Archive of SID

تشخیص سطح اختلالات صوتی در گفتار کودکان دارای پروتز حلزون شنوایی و سمعک توسط سیستم‌های هوشمند مصنوعی

زینب محمودی
کارشناس ارشد مهندسی پزشکی، گروه مهندسی پزشکی دانشگاه آزاد اسلامی واحد مشهد، عضو باشگاه پژوهشگران جوان

سعید راحتی
دکتری برق، استادیار گروه برق دانشگاه آزاد اسلامی واحد مشهد

محمد مهدی قاسمی
متخصص گوش، گلو و بینی، استاد و عضو مرکز تحقیقات گوش، گلو، بینی و جراحی سر و گردن دانشگاه علوم پزشکی مشهد

محسن رجعتی
متخصص گوش، گلو و بینی، استادیار و عضو مرکز تحقیقات گوش، گلو، بینی و جراحی سر و گردن دانشگاه علوم پزشکی مشهد

وحید اسدپور
دکتری برق،استادیار گروه برق دانشگاه غیرانتفاعی سجاد مشهد

حمید طیرانی
کارشناس ارشد گفتار درمانی، مرکز کاشت حلزون، دانشگاه علوم پزشکی مشهد

نگارنده پاسخگو: زینب محمودی
آدرس: مشهد، قاسم آباد، خیابان استاد یوسفی، دانشکده مهندسی دانشگاه آزاد اسلامی واحد مشهد، گروه مهندسی پزشکی
تلفن: ۰۹۱۵۵۲۱۵۶۱۷
نمابر: ۶۶۲۹۴۶۷-۰۵۱۱
پست الکترونیک:
zmbe85@yahoo.com

تاریخ وصول: ۱۳۸۸/۸/۱۸
تاریخ تایید: ۱۳۸۸/۱۰/۱۰

مقدمه
بهبود تولید گفتار در انسان در ارتباط تنگاتنگ با بهبود شنوایی قرار دارد. در افراد ناشنوای بدون سابقه‌ی شنوایی یا کسانی که در زیر سن ۵ سال ناشنوا شده‌اند پس از عمل جراحی کاشت حلزون و یا استفاده از سمعک، شنوایی از دست رفته ترمیم می‌گردد. تولید گفتار در این کودکان یک عمل مهارت پذیر بوده و با به دست آوردن باز خورد شنوایی به تدریج ویژگی‌های آوایی گفتار آن‌ها به افراد طبیعی نزدیک می‌شود.

هدف
هدف از این تحقیق بررسی روش‌هایی است که بر اساس آن بتوان صوت کودک را بر اساس ویژگی‌های پردازشی کمّی و سطح بندی کرد.

مواد و روش‌ها
۴ سطح برای کیفیت صوت این کودکان در نظر گرفته شد که شامل: سطح۱: کودکانی که آواسازی هدفمند دارند و از کلمات به صورت تقلیدی و خود به خودی استفاده می‌کنند. سطح ۲: کودکانی که آواسازی هدفمند دارند، از کلمات استفاده می‌کنند، از جملات و عبارات کوتاه به صورت تقلیدی استفاده می‌کنند. سطح ۳: کودکانی که آواسازی هدفمند دارند، از کلمات استفاده می‌کنند. از جملات و عبارات مختلف به صورت خود انگیخته استفاده می‌کنند و سطح۴: کودکان نرمال بدون سابقه‌ی ناشنوایی که به عنوان مرجع به کار گرفته شده است. در این مطالعه ۳۰ کودک شرکت کردند . ویژگی‌های صوتی شامل فرکانس پایه، فرکانس فرمنت‌های اول تا سوم و نسبت آن‌ها، شدت نسبی صوت، میزان خیشومی شدگی (نازالیتی) و ویژگی‌های غیرخطی شامل بعد فرکتال و آنتروپی تقریبی از ۵ کلمه "ماشین، "مار"، "موش"، "گاو" و "موز" استخراج شد و سپس توسط سیستم شناخت بر اساس اختلاف ویژگی‌های صوتی، سطوح معرفی شده تفکیک شدند.

نتایج
این تحقیق نشان داد که سطوح تعریف شده، تا میزان ۹۳/۱ درصد قابل جداسازی هستند. در میان ویژگی‌های معرفی شده دو ویژگی نسبت فرمنت اول به دوم و نسبت فرمنت سوم به دوم می‌توانند مستقلا روند بهبود کیفیت صوت کودک را با گذشت زمان بعد از کاشت حلزون شنوایی و یا استفاده از سمعک بیان کنند.

نتیجه‌گیری
خروجی این تحقیق می‌تواند به عنوان یک سامانه‌ی مستقل از گوینده عمل کرده و به عنوان یک تست کمک تشخیصی به گفتار درمان در تعیین سطح و وضعیت گفتاری کودک یاری رساند.

واژه‌های کلیدی
Voice abnormalities, Children with cochlear implant, Hearing aid, Frequency features of voice, Nonlinear features of voice, Artificial decision making systems.

www.SID.ir

Archive of SID

مقدمه

بهبود تولید گفتار در انسان در ارتباط تنگاتنگ با بهبود شنوایی قرار دارد. به عبارت دیگر تا زمانی که شخص قادر به شنیدن آن چه خود می‌گوید نباشد، نمی‌تواند به طور صحیح صحبت نماید. فرایند تولید گفتار در افراد ناشنوای بدون سابقه‌ی شنوایی یا کسانی که در زیر سن ۵ سال ناشنوا شده‌اند بسیار ناقص است، زیرا این افراد قادر به شنیدن صدای خود (فیدبک شنوایی) و دیگران نیستند و فاقد هر گونه پیش‌زمینه‌ی ذهنی در مورد اصوات هستند (۲،۱). پس از عمل جراحی کاشت حلزون و یا استفاده از سمعک، شنوایی از دست رفته‌ی این بیماران ترمیم می‌گردد. لذا این انتظار وجود دارد که تولید گفتار آن‌ها با گذشت زمان بهبود یافته و به سمت طبیعی شدن پیش رود. این افراد بعد از بازگشت شنوایی از صدای تولید شده خود برای تنظیم ویژگی‌های گفتاری از قبیل بلندی صدا، لحن گفتار، کشش زمانی آواهای گفتاری و غیره استفاده می‌کنند. برخی از ویژگی‌های صوتی در افراد دارای اختلالات گفتاری در اثر ناشنوایی به این شرح است (۳،۱): ۱- نازال بودن بیش از حد صدا، ۲- عقب بودن زبان در دهان که گاهی اوقات باعث برخی اختلالات تشدیدی نظیر بیش از حد خیشومی شدن گفتار می‌باشد، ۳- صدای افراد ناشنوا معمولا از نظر فرکانس پایه (f_0)، شدت و کشش صدا اشکال دارد، ۴- تنش عضلات حنجره که نهایتا موجب زیرتر شدن صدا و اختلال در فرکانس گام می‌گردد. مشکلات فوق باعث می‌شود که شنونده‌ی معمولی(تعلیم ندیده) تنها بتواند ۲۰ تا ۲۵ درصد از گفتار یک فرد ناشنوا را درک کند. در ارتباطات کودک با دیگران، محتوای کلام تنها ۷ درصد در رساندن احساسات و مفهوم سخن موثر است، ۵۵ درصد حالات چهره و ۳۸ درصد صوت اوست که تاثیر گذار است (۴). بنا بر این بررسی صوت کودک به تنهایی سهم به سزایی در ارتباطات شفاهی او با افراد دارد. از آن جا که تا به حال هیچ سطح‌بندی کمی در مورد صوت این کودکان انجام نشده و قضاوت‌ها در مورد کیفیت صوت این افراد ذهنی بوده و وابسته به قضاوت و مهارت شنونده متفاوت است (۵)، بنا بر این ایجاد یک سیستم خودکار که بتواند با تعیین وضعیت صوتی گفتار کودک و بیان آن به صورت یک سطح مشخص بر اساس ویژگی‌های آوایی تعریف شده به گفتار درمان در ارزیابی وضعیت مددجو کمک کند بسیار لازم و ضروری به نظر می‌رسد. زیرا عدم مشخص شدن دقیق سطح صوت کودک باعث ارایه‌ی آموزش‌های نامناسب و افت روند پیشرفتش می‌گردد. در این تحقیق روش‌هایی بررسی می‌شوند که بر اساس آن بتوان صوت کودک را بر اساس ویژگی‌های پردازشی کمّی و سطح بندی کرد.

روش کار

برای شروع تحقیق، ۴ سطح کیفی زیر برای گفتار کودکان دارای پروتز حلزون شنوایی و سمعک با نظر متخصص گفتار درمان در نظر گرفته شد:

سطح ۱: کودکانی که آواسازی هدفمند دارند و از کلمات به صورت تقلیدی و خود به خودی استفاده می‌کنند.

سطح ۲: کودکانی که آواسازی هدفمند دارند، از کلمات استفاده می‌کنند، از جملات و عبارات کوتاه به صورت تقلیدی استفاده می‌کنند.

سطح ۳: کودکانی که آواسازی هدفمند دارند، از کلمات استفاده می‌کنند. از جملات و عبارات مختلف به صورت خود انگیخته استفاده می‌کنند. بنا بر این این کودکان در بالاترین سطح گفتاری و صوتی هستند.

سطح ۴: کودکان نرمال بدون سابقه‌ی ناشنوایی(این سطح به عنوان مرجع به کار گرفته شده است).

هدف این تحقیق، جداسازی و کمی کردن سطوح فوق بر اساس ویژگی‌های آوایی این کودکان می‌باشد. زیرا همان‌طور که بعد از کاشت حلزون شنوایی و استفاده از سمعک، مهارت‌های گفتاری این کودکان در تولید کلمات و جملات مختلف توسعه می‌یابد، ویژگی‌های آوایی صوت آن‌ها نیز بهبود قابل ملاحظه‌ای پیدا می‌کند. در نتیجه بر مبنای این ویژگی‌های آوایی سطوح فوق قابل جداسازی می‌باشند. تعریف سطوح فوق با استفاده از معیار SIR صورت گرفته است (۶).

www.SID.ir

Archive of SID

این معیار که در جدول (۱)آمده است دارای ۵ سطح است که تقریبا متناظر با ۴ سطح ناهنجاری تعریف شده در این تحقیق می‌باشد. افراد مورد مطالعه در محدوده‌ی سنی بین ۲ تا ۶ سال از کودکان دارای پروتز حلزون شنوایی و سمعک و کودکان نرمال از یکی از مهدهای کودکان طبیعی انتخاب شدند. جداول (۲و۳و۴)، اطلاعات مربوط به این افراد را نشان می‌دهد.

جدول۱: مقایسه‌ی سطوح گفتاری تعریف شده در تحقیق و معیار SIR

سطح اختلالات گفتاری در تحقیق	سطوح قابل درک بودن گفتار در معیار SIR	امتیاز SIR
سطح ۱	گفتار کودک حتی در سطح کلمات مجزا به سختی قابل درک است.	۱
سطح ۲	جملات و گفتار پیوسته‌ای که کودک به طور خود به خودی بیان می‌کند قابل درک نیست اما در سطح کلمات انفرادی قابل درک است.	۲
سطح۲	جملات و گفتار پیوسته‌ای که کودک به طور خود به خودی بیان می‌کند برای شنندنده‌ای که بر روی گفتار تمرکز کند قابل درک است.	۳
سطح ۳	جملات و گفتار پیوسته‌ای که کودک به طور خود به خودی بیان می‌کند برای شنونده‌ای که آشنایی اندکی با گفتار ناشنوایان دارد قابل درک است و نیازی به تمرکز نیست.	۴
سطح ۴	گفتار کودک چه در سطح جملات و چه در سطح کلمات برای هر شنونده‌ای قابل درک است.	۵

جدول۲: مشخصات بیماران مورد مطالعه

سطوح مورد بررسی	نام کودک و جنسیت	نوع پروتز مورد استفاده	سن تقویمی	سن کاشت حلزون شنوایی و یا استفاده از سمعک	مدت زمان گذشته از عمل و یا استفاده از سمعک
سطح ۱ (۶ نفر)	م.پ (دختر)	کاشت حلزون شنوایی	۳ سال و ۱۰ ماه	۲ سال و ۸ ماهگی	۱ سال و ۲ ماه
	ز.ک (دختر)	کاشت حلزون شنوایی	۴ سال و ۸ ماه	۳ سال و ۷ ماهگی	۱ سال و ۱ ماه
	ع.ا (پسر)	کاشت حلزون شنوایی	۴ سال و ۱۱ ماه	۳ سال و ۶ ماهگی	۱ سال و ۵ ماه
	م.ض(پسر)	سمعک	۲ سال و ۱۰ ماه	۶ ماهگی	۲ سال و ۴ ماه
	م.ص(پسر)	کاشت حلزون شنوایی	۵ سال و ۵ ماه	۴ سال و ۵ ماه	۱ سال
	م.ن (دختر)	سمعک	۴ سال و ۴ ماه	۲ سال و ۶ ماهگی	۱ سال و ۱۰ ماه
	ف.خ (دختر)	سمعک	۵ سال و نیم	۴ سال و ۶ ماهگی	۱ سال
سطح ۲ (۶ نفر)	م.ح(پسر)	کاشت حلزون شنوایی	۴ سال	۲ سال و ۱۱ ماهگی	۱ سال و ۱ ماه
	م.خ(دختر)	کاشت حلزون شنوایی	۵ سال و ۱۰ ماه	۵ سال و ۳ ماهگی	۷ ماه
	ن.ح (پسر)	سمعک	۴ سال و ۳ ماه	۱ سالگی	۳ سال و ۳ ماه
	م.ف (پسر)	سمعک	۴ سال و نیم	۲ سالگی	۲ سال و نیم
	م.م (پسر)	کاشت حلزون شنوایی	۵ سال	۳ سال و ۴ ماهگی	۱ سال و ۸ ماه
	ح.پ (پسر)	سمعک	۳ سال و نیم	۱ سال و نیم	۲ سال
	د.ش(پسر)	سمعک	۴ سال	۲ سال	۲ سال
سطح ۳ (۶ نفر)	ن.خ (دختر)	کاشت حلزون شنوایی	۴ سال و نیم	۲ سال و ۹ ماهگی	۱ سال و ۹ ماه
	ر.ا (دختر)	سمعک	۵ سال و ۹ ماه	۱ سال و ۶ ماهگی	۴ سال و ۳ ماه
	س.غ (پسر)	کاشت حلزون شنوایی	۵ سال و ۲ ماه	۳ سال و ۴ ماه	۱ سال و ۱۰ ماه
	ا.پ (دختر)	کاشت حلزون شنوایی	۶ سال	۳ سال و ۲ ماه	۲ سال و ۱۰ ماه

www.SID.ir

Archive of SID

جدول۳: کودکان سالم مورد مطالعه

سطوح مورد بررسی	نام کودک	جنسیت	سن تقویمی
گروه نرمال (سطح ۴)	ا.ز	پسر	۵ سال
	د.ف	دختر	۵ سال و ۶ ماه
	م.ر	دختر	۶ سال
	ک.ب	دختر	۳ سال
	ا.ت	پسر	۳ سال
	ح.ع	پسر	۴ سال
	ع.خ	پسر	۵ سال
	م.ن	پسر	۵ سال
	پ.م	پسر	۴ سال
	س.ب	پسر	۴ سال
	س.ا	پسر	۵ سال
	ز.س	دختر	۶ سال

جدول٤: میانگین مشخصات سنی بیماران در تمام سطوح دارای اختلال

سطح مورد بررسی	میانگین سن تقویمی	میانگین سن کاشت حلزون و یا استفاده از سمعک	میانگین مدت زمان گذشته از عمل و یا استفاده از سمعک
سطح ۱	۵۲ ماه (۴ سال و ۴ ماه)	۴۵/۳ ماه (۳ سال و ۹ ماه)	۱۸ ماه (۱ سال و ۶ ماه)
سطح ۲	۵۸ ماه (۴ سال و ۱۰ ماه)	۳۸ ماه (۳ سال و ۲ ماه)	۲۰/۱۶ ماه (۱ سال و ۸ ماه)
سطح ۳	۵۹ ماه (۵ سال)	۲۸/۵ ماه (۲ سال و ۵ ماه)	۲۹/۳ ماه (۳ سال و ۴ ماه)

کلمات فارسی: "ماشین(/mashin/mɑ:ʃin)"، "مار (/mar /mɑ:r) "، "موش(/moosh / mu:ʃ)"، "گاو (/gav /gɑ:v)" ، "موز (/moz/moʊz) " برای بررسی واج‌های مشخص شده (دارای خط) در گفتار این کودکان انتخاب شدند. معیار انتخاب این کلمات سادگی بیان آن برای همه‌ی سطوح گفتاری، قابلیت نمایش آن‌ها توسط شکلشان و هم‌چنین دارا بودن واج‌های هدف می‌باشد. سه واکه‌ی اصلی زبان فارسی: ای – آ- او در سه کلمه اول (ماشین، مار، موش) و ۳ هم‌خوان واکدار زبان فارسی: گ-و – ز در دو کلمه آخر (گاو، موز)، در نظر گرفته شدند. برای جلوگیری از هر نوع گفتار تقلیدی، برای هر کدام از کلمات فوق شکلی انتخاب شد و در نرم افزار PowerPoint به صورت slide show برای کودک به نمایش در آمد. از کودک خواسته شد تا بعد از دیدن هر شکل نام آن را بگوید و هم‌زمان گفتار او ثبت گردید. به این ترتیب از ۱۸ کودک دارای پروتز و سمعک (۶ کودک در هر سطح) و ۱۲ کودک نرمال بدون سابقه‌ی ناشنوایی، در مجموع ۳۰ نفر در محدوده‌ی سنی ۳ تا ۶ سال، ثبت گرفته شد. هر شکل ۷ بار برای کودک تکرار شد تا بیان کند. فرکانس نمونه برداری ۴۴۱۰۰ هرتز بود. ۶۰ درصد کلمات برای آموزش و ۴۰ درصد برای آزمایش طبقه‌بندی کننده‌ها به کار رفت.

۱- استخراج ویژگی های صوتی، ۱-۱- شدت نسبی صوت[1] مشخص شده است که افراد دارای اختلالات شنیداری تمایل دارند با صدای بلندتر صحبت کنند (۳). در تحقیق حاضر، نسبت انرژی هر هجا به ماکزیمم انرژی موجود در آن (شدت نسبی صوت) به کار گرفته می‌شود. ۱-۲- فرکانس پایه[2] f_0 فرکانسی است که تارهای صوتی با آن نوسان می کنند. گزارش‌های متعددی مبنی بر بالاتر بودن فرکانس گام صوت کودکان دارای پروتز حلزون شنوایی نسبت به گروه کنترل وجود دارد. ۱-۳- فرمانت[3]

[1]Rational Intensity (RI)
[2]Fundamental Frequency
[3]Formant

www.SID.ir

Archive of SID

فرمنت‌ها قله‌های طیف فرکانسی صوت هستند. هر کدام از فرمنت‌های اول تا سوم به تنهایی قادر به توضیح الگویی مشخص در بهبود اختلال گفتاری این کودکان نیستند، در حالی که با بررسی ویژگی f_1/f_2 می‌توان الگوی بهبود اختلاف نمونه‌های بیمار با گروه کنترل را با گذشت زمان بعد از عمل با کاشت پروتز مشاهده کرد (۷).

۱-۴- میزان خیشومی شدن صوت

یک مشکل مشترک در تولید گفتار توسط ناتوانان گفتاری دارای پروتز و سمعک، نازال شدن بی‌اختیار گفتار آنان است. اختلاف دامنه‌ی فرمنت اول با دامنه‌ی قله‌ی طیفی اضافه شده در فرکانس ۱ کیلو هرتز (AF_1-A_{1kh})، معیار معتبری برای تعیین میزان نازال بودن واکه‌ها در ناتوانان شنیداری است (۱). باید توجه داشت که تغییرات نازالیته با ویژگی AF_1-A_{1kh} نسبت عکس دارد، به این معنی که افزایش AF_1-A_{1kh} نشان دهنده‌ی کاهش نازالیته است.

۱-۵- تحلیل‌های غیرخطی

سیگنال‌های آشوبگونه[1] به آن دسته از سیگنال‌هایی خطاب می‌گردد که توسط یک سیستم غیرخطی با رفتار غیرپریودیک و پیچیده تولید می‌شود. یکی از ملزومات این سیستم‌ها غیرخطی بودن آن است که تولید ویژگی‌های غیرخطی می‌کند. از جمله این ویژگی‌های غیرخطی بعد فرکتال و آنتروپی تقریبی هستند که در ادامه معرفی می‌شوند.

۱-۶- بعد فرکتال

محاسبه‌ی بعد فرکتال سیگنال، یکی از معمول‌ترین راهکارهایی است که از آن در به دست آوردن تخمینی از میزان و درجه رفتار آشوب گونه آن استفاده می‌شود. هرچه بعد فرکتال و آنتروپی سیگنال صوت بیشتر باشد نشان دهنده‌ی آن است که سیگنال و منبع تولید کننده‌ی آن (سیستم گویایی کودک) دارای ساختار کامل‌تر و پیچیده‌تر می‌باشد. برای مشاهده‌ی نحوه‌ی محاسبه‌ی این ویژگی رجوع شود به (۸).

۱-۷- آنتروپی تقریبی[2]

آنتروپی تقریبی نیز مانند بعد فرکتال به بررسی میزان پیچیدگی و بی‌نظمی فضای فاز سیگنال می‌پردازد. برای مشاهده‌ی نحوه‌ی محاسبه‌ی آنتروپی تقریبی و مسایل مربوط به آن مراجعه شود به (۱۰،۹).

۲- سیستم‌های شناخت مصنوعی و تصمیم‌گیری در مورد سطح صوتی کودک.

سیستم شناخت[3] به سیستمی می‌گویند که با گرفتن ورودی‌های مناسب، تصمیمات مشخصی را بر اساس ورودی‌ها اتخاذ می‌کند (۱۲،۱۱). این سیستم‌ها از مظاهر هوش مصنوعی می‌باشند. هوش مصنوعی، به معنای سیستم‌های ماشینی هوشمند است و آن را این گونه تعریف می‌کنند: "سیستمی که می‌تواند با توجه به محیط اطرافش تصمیماتی را بگیرد که شانس موفقیتش را افزایش دهد". توضیحات بیشتر در مورد سیتم‌های شناخت مراجع موجود است (۱۳،۱۲). در این تحقیق برای اولین‌بار، سیستم‌های تصمیم گیرنده‌ی مصنوعی برای تصمیم‌گیری در مورد میزان اختلال صوتی کودکانی که از سمعک و یا پروتز حلزون شنوایی استفاده می‌کنند به کار گرفته شده است. هدف نهایی طراحی سیستمی است که با گرفتن ورودی که همان مشخصات صوتی کودک می‌باشد، قادر باشد تصمیمی در مورد سطح اختلال صوت او بگیرد و به این ترتیب در کنار ابزار و روش‌های تشخیصی گفتار درمانی به متخصص گفتار درمان در آموزش این افراد کمک کند.

ورودی‌های این سیستم همان پارامترهای ۲-۱-۱ تا ۲-۱-۷ هستند که بعد از این که صوت کودک ضبط شد توسط نرم افزار Praat و Matlab از سیگنال صوت استخراج می‌شوند. سیستم شناخت نیز توسط نرم افزار Matlab پیاده سازی شد.

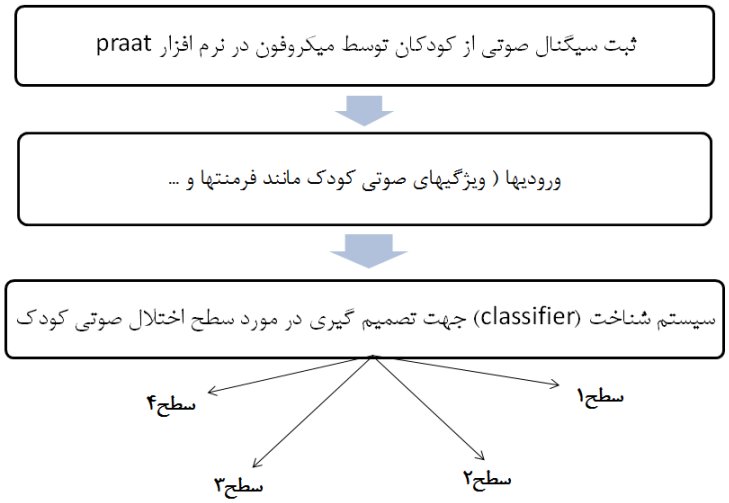
تصویر(۱) بلوک دیاگرام کلی سیستم را نشان می‌دهد. همان طور که بیان شد، در ابتدا کودک کلمات مورد نظر را بیان می‌کند و هم‌زمان صوت او ضبط می‌شود. سپس پارامترهای صوتی از صوت استخراج می‌شوند. این پارامترها به سیستم شناخت رفته و بر اساس آن‌ها تصمیم مقتضی در مورد این که

[1] Chaotic
[2] Approximate Entropy
[3] Classifier

www.SID.ir

Archive of SID

کودک در کدام سطح صوتی قرار دارد گرفته می‌شود. سیستم شناختی که در این تحقیق به کار رفت، مدل مخفی مارکوف (HMM) است که توضیحات و الگوریتم‌های مربوط به آن در مرجع آورده شده است (۱۴،۱۳). پارامترهای (ویژگی‌ها) به کار رفته در این تحقیق در جدول (۵)آورده شده‌اند. البته به دالیل ناپایدار شدن سیستم شناخت از پارامتر f_3/f_2 به عنوان ورودی استفاده نشد و تغییرات آن در کودکان سطوح مختلف به طور جداگانه بررسی شد.



تصویر ۱: بلوک دیاگرام کلی سیستم تعیین سطح صوتی

جدول۵: ویژگی‌های ورودی به سیستم تصمیم گیرنده

ردیف	ویژگی	توضیح
۱	f_0	فرکانس گام سیگنال گفتاری
۲	RI(Rational Intensity)	شدت نسبی صوت
۳	f_1	فرکانس فرمنت اول
۴	f_2	فرکانس فرمنت دوم
۵	f_3	فرکانس فرمنت سوم
۶	f_1/f_2	نسبت فرکانس فرمنت اول به فرکانس فرمنت دوم
۷	f_3/f_2	نسبت فرکانس فرمنت سوم به فرمنت دوم
۸	f_1/f_0	نسبت فرکانس فرمنت اول به فرکانس گام
۹	f_2/f_0	نسبت فرکانس فرمنت دوم به فرکانس گام
۱۰	f_3/f_0	نسبت فرکانس فرمنت سوم به فرکانس گام
۱۱	Nasality (Af1-A1k)	میزان خیشومی شدگی
۱۲	Entropy	آنتروپی تقریبی
۱۳	Fractal dimension	بعد فرکتال

در واقع هر کلمه به عنوان یک طبقه بند تعیین کننده‌ی وضعیت گفتاری فرد عمل می‌کند و سپس تصمیم نهایی طی یک ساختار سلسله مراتبی توسط ترکیب طبقه‌بندها اتخاذ می‌گردد. یعنی ابتدا هر کلمه در مورد ورودی‌هایی که به آن

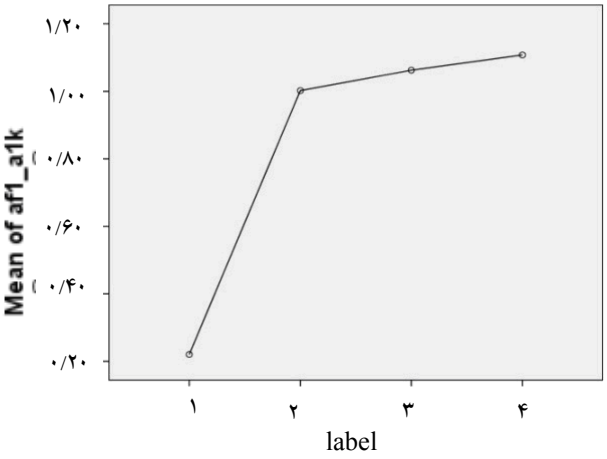
www.SID.ir

Archive of SID

وارد شده تصمیم می‌گیرد و سپس در مرحله‌ی بالاتر این تصمیم‌ها با هم ترکیب می‌شوند (۱۷-۱۵). لزوم این ترکیب از آن جا ناشی می‌شود که هیچ کلمه‌ای به تنهایی نمی‌تواند همه‌ی سطوح را به خوبی تشخیص دهد، بنا بر این ترکیبی از کلمات باید به کارگرفته شود تا بتوان یک تصمیم قابل اطمینان در مورد سطح صوتی کودک گرفت. برای هر سطح در هر کلمه یک مدل مخفی مارکوف در نظر گرفته می‌شود. به این ترتیب در مجموع ۵ کلمه ۲۰ HMM (در هر کلمه ۴ HMM) وجود دارد که احتمال خروجی هر کدام برای تصمیم‌گیری نهایی در مورد صوت کودک به کار می‌رود. هر کلمه با توجه به احتمال‌های خروجی مدل‌های مارکوفش یک سطح مشخص صوتی تعیین می‌کند. این تصمیم گیری‌های محلی باید با هم ترکیب شده و یک تصمیم نهایی در مورد صوت کودک گرفته شود.

نتایج

روند کلی غالب در همه‌ی کلمات در ویژگی بعد فرکتال افزایش و در ویژگی $\frac{f_3}{f_2}$ و $\frac{f_4}{f_2}$ کاهش است و این تغییرات در همه‌ی کلمات شبیه هم می‌باشد (افزایش بعد فرکتال، کاهش نسبت فرمنت اول به دوم و نسبت فرمنت سوم به دوم در همه‌ی کلمات). این مسئله نشان می‌دهد که این ویژگی‌ها می‌توانند مستقل از کلمه‌ی بیان شده تغییرات آوایی گفتار کودک را در پروسه بهبود گفتاری نشان دهند. به عنوان نمونه میانگین اختلاف ویژگی‌ها در پارامتر نازالیته در تصویر (۲) آورده شده است. این شکل میانگین ویژگی اختلاف دامنه‌ی فرمنت اول با قله‌ی طیفی اضافه شده در واکه "ای" از کلمه‌ی ماشین را نشان می‌دهد. محور افقی (label) سطوح گفتاری و محور عمودی، میانگین مقدار این ویژگی در تمام افراد هر سطح می‌باشد. تغییرات ویژگی AF_1-A_{1khz} که اختلاف دامنه‌ی فرمنت اول با دامنه‌ی قله‌ی طیفی در ۱۰۰۰ هرتز است و در این تصاویر فقط در واکه " ای" از کلمه " ماشین" نشان داده شده است یک روند افزایشی می‌باشد. تغییرات نازالیته با ویژگی AF_1-A_{1khz} نسبت عکس دارد، به این معنی که افزایش AF_1-A_{1khz} نشان دهنده‌ی کاهش نازالیته است. با توجه به تصویر مشخص می‌شود که با افزایش سطح صوتی کودک، نازالیته کاهش می‌یابد. افزایش این ویژگی نشان دهنده‌ی کاهش نازالیته در این واکه می‌باشد. جدول (۶) مقادیر آماری مربوط به تصویر (۲) را نشان می‌دهد.



تصویر ۲: اختلاف دامنه‌ی فرمنت اول منهای دامنه‌ی قله‌ی طیفی ایجاد شده در ۱۰۰۰ هرتز استخراج شده از واکه "ای" در کلمه‌ی "ماشین"

جدول٦: مقادیر آماری مربوط به تصویر ۲

label	Mean	Std. Deviation	Variance
۱/۰۰	۰/۲۲۰۶	۰/۶۰۲۰۰	۰/۳۶۲
۲/۰۰	۰/۰۰۲۲	۱/۵۳۶۹۲	۲/۳۶۲
۳/۰۰	۱/۰۶۲۴	۱/۸۹۷۱۰	۳/۵۹۹
۴/۰۰	۱/۱۰۷۷	۲/۵۰۳۴۴	۶/۲۶۷

برای بررسی معنی‌داری تغییرات نازالیته در سطوح مختلف گفتاری از آزمون تی بین هر دو کلاس متوالی استفاده شد. جدول (۷) نتایج آزمون تی را برای ویژگی نازالیته نشان می‌دهد. مشاهده می‌شود که اختلاف بین سطح ۱و۲ با حاشیه‌ی اطمینان ۹۵ درصد معنی‌دار است($P<۰/۰۵$). اما بین ۲ و۳ و نیز بین ۳ و۴ معنی‌دار نیست. این مسئله نشان می‌دهد که بیشترین بهبود در نازالیته در ماه‌های اول بعد از استفاده از پروتز ایجاد می‌شود و از آن به بعد این تغییرات کندتر صورت می‌پذیرد که با آن چه در (۱) به دست آمده مطابقت دارد. در بررسی سایر ویژگی‌ها مشخص شد که اختلاف بین سطوح متوالی در بیشتر آن‌ها با $P<۰/۰۵$ معنی‌دار است.

www.SID.ir

Archive of SID

جدول ۷: نتایج آزمون تی برای ویژگی AF_{1}-A_{1khz} بین گروه‌های متوالی برای واکه 'i' در کلمه‌ی "ماشین"

classes	AF_{1}-A_{1khz}
Class 1 & 2	P<۰/۰۵
Class 2 & 3	P<۰/۰۵
Class 3 & 4	P<۰/۰۵

برای این که بفهمیم آیا سطوح مختلف گفتاری تعیین شده با تعریفی که در بخش ۲-۱ ارایه شد قابل جداسازی هستند یا خیر، از آنالیز LDA استفاده شد. نتایج این تحلیل، تعیین می‌کند که آیا اساسا تفاوت‌های قابل جداسازی بین این سطوح وجود دارد یا خیر. این تحلیل در نرم افزار SPSS به روش Loo (Leave One Out) cross-validation انجام شده و حاصل آن در جدول (۸) نشان داده شده است.

نتایج نشان دادند که سطوح گفتاری مختلف توسط کلمات انتخاب شده قابل جداسازی هستند اما با کمی دقت در این جدول متوجه می‌شویم که هیچ‌کدام از کلمات به تنهایی نمی‌تواند همه‌ی سطوح را به خوبی از هم جدا کند.

به عنوان مثال کلمه‌ی ماشین، تنها سطوح ۲ و۴ را می‌تواند خوب تشخیص دهد و در تفکیک دو سطح دیگر به خوبی عمل نمی‌کند. از این‌جا لزوم ترکیب سیستم‌های شناخت مشخص می‌شود تا نتیجه‌ی نهایی سطح‌بندی بهتر از سطح بندی توسط تک تک کلمات باشد. جدول ۹ خروجی ترکیب سیستم‌های شناخت را نشان می دهد. سطوح گفتاری تعریف شده به طور میانگین با درصد ۹۳/۱ درصد قابل جداسازی هستند.

جدول ۸: درصد تفکیک سطوح توسط کلمات مختلف با استفاده از نرم افزار SPSS

کلمه	سطح ٤	سطح ۳	سطح ۲	سطح ۱
"ماشین"	٪۸۸/۱	٪۵۴/۸	٪۷۳/۸	٪۴۷/۶
"مار"	٪۸۶/۹	٪۴۲/۹	٪۴۷/۶	٪۶۹
"موش"	٪۸۸/۱	٪۵۲/۴	٪۶۴/۳	٪۵۷/۱
"گاو"	٪۷۰/۲	٪۵۲/۴	٪۶۶/۷	٪۷۱/۴
"موز"	٪۷۰	٪۵۲/۴	٪۶۶/۷	٪۷۱/۴

جدول۹ : نتایج سطح‌بندی گفتاری (درصد تشخیص سطوح گفتاری توسط سیستم شناخت)

ویژگی‌ها	میانگین صحت در ٤ سطح	صحت‌تشخیص سطح ٤	صحت تشخیص سطح ۳	صحت‌تشخیص سطح ۲	صحت تشخیص سطح ۱
فرکانس پایه، فرمنت‌های اول تا سوم، نسبت فرمنت اول به دوم، نسبت شدت صوت به مقدار ماکزیمم آن در کلمه (شدت نسبی)، میزان خیشومی شدگی، آنتروپی تقریبی و بعد فرکتال	٪۹۳/۱	٪۹۱	٪۱۰۰	٪۸۱/۲۵	٪۱۰۰

بحث

با مشاهده‌ی جدول (۴) که میانگین مشخصات سنی بیماران را در سطوح دارای اختلال نشان می‌دهد به این نتیجه‌گیری اولیه می‌رسیم که هر چه سطح گفتاری بالاتر (اختلالات کمتر) باشد، مدت زمان گذشته از عمل جراحی کاشت حلزون و یا استفاده از سمعک بیشتر است که در تحقیقات قبلی هم اکثرا به آن اشاره شده است (۳-۱،۱۸،۱۹). اما با توجه به جدول در می‌یابیم که هر چه کودک در سن پایین‌تری شروع به استفاده از پروتز کرده، بهبود سریع‌تر و نتایج بهتری در گفتار او حاصل شده به طوری که در سطح گفتاری بالاتری قرار گرفته است.

این مسئله با نتایج به دست آمده تطابق دارد (۷). هم‌چنین مشاهده می‌شود که هر چه سن تقویمی کودک بالاتر رفته، سطح گفتاری او نیز متکامل‌تر شده. دلیل این امر آن است که کودکان بزرگتر با توجه به کامل‌تر بودن سیستم گویایی‌شان، قدرت بیشتری در کنترل ویژگی‌های آوایی گفتار خود دارند که با نتایج سازگاری دارد (۱۹). آن‌چه در بهبود میزان نازالیتی به دست آمده با نتایج (۱) تطابق دارد. با این تفاوت که در آن تحقیق بهبود این ویژگی را با گذشت زمان مورد مطالعه قرار داده است اما در تحقیق حاضر بهبود این ویژگی با بالا رفتن

www.SID.ir

Archive of SID

سطح گفتاری کودک مورد بررسی قرار گرفته که با توجه به آن چه در پاراگراف اول این بحث گفته شد دو مقوله مشابه می‌باشند. در این تحقیق برای اولین‌بار در مطالعات انجام شده بر روی صوت کودکان دارای پروتز حلزون شنوایی و سمعک، از ویژگی‌های غیرخطی علاوه بر ویژگی‌های فرکانسی معمول، در تحلیل ویژگی‌های آوایی این کودکان استفاده گردید و نتایج به دست آمده قابلیت تحلیل‌های غیرخطی را در بررسی روند بهبود گفتار این افراد ثابت کرد. در این تحقیق نیز مشخص شده است که نقش ویژگی بعد همبستگی سیگنال‌های گفتاری که یک ویژگی غیرخطی فضای فاز می‌باشد (۲۰، ۲۱)، نسبت به ویژگی‌های آکوستیکی کلاسیک در جداسازی نمونه‌های سالم از بیمار از اهمیت بالاتری برخوردار بوده است. نسبت f_3/f_2 ، ویژگی جدیدی است که تا به حال در تحقیقات قبلی که در مورد کودکان دارای پروتز شنوایی انجام شده مورد مطالعه قرار نگرفته است. در این تحقیق نشان دادیم که این ویژگی قادر است به خوبی سطوح مختلف گفتاری را از هم تفکیک کند و روند منظم آن با افزایش سطح گفتاری، شناسه‌ی خوبی برای بررسی بهبود گفتاری کودک می‌باشد. مفید بودن این نسبت از آن جا ناشی می‌شود که تغییرات فرکانس فرمنت‌ها نسبت به هم (formant transition) مشکل رایجی در این بیماران است. با بهبود صوتی این کودکان فرکانس فرمنت‌ها به محدوده‌ی طبیعی خود باز می‌گردد و این نسبت اصلاح می‌شود. همان طور که شکل ۲ نشان می‌دهد و جدول (۷) نیز آن را تایید می‌کند تغییرات ویژگی AF_1-A_{1kh} بیشتر در سطوح پایین است و از سطح ۲ به بعد تغییرات چندانی مشاهده نمی‌شود که با نتایج به دست آمده از (۱) مشابهت دارد. در (۱) برای اولین بار از اختلاف بین دامنه‌ی فرمنت اول با دامنه‌ی قله‌ی طیفی اضافه شده در حدود فرکانس ۱ کیلو هرتز به عنوان معیاری برای اندازه‌گیری میزان نازالیته در افراد دارای پروتز حلزون شنوایی استفاده شد. در تحقیق او این نتیجه به دست آمد که این ویژگی در فاصله‌ی ۶ تا ۱۲ ماه پس از عمل جراحی بیشترین تغییرات را داشته و روند بهبود این ویژگی در فواصل پس از آن با نرخ کندتری صورت گرفته است. ما این روند را در سطوح گفتاری مشاهده می‌کنیم؛ از سطح ۱ به ۲ بیشترین تغییرات را داریم ولی از ۲ به بعد تغییرات با نرخ کندتری ایجاد شده است. در (۷)، کاهش فرکانس پایه و نزدیک شدن فرمنت‌های اول تا سوم به رنج نرمال در واکه /a/ بعد از عمل کاشت پروتز گزارش شده است. در تحقیق حاضر نیز نزدیک شدن فرمنت‌ها به محدوده‌ی طبیعی گزارش شده که لزوما روند افزایشی یا کاهشی ندارد. این مسئله در سایر مطالعات نیز مشاهده می‌شود(۱۲). Hocevar-Boltezar و همکاران در مطالعه‌ی خود، واکه /a/ را در ۳۱ کودک ثبت کردند (۲۲). نتایج ثابت کرد که بعد از عمل جراحی کودک می‌تواند کنترل لحظه به لحظه فرکانس پایه صوت خود راحفظ کند که با نتایج تحقیق حاضر مطابقت دارد. بخشایی و همکاران در مطالعه‌ای کودکان را بر اساس قابل فهم بودن گفتارشان برای شنونده و با استفاده از روش peech Intelligibility Rating (SIR) criteria در یکی از ۵ سطح کیفی قرار دادند، اما در تحقیق مذکور هیچ سطح‌بندی کمّی و عینی صورت نگرفته است (۲۳). در حالی که در تحقیق حاضر برای اولین‌بار سطوح کیفی تعریف شده به صورت کمّی و بر اساس ویژگی‌های پردازش گفتار، دسته‌بندی شدند.

نتیجه‌گیری

در این تحقیق ۳۰ کودک مورد مطالعه قرار گرفتند. با توجه به تنوع کودکان در هر سطح خروجی این تحقیق می‌تواند به عنوان یک سامانه‌ی مستقل از گوینده عمل کرده و برای طیف وسیعی از کودکان در محدوده‌ی سنی مشابه به کار رود. به کارگیری این سامانه می‌تواند یک راهکار موثر در ارزیابی روش آموزش این کودکان باشد. با انجام تست تشخیص سطح، گفتار درمان به این مسئله پی می‌برد که آیا آموزش‌های وارد شده تا به حال موثر بوده‌اند و آیا کودک با توجه به نوع و مدت زمان آموزش‌ها در سطح مناسب خود قرار دارد یا خیر. استراتژی‌های آموزشی مختلفی در مورد کودکانی که از پروتز حلزون شنوایی استفاده می‌کنند، قابل پیاده‌سازی است.

www.SID.ir

Archive of SID

اگرچه مقایسه‌ی مستقیم این استراتژی‌ها بدون در نظر گرفتن یک معیار کمّی مناسب امکان‌پذیر نخواهد بود. کمّی‌سازی اختلالات صوت و گفتار این کودکان و بیان آن به صورت سطح‌بندی شده این امکان را به متخصص می‌دهد تا نتایج روش‌های مختلف آموزشی را با استناد به این سطوح کمی شده مقایسه کند و روش مناسب‌تر را انتخاب نماید. بنا بر این این کمّی‌سازی به گفتار درمان در تعیین و تشخیص اختلالات گفتاری کمک شایانی می‌کند. علاوه بر آن طراحی یک وب سایت مجهز به موتور سامانه طراحی شده در این تحقیق، به منظور فراهم آوردن تسهیلات ویژه‌ای برای بیمارانی که زیر نظر متخصص، تحت گفتار درمانی قرار گرفته‌اند به بیماران این امکان را می‌دهد تا با صلاح دید گفتار درمان، طبق جدول زمان‌بندی به این وب سایت‌ها از هر نقطه‌ای متصل شده و نمونه‌ای از سیگنال صوتی خود را ارسال نمایند. سپس موتور سامانه‌ی تحلیل‌گر تعبیه شده در این سایت‌ها با انجام پردازش‌هایی شامل استخراج ویژگی‌های مناسب، طبقه‌بندی و بهینه‌سازی آن‌ها بتواند به انجام امور تشخیصی در حداقل زمان ممکن پرداخته و روند بهبود و یا عدم بهبود وضعیت گفتاری فرد را به صورت ماهانه به آدرس پست الکترونیک او ارسال کند.

یافته‌ی نوین	کاربرد بالینی
در تحقیق حاضر برای اولین‌بار سطوح کیفی تعریف شده به صورت کمّی و بر اساس ویژگی‌های پردازش گفتار، دسته‌بندی شدند.	خروجی این تحقیق می‌تواند به عنوان یک سامانه‌ی مستقل از گوینده عمل کرده و به عنوان یک تست کمک تشخیصی به گفتار درمان در تعیین سطح و وضعیت گفتاری کودک یاری رساند.

www.SID.ir

Archive of SID

References

1. Amirfattahi R. Introducing and improving speech models for analyzing temporal changes in Persian patients with cochlear implant. Dessertation.Tehran, Amirkabir polytechnic university, College of Electronic Engineering, PhD in Electronic Engineering, 2002.
2. Khalesi M, Amirfattahi R, Sheikh Zadeh H, Khorsandi MT, Motesaddi M, Abdi S. Effect of auditory feedback on speech production after cochlear implantation. Tehran University of medical sciences 2001; 59(5): 1-9.
3. Tayarani H. Effect of Aural-verbal rehabilitation on children's voice under 12 with cochlear implant. Dissertation. Tehran, Tehran University of Rehabilitation science, College of Speech Rehabilitation, MS.c in Speech Rehabilitation. 2002.
4. Perrin E, Berger C, Topouzkhanian A, Truy E, Morgon A. Evaluation of cochlear implanted children's voices. International Journal of Pediatric Otorhinolaryngology 1999; 47(2):181-186.
5. Kunisue K, Fukushima K, Nagayasu R, Kawasaki A, Nishizaki K. Longitudinal formant analysis after cochlear implantation in school-aged children. Int J Pediatr Otorhinolaryngol. 2006; 70(12):2033-42.
6. Bowen C. [Speech Intelligibility from 12 to 48 months]. Retrieved from www.speech-language-therapy.com/intelligibility.htm on (date), 2006.
7. Seifert E, Oswald M, Bruns U, Vischer M, Kompis M, Haeusler R. Changes of voice and articulation in children with cochlear implants. International Journal of Pediatric Otorhinolaryngology 2002; 66 (2): 115-23.
8. Ravari M, Khalilzade MA, Hashemi Golpayegani MR, Rahati S. Processing and analyzing Electroencephalographic signal for predicting reaction time in response to visual stimulation by using fractal dimension and entropy.13th Conference on Biomedical Engineering. Sharif university of Technology, Tehran. Iran. 2006.
9. Pincus SM. Approximate entropy as a measure of system complexity (Statistic/stochastic processes/chaos/dimension)". Mathematics1991; 88(6): 2297–2301.
10. Wellens M, Riihijarvi J, Mahonen P.Evaluation of Spectrum Occupancy using Approximate and Multiscale Entropy Metrics. in Proc of SDR workshop, San Francisco, CA, USA, June 2008.
11. Hilborn RC. Chaos and Nonlinear Dynamics: An Introduction for Scientists and Engineers. Second Edition, London; Oxford University Press:2007.p.233-8.
12. Duda O, Hart RE, Wiley DG. Pattern Classification. Second Edition. LA; John Wiley & Sons Inc: 2000. p. 332-4.
13. Rabiner LR. A tutorial on Hidden Markov Models and Selected Application in Speech Recognition. Proc of the IEEE 1989; 77(2): 257-86.
14. Davarinia F, Rahati S.classifying different mental tasks using Hidden Markov Models and chaotic Neural networks. 2th Conference on data mining 2008. Amirkabir polytechnic university, Tehran, Iran.
15. Huenupán F, Yoma N, Molina C, Garreton C. Confidence based multiple classifier fusion in speaker verification , Pattern Recognition Letters 2008; 29: 957-66.
16. Huenupán F, Yoma N, Molina C, Garreton C. Speaker verification with multiple classifier fusion using Bayes based confidence measure 2008; 29(7):957-66.
17. Ruta D, Gabrys B. An Overview of Classifier Fusion Methods. Journal of Computing and Information Systems, 2000; 7(1): 1-10.
18. Amirfattahi R, Sheikhzade H.Analyzing alterations in vowel articulation after using multichannel cochlear implant in children without hearing antecedent. Amirkabir Journal of science and technology 2001; 46: 209-27.

www.SID.ir

Archive of SID

19. Hocevar-Boltezar I, Vatovec J, Gros A, Zargi M. The influence of cochlear implantation on some voice parameters, International Journal of Pediatric Otorhinolaryngology 2005; 69 (12): 1635-40.
20. Behrouzmand R, Almasganj F, Moradi M.H. Application of Fractal and nonlinear dynamics of speech signals for diagnosing larynx abnormalities in patients with paralyzed vocal cord. 14th Conference on Electronic Engineering, Tehran.Iran. 2005.
21. Behrouzmand R. Diagnosing laryngeal abnormalities and Possibility of classifying them by using speech processing of patients. MS in Biomedical Engineering. Dissertation. Tehran: Amirkabir polytechnic university, College of Biomedical Engineering, 2006.
22. Hocevar-Boltezar I, Boltezar M, Zargi M.The influence of cochlear implantation on vowel articulation. The European Journal of Medicine 2008; 120(7, 8): 228- 33.
23. Bakhshaee M, Ghasemi MM, Shakeri MT, Razmara N, Tayarani H, Tale MR. Speech development in children after cochlear implantation. Eur Arch Otorhinolaryngol. 2007; 264 (11): 1263-6.

www.SID.ir