzın tamamen kapatılması ile oluşan ünsüzlerde telaffuzun zorlaşması nedeni ile tanınma doğruluğunun düştüğü değerlendirilmektedir. Ayrıca kaygan ünsüz harflerin ise bazı ünlü harflere benzemesi sebebi ile ayırt etme doğruluğunun azaldığı düşünülmektedir.

4.2. Kelime Tanıma Temelli Testler (Word Recognition Based Testing)

Kelime Tanıma Temelli olan testlerde sistem tarafından daha yüksek performansta tanıma sağlanabilmesi için,



Şekil 10. Harflerin ortalama tanınma oranları grafiği (Graph of the speech recognition of the letters and their percentages)

sisteme eğitilecek olan kelimelerin sahip olması gerekli olan özelliklerine ilave olarak konuşmacının ses özellikleri de araştırılmıştır. Yazı yazma işlemi harf temelli olduğundan, bu testler girişten MCU kullanılan kontrol biriminin çıkışına kadar olan birimler için yapılmıştır. Bu testlerde sağlanan verilerin yorumlanması ile aynı zamanda yeni çalışmalarda kullanılması düşünülen bu sistemin güvenilirliği de belirlenmeye çalışılmıştır. Sistemin performans testleri, eş sesli olan ve eş sesli olmayan ve Çizelge 3 ve Çizelge 4' de sırası ile verilen eğitim verileri kullanılarak gürültülü ve gürültüsüz ortamda yapılmıştır.

Çizelge 3. Eşsesli kelime tanıma için seçilen eğitim verileri (Selected training data for homonym word recognition)

Eğitim veri No	Eğitim verileri
1	ileri
2	geri
3	ısıt
4	ışık
5	kapı
6	kapat
7	dvd
8	vcd
9	televizyon
10	müzik

Çizelge 4. Eşsesli olmayan kelime tanıma için seçilen eğitim verileri (Selected training data for non homonym word recognition)

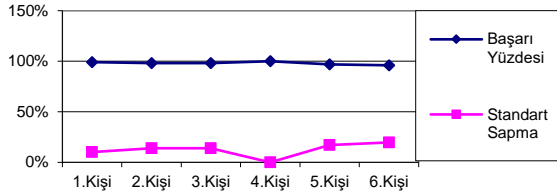
Eğitim veri No	Eğitim verileri
1	perde
2	radio
3	telefon
4	bilgisayar
5	klima
6	sağ
7	lamba
8	aç
9	ısıtıcı
10	garaj

Çizelge 5 ve Şekil 11' de gürültüsüz ortamda eş sesli olmayan kelimelerle gerçekleştirilen tanıma sonuçları, Çizelge 6 ve Şekil 12' de ise gürültüsüz ortamda eş sesli olan kelimelerle gerçekleştirilen tanıma sonuçları verilmiştir. Ayrıca bu iki testin karşılaştırmalı sonuçları

Çizelge 7 ve Şekil 13'te beraber sunulmuştur. Gürültülü ortam testleri ise daha sonraki kısımlarda verilmiştir.

Çizelge 5. Gürültüsüz ortamda eşesli olmayan kelime tanıma oranları (Non-homophones word recognition rates in noiseless environment)

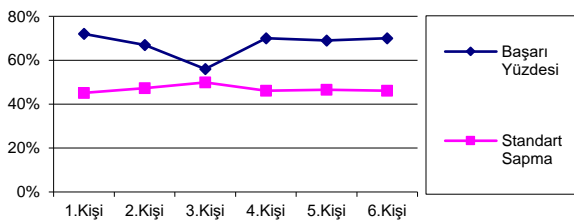
Konuşmacı Ses Örnekleri	Toplam Örnek Kelime Sayısı	Toplam Tekrar Sayısı	Doğru Tanınan Örnek Kelime Sayısı	Yanlış Tanınan Örnek Kelime Sayısı	Başarı Yüzdesi	Standart Sapma
#1	10	10	99	1	%99	,10000
#2	10	10	98	2	%98	,14071
#3	10	10	98	2	%98	,14071
#4	10	10	100	0	%100	,00000
#5	10	10	97	3	%97	,17145
#6	10	10	96	4	%96	,19695



Şekil 11. Gürültüsüz ortamda eşesli olmayan kelime tanıma oranları (Non-homophones word recognition rates in noiseless environment)

Çizelge 6. Gürültüsüz ortamda eşesli olan kelime tanıma oranları (Homophones word recognition rates in noiseless environment)

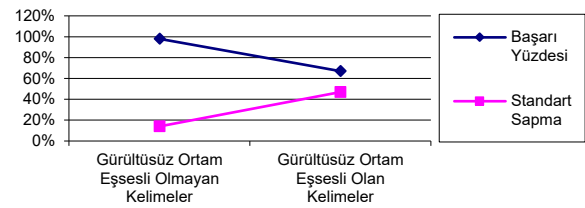
Konuşmacı Ses Örnekleri	Toplam Örnek Kelime Sayısı	Toplam Tekrar Sayısı	Doğru Tanınan Örnek Kelime Sayısı	Yanlış Tanınan Örnek Kelime Sayısı	Başarı Yüzdesi	Standart Sapma
#1	10	10	72	28	%72	,45126
#2	10	10	67	33	%67	,47258
#3	10	10	56	44	%56	,49889
#4	10	10	70	30	%70	,46057
#5	10	10	69	31	%69	,46482
#6	10	10	70	30	%70	,46057



Şekil 12. Gürültüsüz ortamda eşesli olan kelime tanıma oranları (Homophones word recognition rates in noiseless environment)

Çizelge 7. Gürültüsüz ortamda eşesli olan ve eşesli olmayan kelimeler için genel yüzde ve ortalama kelime tanıma oranları (Overall percentage and average word recognition rates for homonyms and non-homonyms in noiseless environments)

Koşullar	Toplam Tekrar Sayısı	Doğru Tanınan Örnek Kelime Sayısı	Yanlış Tanınan Örnek Kelime Sayısı	Başarı Yüzdesi	Standart Sapma
Gürültüsüz Ortam Eşesli Olmayan Kelimeler	600	588	12	%98	,14012
Gürültüsüz Ortam Eşesli Olan Kelimeler	600	404	196	%67	,46939



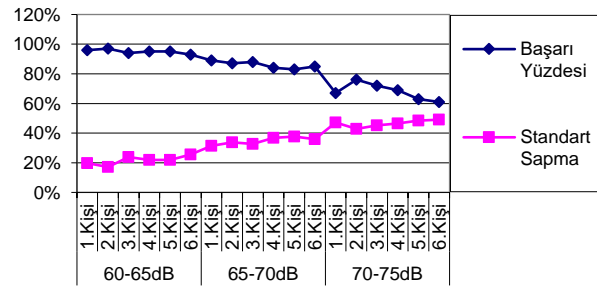
Şekil 13. Eşesli olan ve eşesli olmayan kelimeler için gürültüsüz ortamda genel yüzde ve ortalama kelime tanıma oranları (Overall percentage and average word recognition rates for homonyms and non-homonyms in noiseless environments)

Eşesli olan ve olmayan seslerin gürültüsüz ortamda 1200'e yakın ölçümle elde edilen sonuçları genel olarak incelendiğinde, kelimeler neticede harflerden oluştuğu için, bunların bir önceki aşamada gerçekleştirilen harf temelli testlerin sonuçları ile uyumlu olduğu düşünülmektedir. Bu bağlamda, eş sesli olmayan kelimelerle gerçekleştirilen ses tanımadaki başarının, aynı şartlarda eş sesli olan kelimelerle gerçekleştirilen tanımayla göre daha yüksek olmasının temel sebebinin, söylenişleri birbirine benzer olan eğitim verileri olduğu değerlendirilmektedir. Eşesli olmayan kelimeler nerede ise %100 oranında tanınırken, eş sesli olan kelimelerde bu oran %55-%72 arasındadır. Genel olarak tanınma oranı daha fazla olan harfleri barındıran kelimelerin daha yüksek doğrulukla tanınması beklenmektedir [26,34,35]. Bu sonuçlar üzerinde eğitim verilerinin farklı olmasının da etkili olduğu düşünülebilir.

Eşesli olan ve olmayan seslerin farklı gürültü seviyelerine sahip ortamda tanınmalarına ait ölçümler, ölçüm sayıları ve elde edilen sonuçlar ayrıntılı olarak Çizelge 8 ve 9 ile Şekil 14 ve 15' de verilmiştir. Ayrıca bu testlerin karşılaştırmalı sonuçları Çizelge 10 ve Şekil 16'da beraber sunulmuştur. Bu testlerde Gürültü ölçümleri için RION NL 21 cihazı kullanılmıştır [33].

Çizelge 8. Gürültülü ortamda eşesli olmayan kelime tanıma oranları (Non-homonym word recognition rates in noisy environment)

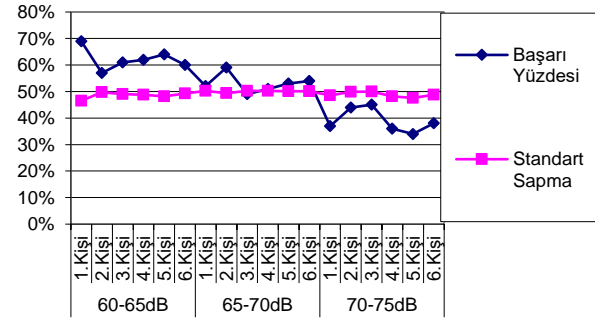
Eşesli Olmayan Kelimeler İçin dB Aralıkları	Konuşmacı Ses Örnekleri	Toplam Örnek Kelime Sayısı	Toplam Tekrar Sayısı	Doğru Tanınan Örnek Kelime Sayısı	Yanlış Tanınan Örnek Kelime Sayısı	Başarı Yüzdesi	Standart Sapma
60-65dB	1.Kişi	10	10	96	4	96%	0,19695
	2.Kişi	10	10	97	3	97%	0,17145
	3.Kişi	10	10	94	6	94%	0,23868
	4.Kişi	10	10	95	5	95%	0,21904
	5.Kişi	10	10	95	5	95%	0,21904
	6.Kişi	10	10	93	7	93%	0,25643
65-70dB	1.Kişi	10	10	89	11	89%	0,31447
	2.Kişi	10	10	87	13	87%	0,338
	3.Kişi	10	10	88	12	88%	0,326
	4.Kişi	10	10	84	16	84%	0,36845
	5.Kişi	10	10	83	17	83%	0,37753
	6.Kişi	10	10	85	15	85%	0,35887
70-75dB	1.Kişi	10	10	67	33	67%	0,47258
	2.Kişi	10	10	76	24	76%	0,42923
	3.Kişi	10	10	72	28	72%	0,45126
	4.Kişi	10	10	69	31	69%	0,46482
	5.Kişi	10	10	63	37	63%	0,48524
	6.Kişi	10	10	61	39	61%	0,49021



Şekil 14. Gürültülü ortamda eşesli olmayan kelime tanıma oranları (Non-homonym word recognition rates in noisy environment)

Çizelge 9. Gürültülü ortamda eşesli olan kelime tanıma oranları (Non-homonym word recognition rates in noisy environment)

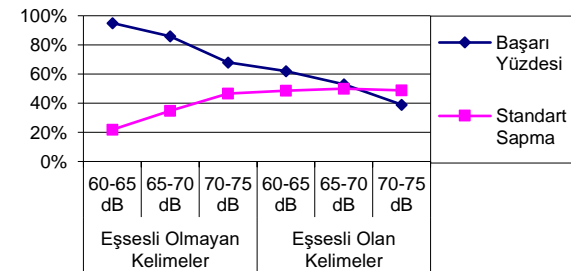
Eşesli Olan Kelimeler İçin dB Aralıkları	Konuşmacı Ses Örnekleri	Toplam Örnek Kelime Sayısı	Toplam Tekrar Sayısı	Doğru Tanınan Örnek Kelime Sayısı	Yanlış Tanınan Örnek Kelime Sayısı	Başarı Yüzdesi	Standart Sapma
60-65dB	1.Kişi	10	10	69	31	69%	0,46482
	2.Kişi	10	10	57	43	57%	0,49757
	3.Kişi	10	10	61	39	61%	0,49021
	4.Kişi	10	10	62	38	62%	0,48783
	5.Kişi	10	10	64	36	64%	0,48242
	6.Kişi	10	10	60	40	60%	0,49237
65-70dB	1.Kişi	10	10	52	48	52%	0,50212
	2.Kişi	10	10	59	41	59%	0,49431
	3.Kişi	10	10	49	51	49%	0,50242
	4.Kişi	10	10	51	49	51%	0,50242
	5.Kişi	10	10	53	47	53%	0,50161
	6.Kişi	10	10	54	46	54%	0,50091
70-75dB	1.Kişi	10	10	37	63	37%	0,48524
	2.Kişi	10	10	44	56	44%	0,49889
	3.Kişi	10	10	45	55	45%	0,5
	4.Kişi	10	10	36	64	36%	0,48242
	5.Kişi	10	10	34	66	34%	0,4761
	6.Kişi	10	10	38	62	38%	0,48783



Şekil 15. Gürültülü ortamda eşesli olan kelime tanıma oranları (Homonym word recognition rates in noisy environment)

Çizelge 10. Gürültülü ortamda eşesli olan ve eşesli olmayan kelimeler için genel yüzde ve ortalama kelime tanıma oranları (Overall percentage and average word recognition rates for homonyms and non-homonyms in noisy environments)

Gürültülü Ortamda Genel Yüzde ve Ortalama Sonuçlar	Koşullar	Toplam Tekrar Sayısı	Doğru Tanınan Örnek Kelime Sayısı	Yanlış Tanınan Örnek Kelime Sayısı	Başarı Yüzdesi	Standart Sapma
Eşesli Olmayan Kelimeler	60-65 dB	600	570	30	95%	0,21813
	65-70 dB	600	516	84	86%	0,34728
	70-75 dB	600	408	192	68%	0,46687
Eşesli Olan Kelimeler	60-65 dB	600	372	228	62%	0,48579
	65-70 dB	600	318	282	53%	0,49952
	70-75 dB	600	234	366	39%	0,48816



Şekil 16. Eşesli olan ve eşesli olmayan kelimeler için gürültülü ortamda genel yüzde ve ortalama kelime tanıma oranları (Overall percentage and average word recognition rates for homonyms and non-homonyms in noisy environments)

Ölçümler 60-65 dB, 65-70 dB ve 70-75 dB aralığında gürültü seviyelerinde gerçekleştirilmiştir. Bu Çizelge ve şekiller ve Çizelge 10 ve Şekil 16' da kıyaslamalı olarak görüldüğü gibi, eşesli olmayan kelimeler için 60-65 dB seviyelerinde gürültülü ortamda %95, 65-70 dB gürültülü ortamda %86 ve 70-75 dB gürültülü ortamda %68 tanıma oranına ulaşılmıştır. Eşesli olan kelimeler için ise 60-65 dB seviyelerinde gürültülü ortamda %62, 65-70 dB gürültülü ortamda %53 ve 70-75 dB gürültülü ortamda %39 tanıma oranı sağlanmıştır.

Bu ölçümlerde alınan sonuçlara göre, eşsesli olmayan kelimeler 65 dB gibi daha düşük gürültülü ortamlarda %95 oranında tanınırken, eş sesli olan kelimelerde aynı şartlarda bu oran %62 seviyelerindedir. Bu durum gürültülü ortamların tanınma oranının düşmesine neden olduğunu göstermektedir. Ayrıca genel olarak tanınma oranı daha fazla olan harfleri barındıran kelimelerin daha yüksek doğrulukla tanınması beklenen bir durumdur [26,34,35]. Bu sonuçlar üzerinde eğitim verilerinin farklı olmasının da etkili olduğu düşünülebilir. Bu sebeple, bu test sonuçları genel olarak değerlendirildiğinde, bunların daha önceki aşamalarda gerçekleştirilen harf temelli testlerin ve gürültüsüz ortamda elde edilen testlerin sonuçları ile uyum sağlamada olduğu görülmektedir. Bu bağlamda çıkarılan diğer bir sonuç ise eş sesli olmayan kelimelerle gerçekleştirilen ses tanımadaki başarının, aynı şartlarda eş sesli olan kelimelerle gerçekleştirilen tanımaya göre daha yüksek olmasının temel sebebinin, söylenişleri birbirine benzer olan eğitim verileri olduğu düşünülebilir.

5. SONUÇ VE ÖNERİLER (RESULTS AND SUGGESTIONS)

Bu çalışmada gerek yerleşik ve gerekse gömülü sistem uygulamalarında kullanılabilecek, düşük maliyetli, yeni bir konuşmacı bağımlı ses tanıma ve kontrol sistemi tasarlanmış, daha yüksek doğrulukla konuşma tanımayı etkileyebilen parametreler araştırılmış ve PC' de yazı yazabilmek için veriler bir bilgisayara aktarılmıştır. Veri iletişimi için birer mikrodenetleyici (MCU) ile kontrol edilen PS/2 ve RS232 veri iletişimi yöntemleri kullanılmıştır.

Sistemin performansını belirlemek için gürültülü ve gürültüsüz ortamlarda testler yapılmıştır. Bu çalışmada yazı yazma işlemi harf tanıma temelli olduğu için Harf Tanıma Temelli Testler ve diğer testler olan Kelime Tanıma Temelli Testler gerçekleştirilmiştir. Testler yapılırken Türkçe konuşma seslerinin özellikleri araştırılmış ve elde edilen sonuçlar bu özellikler göz önüne alınarak yorumlanmıştır.

Harf tanıma temelli testlerde, ünlü sesler, ses tellerinin titreşmesi ile üretilen ve ses yolunda bir engelle uğramayan sesler olduğu için tanınma oranlarının daha yüksek olduğu düşünülmektedir. En düşük %95, en yüksek %100 ve ortalama %99.375 tanınma oranı elde edilmiştir. Ünsüz seslerin ise ses yolunda karşılaştıkları engeller nedeni ile tanınma oranlarının düştüğü değerlendirilmektedir (%93.4). Ünlü ve ünsüz harfler için ortalama %95.2 doğrulukla tanıma gerçekleşmiştir.

Yine bu sonuçlara göre düzenli sesbirimi geçişleri olmayan, gürültü faktörünün hakim olduğu patlamalı ünsüzler ve sürtünmeli ünsüzler için tanınma oranlarının nispeten düşük bulunduğu, ağızın tamamen kapatılması ile oluşan ünsüzlerde telaffuzun zorlaşması nedeni ile tanınma doğruluğunun düştüğü değerlendirilmektedir. Ayrıca kaygan ünsüz harflerin ise bazı ünlü harflere benzemesi sebebi ile ayırt etmede güçlüklerle karşılaşıldığı düşünülmektedir.

Kelime Tanıma Temelli olan testlerde ise sistem tarafından daha yüksek performansta tanıma sağlanabilmesi için, kelimelerin sahip olması gerekli olan özelliklerine ilave olarak konuşmacının ses özellikleri de araştırılmıştır. Bu testler aynı zamanda yeni çalışmalarda kullanılması düşünülen bu sistemin güvenilirliğini belirlemek için de yapılmıştır.

Kelime tanıma temelli olan testler gürültülü ve gürültülü olmayan ortamda eşsesli olan ve olmayan sesler için ayrı ayrı yapılmıştır.

Gürültüsüz ortam için yaklaşık 1200 ölçümle elde edilen veriler genel olarak incelendiğinde, kelimeler neticede harf temelli olduğundan, bir önceki aşamada gerçekleştirilen harf temelli testlerin sonuçları ile uyum sağlamada olduğu değerlendirilmektedir. Bu bağlamda, eş sesli olmayan kelimelerle gerçekleştirilen ses tanımadaki başarının, aynı şartlarda eş sesli olan kelimelerle gerçekleştirilen tanımaya göre daha yüksek olmasının temel sebebinin, söylenişleri birbirine benzer olan eğitim verileri olduğu düşünülmektedir. Eşsesli olmayan kelimeler nerede ise %100 oranında tanınırken, eş sesli olan kelimelerde bu oran %55-%72 arasındadır.

Eşsesli olan ve olmayan kelimelerin farklı gürültü seviyelerine sahip ortamda tanınmalarına ait ölçümler 60-65 dB, 65-70 dB ve 70-75 dB arasındaki gürültü seviyelerinde gerçekleştirilmiştir. Eşsesli olmayan kelimeler için 60-65 dB seviyelerinde gürültülü ortamda %95, 65-70 dB gürültülü ortamda %86 ve 70-75 dB gürültülü ortamda %68 tanıma oranına ulaşılmıştır. Eşsesli olan kelimeler için ise 60-65 dB seviyelerinde gürültülü ortamda %62, 65-70 dB gürültülü ortamda %53 ve 70-75 dB gürültülü ortamda %39 tanıma oranı sağlanmıştır.

Gürültülü ortamda yapılan ölçümlerde eşsesli olmayan kelimeler 65 dB gibi nispeten daha düşük gürültülü ortamlarda %95 oranında tanınırken, eş sesli olan kelimelerde aynı şartlarda bu oran %62 seviyelerindedir. Bu durum gürültülü ortamların tanınma oranının düşmesine neden olduğunu göstermektedir.

Bu test sonuçları hem eş sesli olan ve olmayan eğitim verileri için ve hem de gürültülü ve gürültüsüz ortamın sonuçları için genel olarak değerlendirildiğinde, genel olarak tanınma oranı daha fazla olan harfleri barındıran kelimelerin daha yüksek doğrulukla tanınması beklenmektedir [26,34,35]. Ayrıca bu sonuçlar üzerinde eğitim verilerinin farklı olmasının da etkili olduğu düşünülebilir. Bu bağlamda çıkarılan diğer bir sonuç ise eş sesli olmayan kelimelerle gerçekleştirilen ses tanımadaki başarının, aynı şartlarda eş sesli olan kelimelerle gerçekleştirilen tanımaya göre daha yüksek olmasının temel sebebinin, söylenişleri birbirine benzer kelimelerin kullanıldığı eğitim verileri olduğu yorumlanabilir.

Ayrıca bu çalışmanın amaçlarından birisi olan ses tanımadaki başarıyı artırmak için yapılan testlere göre de çıkarımı yapılan ve ayrıca bunlardan bazılarının ölçümleri yapılmasına rağmen açıklamasının çok uzun olması gerektiği için zor olan durumlardan bazıları aşağıda sunulmaya çalışılmıştır:

1-Eğitilen kelime örneklerinin konuşma ve iletim ortamına duyarlılığı fazladır. Bu sebeple test örneklerinin kayıtları sırasında oluşan gürültü kelime tanıma oranının düşmesine neden olmaktadır.

2-Sesin kalitesi gürültülü ortamda dönüşüm (kayıt) aşamalarında bozulabileceğinden, gürültüsüz ortamda eğitilen sistemde kelimeler daha yüksek performansla tanınmaktadır.

3-Sistemin performansı eğitilecek kelime sayısına da bağlı olduğundan, sisteme kaydedilen eğitilen kelime örneklerinin sayısı azaldıkça, performansın arttığı gözlemlenmiştir.

4- Kullanıcı doğru ses verdiği takdirde ses tanıma devresinde bay ve bayan konuşmacı ses tanıma oranında bir farklılık oluşmamaktadır.

5-Telaffuz farklılıkları, konuşma hızı, ton, vurgu, stres ve heyecan vb. sebepler ile her seferinde aynı ses çıkarılmadığı için sesler arasında farklı tanıma oranları meydana geldiği görülmüştür.

6-Kelimelerin sürelerinin kayıttaki gibi olması tanıma performansını olumlu etkilemektedir.

7-Sistemin hızı kelimelerin söyleniş süresiyle ters orantılıdır.

8-Hafızaya eğitilen kelimelere rasgele erişim özelliği nedeni ile sistemin performansı azalmaktadır. Bu durum, hafızaya yeni kelimeleri tekrar gerekli olan sıra ile yükleyerek belirli ölçülerde düzeltilebilir.

9- Ağızdan önceki bölgelerde de meydana gelen ses değişimlerinden dolayı sistemde her kişide farklı tanıma oranları meydana gelmektedir. Bunlar kontrol edilmeye çalışılarak performans artırılabilir.

10- Eğitilecek kelimelerdeki ses birimlerinin iyi bir şekilde seçilmesinin tanıma performansını olumlu etkilediği düşünülmektedir.

Sonuç olarak, bu çalışmada kelime tanıma için eşsesli olmayan kelimeler için gürültüsüz ortamda %98, 60-65dB gürültülü ortamda %95; eşsesli olan kelimeler için ise gürültüsüz ortamda 67, 60-65dB gürültülü ortamda %62 doğruluk elde edilmiştir. Bu duruma göre sistem gürültülü ve gürültüsüz ortamda ortalama olarak %80,5 tanıma oranı sağlamaktadır. Bu çalışmanın temel konularından birisi olan ünlü ve ünsüz harf tanıma ile PC'ye yazı yazma işlemleri için doğruluk ortalama %95.2 bulunmuştur. Şimdiye kadar ifade edilmeye çalışılan tanıma performansına etki eden sınırlamalara dikkat edilmesinin olumlu etkileri olduğu gözlemlenmiştir.

Yapılan testlerin sonuçları, hafıza ve performansın artırılması ile doğrudan konuşmaların, bu çalışmadaki gibi gömülü sistemler ile doğrudan tanınabileceğini ve çeşitli cihazların kontrolünde kullanılabileceğini göstermektedir. İleride, sistemde ilave olarak Yapay Zeka veya Fuzzy Logic algoritmaları kullanılarak daha doğru tanıma sağlanabilmesi için çalışmalar yapılabilir veya daha önceki konuşma verileri kullanılarak kestirimler yapılmaya ve böylece daha akıllı ve öğrenen bir sistem elde edilmeye çalışılabilir [36-38]. Ayrıca sistem müzik eğitim setlerine adapte edilerek müzik eğitiminin kalitesi artırılmaya çalışılabilir veya müzik

enstrümanlarının akord işlemlerinde kullanılabilir. Ses izi biyometrik bir özellik olduğundan, sistem güvenlik amaçlı uygulamalarda kullanılabilir [26,14,15].

KAYNAKLAR (REFERENCES)

- [1]. <http://www.speech.cs.cmu.edu/comp.speech/Section6/O6.1.html>, 05.10.2017.
- [2]. https://www-03.ibm.com/ibm/history/exhibits/specialprod1/specialprod1_7.html, 16.08.2016.
- [3]. Alsaka Y.A., Doll S.A., Davis S., "Portable speech recognition for the speech and hearing impaired", *Southeastcon '97. Engineering new New Century., Proceedings. IEEE*, Blacksburg, VA, USA, 151-153, (1997).
- [4]. Thiang, "Implementation of speech recognition on MCS51 microcontroller for controlling wheelchair", *International Conference on Intelligent and Advanced Systems*, Kuala Lumpur, Malaysia, 1193-1198, (2007).
- [5]. Mosbah B.B., "Speech recognition for disabilities people", *Information and Communication Technologies, 2006. ICTTA '06. 2nd*, Damascus, Syria, 864-869, (2006).
- [6]. <http://download.springer.com/static/pdf/>, 19.08.2016, "Chapter 9, Speech Technology In Military Application", Springer, fulltext, pp:289-330.
- [7]. McCallum M.C., Cambell J.L., Richman J.B., Brown J.L., "Speech recognition and in-vehicle telematics devices: potential reductions in driver distraction", *International Journal of Speech Technology*, 7(1): 25-33, (2004).
- [8]. Lin P.C., Wang J.F., Lin S.C., Mo M.H., "An embedded system design for ubiquitous speech interactive applications based on a cost effective SPCE061a micro controller", *UIC 2006, Ubiquitous Intelligence and Computing*, Springer-Verlag, Berlin Heidelberg, LNCS 4159: 83-92, (2006).
- [9]. Rabenja R., Amine S., Davrdson J., "Speech recognition for game control", *IEEE, Poster Session 1 : System Level Design, 0-7803-8322-2/04/\$20.00 02004 IEEE*, 97-99, (2004).
- [10]. Haleem M. S., "Voice controlled automation system", *Proceedings of the 12th IEEE International Multitopic Conference*, 508-512, December, 23-24, (2008).
- [11]. Ferrando F., Nouveau G., Philip B., Pradeilles P., Soulenq V., Van-Staen G., Courmontagne P., "A voice recognition system for a submarine piloting", *OCEANS 2009 - EUROPE*, Bremen, Germany, (2009).
- [12]. Türk O., Arslan L.M., "Speech recognition methods for speech therapy", *Signal Processing and Communications Applications Conference, 2004. Proceedings of the IEEE 12th*, 410-413, (2004).
- [13]. Fezari M., Salah M.B., "A voice command system for autonomous robots guidance", *IEEE, AMC'06-Istanbul, Turkey*, pp:261-265, (2006).
- [14]. <http://www.teknoloji.com/2015/06/ses-izi-ve-nefes-izi-parmak-izinin.html>, 06.10.2016.
- [15]. Burunkaya M., Yorulmaz F., "Mikrodenetleyici tabanlı elektronik nota eğitim seti tasarımı ve yapımı", *IATS '09, 5. International Advanced Technologies Symposium*, Karabük, Türkiye, sayfa: 556 - 561, May 13-15, (2009).
- [16]. Ruiz C.B., Tapias F.E.G., del-Brio B.M., Nuez A.B., Marques N.J.M., "Microcontroller implementation of a voice command recognition system for human-machine

- interface in embedded systems”, *IEEE*, 0-7803-9402-X/05/\$20.00©, pp:587-591, (2005).
- [17]. <http://www.imagesco.com/speech/HM2007.pdf>, 19.08.2017.
- [18]. <https://www.lumenvox.com/resources/tips/types-of-speech-recognition.aspx>, 05.10.2017.
- [19]. https://tr.wikipedia.org/wiki/Braille_alfabesi, 16.08.2016.
- [20]. Chan F.Y., Khalid H.M., “Is talking to an automated teller machine natural and fun?”, *Ergonomics*, 46(13-14): 1386–1407, (2003).
- [21]. Yuanyuan S., Jia L., Runsheng L., “Single-chip speech recognition system based on 8051 microcontroller core”, *IEEE Transactions on Consumer Electronics*, 47(1): 149-153, (2001).
- [22]. https://en.wikipedia.org/wiki/Data_transmission, 19.08.2016.
- [23]. http://www.camiresearch.com/Data_Com_Basics/RS232_standard.html, 19.08.2016.
- [24]. <http://www.computer-engineering.org/ps2protocol/>, 19.08.2016.
- [25]. <http://www.elektrotekno.com/about1660.html>, 19.08.2016.
- [26]. Yıldır M., Burunkaya M., “Ses sentezleyici kullanılan ses tanıma düzeni tasarımı ve yapımı”, *UMES’07*, Kocaeli, 1: 395-398, (2007).
- [27]. <http://www.imagesco.com/speech/SR-07.pdf>, 19.08.2017.
- [28]. <http://320volt.com/rs232-seri-portun-pic-mikro-deneticiler-ile-kullanimi/>, 19.08.2016.
- [29]. Burunkaya M., Yıldız S., Karataş İ., “Öğrenci proje çalışmalarında kullanılabilir düşük maliyetli ve genel amaçlı bir veri toplama sisteminin (VTS) gerçekleştirilmesi”, *Bilişim Teknolojileri Dergisi*, 4(3): 21-28, (2011).
- [30]. http://www.lima.com.tr/Support/Lima/PS2Keyboard_TR.PDF 15.08.2016, 12:09.
- [31]. <http://320volt.com/mikrodenetleyici-uygulamaları-bilgiler-pic-c-pic16f877/>, 19.08.2016.
- [32]. <http://www.ume.tubitak.gov.tr/tr/laboratuvarlarımız/laboratuvar-ortam-kosullari-1> 15.08.2016, 17:11.
- [33] NL-21/NL-31 Sound Level Meter, Rion Co., Ltd., Tokyo, Japan.
<http://www.noisemeasurement.com.au/downloads/NL-21.pdf>, 14.08.2017.
- [34] Ertaş F., Eskidere Ö., “Yazılım tabanlı sözcük sentezleyici”, *DEÜ Müh. Fak. Fen ve Mühendislik Dergisi*, 3(1): 1-27, (2001).
- [35]. Efendioğlu S., İşcan A., “Türkçe ses bilgisi öğretiminde ses olaylarının sınıflandırılması”, *A.Ü. Türkiyat Araştırmaları Enstitüsü Dergisi (TAED)*, Erzurum, 43: 121-143, (2010).
- [36]. Goktas, H.H., Yucel, M., "A fuzzy logic based device for the determination of temperature dependence of EDFAs", *Microwave and Optical Technology Letters*, 50(9): 2331-2334, (2008).
- [37]. Celebi, F. V., Yucel, M., Goktas, H.H., “Fuzzy logic based device to implement a single CAD model for a laser diode based on characteristic quantities”, *Optik*, 123(6): 471-474, (2012).
- [38]. Celebi, F. V., Yucel, M., Goktas, H.H., Danisman, K., “Intelligent modelling of alpha (α) parameter; comparison of ANN and ANFIS cases”, *Optoelectronics and Advanced Materials – Rapid Communications*, 7(5-6): 470–474, (2013).