

پژوهش‌نامه‌ی آموزش زبان فارسی به غیر فارسی‌زبانان

(علمی-پژوهشی)

سال سوم، شماره‌ی دوم (پیاپی ۷)، تابستان ۱۳۹۳

آزمون چندگزینه‌ای شنیدن: مقایسه‌ی تأثیر ارائه‌ی گزینه‌ها در دو قالب نوشتاری و گفتاری

بهمن زندی

استاد گروه زبان‌شناسی دانشگاه پیام نور

بلقیس روشن

دانشیار گروه زبان‌شناسی دانشگاه پیام نور

امیررضا وکیلی‌فرد

استادیار گروه آموزش زبان فارسی دانشگاه بین‌المللی امام خمینی (ره)

لیلا گلیپور

دانشجوی دکتری زبان‌شناسی همگانی دانشگاه پیام نور تهران

چکیده

مهارت شنیداری و مهارت گفتاری از لحاظ دانش آموزش‌شکافی زبان ارتباط تنگاتنگی در تعاملات زبانی با یکدیگر دارند. از این‌رو، عموماً اندازه‌گیری درک شنیداری بدون در نظر گرفتن سایر مهارت‌های زبانی فارسی‌آموزان کمی دشوار به نظر می‌رسد. هدف این پژوهش، پاسخگویی به این پرسش است: آزمونی که گزینه‌های آن در قالب نوشتاری ارائه می‌گردد، برای ارزیابی مهارت شنیدن مناسب‌تر است یا آزمونی که گزینه‌ها در آن به صورت گفتاری طراحی و اجرا می‌گردد؟ بدین منظور، برای سنجش درک شنیداری، آزمونی طراحی گردید و برای اجرای این پژوهش، گزینه‌های آزمون در دو قالب نوشتاری و گفتاری ارائه گردید. نمونه‌ی آماری این پژوهش را ۳۶ فارسی‌آموز غیرایرانی مرکز آموزش زبان فارسی دانشگاه بین‌المللی امام خمینی (ره) در رده‌ی سنی ۱۸ تا ۲۵ سال تشکیل می‌دادند که از لحاظ زبانی در سطح پیشرفته و به‌طور تصادفی انتخاب شدند. روش انجام تحقیق از نوع میدانی بوده و برای تجزیه و تحلیل داده‌ها از نرم افزارهای آماری استفاده شده است. یافته‌های پژوهش نشان داد که در مقایسه با قالب گفتاری، دیدگاه آزمودنی‌ها نسبت به دیدگاه نوشتاری، مثبت‌تر و رضایت‌مندی آنان نیز از این نوع قالب ارائه گزینه‌ها بیشتر بوده است، اما آزمودنی‌ها در پاسخ به گزینه‌ها در قالب گفتاری عملکرد بهتری داشته و موفقیت بیشتری به دست آورده‌اند.

کلیدواژه‌ها: آزمون شنیدن، پرسش‌های چندگزینه‌ای، گزینه‌های گفتاری و نوشتاری

تاریخ دریافت مقاله: ۱۳۹۳/۳/۲

تاریخ پذیرش نهایی مقاله: ۱۳۹۳/۴/۱۵

رایانشانی نویسنده‌ی مسئول (لیلا گلیپور):

l.g.135569@gmail.com

۱. مقدمه

سیستم‌ها عموماً دارای سه مؤلفه‌ی ورودی، پردازش و خروجی می‌باشند. در یادگیری زبان دوم، مهارت شنیدن و خواندن به عنوان دو درونداد، آغازگر فرایند یادگیری زبان دوم می‌باشند و مغز، کار تجزیه و تحلیل را بر عهده دارد و در نهایت صحبت کردن و نوشتن، خروجی فرایند یادگیری زبان دوم است. صحبت کردن در نتیجه‌ی بلوغ شنیداری رخ می‌دهد؛ بنابراین، مهارت شنیدن در میان مهارت‌های زبانی از اهمیت ویژه‌ای در آموزش برخوردار است. در فرایند فراگیری زبان اول نیز کودکان مدت‌ها فقط گوش می‌کنند و مقداری از دانش زبانی اولیه‌ی آنها صرفاً با شنیدن کسب می‌شود. کودک پانزده‌ماهه می‌تواند در مقابل گفتارهایی از قبیل «دست زن»، «به من نگاه کن»، «بیا اینجا» و... عکس‌العمل مناسب بروز دهد که نشانه‌ی درک درست وی از مفهوم این گفتارها و نشانه‌ی پردازش این گفتارها بر پایه‌ی دانش زبانی است که به‌وسیله‌ی شنیدن کسب کرده است. شاید به همین سبب باید آموزش زبان دوم/خارجی را نیز با مهارت شنیدن آغاز کرد (ضیاء‌حسینی، ۱۳۸۵: ۱۱۰).

در گذشته، مهارت شنیداری یک مهارت منفعل و غیرپویا به شمار می‌رفت و در برنامه‌های درسی توجه زیادی به این مهارت نمی‌شد، زیرا قابل مشاهده نبود؛ علاوه بر این، تأکید زیاد بر روی مهارت‌های تولیدی و صحبت کردن نیز دلیل دیگری بود که مهارت شنیدن را کم‌اهمیت جلوه می‌داد، اما در انگاره‌های اخیر، مهارت شنیداری یک فرایند فعال و پویا به شمار می‌رود که می‌تواند تأثیر بسزایی در یادگیری زبان داشته باشد؛ از این‌رو، توجه خاصی را از سوی مدرسان زبان دوم طلب می‌کند (کومب و همکاران، ۲۰۰۷: ۹۰).

در آزمون‌سازی استاندارد، چندین آزمون بیان شفاهی جداگانه قابل دسترس هستند، اما به‌ندرت می‌توان یک آزمون شنیدن پیدا کرد. یکی از دلایل این موضوع آن است که شنیدن، اغلب بخشی از صحبت کردن در نظر گرفته می‌شود. چگونه می‌توان بدون داشتن مهارت در شنیدن، به زبانی دیگر صحبت کرد؟ علاوه بر این، چون مهارت صحبت کردن ماهیت کاملاً قابل‌مشاهده‌ای دارد، در مقایسه با مهارت شنیدن عینی‌تر به نظر می‌رسد. شاید جهت‌گیری‌های کلی در قبال صحبت کردن، دلیل بهتری برای این امر باشد. متأسفانه ارزش یک سخنگوی خوب از یک شنونده‌ی خوب بسیار بیشتر است. مردم معمولاً برای این که بی‌رسند آیا فردی در یک زبان دیگر مهارت دارد این سوال را می‌پرسند که «آیا انگلیسی صحبت می‌کنید؟». آنها به‌ندرت می‌پرسند «آیا زبان انگلیسی را درک می‌کنید؟». هر آموزگار زبانی می‌داند که توانایی بیان شفاهی افراد به‌جز در تک‌گویی‌ها، سخنرانی‌ها، قرائت کردن‌ها و امثال آن دقیقاً می‌تواند توانایی درک شنیداری آنها را تقویت کند. در زندگی روزمره، چه در محیط کار و چه در بافت‌های آموزشی و محیط خانه، درک شنیداری از لحاظ زمان، تعداد واژگان، تلاش و توجه - که همگی قابل اندازه‌گیری هستند - به‌مراتب بیشتر از

صحبت کردن است؛ بنابراین، توجه دقیق به شنیدن، به‌عنوان شیوه‌ای از کنش ارزشیابی بسیار ضروری است (براون، ۲۰۰۴: ۱۱۹).

بر همین اساس و با توجه به اهمیت مهارت شنیدن، در این پژوهش سعی شده است که دو نوع آزمون با قالب نوشتاری و گفتاری مورد بررسی و مقایسه قرار گیرد تا مشخص شود که کدام یک برای ارزیابی مهارت درک شنیداری مناسب‌تر است. پرسش‌های پژوهش حاضر به شکل زیر است:

(۱) به منظور ارزیابی مهارت درک شنیداری فارسی‌آموزان، آزمون چندگزینه‌ای نوشتاری مناسب‌تر است یا گفتاری؟

(۲) فارسی‌آموزان در مورد نوع آزمون شنیدن، چه دیدگاه و نظری دارند؟
با توجه به پرسش‌های مطرح شده، فرضیه‌ی این پژوهش بر این پایه استوار است که آزمون گفتاری نسبت به نوشتاری، هم از لحاظ عملکرد و هم از لحاظ دیدگاه آزمودنی‌ها برای ارزشیابی درک شنیداری مناسب‌تر است.

۲. پیشینه‌ی پژوهش

در این بخش، مطالعات و پژوهش‌های خارجی و داخلی که در گذشته انجام شده است، مورد بررسی قرار می‌گیرند.

۱.۲. پژوهش‌های خارجی

در گذشته پژوهش‌هایی در مورد نوع سؤالات آزمون شنیدن مانند چندگزینه‌ای^۱، پرکردن جای خالی^۲، سؤالات کوتاه‌پاسخ^۳ و سؤالات بازپاسخ^۴ انجام شده است؛ به‌عنوان مثال ایکین (۱۹۹۲) به منظور ارزیابی مهارت شنیداری زبان‌آموزانی که فرانسوی یاد می‌گیرند، سؤالات چندگزینه‌ای را با دیگر سؤالات مورد بررسی قرار داد و در نهایت مشخص شد که آزمون‌دهندگان از سؤالات چهارگزینه‌ای نمره‌ی بهتری گرفتند. تینگ (۱۹۹۸) سه نوع سؤالات چندگزینه‌ای، بندش^۵ و سؤالات محدودپاسخ را مورد بررسی قرار داد و نتیجه گرفت که دانشجویان دانشگاهی در پاسخ به سؤالات چندگزینه‌ای موفق‌تر هستند. مطالعه‌ی دیگری که توسط چنگ (۲۰۰۴) انجام گرفت، مشخص کرد که آزمون‌دهندگان در سؤالات چندگزینه‌ای و بندش در مقایسه با سؤالات محدودپاسخ، عملکرد بهتری دارند.

^۱. multiple choice

^۲. filling the blank

^۳. short-answer question

^۴. open-ended

^۵. cloze test

چنگ (۲۰۰۵) در بررسی‌های خود نشان داد که آزمون‌دهندگانی که سطح مهارت زبانی آنها در سطح مقدماتی تا متوسط است، در سؤالات چهارگزینه‌ای نمرات بالاتری کسب کرده‌اند. او همچنین پیشنهاد کرد که آزمون‌هایی که مستلزم نوشتن است با ساختار توانایی شنیدن همسو نیست، چون که تقاضایی به‌جز درک شنیداری دارد.

چنگ (۲۰۰۵) و وو (۱۹۹۸) خاطر نشان کردند که اگرچه سؤالات چندگزینه‌ای نسبت به سؤال‌های بازپاسخ، آسان‌تر هستند، در تحقیقات دیگر نشان داده‌اند که این نوع آزمون برای زبان‌آموزانی که سطح مهارت زبانی آنها پایین است، مفید نیست؛ چون نیازمند مهارت خواندن است. وو بیان کرد که اگر از سؤالات چندگزینه‌ای برای زبان‌آموزان سطح پایین استفاده شود، پایایی آزمون شنیدن تحت تأثیر قرار می‌گیرد.

پژوهش دیگری که توسط فریدل و کوستین (۱۹۹۹) انجام گرفت حاکی از این است که علاوه بر نوع سؤالات (چندگزینه‌ای یا بازپاسخ) عوامل مهم دیگری در سهولت و دشواری آزمون شنیدن مؤثر است؛ به‌عنوان مثال، اگر زبان‌آموز اطلاعات مورد نیاز را در ابتدای متن بشنود، راحت‌تر به سؤالات پاسخ می‌دهد؛ در حالی که اگر همین اطلاعات در وسط یا انتهای متن باشد، پاسخگویی به آنها سخت‌تر می‌شود. علاوه بر این، اگر واژه‌هایی که در سؤال استفاده می‌شوند با واژه‌های متن هم‌پوشانی داشته باشند، ضریب سهولت آزمون نیز بیشتر خواهد شد.

یانگاو و گرین (۲۰۰۸) در پژوهش خود به این نتیجه رسیدند که سؤالات استنباطی^۱ نسبت به سؤالات عینی^۲، در دشواری آزمون، تأثیرگذارتر است. آنا و همکاران (۲۰۱۳) در پژوهشی چنین بیان کرده‌اند که نوع سؤالات آزمون بر عملکرد زبان‌آموزان تأثیر بسیاری دارد. آنها با مقایسه‌ی سؤالات چندگزینه‌ای شفاهی و کتبی در آزمون درک شنیداری به این نتیجه رسیدند که زبان‌آموزانی که در سطح پایین‌تری قرار دارند در سؤالات چندگزینه‌ای کتبی نمره‌ی بالاتری گرفتند، ولی دانشجویان پیشرفته در هر دو حالت نمره‌ی یکسانی کسب کردند؛ این در حالی بود که طبق نظرسنجی، ۷۵ درصد زبان‌آموزان سؤالات چندگزینه‌ای کتبی را ترجیح می‌دادند.

۲.۲. پژوهش‌های داخلی

قنسولی (۱۳۸۹) در مقاله‌ای با عنوان «طراحی و روایی‌سنجی آزمون بسندگی زبان فارسی»، مدل چندبعدی زبان را از میان مدل‌های مختلف اجزای بسندگی زبان انتخاب کرده است و دلیل آن را نیز این‌گونه بیان می‌کند که این مدل با شرایط آزمون واقعی تناسب دارد. در این مدل به دو مهارت اصلی زبان یعنی خواندن، شنیدن و مقوله‌های دستور و واژگان توجه شده است. این مدل از لحاظ ظاهری با مدل‌های

1. inferential questions

2. objective questions

استاندارد کنونی همانند تافل و آیلتس شباهت دارد. در طراحی این آزمون تنها به دو مهارت اصلی خواندن و شنیدن توجه شده و مهارت‌های گفتاری و نوشتاری به دلیل محدودیت‌های تحقیق لحاظ نگردیده است. در این پژوهش ضمن ساخت یک نمونه آزمون بسندگی زبان فارسی با تأکید بر آزمون‌های شنیدن، خواندن و دستور سعی شده است که روایی و پایایی آزمون نیز مورد ارزیابی قرار گیرد. این آزمون که به منظور ارزیابی مهارت شنیدن طراحی شده، از ۳۰ سؤال چهارگزینه‌ای نوشتاری در قالب گفتگوهای کوتاه، متوسط و بلند تشکیل شده است.

جلیلی (۱۳۹۰) در پژوهشی با عنوان «آزمون مهارتی فارسی بر پایه‌ی چهار مهارت اصلی زبانی» تلاش کرده است یک آزمون مهارتی برای زبان فارسی طراحی نماید. این آزمون بر پایه‌ی آیلتس طراحی شده و به اختصار «آمفا» نامیده شده است. آزمون درک شنیداری آمفا دارای ۴۰ پرسش و مدت زمان آن نیز ۴۰ دقیقه است. پرسش‌ها از نوع تکمیل‌کردنی، بازپاسخ، پر کردن جدول، مشخص کردن اجزای یک تصویر و چندگزینه‌ای می‌باشند.

شقاقی و همکاران (۱۳۹۰) در مقاله‌ای با عنوان «هنجاریابی آزمون بین‌المللی زبان فارسی»، آزمون استاندارد کانون زبان را از نظر هنجاریابی مورد بررسی قرار داده‌اند. این آزمون به ترتیب شامل بخش‌های شنیدن (۳۰ سؤال)، دستور (۳۰ سؤال)، و واژگان و خواندن (هر یک ۲۰ سؤال) است اما مهارت گفتاری در این آزمون لحاظ نگردیده و همانند آزمون «تافل کاغذی»^۱ برای دستور بخش جداگانه‌ای در نظر گرفته شده است. سؤالات مهارت شنیدن در این آزمون نیز از نوع چهارگزینه‌ای است و به صورت کتبی طراحی شده است.

بررسی پژوهش‌های داخلی بیانگر این موضوع است که همه‌ی آزمون‌هایی که به منظور ارزیابی مهارت شنیداری زبان فارسی طراحی شده‌اند در قالب آزمون نوشتاری بوده‌اند و هیچ یک از آزمون‌های شنیداری به صورت گفتاری طراحی نشده است.

۳.۲. مروری بر آزمون‌های استاندارد درک شنیداری

تاکنون آزمون‌های استاندارد مختلفی برای زبان‌های دوم/خارجی طراحی و برگزار شده است که در این جا به‌طور خلاصه راهکارهای ارزیابی مهارت درک شنیداری چهار نوع آزمون استاندارد شامل تافل، آیلتس، میشیگان و پیرسون مورد بررسی قرار می‌گیرد.

۱. Paper-Based TOEFL (=PBT)

۱،۳،۲. آزمون تافل^۱

این آزمون میزان توانایی افراد را در زمینه‌ی برقراری ارتباط و سطح مهارت زبان انگلیسی گویشوران غیرانگلیسی‌زبان در سطح دانشگاه‌ها و مؤسسات آموزشی می‌سنجد. آزمون تافل توسط مؤسسه‌ی «ای‌تی‌اس»^۲ در سه نوع کاغذی، رایانه‌ای و اینترنتی برگزار می‌شود. تافل اولین بار در سال ۱۹۶۴ برگزار شد. مدرک تافل تنها دو سال اعتبار دارد و این آزمون نمره‌ی ردی یا قبولی ندارد و هر دانشگاه یا مؤسسه نمره‌ی خاص خودش را طلب می‌کند (براون، ۲۰۰۴: ۸۴-۷۲).

الف) بخش شنیداری تافل کاغذی: این بخش از تافل توانایی افراد را در درک مکالمه‌های انگلیسی مورد سنجش قرار می‌دهد. این بخش خود به سه قسمت تقسیم می‌شود؛ قسمت اول این شامل مکالمه‌هایی کوتاه می‌باشد که بین دو نفر صورت می‌گیرند و پس از پایان هر مکالمه، پرسشی مبنی بر این که چه چیزی بین این دو نفر رد و بدل شده است و یا پرسشی درباره‌ی شرایط رخداد مکالمه مطرح می‌شود. در قسمت دوم، مکالمه‌های گسترده‌تری بین دو یا چند نفر انجام می‌شود و قسمت سوم نیز در برگرفته‌ی نطق‌ها و سخنرانی‌هایی است که بیشتر درباره‌ی موضوعات دانشگاهی است. این بخش از آزمون دارای ۵۰ پرسش است و مدت پاسخ‌گویی به آن ۳۵ دقیقه است (همان).

ب) بخش شنیداری تافل رایانه‌ای: برخلاف آزمون کاغذی، زبان‌آموز در آزمون شنیدن تافل رایانه‌ای از گوشی استفاده می‌کند. همچنین، در این آزمون علاوه بر صدایی که شنیده می‌شود، یک تصویر هم به نمایش درمی‌آید. به همراه هر پرسش، گزینه‌هایی در صفحه نمایش داده می‌شوند که باید کنار گزینه‌ی صحیح کلیک کرد. در آزمون تافل رایانه‌ای گاهی بیش از یک پاسخ صحیح وجود دارد؛ در چنین حالتی به شما گفته می‌شود که چند پاسخ صحیح را باید انتخاب کنید. در کل، بخش شنیداری تافل رایانه‌ای به دو قسمت تقسیم می‌شود؛ قسمت اول شامل مکالمه‌های کوتاه است و در قسمت دوم هم مکالمه‌ها و سخنرانی‌های طولانی‌تری پخش می‌شود. کل بخش شنیداری آزمون دارای ۳۰ تا ۵۰ پرسش است و مدت پاسخ‌گویی به آن ۴۰ تا ۶۰ دقیقه است (همان).

پ) بخش شنیداری تافل اینترنتی: بخش شنیداری آزمون تافل اینترنتی توانایی داوطلب را در درک مکالمات انگلیسی ارزیابی می‌کند. تفاوت چندانی در بخش شنیداری آزمون تافل اینترنتی با نسخه‌ی رایانه‌ای آن وجود ندارد. این بخش شامل دو مکالمه و چهار سخنرانی (بیشتر در حوزه‌ی دانشگاهی) همراه با پرسش‌های متعدد است. سؤالات آزمون تافل اینترنتی در قالب نوشتاری و شامل انواع چندگزینه‌ای، درست-نادرست، تکمیل کردنی و محدود پاسخ می‌باشد (همان).

^۱. Test of English as a Foreign Language (=TOEFL)

^۲. Educational Testing Service(=ETS)

۲،۳،۲. آزمون آیلتس^۱

در حال حاضر آیلتس یکی از معتبرترین آزمون‌های زبان انگلیسی است که در سراسر دنیا برگزار می‌شود. آیلتس در واقع آزمونی بریتانیایی است و از این نظر در مقابل تافل که آزمونی آمریکایی است، قرار می‌گیرد. این آزمون با همکاری مشترک دانشگاه کمبریج انگلستان، سفارت انگلستان و حوزه‌ی گسترش برنامه‌های بین‌الملل سفارت استرالیا برگزار می‌شود. آیلتس دارای دو نوع دانشگاهی^۲ و عمومی^۳ است و سطح‌بندی این آزمون از ۰ تا ۹ می‌باشد. نمره‌ی قبولی برای این آزمون وجود ندارد اما به‌طور معمول نمره‌ی ۵، سطح متوسط، و نمره‌ی ۶ به بالا، سطح خوب تلقی می‌شود. بخش ارزیابی مهارت شنیدن در آزمون آیلتس در چهار بخش و دارای ۴۰ سؤال است؛ دو بخش اول شامل موضوعاتی اجتماعی است و مکالمه بین دو یا چند نفر می‌باشد و دو بخش دوم شامل موضوعات دانشگاهی است. یک مکالمه‌ی حداکثر چهارنفره و یک کنفرانس، تشکیل‌دهنده‌ی این بخش‌ها هستند. زمان این آزمون حدود ۳۰ دقیقه است و فایل صوتی فقط یک بار پخش می‌شود و قبل و بعد از پخش فایل صوتی، آزمون‌دهنده زمان دارد تا سؤالات را بررسی کند و در پایان آزمون نیز ۱۰ دقیقه فرصت دارد تا جواب‌ها را از دفترچه‌ی سؤالات به پاسخ‌نامه منتقل نماید. سؤالات به صورت چهارگزینه‌ای و کامل‌کردنی است و در قالب نوشتاری طراحی شده‌اند (همان).

۳،۳،۲. آزمون مهارتی زبان انگلیسی میشیگان^۴

این آزمون توسط مؤسسه‌ی زبان انگلیسی دانشگاه میشیگان برگزار می‌شود و هدف آن سنجش مهارت کلی و توانایی‌های زبانی داوطلبان است. این آزمون بیشتر در برنامه‌ها و کالج‌های زبان در آمریکا و کانادا مورد استفاده قرار می‌گیرد و از سه بخش تشکیل شده است (براون، ۲۰۱۱: ۱۱۹). ارزیابی درک مهارت شنیداری این آزمون از سه بخش تشکیل شده است؛ در بخش اول، آزمون‌دهنده باید بعد از شنیدن سؤال در فایل صوتی، پاسخ سؤال را از میان سه گزینه‌ی موجود انتخاب نماید. در بخش دوم، آزمون‌دهنده بعد از شنیدن یک مکالمه باید گزینه‌ی درست را از بین گزینه‌ها انتخاب نماید و در بخش سوم نیز سه مصاحبه مانند مصاحبه‌های رادیویی وجود دارد که آزمون‌دهنده باید پاسخ درست را از میان گزینه‌ها انتخاب نماید. زمان این آزمون ۳۰ تا ۴۰ دقیقه است و قالب سؤالات نیز به صورت نوشتاری طراحی شده است (همان).

1. International English Language Testing System (=IELTS)

2. academic

3. general

4. Michigan English Language Assessment Battery (=MELAB)

۴.۳.۲. آزمون مهارت پیرسون^۱

مهارت شنیدن در این آزمون همراه با مهارت‌های نوشتن و خواندن ارزیابی می‌شود و از دو بخش تشکیل شده است. آزمون پیرسون به صورت رایانه‌ای برگزار می‌شود و هر فایل صوتی فقط یک بار پخش می‌شود و امکان تنظیم صدا برای آزمون‌دهنده وجود دارد و یادداشت‌برداری در حین آزمون نیز مجاز است. در بخش اول، دو یا سه فایل صوتی شنیده می‌شود و آزمون‌دهنده باید خلاصه‌ای ۵۰ تا ۷۰ کلمه‌ای از متن را بنویسد. مدت زمان این بخش در حدود ۲۰ تا ۳۰ دقیقه است. سؤالات بخش دوم شامل مواردی چون سؤالات چندگزینه‌ای با بیش از یک پاسخ درست، پرکردن جاهای خالی طبق فایل صوتی، انتخاب خلاصه‌ی متن نوشتاری مرتبط با فایل شنیداری، سؤالات چندگزینه‌ای با تنها یک پاسخ درست، انتخاب واژه یا عبارت آخر فایل شنیداری از میان گزینه‌ها، نوشتن جمله‌ی شنیده شده مانند دیکته و مشخص کردن واژه‌های نادرست در حین شنیدن فایل صوتی را در برمی‌گیرد (همان).

بررسی آزمون‌های تافل، آیلتس، میشیگان و پیرسون مشخص می‌کند که قالب این آزمون‌ها به صورت نوشتاری طراحی شده است.

۴. مبانی نظری

در این بخش موضوعاتی از قبیل ساختار آزمون شنیداری، زمان در آموزش مهارت شنیداری و عوامل مؤثر در سهولت و دشواری این نوع آزمون مورد بررسی قرار می‌گیرد که باید برای طراحی و تهیه‌ی آزمون شنیداری به آن توجه نمود.

۱.۴. ساختار آزمون شنیداری

بهتر است آزمون با یک سؤال آسان شروع شود چون باعث کاهش اضطراب آزمون‌دهندگان در آغاز آزمون می‌گردد. ترتیب چینش سؤال‌ها و ترتیب متن شنیداری باید یکسان باشد. همچنین سؤال‌ها باید در طول متن شنیداری پخش شود و توزیع آن تقریباً یکنواخت باشد. توصیه می‌شود که از ۱۵ تا ۲۰ ثانیه‌ی اول متن شنیداری سؤال مطرح نشود تا آزمون‌دهندگان بتوانند خود را با فضای متن تطبیق دهند. نباید تنها به آزمون درک جزئیاتی مانند تاریخ و اعداد اکتفا کرد، بلکه باید مهارت‌های شنیداری مختلف مانند درک موضوع اصلی، استنتاج، نتیجه‌گیری و توانایی پاسخ به سؤالات را هم مورد ارزیابی قرار داد (کومب و همکاران، ۲۰۰۷: ۹۶).

^۱. Pearson English Test

غالباً در آزمون‌های شنیداری از سؤال‌های عینی مانند سؤال‌های چندگزینه‌ای و درست و نادرست استفاده می‌شود؛ زیرا هم مطمئن‌تر است و هم تصحیح و تحلیل کردن آن راحت‌تر است. در استفاده از این نوع سؤال‌ها به جای چهار گزینه بهتر است که سه گزینه ارائه شود، چون حافظه‌ی کوتاه‌مدت نقش مهمی در درک شنیداری دارد و از آن‌جا که امکان گوش دادن چندباره به متن شنیداری وجود ندارد، تعداد گزینه‌های زیاد در آزمون شنیداری به بار حافظه‌ی آنها می‌افزاید و پاسخ دادن به سؤال‌های آزمون را دشوارتر می‌کند. تصاویر هم معمولاً به‌عنوان بخشی از آزمون شنیداری به‌کار می‌روند. در صورت استفاده از تصاویر، باید تصاویر واضح و مشخص باشند (همان: ۹۶).

برای دانشجویانی که می‌خواهند در فضایی ادامه‌ی تحصیلی بدهند که تدریس به زبان مورد نظر آزمون، انجام می‌گیرد، درک و فهم صحبت‌ها و سخنرانی‌های رسمی و غیررسمی بسیار ضروری است؛ بنابراین، در آزمون‌های شنیداری باید از این متون نیز استفاده شود (فرهادی و همکاران، ۲۰۰۷: ۲۲۰).

۲.۴. زمان در آزمون شنیداری

به اعتقاد باک (۲۰۰۱، نقل از کومب، ۲۰۰۷: ۹۷) مدت زمان آزمون شنیداری به دو عامل «طول نوار ضبط شده و دفعات تکرار متن شنیداری» بستگی دارد. آزمون‌های مهارتی تافل و آیلتس، به آزمون‌دهندگان اجازه می‌دهند که تنها یک بار متن شنیداری را بشنوند، در حالی که در آزمون‌های پیشرفت، متن مورد نظر دو بار پخش می‌شود. باک (نقل از همان) پیشنهاد می‌کند که اگر هدف شما ارزیابی ایده‌ی اصلی متن است، بهتر است یک بار پخش شود، اما اگر هدف شما ارزیابی جزئیات است بهتر است متن شنیداری دو بار پخش شود. طبق گفته‌ی کرول (۱۹۷۲؛ نقل از همان) «آزمون‌های شنیدن نباید بیشتر از ۳۰ دقیقه طول بکشند». همچنین، باید زمانی برای پیش‌خوانی سؤال‌ها در اختیار آزمون‌دهندگان قرار گیرد و آنها در خلال شنیدن متن شنیداری به سؤال‌ها پاسخ دهند؛ علاوه بر این، اگر قرار است که پاسخ سؤالات از برگه‌ی سؤال به پاسخ‌نامه منتقل شود، باید مقداری زمان اضافی نیز در نظر گرفته شود (کومب، ۲۰۰۷: ۹۷). در آزمون‌های مهارتی تافل و آیلتس، زمان آزمون شنیداری ۳۰ تا ۳۵ دقیقه است و ۱۰ دقیقه زمان اضافی نیز برای انتقال پاسخ‌ها به پاسخ‌نامه در نظر گرفته می‌شود.

۳.۴. عوامل مؤثر در سهولت و دشواری درک مطلب شنیداری

سرعت صحبت کردن یکی از عواملی است که باعث می‌شود درک مطلب شنیداری برای شنونده سخت یا آسان باشد. سرعت صحبت کردن شرکت‌کنندگان در یک متن شنیداری باید طبیعی باشد. سرعت طبیعی متن شنیداری ۱۶۵-۱۸۰ واژه در دقیقه است هلگسن (۲۰۰۳: ۳۳). وی علاوه بر متغیر سرعت که باعث سختی یا آسانی متن شنیداری می‌شود، مسأله‌ی بار شناختی را هم مطرح می‌کند.

براون (۱۹۹۵؛ نقل از همان) شش عامل را که در افزایش یا کاهش راحتی درک مطلب شنیداری مؤثر هستند به شرح زیر معرفی می‌کند:

- ۱) تعداد صداهای افراد یا اشیا در یک متن: مثلاً تعداد صداهای بیشتر باعث دشواری متن می‌شود.
- ۲) میزان تمایز صداهای افراد یا اشیا از یکدیگر: مثلاً وجود صدای یک زن و یک مرد در مقایسه با صدای دو مرد و یا دو زن موجب آسان‌تر شدن درک متن می‌شود.
- ۳) ارتباطات فضایی^۱ ساده و خاص، آسان‌تر از ارتباطات پیچیده فهمیده می‌شوند: مثلاً یک جمله‌ی امری مانند «مستقیم برو جلو» آسان‌تر از جمله‌ی «همین‌طور به راحت ادامه بده» درک خواهد شد.
- ۴) ترتیب وقایع: وقتی اطلاعات دریافتی به گونه‌ای تنظیم شده باشد که وقایع در آن به ترتیب روی داده‌اند، مطلب راحت‌تر فهمیده می‌شوند.
- ۵) تعداد استنباط‌های مورد نیاز شنونده هر قدر کمتر باشد، متن را راحت‌تر می‌فهمد.
- ۶) اطلاعات و تجربیات پیشین شنونده: اگر شنونده در زمینه‌ی متن شنیداری مطالبی را از پیش خوانده باشد، متن را راحت‌تر درک خواهد کرد.

۴.۴. راه‌های ارتقای کیفی آزمون شنیدن

انتخاب تصاویر در صورت استفاده باید با دقت بسیاری انجام گیرد. تصاویر باید ساده و قابل فهم باشند و برای تمام آزمون‌دهندگان تفسیر یکسانی داشته باشند (فرهادی و همکاران، ۲۰۰۷: ۲۲۳).

صحبت‌های موجود در متن شنیداری باید نمونه‌ای از فعالیت‌های گفتاری در زندگی واقعی باشد (همان). متون شنیداری نباید به نحوی باشد که فهمیدن آنها مستلزم داشتن دانشی خاص (مانند داشتن تحصیلات در یک رشته‌ی دانشگاهی) باشد؛ به عبارت دیگر، متون آزمون باید عمومی باشند.

در بسیاری از آزمون‌های شنیداری به‌ویژه در آزمون‌های شنیداری دانشگاهی، ممکن است آزمون‌دهندگان بخواهند در حین گوش دادن به یادداشت‌برداری نیز بپردازند. نکته‌ی شایان ذکر این است که ساختار و اهداف آزمون تعیین خواهد کرد که چنین امکانی وجود دارد یا خیر. هیل و کورتنی (۱۹۹۴) دریافتند که دادن اجازه‌ی یادداشت‌برداری به آزمون‌دهندگان، نه تنها عملکرد آزمون‌دهندگان را در آزمون شنیداری تافل بهبود نمی‌بخشد بلکه عملکردشان را نیز تضعیف می‌نماید (باک، ۲۰۰۱: ۱۳۸).

بریندلی (۱۹۹۸؛ نقل از باک، ۲۰۰۱: ۱۳۸) اشاره می‌کند که فاصله‌ی بین سؤال‌ها در متن شنیداری نباید به نحوی باشد که آزمون‌دهندگان مجبور باشند در حین پاسخ‌گویی به سؤال بعدی نیز گوش دهند.

^۱. spatial relationships

۵.۴. انواع ارزشیابی‌کنش‌های شنیداری

(۱) شنیدن فشرده:^۱ آزمون شنیدن فشرده در واقع، ارزشیابی‌شناسایی عناصر واجی و ساخت‌واژی زبان است. در این نوع ارزیابی، محرکی گفتاری ارائه می‌گردد و از آزمون‌دهندگان خواسته می‌شود که از میان دو یا چند گزینه، این محرک را تشخیص دهند. این نوع فعالیت‌ها به دلیل نبود بافت مناسب چندان واقعی به نظر نمی‌رسند (براون، ۲۰۱۱: ۱۶۷).

(۲) تشخیص دگرگفت:^۲ سنجش واژگان، عبارت و جملات در مهارت‌های خُرد درک شنیداری اغلب از طریق ارائه‌ی یک جمله‌ی محرک و فراخواندن آزمون‌دهندگان به انتخاب دگرگفتِ درست از میان چندگزینه انجام می‌شود.

(۳) شنیدن واکنشی:^۳ شنیدن رشته‌کلامی نسبتاً کوتاه (احوال‌پرسی، پرسش، دستور، اطمینان از درک و...) به منظور بیان پاسخی کوتاه می‌باشد. یک قالب پرسش و پاسخ می‌تواند تعاملی را در این فعالیت‌های شنیداری به وجود آورد؛ از این‌رو، واکنش آزمون‌دهنده در حقیقت، پاسخی مناسب به یک پرسش است.

(۴) شنیدن گزینشی:^۴ این نوع شنیدن، تحلیل رشته‌گفتمان‌هایی مانند تک‌گویی‌های کوتاه به مدت چند دقیقه را شامل می‌شود که با هدف دستیابی به اطلاعاتی خاص انجام می‌گیرد. هدف چنین کنشی الزاماً جستجوی معناهای کلی و جامع نیست، بلکه کسب توانایی در درک اطلاعات به‌ویژه در بافتی متشکل از رشته‌کلام‌های طولانی‌تر (مانند توضیحات یک آموزگار در کلاس، خبرهای تلویزیونی، رادیویی یا داستان) است. فعالیت‌های ارزشیابی در شنیدن گزینشی از زبان‌آموزان می‌خواهند که به‌عنوان مثال، به اسامی، اعداد، توضیحات و یا واقعیت‌ها و حوادثی خاص دقت کنند.

(۵) تکمیل کردن جملات: نوعی از شنیدن گزینشی است که نیمی از جمله با تأکید ویژه‌ای شنیده می‌شود و آزمون‌دهنده باید از بین گزینه‌ها، ادامه‌ی منطقی جمله را انتخاب کند.

(۶) بندش شنیداری:^۵ فعالیت‌های بندش شنیداری که گاهی دیکته‌های بندش و یا دیکته‌های ناقص نیز نامیده می‌شوند، نوعی ارزیابی شنیدن گزینشی است که از آزمون‌دهنده می‌خواهند که به یک داستان، تک‌گویی یا یک مکالمه گوش کنند و به‌طور هم‌زمان متنی نوشتاری را بخوانند که واژگان یا عبارت‌هایی از آن حذف شده‌اند. این آزمون شامل متنی است که واژگانی به فواصل مشخص (معمولاً هفت واژه در میان) از آن حذف شده‌اند و از آزمون‌دهندگان خواسته می‌شود که با واژه‌های شنیده شده جاهای خالی را پر کنند. یکی از نقاط ضعف روش بندش شنیداری این است که به راحتی به فعالیت‌های درک مطلب

1. intensive listening

2. paraphrase recognition

3. responsive listening

4. selective listening

5. listening cloze

تبدیل می‌شود؛ به عبارت دیگر ممکن است آزمون‌دهنده بدون گوش کردن، موفق به کامل کردن واژگان یا عبارات مناسب شود. برای جلوگیری از چنین مشکلی باید واژه‌هایی برای جاهای خالی در نظر گرفته شود که بار اطلاعاتی زیادی دارند؛ چون در این شیوه، هدف سنجش درک شنیداری است نه حدس زدن دستور یا واژگان.

۷) انتقال اطلاعات: ^۱ می‌توان شنیدن گزینشی را نیز به کمک تکنیک انتقال اطلاعات، ارزشیابی نمود. در این شیوه، اطلاعات شنیداری پردازش شده باید در قالب ارائه‌ی تصویر، نام‌گذاری اجزای یک نمودار، تشخیص یک شی در یک عکس، پر کردن فرم و یا نشان دادن مسیرها در یک نقشه منتقل شوند. فعالیت‌های انتقال اطلاعات می‌توانند با استفاده از نمودار، نقشه، جدول، جدول زمانی و دیگر اطلاعات زندگی روزمره‌ی واقعی‌تر به نظر برسند؛ به‌عنوان مثال می‌توان به شنیدن برنامه‌ی هفتگی و پر کردن جدول اشاره کرد.

۸) تکرار جملات: ^۲ تکرار یک جمله یا شبه‌جمله یا چند جمله به‌عنوان نوعی از ارزشیابی درک شنیداری به‌کار گرفته می‌شود؛ مانند آزمون دیکته. آزمون‌دهندگان باید بتوانند یک رشته‌ی زبانی را برای مدتی در ذهن خود نگاه دارند تا قادر باشند آن را بیان کنند و سپس با تکرار شفاهی محرک، واکنش نشان دهند. درک شنیداری نادرست، چه در سطح واجی و چه در سطح گفتمان را می‌توان در تکرار نادرست مشاهده کرد. در واقع، نادرستی در تکرار مشابه، نادرستی در درک شنیداری است. به‌عنوان مثال، آزمون «فون پاس» ^۳ یک نوع آزمون تجاری است که در ارزشیابی بیان شفاهی و درک شنیداری به شدت بر تکرار جملات تأکید می‌کند. بدیهی است که ارزشیابی با روش تکرار جملات خالی از اشکال نیست، چون این گونه فعالیت‌ها، آزمون مهارت‌های شفاهی هستند و نه صرفاً آزمون‌های شنیداری (باک، ۲۰۰۱: ۷۹). علاوه بر این، ممکن است که این آزمون فقط ارزشیابی تشخیص صداها باشد و همین عدم تکیه بر حافظه‌ی کوتاه‌مدت می‌تواند از ارزش آن بکاهد و اعتبار آن را به‌عنوان شیوه‌ای برای ارزشیابی درک شنیداری زیر سؤال ببرد؛ علاوه بر این معلم ممکن است نتواند خطای درک شنیداری را از خطای بیان شفاهی تشخیص دهد.

۹) شنیدار گسترده: ^۴ هدف از این نوع شنیدن، درک بالا به پایین است. شنیدار گسترده، از یک سو شنیدن سخنرانی‌های طولانی و از سوی دیگر، شنیدن یک مکالمه و استنباط و هدفی جامع را در برمی‌گیرد.

1. information transfer

2. sentence repetition

3. phone pass

4. extensive listening

شنیدن با هدف پی بردن به محتوای کلام و پیام اصلی یا استنباط کردن، همگی از انواع شنیدن گسترده هستند.

۱۰) دیکته: ^۱ یکی از ابزارهای قدیمی ارزشیابی زبان خصوصاً برای درک مهارت شنیداری، استفاده از ابزار دیکته است. در دیکته، آزمون‌دهندگان به متنی که عموماً از ۵۰ تا ۱۰۰ واژه تشکیل شده است و سه بار تکرار می‌شود، گوش می‌دهند؛ این نوع شنیدن با سه نوع سرعت انجام می‌شود؛ نخست با سرعت معمولی، سپس با مکث‌هایی بلند بین عبارت‌ها و واژگان که در آن آزمون‌دهندگان آنچه را که لحظاتی قبل شنیده‌اند می‌نویسند و سرانجام بار دیگر با سرعت طبیعی تا آزمون‌دهندگان نوشته‌هایشان را بررسی و ویرایش کنند. «هیوز (۱۹۸۹)، کوهن (۱۹۹۴)، بیلی (۱۹۹۸) و باک (۲۰۰۱) همگی از منطقی بودن دیکته به‌عنوان آزمونی که لازمه‌ی آن سطح مناسبی از مهارت زبانی به منظور پردازش و نگارش درست تمامی اجزا است، حمایت کرده‌اند. سطح دشواری فعالیت دیکته را می‌توان به کمک طول گروه‌های واژگانی، طول مکث‌ها، سرعت خوانده شدن متن، پیچیدگی گفتمان، دستور و واژگان کنترل کرد. البته قوه‌ی حافظه و پردازش معناشناختی هم در این فعالیت مورد تأکید قرار می‌گیرند» (باک، ۲۰۰۱: ۷۸). علاوه بر این، کاربردپذیری برگزاری آزمون‌های دیکته، میزان مناسب پایایی در یک سیستم نمره‌گذاری منسجم و ارتباط قوی با دیگر توانایی‌های زبانی سبب می‌شوند که از آزمون دیکته در ارزشیابی‌های گسترده برای سنجش درک شنیداری استفاده شود.

۱۱) بازگویی: ^۲ نوع دیگری از فعالیت‌های شنیداری گسترده، بازگویی است. در این فعالیت، آزمون‌دهندگان به یک داستان یا رویداد خبری گوش می‌کنند و فقط آن را به صورت شفاهی و یا کتبی بازگو و یا خلاصه می‌نمایند. آزمون‌دهندگان در حین پاسخ دادن باید پیام، مفهوم اصلی، هدف، دلایل و/یا نتیجه‌گیری را تشخیص دهند تا نشان دهند که به‌طور کامل متن را درک کرده‌اند. نمره‌گذاری تا حدودی طبق معیاری از پیش تعیین شده صورت می‌گیرد؛ زیرا می‌توان عناصر ضروری در بازگویی کردن داستان را مشخص کرد. پایایی این نوع فعالیت نیز ممکن است چندان بالا نباشد؛ علاوه بر این، زمان و زحمتی که صرف خواندن و ارزیابی پاسخ می‌شود، از میزان کاربردپذیری می‌کاهد. روایی، پردازش شناختی، توانایی ارتباطی و واقعی بودن در این فعالیت به خوبی مشهود است (براون، ۲۰۰۴: ۱۳۸-۱۲۲).

1. dictation
2. retelling

۵. روش پژوهش

پژوهش حاضر به منظور ارزیابی درک شنیداری فارسی‌آموزان طراحی شده است. از آن‌جا که در آزمون درک شنیداری، به جای پاسخ‌های چهارگزینه‌ای بهتر است که از سه گزینه استفاده شود، ساختار این آزمون، سه‌گزینه‌ای و قالب آن به دو صورت گفتاری و نوشتاری می‌باشد. در قالب گفتاری، آزمودنی در برگی آزمون فقط پایه‌ی سؤال را به صورت کتبی دارد و گزینه‌های سؤال را از فایل صوتی می‌شنود و گزینه‌ی صحیح را باید در پاسخ‌نامه علامت بزند، اما در قالب نوشتاری، هم پایه‌ی سؤال و هم گزینه‌ها به صورت نوشتاری در برگی آزمون ارائه شده است و مهارت درک خوانداری آزمودنی نقش بسزایی در نوع نوشتاری دارد.

جامعه‌ی آماری این پژوهش را فارسی‌آموزان مرکز آموزش زبان فارسی دانشگاه بین‌المللی امام خمینی (ره) تشکیل می‌دهند؛ تعداد کل این جامعه‌ی آماری ۱۵۰ نفر بوده است. حجم نمونه در این پژوهش، ۳۶ فارسی‌آموز چینی‌زبان و عرب‌زبان می‌باشند که به‌طور تصادفی انتخاب شده‌اند؛ همه‌ی آزمودنی‌ها از لحاظ زبانی در سطح پیشرفته بوده و به منظور تحصیل در دانشگاه‌های ایران در حدود ۶ ماه در این مرکز مشغول به یادگیری زبان فارسی بوده‌اند. این فارسی‌آموزان در رده‌ی سنی ۱۸ تا ۲۵ سال قرار داشته و متشکل از هر دو جنس زن و مرد بوده‌اند.

در این پژوهش، ابتدا به منظور ارائه‌ی مبانی و مفاهیم نظری و جهت مطالعه‌ی پژوهش‌های قبلی از روش توصیفی-تحلیلی استفاده شده است. برای ساخت آزمون و به منظور تعیین روایی محتوایی و صوری آزمون از قضاوت متخصصان حوزه‌ی آموزش زبان استفاده شده است و از روش پیمایشی توافق‌محور^۱ استفاده گردیده است. برای دستیابی به این مهم، فرم نظرسنجی در اختیار مدرسان زبان فارسی قرار گرفت و یکسان بودن سطح آزمون گفتاری و نوشتاری از لحاظ ساختارهای دستوری و واژگانی، مناسب بودن متن‌ها، واضح بودن فایل صوتی، مناسب بودن سرعت گفتار، زمان بین سؤالات و طول هر سؤال مورد بررسی قرار گرفت و به منظور تعیین اعتباربخشی آزمون از روش میدانی استفاده شد.

۱.۵. ابزار پژوهش

الف) آزمون محقق‌ساخته: یکی از ابزارهای این پژوهش است که از ۲۰ سؤال تشکیل شده است؛ ۱۰ سؤال آن با قالب گفتاری و ۱۰ سؤال دیگر با قالب نوشتاری طراحی شده است و تا حد امکان سعی شده است که متون شنیداری و نوع سؤالات از لحاظ ویژگی‌های زبانی، کاربردی و ساختار در هر دو نوع

۱. focal group

قالب یکسان باشند. هر قالب، از سه مکالمه‌ی کوتاه و شش متن کوتاه تشکیل شده است. بخش‌های مورد سنجش در این آزمون شامل درک موضوع اصلی متن، نتیجه‌گیری، درک جزئیات (مکان و زمان)، درک توالی یک رویداد، درک مقایسه کردن و دگرگفت بوده است.

ب) پرسش‌نامه: ابزار دیگر این پژوهش است. به منظور تحلیل کیفی روایی آزمون و تعیین ارتباط سؤالات آن با سازه‌های دانش زبان، ناگزیر از به‌کارگیری روش بازنگری^۱ شدیم. در پایان آزمون، پرسش‌نامه‌ای به آزمودنی‌ها ارائه شد تا دیدگاه آزمودنی نیز درباره‌ی قالب آزمون گفتاری و نوشتاری مورد بررسی قرار گیرد.

پ) نرم افزار اس‌پی‌اس/س به منظور تجزیه و تحلیل داده‌ها

۶. یافته‌های پژوهش

جهت ارزیابی فرضیه‌های این پژوهش، از آزمون فرض «مقایسه‌ی زوج‌ها» استفاده می‌شود. زمانی از این آزمون فرضیه استفاده می‌شود که دو جامعه از هم مستقل نباشند. می‌توان فرض نمود که مشاهده‌های دسته‌ی اول مربوط به جامعه‌ی اول و مشاهده‌های دسته‌ی دوم مربوط به جامعه‌ی دوم هستند که به ترتیب با « X_1, \dots, X_n و « Y_1, \dots, Y_n » نمایش داده می‌شوند. واضح است که این دو جامعه از هم مستقل نمی‌باشند؛ زیرا در این روش از هر فرد جامعه یک زوج داده حاصل شده است. اگر میانگین جامعه‌ی اول « μ_1 » و میانگین جامعه‌ی دوم « μ_2 » در نظر گرفته شوند، به این دلیل که دو جامعه در واقع یک جامعه هستند، می‌توان نوشت « $\mu_1 - \mu_2 = \mu_d$ » که « d » تفاوت دو مشاهده‌ی « $d_i = X_i - Y_i$ » برای « $i=1, \dots, n$ » است. همچنین، در این آزمون فرض می‌شود که دو جامعه دارای توزیع نرمال هستند.

در این پژوهش، هر دانشجو دارای دو نمره‌ی شفاهی و کتبی است. اگر مشاهدات جامعه‌ی اول، نمرات شفاهی (X) و مشاهدات جامعه دوم، نمرات کتبی (Y) در نظر گرفته شود، این دو نمره، یک زوج را به صورت (X, Y) تشکیل می‌دهند؛ یعنی برای هر دانشجو یک زوج داده وجود دارد که از هم مستقل نیستند.

آزمون کولموگوروف-اسمرینوف^۲ جهت بررسی نرمال بودن متغیرهای تحقیق: آزمون کولموگوروف-اسمرینوف به منظور بررسی توزیع داده‌های یک متغیر کمی مورد استفاده قرار می‌گیرد. آزمون فرضیه‌ی آماری مربوط به نرمال بودن توزیع یک متغیر کمی به صورت زیر مطرح می‌شود.

H_0 : داده‌ها دارای توزیع نرمال هستند.

H_1 : داده‌ها دارای توزیع نرمال نیستند.

1. retrospection

2. Kolmogorov-Smirnov Test

اگر داده‌ها دارای توزیع نرمال باشند، می‌توان از آزمون‌های پارامتریک (آزمون تفاوت میانگین دو جامعه غیرمستقل) استفاده نمود اما اگر داده‌ها دارای توزیع نرمال نباشند از آزمون‌های ناپارامتریک (آزمون علامت زوج-نمونه‌ای) استفاده می‌شود. با توجه به این که نمونه‌ی آماری مربوط به این تحقیق شامل ۳۶ دانشجوی می‌باشد، برای بررسی آزمون فرض مربوط به متغیرهای مورد نظر، یعنی نمره‌ی شفاهی و کتبی، باید نرمال بودن توزیع این متغیرها مورد بررسی قرار گیرد.

معنی‌داری که به اختصار آن را با (sig) نشان می‌دهیم، میزان خطایی است که در رد فرضیه‌ی صفر (H_0) مرتکب می‌شویم و آلفا (α) سطح خطایی است که پژوهشگر در تحقیق خود در نظر می‌گیرد و معمولاً مقدار ۵ درصد برای α در نظر گرفته می‌شود.

با توجه به جدول زیر و به این دلیل که سطح معنی‌داری آزمون کولموگروف-اسمیرنوف برای متغیرهای تحقیق یعنی نمره‌ی شفاهی و نمره‌ی کتبی بزرگتر از ۰/۰۵ است، فرض صفر تأیید و ادعای نرمال بودن توزیع این دو متغیر پذیرفته می‌شود.

جدول ۱. نتیجه‌ی آزمون کولموگروف-اسمیرنوف برای متغیرهای تحقیق

وضعیت	سطح معنی‌داری	آماره‌ی آزمون	متغیر تحقیق
نرمال	۰/۲۷۴	۰/۹۹۶	نمره‌ی شفاهی
نرمال	۰/۰۹۸	۱/۲۲۸	نمره‌ی کتبی

حال به بررسی فرضیه‌ی پژوهش می‌پردازیم.

فرضیه‌ی پژوهش: بهتر بودن ارائه‌ی گزینه‌های آزمون به صورت گفتاری نسبت به ارائه‌ی گزینه‌های آزمون به صورت نوشتاری

$$H_0: \mu_d \geq 0 \text{ (میانگین نمرات گفتاری دانشجویان بیشتر از میانگین نمرات نوشتاری آنهاست)}$$

$$H_1: \mu_d < 0 \text{ (میانگین نمرات نوشتاری دانشجویان بیشتر از میانگین نمرات گفتاری آنهاست)}$$

نتیجه‌ی آزمون مقایسه‌ی زوجی برای دو جامعه‌ی غیرمستقل در فرضیه‌ی مورد نظر شامل سه ستاده^۱ است. در ستاده‌ی اول (جدول ۲) آمار توصیفی مربوط به متغیرهای نمره‌ی گفتاری و نمره‌ی نوشتاری آورده شده است که به ترتیب شامل تعداد داده‌ها و شاخص‌های توصیفی این متغیرها (میانگین و انحراف معیار) می‌باشد.

۱. data

جدول ۲. آمار توصیفی متغیرهای نمره‌ی شفاهی و نمره‌ی کتبی

متغیر	تعداد	میانگین	انحراف معیار
نمره‌ی گفتاری	۳۶	۷,۸۹	۱,۶۵۲
نمره‌ی نوشتاری	۳۶	۷,۴۲	۱,۸۸۸

برای بررسی رابطه (همبستگی) دو متغیر می‌توان فرضیه‌های آماری را به صورت زیر تعریف کرد:

H_0 : همبستگی معنی‌داری بین دو متغیر وجود ندارد.

H_1 : همبستگی معنی‌داری بین دو متغیر وجود دارد.

ستاده‌ی دوم (جدول ۳) به ترتیب تعداد داده‌ها، ضریب همبستگی و ضریب معناداری را برای ضریب همبستگی ارائه می‌کند. ضریب همبستگی برای ۳۶ داده‌ی زوجی برابر $0/601$ است. با توجه به این جدول، ضریب معناداری ($0/000$) کمتر از سطح معنی‌داری $0/05$ است، لذا فرض H_0 رد می‌شود. در نتیجه بین این دو متغیر (نمرات گفتاری و نوشتاری) همبستگی معنی‌داری وجود دارد.

جدول ۳. ضریب همبستگی متغیرهای نمره‌ی شفاهی و کتبی

متغیر	تعداد	ضریب همبستگی	ضریب معناداری
نمره گفتاری و نوشتاری	۳۶	$0/601$	$0/000$

ستاده‌ی سوم (جدول ۴) که مربوط به آمار استنباطی است، حاوی نتایج آزمون مقایسه‌ی زوجی می‌باشد و شامل میانگین و انحراف معیار برای تفاوت میانگین نمرات شفاهی و کتبی، آماره‌ی t ، درجه‌ی آزادی و ضریب معناداری است.

جدول ۴. نتایج آزمون مقایسه‌ی زوجی

متغیر	میانگین	انحراف معیار	آماره‌ی t	درجه‌ی آزادی	ضریب معناداری
نوشتاری و گفتاری	$0/472$	$1/594$	$1/777$	۳۵	$0/084$

با استفاده از جدول t و سطح معنی‌داری $\alpha=0/05$ ، $t_{35,0.05} = -1/64$ است. در این جا چون $-1/64 < 1/777$ است، با ۹۵ درصد اطمینان قضاوت می‌کنیم که فرضیه‌ی H_0 پذیرفته می‌شود؛ یعنی میانگین نمرات آزمون گفتاری دانشجویان بیشتر از میانگین نمرات نوشتاری آنها است و در نتیجه، آزمون گفتاری بهتر است.

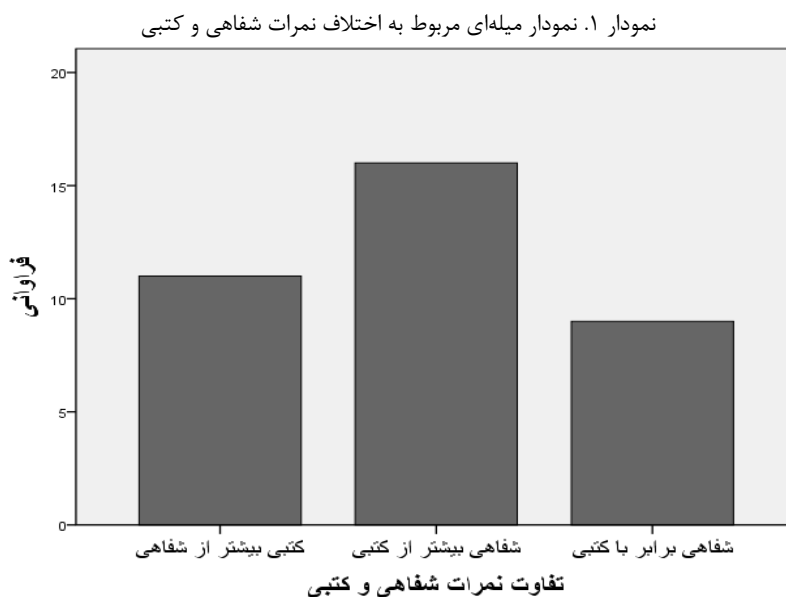
جدول (۵) اختلاف نمره‌ی آزمون گفتاری و نمره آزمون نوشتاری را نشان می‌دهد. با توجه به این جدول، تعداد اختلاف منفی نمرات ۹، اختلاف مثبت نمرات ۱۶ و در نهایت، ۱۱ مورد هم بدون تفاوت می‌باشد؛ این

بدان معنی است که نمره‌ی آزمون گفتاری ۹ نفر از ۳۶ نفر، از نمره‌ی آزمون نوشتاری آنها کمتر است. همچنین ۱۶ نفر از دانشجویان دارای نمره‌ی گفتاری بیشتر از نمره‌ی نوشتاری هستند و نمره‌ی ۱۱ نفر هم در هر دو نوع آزمون برابر است.

جدول ۵. تفاوت نمرات شفاهی و کتبی

تعداد	نمره‌ی گفتاری - نمره‌ی نوشتاری
۹	نمره‌ی نوشتاری < نمره‌ی گفتاری
۱۶	نمره‌ی نوشتاری > نمره‌ی گفتاری
۱۱	نمره‌ی گفتاری = نمره‌ی نوشتاری

نمودار (۱) اختلاف نمرات شفاهی و کتبی را نمایش می‌دهد.



نتایج فرم نظرسنجی فارسی‌آموزان برای ارزیابی مهارت درک شنیداری به شرح زیر است: ۲۱ نفر از ۳۶ آزمودنی، آزمون نوشتاری را بهتر از گفتاری دانسته‌اند در حالی که نتیجه‌ی آزمون نشان داد که در عمل ۱۰ نفر از این گروه نمره‌ی گفتاری بالاتری دارند. گروهی از زبان‌آموزان که نوع نوشتاری را ترجیح می‌دادند دلایل خود را این گونه ذکر کردند:

- در نوع نوشتاری، گزینه‌ها را می‌توان دید و در نتیجه، درک آنها راحت‌تر است.
- در نوع نوشتاری، می‌توان گزینه‌ی درست را حدس زد.
- در نوع نوشتاری، قبل از شنیدن متن شنیداری می‌توان گزینه‌ها را خواند و با موضوع متن آشنا شد.

- در نوع گفتاری، سطح زبانی باید پیشرفته باشد.
 - در نوع گفتاری گزینه‌ها فراموش می‌شوند.
 - در نوع گفتاری مهارت درک شنیدن باید خیلی خوب باشد.
 - همچنین، از نظر ۱۵ آزمودنی، آزمون گفتاری بهتر از آزمون نوشتاری بود و با بررسی نتایج آزمون مشخص شد که از این گروه فقط ۲ نفر نمره‌ی نوشتاری آنان بیشتر بوده است. گروهی از زبان‌آموزان که نوع گفتاری را ترجیح می‌دادند دلایل خود را این گونه ذکر کردند:
 - نوع گفتاری برای تقویت درک شنیداری مفیدتر است.
 - در نوع نوشتاری، مهارت درک نوشتاری هم باید خوب باشد.
 - نوع گفتاری، به دقت بیشتری نیاز دارد و بهتر است.
 - اگر معنی متن کامل درک شود، به صورت گفتاری هم می‌توان به سؤالات جواب داد و نیازی به متن نوشتاری نیست.
 - در نوع گفتاری، تمام تمرکز برای خوب گوش کردن صرف می‌شود.
 - در نوع گفتاری اضطراب کمبود وقت برای خواندن گزینه‌ها وجود ندارد.
 - در زندگی واقعی از نوع گفتاری استفاده می‌شود.
- همان‌طور که از مطالب بالا استنباط می‌شود، نتایج فرم نظرسنجی و دیدگاه فارسی‌آموزان با عملکرد واقعی زبان‌آموز مغایرت دارد.

۷. بحث و نتیجه‌گیری

با توجه به نتایج فرم نظرسنجی مشخص شد که رضایت‌مندی آزمودنی‌ها از آزمون نوشتاری، بیشتر از آزمون گفتاری است؛ زیرا از لحاظ روان‌شناسی، دیدن گزینه‌ها احساس آرامش و تأثیر مثبتی برای آزمودنی به همراه دارد. همچنین، آزمون نوشتاری به آزمودنی این اجازه را می‌دهد که کنترل بیشتری بر روی راهبردهای آزمون داشته باشد؛ به‌عنوان مثال، امکان برگشت برای مرور و/یا تغییر گزینه‌ها در نوع نوشتاری وجود دارد و گزینه‌ها سرنخی را برای درک متن به دست می‌دهند. با این حال، موفقیت در این نوع آزمون مستلزم توانایی درک خوانداری بالایی است. این امر را می‌توان از مطالعات زیادی که در زمینه‌ی آسیب‌شناسی ارزیابی مهارت شنیدن انجام شده، درک کرد. نتایج حاکی از این است که در این نوع آزمون، تطبیق هم‌زمان متن نوشتاری و گفتاری اغلب با کندی صورت می‌گیرد و همین امر در نتایج آزمون تأثیر منفی دارد (خصوصاً اگر واژه‌های موجود در متن و گزینه‌ها از لحاظ سطح و آشنایی یکسان نباشند).

با وجود این که دیدگاه آزمودنی‌ها نسبت به آزمون نوشتاری نسبت به آزمون گفتاری، مثبت‌تر بود، نتایج نشان داد که آنها در آزمون گفتاری بهتر عمل کرده‌اند و نمرات بالاتری به دست آورده‌اند؛ زیرا مطالعات و بررسی‌های بیولوژیکی مغز و حافظه بیانگر این مطلب هستند که در حافظه‌ی کوتاه‌مدت، رمزگذاری آوایی^۱ بر رمزگذاری دیداری^۲ برتری دارد؛ بدین معنا که نگهداری موقت اطلاعات در حافظه‌ی کوتاه‌مدت اساساً به صورت شنیداری است و ارائه‌ی شنیداری نسبت به ارائه‌ی دیداری فضای بیشتری از فراخوانی حافظه‌ی کوتاه‌مدت را به خود اختصاص می‌دهد. زبان‌شناسان هم معتقدند که برتری تأثیر حافظه در مطالب شنیداری نسبت به دیداری، مدیون ویژگی‌های زبرنجیری زبان گفتاری است و به همین دلیل است که در ابتدای یادگیری نیز بیشتر از حس شنیداری استفاده می‌شود (گشمردی، ۱۳۸۴: ۱۴۰).

بنابراین، هم یافته‌های پژوهش با نتایج علمی هم‌سوایی دارد و هم فرضیه‌ی این پژوهش را اثبات می‌کند، اما نکته‌ای که در طراحی آزمون گفتاری باید بدان توجه کرد این است که حافظه‌ی کوتاه‌مدت نقش بسزایی در ارزیابی درک مهارت شنیداری دارد؛ به عبارت دیگر، «فراخوانی حافظه» - یعنی تعداد اقلامی که امکان دارد بلافاصله بعد از ارائه‌ی شنیداری یا دیداری به یاد آورده شود - معمولاً برای ارقام، هفت ماده، برای حروف، شش ماده و برای واژه‌ها، پنج ماده است. البته موارد استثنایی هم وجود دارد. از این‌رو، هنگام رویارویی با یک زبان خارجی، فراخوانی حافظه محدود است و این ضعف نیز از آن ناشی می‌شود که زبان‌آموزان، پایه‌ی شناختی و اطلاعات مفهومی را که بتوانند نگهداری واژه‌های زبان خارجی را در حافظه‌ی کوتاه‌مدت سهولت بخشد، در اختیار ندارند و قادر به انجام فعالیت‌های ذهنی لازم در آن واحد نیستند. در همین زمینه، در طراحی آزمون نیز می‌توان به پدیده‌ی تأخر^۳ و تقدم^۴ اشاره کرد؛ زیرا در به‌یادسپاری و به‌یادآوری وقایع زبانی، مطالب آغازین و پایانی بهتر به یاد می‌مانند و زودتر هم به خاطر آورده می‌شوند (روشن و گلیپور، ۱۳۹۰: ۳۹۲).

در پایان می‌توان نتیجه گرفت که اگر با توجه به اهداف و برنامه‌های آموزشی ناگزیر به طراحی آزمون شنیداری از نوع نوشتاری هستیم، باید تا حد امکان نقش درک خوانداری را کم‌رنگ‌تر کنیم و از عبارات و جملات کوتاه‌تر استفاده نماییم؛ زیرا سطح مهارت خواندن بر نتیجه‌ی آزمون شنیداری تأثیر مستقیمی دارد و بر روی پایایی و روایی آزمون تأثیر می‌گذارد؛ همچنین در طراحی آزمون گفتاری نیز باید نقش و کارکرد حافظه را در نظر گرفت. لازم به ذکر است که متغیرهایی مانند زبان مادری آزمودنی، روش‌های به‌کار گرفته

1. phonological coding

2. visual coding

3. recently effect

4. primary effect

شده در آموزش مهارت شنیداری، آزمودنی، و سطح آزمودنی نیز ممکن است بر یافته‌های این پژوهش تأثیرگذار باشند که از کنترل پژوهشگر خارج بوده و نیازمند پژوهش‌های دیگری است.

منابع:

- براون، د. (۲۰۱۱). *ارزشیابی زبان: اصول و عملکردهای کلاسی*. ترجمه‌ی علیرضا فرحبخش. تهران: رهنما.
- جلیلی، ا. (۱۳۹۰). *آزمون مهارتی فارسی (آمفا) بر پایه‌ی چهار مهارت اصلی زبانی*. رساله‌ی کارشناسی ارشد. تهران: دانشگاه علامه طباطبائی.
- روشن، ب. و گلیپور، ل. (۱۳۹۰). راهکارهای شناختی افزایش حافظه در آموزش زبان فارسی. *مجموعه مقالات نخستین همایش بین‌المللی آموزش زبان فارسی، چشم‌انداز، راهکارها و موانع*، صص: ۴۰۶-۳۸۸.
- شفاق، ن. و همکاران. (۱۳۹۰). *هنجاریابی آزمون بین‌المللی زبان فارسی*. *فصلنامه‌ی پازند*، شماره ۹۱، صص: ۹۹-۸۷.
- ضیاحسینی، م. (۱۳۸۵). *اصول و نظریه‌های آموزش زبان فارسی به غیرفارسی‌زبانان*. تهران: انتشارات سخن.
- قنسولی، ب. (۱۳۸۹). *طراحی و روایی‌سنجی آزمون بسندگی زبان فارسی*. *مجله‌ی پژوهش زبان‌های خارجی*، شماره ۵۷، صص: ۱۲۹-۱۱۵.
- گشمردی، م. (۱۳۸۴). *نقش حافظه در فرایند یادگیری زبان*. *پژوهش ادبیات معاصر جهان*، دوره ۱۰، شماره ۲۸، صص: ۱۵۶-۱۳۳.

- Anna, C. et al.** (2013) Investigating the Effects of Multiple-Choice Listening Test Items in the Oral Versus Written Mode on L2 Listeners Performance and Perception, *SciVerse Science Direct. system*,41: 575-586
- Brown, D. H.** (2004). *Language Assessment Principles and Classroom Practice*. Pearson Education Inc. Longman.
- Buck, G.** (2001). *Assessing Listening*. Cambridge University Press.
- Chang, A. C. S.** (2005). The Perceived Effectiveness of Question Preview in Listening Comprehension Tests by EFL Learners. *New Zealand Study of Applied Linguistics*, 11: 75-99.
- Cheng, H. F.** (2004). A Comparison of Multiple-Choice and Open-Ended Response Formats for the Assessment of Listening Proficiency in English. *Foreign Language Annals*, 37(4): 544-555
- Coombe, C. et al.** (2007). *A Practical Guide to Assessing English Language Learners*. US: The University of Michigan Press
- Eykin, L. B.** (1992). *The Effects of Listening Guides on the Comprehension of Authentic Texts by Novice Learners of French as a Second Language*. Diss, University of South Carolina.
- Farhady, H. et al.** (2007). *Testing Language Skills from Theory to Practice*. Tehran: SAMT.

- Freedle, R. & Kostin, I.** (1999). Does the Text Matter in a Multiple-Choice Test of Comprehension? The Case for the Construct Validity of TOEFL's Minitalks. *Language Testing*, 16(1): 2-32.
- Helgesen, M.** (2003). Listening. In D. Nunan (Ed.), *Practical English Language Teaching*. US: McGraw-Hill
- Teng, H. C.** (1998). The Effects of Text Types and Task Types on English Listening Comprehension [in Chinese]. *English Teaching*, 23(19): 5-18.
- Wu, Y.** (1998). What do Tests of Listening Comprehension Test? A Retrospection Study of EFL Test-Takers Performing a Multiple-Choice Task. *Language Testing*, 15(1): 21-44.
- Yanagawa, K. & Green, A.** (2008). To Show or Not to Show: The Effects of Items Stems and Answer Options on Performance on a Multiple-Choice Comprehension Test. *System*, 36: 107-122.