

ساخت و ارزشیابی آزمون بازشناسی گفتار بزرگسالان فارسی زبان

محمد مصلح - کارشناس ارشد شنوایی شناسی

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 chosen according to judgement of 5 adult native speakers of Persian with high school diploma.

12 open set 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 باهدف ساخت و ارزشیابی فهرست‌های ۲۵ واژه‌ای تک‌هجایی متوازن واجی برای بزرگسالان فارسی‌زبان در دو مرحله صورت گرفته است. در مرحله ساخت، به منظور اجرای معیار توازن واجی در فهرست‌ها، بسامد وقوع هر یک از ۲۹ واج زبان فارسی (شامل ۶ واکه و ۲۳ همخوان) در گفتار بزرگسالان تعیین شد که نشان‌دهنده وجود اختلاف معنی‌دار بین برخی واج‌ها بود. پس از استخراج کلیه واژه‌های تک‌هجایی موجود در ۴ جلد اصلی فرهنگ فارسی معین و صرف نظر کردن از واژه‌های دشوار (باتوجه به معیار ساخت آزمون‌ها) باقضاوت ۵ فرد بزرگسال دارای دیپلم، واژه‌های مناسب موجود در گنجینه واژگان افراد بزرگسال انتخاب شدند. این واژه‌ها باتوجه به معیار توازن واجی در فهرست‌های ۲۵ واژه‌ای گنجانده شدند. بدین ترتیب ۱۲ فهرست ۲۵ واژه‌ای شامل ۵ فهرست با الگوی هجایی CVC، ۴ فهرست با الگوی CVCC و ۳ فهرست به صورت ترکیبی (CVC و CVCC) ساخته شدند. فهرست‌های متوازن شده در استودیوی صدا برداری با صدای یک گوینده ضبط گردید.

در مرحله ارزشیابی، برای هنجاریابی آزمون ۶۰ نفر (۳۰ زن و ۳۰ مرد) دارای شنوایی طبیعی در محدوده سنی ۲۵-۱۸ سال به صورت تصادفی مورد ارزیابی با فهرست های ساخته شده قرار گرفتند. نتایج و یافته های این پژوهش عبارتند از:

- ۱- همه افراد طبیعی با استفاده از تمام فهرست ها، ۱۰۰-۹۲٪ امتیاز را در راحت ترین سطح بلندی ارائه آزمون بدست آوردند.
- ۲- بین امتیازهای بدست آمده از تمام فهرست ها در آزمون نخستین و آزمون دوباره اختلاف معنی داری مشاهده نمی شود ($p > 0.05$).
- ۳- بین امتیازهای بدست آمده از ۱۲ فهرست اختلاف معنی داری وجود ندارد ($p > 0.05$).
- ۴- بین امتیازهای بدست آمده از ۱۲ فهرست بر حسب جنس اختلاف معنی داری وجود ندارد ($p > 0.05$).

باتوجه به نحوه ساخت و نتایج حاصل از آن، آزمون بازنشاسی گفتار بزرگسالان پایا و از نظر توازن واجی و همسانی فهرست ها از نظر دشواری معنایی وازه ها معتبر است و برای ارزیابی بازنشاسی گفتار بزرگسالان فارسی زبان کارایی دارد.

مقدمه

- ۱- تعیین فراوانی هریک از واج های زبان فارسی
- ۲- استخراج وازه های تک هجایی زبان فارسی
- ۳- ساختن فهرست های ۲۵ وازه ای
- ۴- متوازن کردن فهرست ها از لحاظ واجی
- ۵- ضبط کردن فهرست ها
- ۶- هنجاریابی آزمون در افراد طبیعی از لحاظ شنوایی

روش پژوهش

در طرح «ساخت و ارزشیابی آزمون بازنشاسی گفتار برای بزرگسالان فارسی زبان» به منظور گردآوری داده های مورد نیاز برای اجرا، ساخت و ارزشیابی در مراحل مختلف راهکارهای متفاوتی اتخاذ گردید که توضیح داده می شود.

فراوانی واج ها یا بسامد رخداد آواها یک متغیر کیفی و وابسته است و تعیین آن یکی از مهم ترین بخش های این پژوهش بود. در طراحی آزمون بازنشاسی گفتار به این نکته که مواد گفتاری آزمون باید از نظر ترکیب آوایی معادل با گفتار روزمره باشند، توجه شده است (مارتین، ۱۹۹۷). در این پژوهش برای تعیین این فراوانی از برنامه های تلویزیونی دارای بیشترین بیننده استفاده شده است. در این انتخاب با نظر روابط عمومی صدا و سیما و مجله سروش، از برنامه های علمی و تخصصی ویژه صرف نظر گردیده، برنامه هایی با محتوای اجتماعی، خانوادگی و اخبار گزینش و به مدت ۶ ساعت ضبط و بازنویسی شدند. پس از آوانگاری، لازم بود فراوانی تمام آواها و نسبت بین آنها محاسبه و شمارش شود. این کار توسط یک نرم افزار در پژوهشکده الکترونیک دانشگاه صنعتی شریف انجام شد. کل آواهای شمارش شده ۵۲۶۸۷ آوا بود. همچنین فراوانی وقوع هر آوا پس از آوای دیگر محاسبه شده که پژوهشگر را در زمینه گزینش وازه های تک هجایی از نظر بافت آغازی و یا پایانی وازه با یک آوا کمک کرده است.

گفتار، بخش مهمی از ارتباط را تشکیل می دهد و شامل مجموعه بسیار پیچیده و ظریف از متغیرهایی است که برای آشکار ساختن افکار انسانی عمل می کند. 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 آزمون ها «آزمون بازنشاسی گفتار» می باشد. این آزمون به منظور مطالعه برقراری درست ارتباط و میزان تمایز وازه ها از طرف شنونده، بررسی کارکرد درست سیستم های تقویتی و نیاز برای بازخوانی افراد دچار آسیب شنوایی به کار می رود. برای این که آزمون در اجرای موارد فوق کارایی داشته باشد باید به صورت استاندارد شده، مطابق معیارهای زیر بکار رود:

- ۱- ساختار و محتوای آزمون از «واژه های تک هجایی معنی دار» تشکیل شده باشد.
 - ۲- مواد گفتاری آزمون از لحاظ معنایی برای شنونده آشنا باشد.
 - ۳- در فهرست وازگان آزمون، توازن واجی وجود داشته باشد.
 - ۴- آزمون به صورت ضبط شده بکار رود.
- رعایت این معیارها در اجرای آزمون و تفسیر نتایج حاصل از آن اهمیت زیادی دارد (مارتین، ۱۹۹۷). باتوجه به اهمیت تشخیصی و توان بخشی بسیار بالای این آزمون، تاکنون آزمون استاندارد دی به زبان فارسی برای بزرگسالان در دسترس مراکز شنوایی شناسی قرار نگرفته بود. لذا ساخت و استاندارد نمودن آن با معیارهای مربوطه از مهم ترین نیازهای مراکز شنوایی شناسی به شمار می آمد.

اهداف

هدف کلی و نهایی این پژوهش ساخت آزمون بازنشاسی گفتار برای بزرگسالان فارسی زبان و روا و معتبر نمودن آن طبق معیارهای مربوطه بود. اهداف دیگر پژوهش که برای رسیدن به هدف نهایی باید به آنها نیز پرداخته می شد بدین قرار است:

بر پایه بررسی این مجموعه از داده ها، درصد وقوع ۲۹ واج شامل ۲۳ همخوان و ۶ واکه (نمره، ۱۳۶۴: ۳۹) تعیین شده است. در این بررسی واج /æ/ با ۱۰/۸۱٪ بیشترین و واج /z/ با ۰/۰۴٪ کمترین درصد وقوع را در گفتار فارسی زبانان بزرگسال دارند (جدول ۱).

جدول ۱- درصد وقوع واج‌ها در گفتار بزرگسالان فارسی زبان تهرانی

شماره	واج	درصد وقوع
۱	/æ/	۱۰/۸۱
۲	/e/	۱۰/۳۹
۳	/a/	۸/۲۰
۴	/ɪ/	۶/۳۹
۵	/i/	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	/S/	۲/۳۶
۱۷	/k/	۲/۱۵
۱۸	/u/	۲/۰۳
۱۹	/ʃ/	۱/۸۸
۲۰	/z/	۱/۸۶
۲۱	/x/	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	/ʒ/	۰/۰۴
جمع		۱۰۰

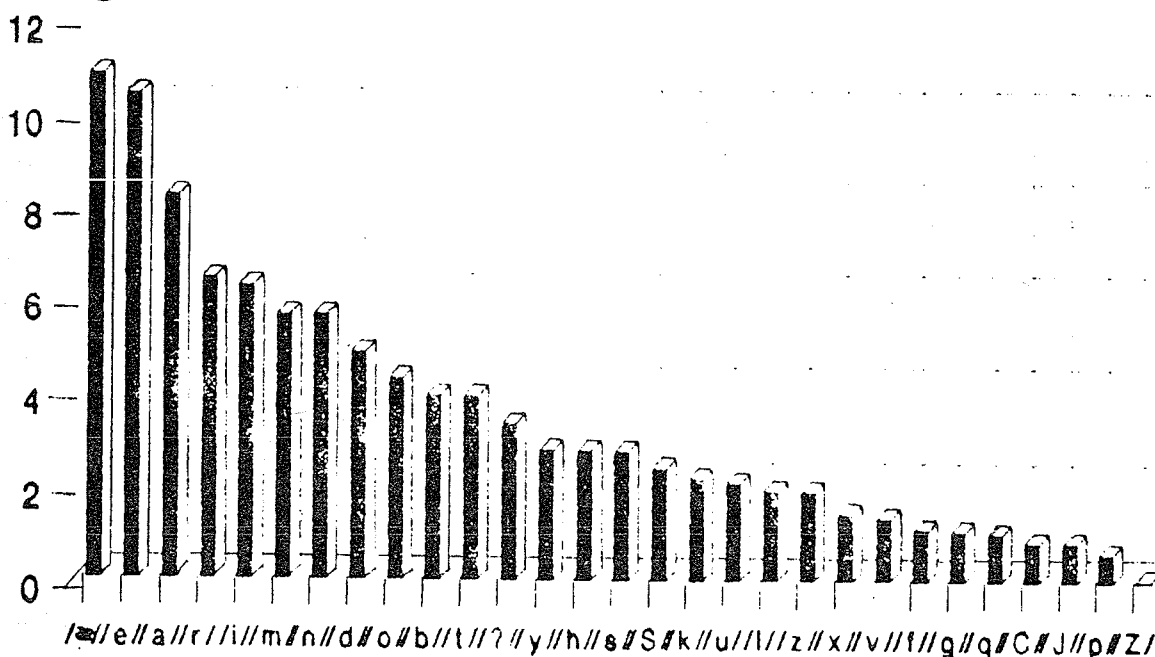
در گام بعدی لازم بود واژه‌های تک‌هجایی مناسب در دسترس باشند تا بتوان از آنها در تنظیم فهرست‌های آزمون استفاده نمود. برای این کار کلیه واژه‌های تک‌هجایی زبان فارسی که در ۴ جلد اصلی فرهنگ فارسی معین وجود دارند، استخراج شدند (۳۱۰۵ واژه). از این تعداد ۱۶۴۴ واژه دارای الگوی CVCC، ۱۳۸۸ واژه به صورت CVC و ۷۳ واژه به صورت CV هستند. از آنجا که این آزمون باید برای عموم بزرگسالان فارسی‌زبان قابلیت اجرا داشته باشد، واژه‌های مورد استفاده، بایستی در گنجینه واژگان این افراد وجود داشته و از لحاظ معنایی برای آنها دشوار نباشد. برای تعیین دشواری معنایی واژه‌های تک‌هجایی استخراج شده، از ۵ نفر فارسی‌زبان دارای دیپلم، در محدوده سنی ۱۸-۲۵ سال استفاده گردید. برپایه قضاوت آنها واژه‌هایی با سطح دشواری معنایی متوسط، گزینش شدند. اگر ۳ نفر آنان یا بیشتر، واژه‌ای را به عنوان متوسط انتخاب می‌کردند،

آن واژه برگزیده می‌شد و در این صورت حذف می‌گردید. در مرحله بعد واژه‌های تک‌هجایی گزینش شده با توجه به فراوانی رخداد آواها در جمعیت نمونه در فهرست‌های ۲۵ واژه‌ای گنجانده شدند. فراوانی وقوع و درصد تقدم و تأخر هر کدام از واج‌ها در این فهرست رعایت شده است. با توجه به 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 CVC، ۳ 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 مخدوش‌کننده نظیر 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 استاندارد 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 شنوایی با افراد طبیعی قابل تفسیر باشد. جهت هنجاریابی این آزمون ۶۰ نفر از فارسی‌زبانانی که از لحاظ

درصد وقوع



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

جدول ۲- مقایسه میانگین امتیازهای آزمودنی‌ها برای کلیه فهرست‌ها به تفکیک جنس

شماره فهرست	جنس	خطای معیار میانگین	انحراف معیار	تفاوت میانگین‌ها	p	درجه آزادی	t
۱	مرد	۰/۵۰۵	۲/۷۶۶	۰/۲۶۷	۰/۶۰۱	۲۹	۰/۵۳
	زن	۰/۴۶۷	۲/۵۵۹	۰/۲۶۷	۰/۵۷۳	۲۹	۰/۵۷
۲	مرد	۰/۳۸۴	۲/۱۰۱	۰/۰۰۰	۱/۰۰۰	۲۹	۰/۰۰۰
	زن	۰/۵۰۵	۲/۷۶۶	۰/۲۶۷	۰/۶۰۱	۲۹	۰/۵۳
۳	مرد	۰/۲۷۱	۱/۴۸۶	۰/۰۰۰	۱/۰۰۰	۲۹	۰/۰۰۰
	زن	۰/۴۰۶	۲/۲۲۴	۰/۱۳۳	۰/۷۴۵	۲۹	۰/۳۳
۴	مرد	۰/۴۷۶	۲/۵۵۹	۰/۲۶۷	۰/۵۷۳	۲۹	۰/۵۷
	زن	۰/۵۰۵	۲/۷۶۶	۰/۲۶۷	۰/۶۰۱	۲۹	۰/۵۳
۵	مرد	۰/۴۷۶	۲/۵۵۹	۰/۲۶۷	۰/۵۷۳	۲۹	۰/۵۷
	زن	۰/۵۲۵	۲/۸۷۴	۰/۱۳۳	۰/۸۰۱	۲۹	۰/۲۵
۶	مرد	۰/۳۸۴	۲/۱۰۱	۰/۰۰۰	۱/۰۰۰	۲۹	۰/۰۰۰
	زن	۰/۳۸۴	۲/۱۰۱	۰/۰۰۰	۱/۰۰۰	۲۹	۰/۰۰۰
۷	مرد	۰/۵۲۵	۲/۸۷۴	۰/۱۳۳	۰/۸۰۱	۲۹	۰/۲۵
	زن	۰/۳۵۸	۱/۹۶۱	۰/۱۳۳	۰/۷۱۲	۲۹	۰/۳۷
۸	مرد	۰/۴۰۶	۲/۲۲۴	۰/۱۳۳	۰/۷۴۵	۲۹	۰/۳۳
	زن	۰/۵۹۱	۳/۲۳۵	۰/۱۳۳	۰/۸۲۳	۲۹	۰/۲۳
۹	مرد	۰/۵۹۱	۳/۲۳۵	۰/۱۳۳	۰/۸۲۳	۲۹	۰/۲۳
	زن	۰/۴۰۶	۲/۲۲۴	۰/۱۳۳	۰/۷۴۵	۲۹	۰/۳۳
۱۰	مرد	۰/۵۲۵	۲/۸۷۴	۰/۱۳۳	۰/۸۰۱	۲۹	۰/۲۵
	زن	۰/۴۸۳	۲/۶۴۷	۰/۳۰۰	۰/۵۱۵	۲۹	۰/۷۳
۱۱	مرد	۰/۵۰۷	۲/۷۷۹	۰/۰۰۰	۱/۰۰۰	۲۹	۰/۰۰۰
	زن	۰/۴۲۶	۲/۳۳۳	۰/۲۶۷	۰/۵۳۶	۲۹	۰/۶۳
۱۲	مرد	۰/۵۰۷	۲/۷۷۹	۰/۰۰۰	۱/۰۰۰	۲۹	۰/۰۰۰
	زن	۰/۵۲۵	۲/۸۷۴	۰/۱۳۳	۰/۸۰۱	۲۹	۰/۲۵

شنوایی طبیعی به شمار می‌آیند (۳۰ مرد و ۳۰ زن) در محدوده سنی ۱۸-۲۵ سال مورد آزمون قرار گرفتند. مطابق نظر مارتین حساسیت شنوایی این افراد تا بسامد ۴۰۰۰ هرتز زیر ۱۰ دسی‌بل و در بسامد ۶۰۰۰ تا ۸۰۰۰ هرتز حداکثر ۱۵ دسی‌بل بود. این افراد مراجعین کلینیک سنجش شنوایی بیمارستان ساسان در تهران بودند که به طریق نمونه‌گیری تصادفی گزینش شدند. سطح شدتی ارائه آزمون برای همه افراد سطح راحتی شنوایی (MCL) آنها بود.

نتایج آماری

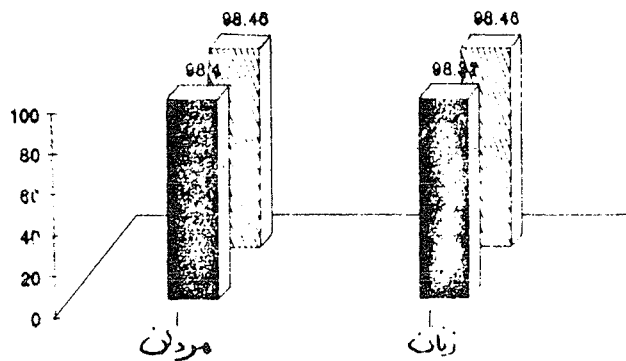
تجزیه و تحلیل داده‌های حاصل از اجرای آزمون نخستین و آزمون دوباره که برای تعیین و اثبات پایایی آزمون انجام شده‌اند برای هر فهرست به تفکیک جنس در جدول ۲ آمده است. برای مقایسه میانگین امتیازهای دو مرحله آزمون نخستین و آزمون دوباره در تمام فهرست‌ها از آزمون t زوجی استفاده گردیده است.

مقایسه میانگین امتیازها طی آزمون نخستین و آزمون دوباره در کلیه فهرست‌ها و در آزمودنی‌های زن و مرد هیچ‌گونه اختلاف معنی‌داری را نشان نداده است. نمودار ۲ به شکل ستونی میانگین امتیازات در آزمون نخستین و آزمون دوباره تمام فهرست‌های آزمودنی‌های زن و مرد را نشان می‌دهد.

میانگین امتیازهای به دست آمده از ۱۲ فهرست طی مراحل آزمون نخستین و آزمون دوباره به‌طور جداگانه جهت آزمون تحلیل یکسویه پرش one-way رتبه‌گذاری گردیدند و توسط این آزمون بایکدیگر مقایسه شدند (جدول ۳ و ۴).

همان‌گونه که مشاهده می‌شود بین امتیازهای حاصل از ۱۲ فهرست طی آزمون نخستین و آزمون دوباره اختلاف معنی‌داری وجود ندارد. نتایج حاصل از این مرحله همسانی فهرست‌ها را مجدداً تأیید می‌نماید.

میانگین امتیازات هر دو جنس



نمودار ۲- میانگین امتیازات در آزمون نخستین و آزمون دوباره در تمام فهرست‌ها

جدول ۳- نتایج تجزیه و تحلیل یکسویه پرش بر اختلاف امتیازات بارشناسی گفتار در دو جنس طی آزمون نخستین

منبع تغییرات	مجموع مربعات	درجه آزادی	میانگین مربعات	نسبت پرش	p
بین گروه‌ها	۱۹/۸۷۸	۱۱	۱/۸۱۶۲	۰/۳۶۰۶	۰/۹۷۰۵
داخل گروه‌ها	۳۵۶۶/۱۳۳	۷۰۸	۵/۰۳۶۹		
جمع کل	۳۵۸۶/۱۱۱	۷۱۹			

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

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

همچنین امتیازهای افراد مورد مطالعه برای هر فهرست برحسب جنس توسط آزمون t 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 بارشناسی گفتار در دو جنس طی آزمون نخستین

گروه	تعداد نمونه	میانگین	انحراف استاندارد	خطای معیار میانگین
مردان	۳۶۰	۹۸/۴۰۰	۲/۲۵۸	۰/۱۱۹
زنان	۳۶۰	۹۸/۳۲۲	۲/۲۱۱	۰/۱۱۷

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

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

گروه	تعداد نمونه	میانگین	انحراف استاندارد	خطای معیار میانگین
مردان	۳۶۰	۹۸/۴۶۶	۲/۱۶۴	۰/۱۱۴
زنان	۳۶۰	۹۸/۴۶۶	۲/۱۴۴	۰/۱۱۳

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

نتایج نشان می‌دهد بین امتیازهای دو گروه مردان و زنان برای هیچ‌یک از فهرست‌ها طی دومرحله اجرای آزمون (نخستین و دوباره) اختلاف معنی‌داری وجود ندارد. هدف کلی از این پژوهش ساخت و ارزشیابی فهرست‌های ۲۵واژه‌ای تک‌هجایی متوازن واجی جهت تهیه آزمون بارشناسی گفتار برای بزرگسالان فارسی‌زبان بوده است. بررسی نتایج این پژوهش مؤید این نکته است که آزمون بارشناسی گفتار تهیه‌شده در دستیابی به اهداف کلی و جزئی خود موفق بوده است. این آزمون برای ارزیابی بارشناسی گفتار بزرگسالان به دلیل عدم وجود تفاوت معنی‌دار بین امتیازهای آزمون‌های نخستین و دوباره فهرست‌ها از پایداری لازم برخوردار است. آزمون از نظر توازن واجی و همسانی فهرست‌ها کاملاً معتبر است، زیرا بین درصد وقوع واژه‌ها و همخوان‌ها در فهرست‌ها و درصد وقوع آنها در مجموعه مربوط به هر یک اختلاف معنی‌داری مشاهده نمی‌شود. نتایج حاصل از مقایسه امتیازهای ۱۲فهرست با همدیگر اختلاف معنی‌داری را نشان نمی‌دهد و مؤید همسانی فهرست‌های دوازده‌گانه است. بدین ترتیب آزمون بارشناسی گفتار ساخته شده در این پژوهش برای بزرگسالان فارسی‌زبان مناسب بوده و برای ارزیابی بارشناسی گفتار آنها پایا و معتبر است.

منابع

۱- ثمره، یدالله. ۱۳۶۸. *آواشناسی زبان فارسی*. چاپ دوم. مرکز نشر دانشگاهی. تهران

2- Martin M., 1997. *Speech Audiometry*. Whurr Publishers. London.