

ساخت و ارزشیابی آزمون بازشناسی گفتار بزرگسالان فارسی زبان

محمد مصلح - کارشناس ارشد شنایی‌شناسی

ABSTRACT

Title: Development and Evaluation of a Speech Recognition Test for Persian Speaking Adults

Method and Materials: This research is carried out for development and evaluation of 25 phonemically balanced word lists for Persian speaking adults in two separate stages: development and evaluation.

In the first stage, in order to balance the lists phonemically, frequency -of- occurrences of each 29 phonemes (6 vowels and 23 consonants) of the Persian language in adults speech are determined. This section showed some significant differences between some phonemes' frequencies. Then, all Persian monosyllabic words extracted from the Mo'in persian dictionary. The semantically difficult words were refused and the appropriate words choosed according to judgement of 5 adult native speakers of Persian with high school diploma.

12 openset 25 word lists are prepared. The lists were recorded on magnetic tapes in an audio studio by a professional speaker of IRIB.

In the second stage, in order to evaluate the test's validity and reliability, 60 normal hearing adults (30 male, 30 female), were randomly selected and evaluated as test and retest.

Findings:

1- Normal hearing adults obtained 92-100% scores for each list at their MCL through test-retest.

2- No significant difference was observed a/ in test-retest scores in each list ($P>0.05$), b/ between the lists at test or retest scores ($P>0.05$), c/ between sex ($P>0.05$).

Conclusion: This research is reliable and valid, the lists are phonemically balanced and equal in difficulty and valuable for evaluation of Persian speaking adults speech recognition.

چکیده

این پژوهش با هدف ساخت و ارزشیابی فهرست‌های ۲۵ واژه‌ای تک‌هنجایی متوازن واژی برای بزرگسالان فارسی زبان در دو مرحله صورت گرفته است. در مرحله ساخت، به منظور اجرای معیار توازن واژی در فهرست‌ها، بسامد و قوع هر یک از ۲۹ واژه زبان فارسی (شامل ۶ واکه و ۲۳ همخوان) در گفتار بزرگسالان تعیین شد که نشانده‌نده وجود اختلاف معنی دار بین برخی واژه‌ها بود. پس از استخراج کلیه واژه‌های تک‌هنجایی موجود در ۴ جلد اصلی فرهنگ فارسی معین و صرفظیر کردن از واژه‌های دشوار (با توجه به معیار ساخت آزمون‌ها) با قضاوت ۵ فرد بزرگسال دارای دیلم، واژه‌های مناسب موجود در گنجینه واژگان افراد بزرگسال انتخاب شدند. این واژه‌ها با توجه به معیار توازن واژی در فهرست‌های ۲۵ واژه‌ای گنجانده شدند. بدین ترتیب ۱۲ فهرست ۲۵ واژه‌ای شامل ۵ فهرست با الگوی هنجایی CVC، ۴ فهرست با الگوی CVCC و ۳ فهرست به صورت ترکیبی (CVC و CVCC) ساخته شدند. فهرست‌های متوازن شده در استودیوی صدابرداری با صدای یک گوینده ضبط گردید.

مقدمه

گفتار، بخش مهمی از ارتباط را تشکیل می‌دهد و شامل مجموعه سیار پیچیده و ظریف از متغیرهایی است که برای آشکار ساختن افکار انسانی عمل می‌کند. جهت برقراری این ارتباط باید گفتار توسط شنونده به خوبی دریافت و درک شود. ابزارهای مختلفی برای سنجش دریافت درست اصوات در دست است. برای بررسی درک گفتار توسط شنونده یکی از مهمترین آزمون‌ها «آزمون بازشناسی گفتار» می‌باشد. این آزمون بهمنظور مطالعه برقراری درست ارتباط و میزان تمایز واژه‌ها از طرف شنونده، بررسی کارکرد درست سیستم‌های تقویتی و نیاز برای بازتوانی افراد دچار آسیب شناوی به کار می‌رود. برای این که آزمون در اجرای موارد فوق کارایی داشته باشد باید به صورت استاندارد شده، مطابق معیارهای زیر بکار رود:

- ساختار و محتوای آزمون از «واژه‌های تک‌هنجایی معنی‌دار» تشکیل شده باشد.
 - مواد گفتاری آزمون از لحاظ معنایی برای شنونده آشنا باشد.
 - در فهرست واژگان آزمون، توازن واجی وجود داشته باشد.
 - آزمون به صورت ضبط شده بکار رود.
- رعایت این معیارها در اجرای آزمون و تفسیر نتایج حاصل از آن اهمیت زیادی دارد (مارtin، ۱۹۹۷). با توجه به اهمیت تشخیصی و توانبخشی سیار بالای این آزمون، تاکنون آزمون استانداردی به زبان فارسی برای بزرگسالان در دسترس مراکز شناوی شناسی قرار نگرفته بود. لذا ساخت و استاندارد نمودن آن با معیارهای مربوطه از مهم‌ترین نیازهای مراکز شناوی شناسی به شمار می‌آمد.

اهداف

هدف کلی و نهایی این پژوهش ساخت آزمون بازشناسی گفتار برای بزرگسالان فارسی‌زبان و روا و معبر نمودن آن طبق معیارهای مربوطه بود. اهداف دیگر پژوهش که برای رسیدن به هدف نهایی باید به آنها نیز پرداخته می‌شد بدین قرار است:

در مرحله ارزشیابی، برای هنجاریابی آزمون، عنصر (۳۰ و ۳۰ مرد) دارای شناوی طبیعی در محدوده سنی ۲۵-۱۸ سال به صورت تصادفی مورد ارزیابی با فهرست‌های ساخته شده قرار گرفتند. نتایج و یافته‌های این پژوهش عبارتند از:

- همه افراد طبیعی باستفاده از تمام فهرست‌ها، ۹۲-۱۰۰٪ امتیاز را در راحت‌ترین سطح بلندی ارائه آزمون بدست آوردند.
- بین امتیازهای بدست آمده از تمام فهرست‌ها در آزمون نحسین و آزمون دوباره اختلاف معنی‌داری مشاهده نمی‌شود (۰/۰۵ p).
- بین امتیازهای بدست آمده از ۱۲ فهرست بر حسب جنس اختلاف معنی‌داری وجود ندارد (۰/۰۵ p).
- بین امتیازهای بدست آمده از ۱۲ فهرست بر حسب جنس اختلاف معنی‌داری وجود ندارد (۰/۰۵ p).

باتوجه به نحوه ساخت و نتایج حاصل از آن، آزمون بازشناسی گفتار بزرگسالان پایا و از نظر توازن واجی و همسانی فهرست‌ها از نظر دشواری معنایی واژه‌ها معتر است و برای ارزیابی بازشناسی گفتار بزرگسالان فارسی زبان کارایی دارد.

- تعیین فراوانی هریک از واژه‌های زبان فارسی
- استخراج واژه‌های تک‌هنجایی زبان فارسی
- ساختن فهرست‌های ۲۵ واژه‌ای
- متوازن کردن فهرست‌ها از لحاظ واجی
- ضبط کردن فهرست‌ها
- هنجاریابی آزمون در افراد طبیعی از لحاظ شناوی

روش پژوهش

در طرح «ساخت و ارزشیابی آزمون بازشناسی گفتار برای بزرگسالان فارسی‌زبان» به منظور گردآوری داده‌های مورد نیاز برای اجرا، ساخت و ارزشیابی در مراحل مختلف راهکارهای متفاوتی اتخاذ گردید که توضیح داده می‌شود. فراوانی واژه‌ایا بسامد رخداد آواهای یک متفیر کفی و وابسته است و تعیین آن یکی از مهم‌ترین بخش‌های این پژوهش بود. در طراحی آزمون بازشناسی گفتار به این نکته که مواد گفتاری آزمون باید از نظر ترکیب آوایی معادل با گفتار روزمره باشند، توجه شده است (مارtin، ۱۹۹۷). در این پژوهش برای تعیین این فراوانی از برنامه‌های تلویزیونی دارای ییشترین بینده استفاده شده است. در این انتخاب با نظر روابط عمومی صدا و سیما و مجله سروش، از برنامه‌های علمی و تخصصی ویژه صرف‌نظر گردیده، برنامه‌هایی با محتوای اجتماعی، خانوادگی و اخبار گزینش و به مدت ۶ ساعت ضبط و بازنویسی شدند. پس از آوانگاری، لازم بود فراوانی تمام آواهای و نسبت بین آنها محاسبه و شمارش شود. این کار توسط یک نرم‌افزار در پژوهشکده الکترونیک دانشگاه صنعتی شریف انجام شد. کل آواهای شمارش شده ۵۲۶۸۷ آوا بود. همچنین فراوانی وقوع هر آوا پس از آواز دیگر محاسبه شده که پژوهشگر را در زمینه گزینش واژه‌های تک‌هنجایی از نظر بافت آغازی و یا پایانی واژه با یک آوا کمک کرده است.

برپایه بررسی این مجموعه از داده‌ها، درصد وقوع ۲۹ واژ شامل ۲۳ همخوان و ۶۴ واکه (ثمره، ۱۳۶۴: ۳۹) تعیین شده است. در این بررسی واژ /æ/ با ۱۰٪، واژ /e/ با ۱۰٪، واژ /ɪ/ با ۱۰٪، واژ /ʊ/ با ۱۰٪، واژ /ɒ/ با ۱۰٪ و واژ /ə/ با ۱۰٪ مترین درصد وقوع را در گفتار فارسی زبان بزرگسال دارند (جدول ۱).

جدول ۱- درصد وقوع
واژه‌های در گفوار بزرگ‌سالان
تک‌هنجایی مناسب در دسترس باشد تا
توان از آنها در تنظیم فهرست‌های
آزمون استفاده نمود. برای این کار کلیه

شماره	واژه	درصد وقوع
۱	/æ/	۱۰/۸۱
۲	/e/	۱۰/۳۹
۳	/a/	۸/۲۰
۴	/ɪ/	۶/۳۹
۵	/ʊ/	۶/۲۳
۶	/m/	۵/۶۳
۷	/n/	۵/۶۲
۸	/d/	۴/۸۲
۹	/o/	۴/۲۵
۱۰	/b/	۳/۸۹
۱۱	/v/	۳/۸۹
۱۲	/ʔ/	۳/۳۰
۱۳	/y/	۲/۷۵
۱۴	/h/	۲/۷۳
۱۵	/s/	۲/۷۲
۱۶	/ʃ/	۲/۳۶
۱۷	/k/	۲/۱۵
۱۸	/u/	۲/۰۳
۱۹	/j/	۱/۸۸
۲۰	/z/	۱/۸۶
۲۱	/χ/	۱/۳۹
۲۲	/v/	۱/۳۱
۲۳	/f/	۱/۰۸
۲۴	/g/	۱/۰۵
۲۵	/q/	۱/۰۰
۲۶	/ʃ/	۰/۸۲
۲۷	/c/	۰/۸۲
۲۸	/p/	۰/۶۰
۲۹	/Z/	۰/۰۴
۱۰۰	جمع	

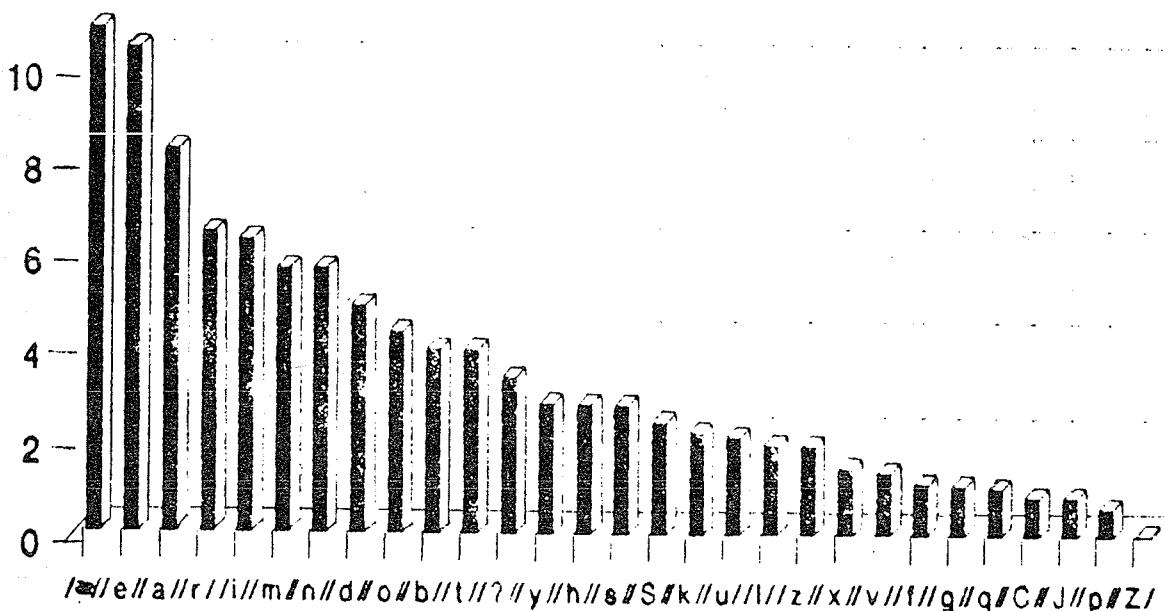
در گام بعدی لازم بود واژه‌های استخراج شده و درجه‌بندی آنها از لحاظ دشواری معنایی^۳ نوع فهرست ایجاد گردید. از مجموع ۱۲ فهرست ۲۵ واژه‌ای ساخته شده، ۵ فهرست شامل واژه‌های CVCC و ۴ فهرست ترکیبی شامل واژه‌های CVC و CVCC می‌باشد.

در گام بعدی برای این که اجرای آزمون برای همه آزمودنی‌ها در شرایط یکسان صورت گیرد و متغیرهای مخدوش کننده نظری تغییرات کیفیت صدای گوینده، شدت صوتی ارائه فهرست‌ها و نسبت سیگنال به نویز، تأثیری در تفسیر نتایج آزمون به دنبال نداشته باشد (مارتین، ۱۹۹۷)، آزمون ساخته شده طبق استاندار DIN ۴۵۶۲۱ در استودیوی صدابرداری مجهز، با شرایط مطلوب از لحاظ فنی و با گویندگی مجری با سابقه ضبط شد.

پس از مرحله ساخت آزمون، ارزشیابی آن از اهمیت ویژه‌ای برخوردار است. فهرست‌های واژگان ضبط شده آزمون جهت اجرا و تفسیر نتایج حاصله باید هنجاریابی شوند. بدین معنا که نتایج آن در افراد طبیعی از لحاظ شنوایی باید مشخص گردد تا تفاوت امتیازات افراد دچار آسیب شنایی با افراد طبیعی قابل تفسیر باشد. جهت هنجاریابی این آزمون، ۶۰ نفر از فارسی‌زبانانی که از لحاظ

درصد وقوع

۱۲-



نمودار ۱- درصد وقوع واژه‌های زبان فارسی- تهران- ۱۳۷۷

شناوی طبیعی به شمار می آیند (۳۰ مرد و ۳۰ زن) در محدوده سنی ۱۸-۲۵ سال مورد آزمون قرار گرفتند. مطابق نظر مارتین حساسیت شناوی این افراد تا بسامد ۴۰۰ هرتز زیر ۱۰ دسی بل و در بسامد ۶۰۰۰ هرتز حداکثر ۱۵ دسی بل بود. این افراد مراجعین کلینیک سنجش شناوی ییمارستان ساسان در تهران بودند که به طریق نمونه گیری تصادفی گرینش شدند. سطح شدتی ارائه آزمون برای همه افراد سطح راحتی شناوی (MCL) آنها بود.

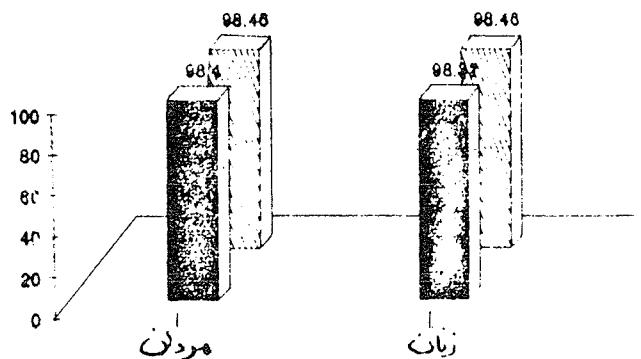
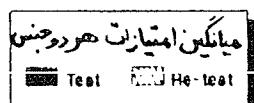
نتایج آماری

تجزیه و تحلیل داده های حاصل از اجرای آزمون نخستین و آزمون دوباره که برای تعیین و اثبات پایایی آزمون انجام شده اند برای هر فهرست به تفکیک جنس در جدول ۲ آمده است. برای مقایسه میانگین امتیاز های دو مرحله آزمون نخستین و آزمون دوباره در تمام فهرست ها از آزمون α زوجی استفاده گردیده است.

مقایسه میانگین امتیاز ها طی آزمون نخستین و آزمون دوباره در کلیه فهرست ها و در آزمودنی های زن و مرد هیچ گونه اختلاف معنی داری را نشان نداده است. نمودار ۲ به شکل سوتی میانگین امتیازات در آزمون نخستین و آزمون دوباره تمام فهرست های آزمودنی های زن و مرد را نشان می دهد.

میانگین امتیاز های به دست آمده از ۱۲ فهرست طی مراحل آزمون نخستین و آزمون دوباره به طور جداگانه جهت آزمون تحلیل یکسویه پراش one-way رتبه گذاری گردیدند و توسط این آزمون با یکدیگر مقایسه شدند (جداول ۳ و ۴).

همان گونه که مشاهده می شود بین امتیاز های حاصل از ۱۲ فهرست طی آزمون نخستین و آزمون دوباره اختلاف معنی داری وجود ندارد. نتایج حاصل از این مرحله همسانی فهرست ها را مجددأ تأیید می نماید.



نمودار ۲- میانگین امتیازات در آزمون نخستین و آزمون دوباره در تمام فهرست ها

جدول ۲- مقایسه میانگین امتیاز های آزمودنی ها برای کلیه فهرست های به تفکیک جنس

۱	درجه آزادی	P	تفاوت میانگین ها	انحراف معیار	خطای معیار میانگین	جنس	شماره فهرست
۰/۵۳	۲۹	۰/۶۰۱	۰/۲۶۷	۲/۷۶۶	۰/۵۰۵	مرد	۱
۰/۵۷	۲۹	۰/۵۷۳	۰/۲۶۷	۲/۵۵۹	۰/۴۶۷	زن	
۰/۷۰	۲۹	۱/۰۰۰	۰/۰۰۰	۲/۱۰۱	۰/۳۸۴	مرد	۲
۰/۵۳	۲۹	۰/۶۰۱	۰/۲۶۷	۲/۷۶۶	۰/۵۰۵	زن	
۰/۷۰	۲۹	۱/۰۰۰	۰/۰۰۰	۱/۴۸۶	۰/۲۷۱	مرد	۳
۰/۳۳	۲۹	۰/۷۴۵	۰/۱۳۳	۲/۲۲۴	۰/۴۰۶	زن	
۰/۵۷	۲۹	۰/۵۷۳	۰/۲۶۷	۲/۵۵۹	۰/۴۷۶	مرد	۴
۰/۵۳	۲۹	۰/۶۰۱	۰/۲۶۷	۲/۷۶۶	۰/۵۰۵	زن	
۰/۵۷	۲۹	۰/۵۷۳	۰/۲۶۷	۲/۵۵۹	۰/۴۷۶	مرد	۵
۰/۲۵	۲۹	۰/۸۰۱	۰/۱۳۳	۲/۸۷۴	۰/۵۲۵	زن	
۰/۷۰	۲۹	۱/۰۰۰	۰/۰۰۰	۲/۱۰۱	۰/۳۸۴	مرد	۶
۰/۷۰	۲۹	۱/۰۰۰	۰/۰۰۰	۲/۱۰۱	۰/۳۸۴	زن	
۰/۲۵	۲۹	۰/۸۰۱	۰/۱۳۳	۲/۸۷۴	۰/۵۲۵	مرد	۷
۰/۳۷	۲۹	۰/۷۱۲	۰/۱۳۳	۱/۹۶۱	۰/۳۵۸	زن	
۰/۳۳	۲۹	۰/۷۴۵	۰/۱۳۳	۲/۲۲۴	۰/۴۰۶	مرد	۸
۰/۲۳	۲۹	۰/۸۲۳	۰/۱۳۳	۳/۲۲۵	۰/۵۹۱	زن	
۰/۲۳	۲۹	۰/۸۲۳	۰/۱۳۳	۳/۲۲۵	۰/۵۹۱	مرد	۹
۰/۳۳	۲۹	۰/۷۴۵	۰/۱۳۳	۲/۲۲۴	۰/۴۰۶	زن	
۰/۲۵	۲۹	۰/۸۰۱	۰/۱۳۳	۲/۸۷۴	۰/۵۲۵	مرد	۱۰
۰/۷۳	۲۹	۰/۵۱۵	۰/۳۰۰	۲/۶۴۷	۰/۴۸۳	زن	
۰/۷۰	۲۹	۱/۰۰۰	۰/۰۰۰	۲/۷۷۹	۰/۵۰۷	مرد	۱۱
۰/۶۳	۲۹	۰/۵۲۶	۰/۲۶۷	۲/۳۳۲	۰/۴۲۶	زن	
۰/۷۰	۲۹	۱/۰۰۰	۰/۰۰۰	۲/۷۷۹	۰/۵۰۷	مرد	۱۲
۰/۲۵	۲۹	۰/۸۰۱	۰/۱۳۳	۲/۸۷۴	۰/۵۲۵	زن	

جدول ۳- نتایج تجزیه و تحلیل یکسویه پراش بر اختلاف امتیازات بارشناستی گکتار در دو جنس طی آزمون نخستین

P	نسبت پراش	میانگین مرتعات	درجه آزادی	مجموع مرتعات	منع تغییرات
۰/۹۷۰۵	۰/۳۶۶	۱/۸۱۶۲	۱۱	۱۹/۸۷۸	بین گروه ها
		۵۰/۰۳۶۹	۷۰۸	۳۵۶۶/۱۳۳	داخل گروه ها
			۷۱۹	۳۵۸۶/۱۱۱	جمع کل

جدول ۴- نتایج تجزیه و تحلیل یکسویه پراش بر اختلاف امتیازات بازشناسی
گفتار در دو جنس طی آزمون دوباره

p	نسبت پراش	میانگین مریعت	درجه آزادی	مجموع مریعت	منبع تغیرات
۰/۹۹۸۷	۰/۱۷۵۷	۰/۸۲۴۲	۱۱	۹/۰۶۷	بین گروه‌ها
		۴/۶۹۲۳	۷۰۸	۳۳۲۲/۱۳۳	داخل گروه‌ها
			۷۱۹	۳۳۳۱/۲	جمع کل

نتایج نشان می‌دهد بین امتیازهای دو گروه مردان و زنان برای هیچ یک از فهرست‌ها طی دو مرحله اجرای آزمون (نخستین و دوباره) اختلاف معنی‌داری وجود ندارد. هدف کلی از این پژوهش ساخت و ارزشیابی فهرست‌های ۲۵ واژه‌ای تک‌هنجایی متوازن واجی جهت تهیه آزمون بازشناسی گفتار برای بزرگسالان فارسی‌زبان بوده است. بررسی نتایج این پژوهش مؤید این نکته است که آزمون بازشناسی گفتار تهیه شده در دستیابی به اهداف کلی و جزئی خود موفق بوده است. این آزمون برای ارزیابی بازشناسی گفتار بزرگسالان بهدلیل عدم وجود تفاوت معنی‌دار بین امتیازهای آزمون‌های نخستین و دوباره فهرست‌ها از پایایی لازم برخوردار است. آزمون از نظر توان واجی و همسانی فهرست‌ها کاملاً معین است، زیرا بین درصد وقوع واکه‌ها و همخوان‌ها در فهرست‌ها و درصد وقوع آنها در مجموعه مربوط به هر یک اختلاف معنی‌داری مشاهده نمی‌شود. نتایج حاصل از مقایسه امتیازهای ۱۲ فهرست با همیگر اختلاف معنی‌داری را نشان نمی‌دهد و مؤید همسانی فهرست‌های دوازده گانه است. بدین ترتیب آزمون بازشناسی گفتار ساخته شده در این پژوهش برای بزرگسالان فارسی‌زبان مناسب بوده و برای ارزیابی بازشناسی گفتار آنها پایا و معین است.

همچنین امتیازهای افراد مورد مطالعه برای هر فهرست بر حسب جنس توسط آزمون ۱ مستقل مقایسه شد. نتایج بدست آمده از این آزمون‌ها در دو جدول ۵ و ۶ به ترتیب طی آزمون نخستین و آزمون دوباره مقایسه گردیده است.

جدول ۵- مقایسه نتایج امتیازات بازشناسی گفتار در دو جنس طی آزمون نخستین

گروه	تعداد نمونه	میانگین انحراف استاندارد	خطای معیار میانگین
مردان	۳۶۰	۹۸/۴۰۰	۲/۲۵۸
زنان	۳۶۰	۹۸/۳۲۲	۲/۲۱۱

p	درجه آزادی	t	خطای معیار اختلاف
۰/۱۶۱	۰/۶۴۱	۷۱۸	۰/۴۷

جدول ۶- مقایسه نتایج امتیازات بازشناسی گفتار در دو جنس طی آزمون دوباره

گروه	تعداد نمونه	میانگین انحراف استاندارد	خطای معیار میانگین
مردان	۳۶۰	۹۸/۴۶۶	۲/۱۶۴
زنان	۳۶۰	۹۸/۴۶۶	۲/۱۴۴

p	درجه آزادی	t	خطای معیار اختلاف
۰/۱۶۷	۱/۰۰	۷۱۸	۰/۰۰۰

منابع

- ۱- ثمره، یدالله. ۱۳۶۸. آواشناسی زبان فارسی. چاپ دوم. مرکز نشر دانشگاهی. تهران
- 2- Martin M., 1997. Speech Audiometry . Whurr Publishers. London.