



تحلیل صوت‌شناختی پارامترهای تأثیرگذار بر گوناگونی‌های بین-گوینده در گویشوران دوزبانه فارسی-انگلیسی^{۱*}

هما اسدی^۲، مارال آسیائی^۳، بتول علی‌نژاد^۴

تاریخ دریافت: ۱۴۰۱/۰۲/۲۲

تاریخ پذیرش: ۱۴۰۱/۰۸/۱۱

چکیده

پژوهش حاضر بر آن است بر مبنای رویکرد صوت‌شناختی و در چارچوب مقایسه قضایی گوینده به بررسی مجموعه‌ای از پارامترهای صوت‌شناختی تأثیرگذار بر گوناگونی‌های بین-گوینده در گویندگان دوزبانه فارسی-انگلیسی بپردازد تا از میان آن‌ها بهترین و مناسب‌ترین پارامترهای صوت‌شناختی فردویژه مشخص شود. به این منظور، استخراج پارامترهای صوت‌شناختی فرکانس پایه و فرکانس سازه‌های اول تا چهارم از طریق استخراج همه‌واکه‌های زنجیره‌های آوایی از پیکره آوایی میراث که شامل صدای افراد دوزبانه فارسی-انگلیسی در دو سبک خوانداری و بداهه است، انجام پذیرفت. سنجش مقادیر فرکانس پایه و فرکانس سازه‌های اول، دوم، سوم و چهارم به شیوه بلندمدت انجام شد. نمونه‌های آوایی با استفاده از برنامه پرات (ویرایش ۶,۲,۰۹) مورد تجزیه و تحلیل صوت‌شناختی قرار گرفت. تحلیل آماری داده‌ها با استفاده از نرم‌افزار آر (ویرایش ۴,۱,۰) انجام گرفت. یافته‌ها نشان داد که تفاوت معناداری میان پارامترهای صوت‌شناختی مورد بررسی در دو زبان فارسی و انگلیسی و نیز دو سبک گفتاری خوانداری و بداهه وجود

^۱ شناسه دیجیتال (DOI): 10.22051/jlr.2022.40224.2174

شناسه دیجیتال (DOR): 20.1001.1.20088833.1402.15.47.5.0

^{*} این پژوهش بخشی از طرح شماره ۹۹۰۲۹۵۸۰ است که در «صندوق حمایت از پژوهشگران و فناوران کشور» تصویب و مورد حمایت مالی قرار گرفته است.

^۲ استادیار گروه زبان‌شناسی، دانشگاه اصفهان، اصفهان، ایران (نویسنده مسئول)؛ h.asadi@fgn.ui.ac.ir

^۳ دکترای تخصصی زبان‌شناسی، دانشگاه الزهراء (س)، تهران، ایران؛ m.asiaee@alzahra.ac.ir

^۴ دانشیار گروه زبان‌شناسی، دانشگاه اصفهان، اصفهان، ایران؛ b.alinezhad@fgn.ui.ac.ir

دارد. با این حال، این پارامترها همچنان در نشان دادن گوناگونی‌های بین-گوینده عملکرد مناسبی داشته‌اند. فرکانس پایه، فرکانس سازه اول و فرکانس سازه سوم در هر دو گروه زنان و مردان بهتر توانسته‌اند گویشوران دوزبانه فارسی-انگلیسی را از یک‌دیگر متمایز کنند. همچنین همبستگی اندک میان فرکانس پایه با فرکانس سازه‌های اول و سوم نمایانگر آن است که این پارامترها اطلاعات متفاوتی در مورد صدای گویندگان منتقل می‌کنند؛ در نتیجه، ترکیب آن‌ها می‌تواند در امر تشخیص هویت گوینده مؤثر باشد.

واژه‌های کلیدی: آواشناسی صوت‌شناختی، تشخیص هویت گویشوران دوزبانه، فرکانس پایه، فرکانس سازه‌ای

۱. مقدمه

زبان در یکی از برون‌ترین حالت‌های خود یعنی گفتار، افزون بر ابزاری برای انتقال پیام‌های زبان‌شناختی، بخشی از هویت هر فرد نیز به شمار می‌رود. اگر افراد گوناگون آواهای یکسانی داشتند، در این صورت، شناسایی و تمایزگذاری میان آن‌ها فقط به وسیله سرنخ‌های دیداری و تنها از طریق گفت‌وگوی چهره‌به‌چهره با این افراد امکان‌پذیر بود. ولی از آن‌جاکه ویژگی‌های جسمانی افراد مختلف با یک‌دیگر متفاوت است، هر فرد ویژگی‌های منحصربه‌فردی^۱ در آوای خود دارد. این ویژگی‌ها افراد را قادر به شناسایی و تشخیص آواهای آشنا و ناآشنا از یک‌دیگر می‌سازد. ولی موج‌های گفتاری دربرگیرنده چه اطلاعاتی هستند که شنونده را قادر به شناسایی فرد گوینده می‌کنند؟ ابرکرومبی (Abercrombie, 1967) بر این باور است که در هر کنش ارتباطی و گفتاری دو نوع اطلاعات وجود دارد: یکی اطلاعات زبان‌شناختی^۲ و دیگری اطلاعات فردویژه^۳ که هر دو نیز به صورت همزمان به مخاطب منتقل می‌شوند. اطلاعات زبان‌شناختی شامل آن دسته از اطلاعاتی است که گویشور می‌کوشد آن‌ها را انتقال دهد (آنچه که گفته می‌شود) و اطلاعات فردویژه در برگیرنده اطلاعات فرازبانی است که ویژگی‌های فردی گویشور از جمله سن، جنسیت و مانند آن را به مخاطب انتقال می‌دهد (چگونه پیامی منتقل می‌شود). پس در موج‌های گفتار سرنخ‌هایی وجود دارد که ذهن به‌وسیله آن‌ها و شناسایی مشخصه‌های موجود در آن موج‌ها، هویت گوینده را شناسایی می‌کند. به دیگر سخن، این سرنخ‌های صوت‌شناختی همان اطلاعات فردویژه پنهان در موج‌های گفتاری‌اند که به شنونده کمک می‌کنند تا هویت فرد گوینده را از روی صدایش بازشناسد. این اطلاعات فردویژه معمولاً از دو منبع سرچشمه می‌گیرند: (۱) منبع

¹ idiosyncratic

² linguistic

³ indexical

طبیعی و ۲) منبع اکتسابی. تفاوت‌های برآمده از منبع طبیعی اساساً با ساختار فیزیکی متفاوت دستگاه گفتار انسان‌ها، در ارتباط است. تغییرات فیزیکی در طول و اندازه دستگاه گفتار منجر به ایجاد تغییر در پارامترهای صوت‌شناختی گفتار می‌گردد. از سوی دیگر، سوای عوامل طبیعی، گویندگان در چگونگی حرکت اندام‌های گویایشان مانند هم عمل نمی‌کنند. گویندگان تحت تأثیر عامل‌های گوناگونی از جمله محیطی که در آن پرورش یافته‌اند، شیوه‌های منحصر به فرد خود را در حرکت‌های تولیدی آواها دارند که این امر نیز خود سبب تغییراتی در پارامترهای صوت‌شناختی گفتار می‌شود. یافتن پارامترهایی که نزدیکترین تعامل میان منبع‌های یادشده را داشته باشند و منجر به نشان دادن گوناگونی‌های بین-گوینده بیشتر و در عین حال گوناگونی‌های درون-گوینده کمتر شود در بازشناسی هویت گوینده بسیار اهمیت دارد.

تشخیص هویت گوینده، به عنوان یکی از مهم‌ترین هدف‌های آواشناسی قضایی^۱، فرایندی است که طی آن تشخیص هویت فردی ناشناخته، از روی ویژگی‌های آوایی وی انجام می‌پذیرد. شناسایی درست افراد از روی آوای آن‌ها، امری است که همه افراد در زندگی روزمره خود انجام می‌دهند و افزون بر اینکه یک توانایی مهم شناختی به شمار می‌آید، مهارتی اجتماعی نیز هست. ولی باید این نکته را نیز در نظر داشت که همیشه شناسایی افراد از روی آوای آن‌ها، حتی اگر فرد شنونده با فرد گوینده آشنایی داشته باشد، امکان‌پذیر نخواهد بود (Foulkes & Barron, 2000). عامل‌های گوناگونی در پیچیده‌تر کردن فرایند تشخیص هویت گوینده نقش دارند. جمله این عامل‌ها می‌توان به شرایط سلامتی، هیجان گوینده، کیفیت گفتار، سبک گفتاری و پیشینه زبانی اشاره کرد. دوزبانگی نیز یکی از عامل‌های مهم و تأثیرگذار در پیدایش تغییرات صوت‌شناختی در گفتار است. گروسجین (Grosjean, 2010) دوزبان‌ها را افرادی معرفی می‌کند که در «زندگی روزمره خود از دو یا چند زبان بهره می‌گیرند» (Grosjean, 2010, p. 4). ولی آیا آوای افراد دوزبان هنگام صحبت کردن به زبان دوم تغییر می‌کند؟

پژوهش‌های پیشین تأثیرگذاری، کارکرد همزمان و تعامل دوسویه نظام‌های زبانی در افراد دوزبان را تأیید کرده‌اند (Grosjean, 1989; Keshavarz & Ingram, 2002; Mok et al., 2015; Paradis, 2001). نظام آوایی گفتار این افراد تحت تأثیر این تعامل قرار می‌گیرد که به دنبال آن ویژگی‌های صوت‌شناختی آواهای گفتار نیز به هنگام تغییر کاربری از یک زبان به زبان دیگر دچار دگرگونی می‌شود. شنونده‌ها نیز نسبت به این تغییرات فیزیکی در موج‌های آوایی گفتار حساس‌اند و برخی از پارامترهای صوت‌شناختی مربوط به این تغییرات، سرخ‌های آوایی

¹ forensic phonetics

مهمی در درک گفتار و تشخیص هویت گوینده به شمار می‌روند. با این وجود، هنوز مشخص نیست که شنونده‌ها به چه جنبه‌هایی از گفتار توجه بیشتری دارند و چه مشخصه‌هایی از گفتار سبب آسان‌سازی تشخیص گویندگان زبان می‌شود. از سوی دیگر، ویژگی‌های صوت‌شناختی موج‌های گفتاری افزون بر اینکه تحت تأثیر مشخصه‌های فردی‌شانند، از جنبه‌ی زبان‌شناختی نیز به دلیل تفاوت نظام‌های واجی بین زبان‌ها تحت تأثیر قرار می‌گیرند. در نتیجه، برخی از پارامترها در یک زبان فردی‌شانند در حالی که ممکن است در زبان دیگر نقش تمایزدهنده میان افراد نداشته باشند (Asadi et al., 2019; Kinoshita, 2001). حال این پرسش پیش می‌آید که آیا پارامترهای صوت‌شناختی در افرادی که به دو زبان صحبت می‌کنند نیز تغییر می‌کنند؟ و آیا پارامترهای صوت‌شناختی گفتار تحت تأثیر ویژگی‌های آوایی زبان اول یا دوم قرار دارند؟ یافتن پارامترهایی که بتوانند بدون توجه به زبان گوینده، مشخصه‌های فردی‌شان را در گویندگان نشان بدهند از اهمیت بسزایی در آواشناسی قضایی برخوردار است. بر این اساس، در پژوهش حاضر بر آنیم گفتار افراد دو زبانه فارسی-انگلیسی را از جنبه‌ی دگرگونی‌های صوت‌شناختی بین-زبانی و بین‌گوینده بررسی کنیم تا از میان آن‌ها بهترین و مناسب‌ترین پارامترهای صوت‌شناختی فردی‌شان را روشن نماییم.

ویژگی‌های صوت‌شناختی گفتار می‌توانند تحت تأثیر سبک گفتاری گوینده نیز تغییر یابند. گویندگان زبان بر اساس نیازها و هدف‌های ارتباطی خود سبک گفتاری مورد نظرشان را بر می‌گزینند. حتی شنوندگان نیز به هنگام گوش دادن به کلام فرد گوینده می‌توانند سبک گفتاری او را شناسایی کنند. معمولاً تفاوت سبک‌ها در پیکر مشخصه‌های نحوی، انتخاب واژگانی، سرعت کلام، درنگ، آهنگ و چگونگی بیان کلام نمودار می‌شود. از جمله سبک‌های گفتاری که در پژوهش‌های آواشناسی قضایی بسیار مورد توجه قرار گرفته‌اند، می‌توان به سبک گفتاری خواننداری و سبک گفتاری بداهه اشاره کرد. در سبک خواننداری، گوینده یک متن از پیش نوشته شده و آماده شده در اختیار دارد و متن را با سرعت معمولی خود می‌خواند. در حالی که در سبک بداهه، متن از پیش آماده شده‌ای وجود ندارد و گوینده در دم و بدون اینکه متنی در اختیار داشته باشد در ارتباط با موضوعی شروع به صحبت کردن می‌کند. به بیانی، در سبک بداهه، گوینده فضای بیشتری در اختیار دارد و احتمال اینکه تفاوت‌های آکوستیکی بیشتری در این سبک گفتاری میان گویندگان دیده شود بیشتر است. پارامترهای آکوستیکی گفتار نیز تحت تأثیر سبک گفتار هستند و در پی تغییر سبک گفتار این پارامترها نیز تغییر می‌کنند. برای نمونه لان (Laan, 1997) در پژوهش خود نشان داد که گفتار خواننداری در مقایسه با گفتار بداهه سرعت

تولید کمتر، کاهش واکه‌ای کمتر، دامنه‌پیریشی^۱ کمتر و تغییرات بالاتری در فرکانس پایه دارد. هولم (Holm, 2003) نیز نشان داد که میانگین فرکانس پایه در گفتار بداهه نسبت به گفتار خوانداری بیشتر است. این تفاوت‌های صوت‌شناختی بر درک شنوندگان نیز تأثیر می‌گذارد تا آن جایی که پژوهش‌های درکی ثابت کرده‌اند که عملکرد شنوندگان در تشخیص هویت گوینده زمانی که در معرض گفتار با سبک‌های متفاوت قرار می‌گیرند کاهش می‌یابد (Afshan et al., 2018; Park et al., 2020). بنابراین از آن‌جا که هیچ دو نمونه گفتاری از یک گویشور یکسان نبوده و تغییر سبک گفتار نیز می‌تواند یکی از عوامل تأثیرگذار در دقت شنوندگان در تشخیص هویت گوینده باشد، از این رو در این پژوهش، افزون بر متغیر دوزبانگی، سبک گفتار نیز به عنوان یکی از عوامل تأثیرگذار در گوناگونی‌های بین-گوینده مورد بررسی قرار خواهد گرفت.

برای بررسی تغییرات صوت‌شناختی در گفتار افراد دوزبانه فارسی-انگلیسی، پارامترهای صوت‌شناختی فرکانس پایه و فرکانس سازه‌های اول تا چهارم در پیکره‌ای آوایی که دربرگیرنده صدای افراد دوزبانه فارسی-انگلیسی در دو سبک خوانداری و بداهه است، مورد تجزیه و تحلیل صوت‌شناختی قرار خواهد گرفت. فرکانس پایه و فرکانس سازه‌های اول تا چهارم در پژوهش‌های آواشناسی قضایی کاربرد فراوانی دارند و به عنوان پارامترهایی کارآمد و مناسب در تشخیص هویت گوینده به شمار می‌روند (Baldwin & French, 1990; Rose & Clermont, 2001; Sambur, 1975). فرکانس سازه‌ها، به مثابه همبسته‌های صوت‌شناختی بازخوانی دستگاه گفتار، بارتاب دهنده اطلاعات فرودیژه بسیاری هستند. از جمله این اطلاعات می‌توان به ویژگی‌های کالبدشناختی همچون طول مجرای صوتی، ویژگی‌های اجتماعی مانند طبقه اجتماعی-اقتصادی فرد و عادت‌های زبانی مانند کامی‌شدگی اشاره نمود (Gold et al., 2013; Nolan & Grigoras, 2005). همچنین، گویندگان از جنبه پخش انرژی در فرکانس پایه گفتارشان با هم تفاوت دارند. فرکانس پایه نیز به مثابه همبسته صوت‌شناختی میزان ارتعاش تارآواها، به عامل‌های گوناگون جسمانی مانند طول و حجم تارآواها بستگی دارد. این پارامتر در عین حال ویژگی‌های شخصیتی فرد را نیز می‌تواند رمزگذاری نماید (Debruyne et al., 2002) که این امر به دلیل تفاوت‌های کالبدشناختی میان دستگاه‌های گفتار افراد و نیز تفاوت‌های آن‌ها در چگونگی حرکت دادن اندام‌های گویایی‌شان است. با توجه به ارتباط پارامترهای صوت‌شناختی فرکانس پایه و فرکانس سازه‌ها با ساختار حنجره و دستگاه گفتار انتظار می‌رود که این پارامترها کمتر تحت تأثیر زبان گوینده باشند و از این روی بهتر بتوانند افراد دوزبانه را بدون در نظر گرفتن زبانی که به

¹ shimmer

آن سخن می‌گویند، از هم متمایز کنند. با توجه به مبحث‌هایی که تاکنون بیان شد پرسش‌های زیر در پژوهش حاضر مطرح می‌شود: نخست اینکه، آیا پارامترهای صوت‌شناختی فرکانس پایه و فرکانس سازه‌های اول تا چهارم در تشخیص هویت گویشوران دو زبانه تأثیر گذارند؟ دوم آنکه، از میان پارامترهای صوت‌شناختی موردبررسی، کدام پارامتر در تشخیص هویت گویشوران دوزبانه مناسب‌تر است؟

اگر چه هدف اصلی پژوهش تمرکز بر تغییراتی است که گویندگان دوزبانه را از هم متمایز می‌سازد ولی با این وجود دستاوردهای برآمده از آن، افزون بر قابلیت به‌کارگیری در روال‌های دادگاهی برای تشخیص هویت گویندگان و نیز سیستم‌های شناسایی خودکار گوینده، در آواشناسی فیزیکی و شنیداری نیز کاربرد دارد.

۲. پیشینه پژوهش

تاکنون پژوهش‌های اندکی، تشخیص هویت افراد دوزبانه را مورد بررسی قرار داده‌اند. یافته‌های این پژوهش‌ها، در گویشوران دوزبانه گوناگون، که در بیشتر موارد، یکی سوی آن زبان انگلیسی است، دارای یافته‌های مشابهی نیستند. از سوی دیگر، جستجوهای نگارندگان در پیوند با این پدیده در زبان فارسی نتیجه‌ای در بر نداشت.

دسته‌ای از پژوهش‌ها با در محوریت قراردادن فرکانس پایه به عنوان یکی از پارامترهای پر کاربرد در نشان‌دادن گوناگونی‌های بین-گوینده به بررسی تفاوت‌های صوت‌شناختی در گفتار افراد دوزبانه پرداختند. آلتبرگ و فراند (Altenberg & Ferrand, 2006) در پژوهش خود به بررسی تغییرات میانگین فرکانس پایه در تک‌زبان‌های انگلیسی با دوزبانه‌های انگلیسی-روسی و انگلیسی-کانتونیز پرداختند. یافته‌های این پژوهش نشان داد که دوزبانه‌های انگلیسی-روسی هنگامی که به زبان روسی صحبت می‌کردند میانگین فرکانس پایه بالاتری نسبت به زبان انگلیسی داشتند، درحالی که چنین تفاوتی برای دوزبانه‌های انگلیسی-کانتونیز دیده نشد. یاروینن و همکاران (Järvinen et al., 2013) میانگین و بازه فرکانس پایه را در فرایند یادگیری زبان خارجی در زبان آموزان انگلیسی و فنلاندی بررسی کردند. نتیجه این پژوهش نشان داد که مقدار فرکانس پایه در گویندگان فنلاندی هنگام صحبت کردن به زبان انگلیسی بالاتر است، درحالی که چنین تفاوتی برای گویندگان انگلیسی زبان به هنگام صحبت کردن به زبان فنلاندی مشاهده نشد. دورین (Dorreen, 2017) با این فرضیه که فرکانس پایه پارامتری زبان‌ویژه است (Gold, 2014; Rose, 2002) به بررسی تغییرات این پارامتر در دو پیکره آوایی همگن و ناهمگن پرداخت. پیکره اول شامل افراد دوزبانه مائوری-انگلیسی بود. درحالی که پیکره دوم با نام

«QuakBox» که پس از زلزله‌های سال ۲۰۱۰ و ۲۰۱۱ در کانتربری^۱ گردآوری شده بود دربرگیرنده روایت‌های افراد از تجربه شخصی خود از زلزله بود. تعداد ۱۷ گوینده دوزبانه که هر یک زبان دوم متفاوتی داشتند از این پیکره برگرفته شد. یافته‌ها نمایانگر آن بود که فرکانس پایه در هر دو پیکره بین دو زبان متفاوت گویندگان تغییر ویژه‌ای نداشته است. دورین (Dorreen, 2017) بر این باور است که احتمالاً ویژگی‌های زیربومی از زبان اول به زبان دوم منتقل می‌شود و بنابراین تفاوتی در مقادیر این پارامتر در گفتار افراد دوزبانه به دست نمی‌آید. با این حال، در پژوهش چنگ (Cheng, 2020) فرکانس پایه در دو زبان انگلیسی و کره‌ای در گفتار افراد دوزبانه انگلیسی-کره‌ای تفاوت معناداری داشته است.

افزون بر آن دسته از پژوهش‌هایی که تغییرات فرکانس پایه را در گفتار افراد دوزبانه بررسی کرده‌اند، دسته دیگر پژوهش‌ها به بررسی دیگر پارامترهای صوت‌شناختی از جمله فرکانس سازه‌ها در گفتار افراد دوزبانه پرداختند. هیرن و همکاران (Heeren et al., 2014) در پژوهش خود به بررسی فرکانس سازه‌های اول تا چهارم به روش بلندمدت در گویندگان دوزبانه هلندی-ترکی پرداختند. بر مبنای یافته‌های این پژوهش، میانگین فرکانس سازه بلندمدت دوم در زبان هلندی نسبت به زبان ترکی بیشتر بود، در حالی که میانگین فرکانس سازه سوم در زبان ترکی نسبت به زبان هلندی بیشتر بود. این بررسی نشان داد که نسبت گوناگونی‌های بین-زبانی در فرکانس سازه‌ها بویژه فرکانس سازه دوم و سوم نسبت به گوناگونی‌های بین-گوینده بسیار کمتر است. از این روی، هیرن و همکاران (Heeren et al., 2014) بر این باورند که فرکانس سازه‌ها می‌تواند پارامتر سودمندی در تشخیص هویت گوینده به‌ویژه افراد دوزبانه باشند. جانسون و همکاران (Johnson et al., 2020) در پژوهش خود مجموعه‌ای از پارامترهای مربوط به منبع و صافی از جمله فرکانس پایه و فرکانس سازه‌ها را در گفتار افراد دوزبانه انگلیسی-کانتونیز بررسی کردند. یافته‌ها نشان داد الگوی تغییرات فرکانس پایه در دو زبان مورد بررسی ثابت بوده است و دیگر پارامترها نیز فقط در شمار اندکی از گویندگان تغییر کرده‌اند. بر مبنای یافته‌های این پژوهش، مشخصه‌های طیفی با وجود تفاوت‌های زنجیری و زبرزنجیری موجود میان دو زبان مورد بررسی یعنی انگلیسی و کانتونیز عملکرد ثابتی در گفتار افراد دوزبانه داشته‌اند. در جدیدترین پژوهش انجام‌شده در این زمینه، لو (Lo, 2021) در چارچوب آواشناسی قضایی به بررسی ویژگی‌های زبان‌ویژه و فردویژه در فرکانس سازه‌های بلندمدت در دوزبانه‌های انگلیسی-فرانسوی پرداخته است. مقایسه‌ها در سه شرایط زبانی انگلیسی-انگلیسی، فرانسوی-فرانسوی و انگلیسی-

¹ Canterbury

فرانسوی انجام پذیرفت. بر پایه یافته‌های به‌دست آمده از این پژوهش، به جز فرکانس سازه اول، دیگر فرکانس سازه‌ها بین دو زبان تفاوت معناداری داشته‌اند که نشان می‌داد مقدار فرکانس سازه دوم، سوم و چهارم در زبان فرانسوی نسبت به زبان انگلیسی بیشتر بوده‌است. در حالی که ویژگی‌های فردویژه در شرایط زبان مشابه همچنان پایدار بود ولی عملکرد پارامترها در شرایط ناهمسانی زبان‌ها کاهش یافت. با این حال، بر مبنای یافته‌های این پژوهش فرکانس سازه بلندمدت همچنان می‌تواند پارامتر مناسبی در نشان دادن گوناگونی‌های بین-گوینده در افراد دوزبانه باشد.

۳. روش‌شناسی پژوهش

در بخش‌های زیر اطلاعات مربوط به شرکت‌کنندگان، چگونگی ضبط داده‌های آوایی، تقطیع داده‌ها و نیز پارامترهای برگزیده شده برای پژوهش حاضر شرح داده می‌شود.

۳.۱. شرکت‌کننده‌ها و داده‌های آوایی

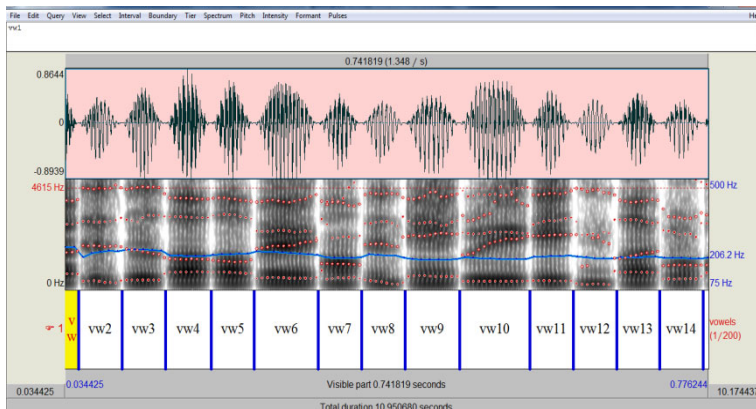
داده‌های پژوهش حاضر از پیکره آوایی میراث (Vaheb et al., 2018a) برگرفته شده‌است. پیکره آوایی میراث به عنوان یکی از بزرگترین پیکره‌های آوایی دوزبانه موجود در ایران با هدف شناسایی گفتار و گوینده گردآوری شده‌است. این پیکره شامل گفتار ۵۰ گویشور دوزبانه فارسی-انگلیسی است که هر یک ۲۵۰ واژه، ۶۳ جمله، ۸۰ عدد و ۱۷ پرسش (پرسش و پاسخ) را به دو زبان فارسی و انگلیسی تلفظ کرده‌اند. پس از ارزیابی‌های اولیه، روی هم رفته، ۲۶ ساعت از داده‌های این پیکره که در برگزیده گفتار ۴۰ گویشور (۲۰ زن و ۲۰ مرد) بود، در این پژوهش به کار گرفته شد. داده‌های آوایی از دو بخش خوانش جمله و پرسش و پاسخ برگزیده شدند. در سبک خوانداری، جمله‌های یکسانی از فایل‌های صوتی گویشوران برگزیده شدند. در گزینش این جمله‌ها، تلاش شد تا جمله‌ها در آغاز و یا پایان فهرست جمله‌هایی که در اختیار گویشوران قرار گرفته بوده‌است، نباشد تا به این ترتیب اثر آهنگ که معمولاً در واژه‌ها و جمله‌های پایانی فهرستی از واژه‌ها و یا جمله‌ها آشکارتر است از بین برود. در سبک بداهه نیز پاسخ گویشوران به پرسش‌هایی مشابه، از کل داده‌های آوایی استخراج شدند. پس از آن مرز گروه‌های آهنگ تولیدشده در هر پاره گفتارها روشن شد؛ استخراج واژه‌ها از فاصله‌های میان گروه‌های آهنگ انجام شد. به این ترتیب از تأثیر احتمالی تولید واج‌ها به صورت نفسی و یا جیرجیری در پایان گروه کاسته شد. همچنین، در بخش پرسش و پاسخ، برای پرهیز از تأثیر گذاری آهنگ جمله‌های پرسشی، فقط زنجیره‌های آوایی مربوط به بخش‌های پاسخ که گوینده در آن به صورت بداهه صحبت می‌کرد، استخراج شده‌است. در پایان، داده‌های به‌دست آمده به دو سبک خوانداری و بداهه دسته‌بندی شدند و وارد مرحله پردازش گفتاری شدند.

۲.۳. شیوه ضبط و تقطیع داده‌ها

در پیکره آوایی میراث، برای ضبط صدای آزمودنی‌ها از میکروفون «Blue» مدل «Yeti» بهره گرفته شده است. این میکروفون به صورت مستقیم به رایانه پیوسته شده بود و فایل‌های صوتی را با فرمت wav با فرکانس نمونه‌برداری ۴۸۰۰۰ هرتز ضبط و ثبت می‌کرد. پس از استخراج داده‌های مورد نظر از دو سبک گفتاری خوانداری و بداهه، داده‌ها وارد مرحله پردازش شدند. داده‌های آوایی با بهره‌گیری از نرم افزار تحلیل گفتار پرات^۱ ویرایش^۱ ۶,۲,۰۹ (Boersma & Weenink, 2022) تقطیع و لایه‌بندی شدند.

۳.۳. پارامترهای صوت‌شناختی و چگونگی استخراج آن‌ها

استخراج پارامترهای صوت‌شناختی با بهره‌گیری از برنامه پردازش گفتار «Vocal Toolkit» که به صورت یک افزونه به نرم‌افزار پرات افزوده می‌شود، انجام پذیرفت (Corretge, 2022). ابتدا واژه‌های موجود در گفتار خوانداری و بداهه با استفاده از زبانه «Extract Vowels» در این افزونه استخراج و در کنار هم قرار داده شدند. در شکل (۱)، نمونه‌ای از طیف‌نگاشت و شبکه متنی مربوط به واژه‌های پیوسته شده را می‌بینید. پس از پیوسته‌سازی واژه‌ها به صورت ردیفی، مقادیر مربوط به فرکانس سازه‌های اول، دوم، سوم و چهارم و نیز فرکانس پایه هر ۵ میلی ثانیه یک بار سنجیده شد. سنجش مقادیر فرکانس سازه‌های بلندمدت و فرکانس پایه بلندمدت با استفاده از برنامه خودکاری که توسط رادک اسکارنیتزل^۲ نوشته شده است انجام گرفت.



شکل ۱: نمونه‌ای از واژه‌های متوالی استخراج شده از پیکره آوایی میراث

^۱ Praat

^۲ Radek Skarnitzl

۳. ۴. تحلیل آماری

تحلیل توصیفی و استنباطی داده‌های آوایی با استفاده از نرم‌افزار آماری آر ویرایش ۴,۱,۰ انجام شد (R Core Team, 2021). در مرحله نخست، با بهره‌گیری از آمار توصیفی مقادیر هر یک از پارامترهای صوت‌شناختی به تفکیک زبان و سبک گفتاری مورد نظر بررسی شد. سپس، برای بررسی معناداری تفاوت‌های بین-زبانی و بین-سبکی از مدل خطی اثر مخلوط^۱ بهره‌گرفته شد. گوینده به‌عنوان اثر ثابت^۲ و زبان، سبک و تعامل میان زبان و سبک (زبان*سبک) به‌عنوان اثر تصادفی^۳ وارد این مدل شدند. در گام پسین به منظور بررسی اهمیت پارامترها و نقش آن‌ها در دسته‌بندی متغیر گوینده از مدل جنگل‌های تصادفی^۴ بهره‌گرفتیم. جنگل‌های تصادفی یک مدل دسته‌بندی است که بر اساس آن مجموعه‌ای از درخت‌های تصمیم^۵ ایجاد می‌شود که هر یک دربرگیرنده شمار بسیاری از درخت‌های تصمیم با زیرمجموعه‌های گوناگون‌اند. هر درخت جداگانه در جنگل‌های تصادفی نشان‌دهنده یک طبقه پیش‌بین است و در نهایت طبقه‌ای که بیشترین آرا را دارد، به‌عنوان پیش‌بین مدل قرار می‌گیرد. در گام پایانی نیز برای سنجش بهترین ترکیب پارامترها و پاسخ به اینکه چه ترکیبی از پارامترها می‌توانند گوناگونی‌های بین-گوینده را بهتر نشان دهند، از آزمون ضریب همبستگی پیرسون^۶ بهره‌گرفتیم. پارامترهایی که کمترین میزان همبستگی را با یک‌دیگر داشته باشند، قابلیت ترکیب شدن را دارا هستند و می‌توان از ترکیب آن‌ها در تشخیص هویت گوینده بهره‌برد.

۴. گزارش یافته‌ها

در بخش‌های زیر و با توجه به پرسش‌های پژوهش در بخش مقدمه، به اجرای آزمون‌های آماری مرتبط پرداخته و یافته‌های مربوط به تجزیه و تحلیل داده‌ها گزارش خواهد شد. نخست، یافته‌های آمار توصیفی مربوط به مقایسه پارامترهای صوت‌شناختی فرکانس پایه و فرکانس سازه‌ها در گفتار گویشوران دوزبانه فارسی-انگلیسی ارائه خواهد شد. سپس، با بهره‌گیری از آمار استنباطی به بررسی تفاوت‌های معنی‌دار در سطح پارامترهای مورد بررسی با توجه به متغیرهای زبان و سبک گفتاری می‌پردازیم.

¹ Linear mixed effects model

² fixed factor

³ random factor

⁴ random forest

⁵ decision trees

⁶ Pearson correlation coefficient

۴. ۱. بررسی توصیفی پارامترهای صوت‌شناختی فرکانس پایه و فرکانس سازه‌های اول تا چهارم در گفتار گویشوران دوزبانه فارسی-انگلیسی

در این بخش با بهره‌گیری از آمار توصیفی پارامترهای صوت‌شناختی فرکانس پایه و فرکانس سازه‌های اول تا چهارم را در دو زبان فارسی و انگلیسی و نیز در سبک‌های خوانداری و بداهه ارائه خواهیم کرد. جدول (۱) و (۲)، میزان میانگین و انحراف معیار را برای هر یک از پارامترهای صوت‌شناختی مورد نظر در هر دو زبان و سبک گفتاری مورد بررسی در دو گروه مردان و زنان نشان می‌دهد.

جدول ۱: میانگین و انحراف معیار فرکانس پایه و فرکانس سازه‌های اول تا چهارم در زبان‌های فارسی و انگلیسی به تفکیک سبک بداهه و خوانداری در گویشوران مرد

پارامتر	فارسی		انگلیسی	
	بداهه	خوانداری	بداهه	خوانداری
F0	۵۰/۱۲۵ (۳۲/۷۵)	۱۹/۱۱۹ (۲۰/۲۴)	۵۱/۱۲۷ (۶۱/۳۵)	۵۸/۱۲۵ (۱۴/۲۳)
F1	۵۹/۴۴۱ (۳۷/۱۳۰)	۸۱/۴۴۶ (۲۱/۱۲۳)	۷۵/۴۱۹ (۱۳/۱۳۲)	۴۳/۴۰۴ (۲۶/۱۲۶)
F2	۱۴۶۵ (۶۷/۳۳۰)	۷۶/۱۴۳۶ (۶۶/۳۰۸)	۶۸/۱۴۶۵ (۰۴/۳۸۳)	۴۱/۱۴۷۸ (۳۷/۳۹۸)
F3	۶۹/۲۵۷۶ (۸۲/۳۲۸)	۵۸/۲۵۷۴ (۸۸/۲۹۸)	۳۲/۲۵۷۳ (۹۰/۳۳۷)	۴۴/۲۵۶۴ (۷۸/۳۴۶)
F4	۸۵/۳۶۰۴ (۰۳/۳۵۶)	۷۳/۳۵۶۱ (۶۴/۳۳۹)	۴۴/۳۵۹۶ (۷۳/۳۷۰)	۴۶/۳۵۷۴ (۷۴/۳۵۱)

بر مبنای مقادیر موجود در جدول (۱)، مقایسه بین-زبانی میانگین‌های مربوط به هر یک از پارامترهای صوت‌شناختی مورد بررسی نشان می‌دهد که در گروه مردان و در سبک خوانداری مقادیر فرکانس سازه‌های اول و سوم در زبان فارسی نسبت به زبان انگلیسی بیشتر است. این در حالی است که مقدار فرکانس پایه، فرکانس سازه‌های دوم و چهارم در زبان فارسی نسبت به زبان انگلیسی کمتر بوده است. همچنین در سبک گفتاری بداهه، مقدار فرکانس سازه‌های اول، سوم و چهارم در زبان فارسی در مقایسه با زبان انگلیسی بیشتر بوده است. فرکانس سازه دوم تغییری در میان دو زبان گفتاری نداشته است. ولی مقدار فرکانس پایه در زبان فارسی نسبت به زبان انگلیسی کمتر بوده است.

جدول ۲: میانگین و انحراف معیار فرکانس پایه و فرکانس سازه‌های اول تا چهارم در زبان‌های فارسی و انگلیسی به تفکیک سبک بداهه و خوانداری در گویشوران زن

انگلیسی		فارسی		پارامتر
خوانداری	بداهه	خوانداری	بداهه	
۹۲/۲۲۷ (۱۱/۳۶)	۵۸/۲۲۴ (۸۱/۴۰)	۸۸/۲۱۴ (۴۱/۳۹)	۹۴/۲۱۸ (۴۱/۳۹)	F0
۲۹/۵۱۲ (۲۰/۲۲۳)	۷۵/۵۱۹ (۶۵/۲۲۸)	۵۱/۶۱۶ (۳۳/۲۴۴)	۱۶/۵۷۴ (۵۳/۲۴۱)	F1
۰۷/۱۷۰۳ (۸۰/۵۲۰)	۴۷/۱۶۶۷ (۶۲/۴۴۸)	۶۸/۱۶۵۵ (۴۱۹/۸۷)	۵۷/۱۶۸۴ (۸۹/۴۳۵)	F2
۹۷/۲۸۴۳ (۸۴/۳۶۹)	۱۵/۲۸۲۹ (۹۸/۳۸۷)	۳۸/۲۸۵۲ (۵۱/۳۴۸)	۵۴/۲۸۴۹ (۸۷/۳۵۸)	F3
۹۴/۳۹۸۲ (۸۳/۳۹۸)	۴۰/۳۹۵۴ (۹۳/۳۷۳)	۵۰/۳۹۸۶ (۵۱/۳۸۹)	۰۷/۳۹۸۳ (۳۲/۳۷۱)	F4

بر مبنای داده‌های موجود در جدول (۲)، در گروه زنان و در سبک خوانداری مقدار فرکانس سازه اول، فرکانس سازه سوم و فرکانس سازه چهارم در زبان فارسی نسبت به زبان انگلیسی بیشتر است. فرکانس پایه و فرکانس سازه دوم در زبان فارسی در مقایسه با زبان انگلیسی کمتر بوده است. در سبک گفتاری بداهه گروه زن‌ها، مقادیر فرکانس سازه‌های اول تا چهارم در زبان فارسی نسبت به زبان انگلیسی بیشتر بوده است. فقط مقدار فرکانس پایه در زبان انگلیسی نسبت به زبان فارسی بیشتر گزارش شده است.

همان‌گونه که داده‌های به‌دست آمده از آمار توصیفی در جدول‌های (۱) و (۲) نشان می‌دهد مقادیر فرکانس پایه و فرکانس سازه‌ها هم تفاوت‌های بین-زبانی و هم تفاوت‌های بین-سبکی داشته‌اند. در ادامه، به بررسی معناداری این تفاوت‌ها با جداسازی هر پارامتر در زبان و سبک‌های گفتاری مورد بررسی می‌پردازیم.

۴.۲. گوناگونی‌های بین-گوینده در گفتار گویشوران دوزبانه فارسی-انگلیسی

بررسی تمایز میان پارامترهای مورد سنجش در این پژوهش در میان گویشوران با بهره‌گیری از مدل خطی اثر مخلوط و در نرم‌افزار آماری آر و بسته نرم‌افزاری «lme4» انجام پذیرفت. در این بررسی گوینده به عنوان اثر ثابت و زبان، سبک و تعامل میان زبان و سبک (زبان*سبک) به عنوان اثر تصادفی وارد مدل شدند. اثر ثابت و اثر تصادفی با مقایسه میان مدل کامل^۱ و مدل کاهش‌یافته^۲ مورد واکاوی قرار گرفتند. در مدل کامل، اثر عامل مورد سنجش حضور دارد؛ این در حالی است که این اثر در مدل کاهش‌یافته از مدل خارج می‌شود. ارزش پی^۳ نیز با استفاده از آزمون آماری

^۱ full model

^۲ reduced model

^۳ P Value

آنوا^۱ و با مقایسه میان مدل کامل و مدل کاهش یافته اندازه گیری می گردد.

کد آر به کاررفته در این پژوهش به شرح زیر است:

1. Model-full (lmer (dependent-variable ~ fixed-factor + (1|random-factor), data=data, REML = FALSE)

Model-reduced (lmer (dependent-variable ~1 + (1|random-factor), data=data, REML = FALSE)

anova(Model-full, Model-reduced, test="Chisq")

جدول (۳)، یافته‌های مربوط به اجرای مدل خطی اثر مخلوط در گویشوران زن و مرد را به

تفکیک تمامی پارامترهای مورد بررسی در این پژوهش نشان می‌دهد.

جدول ۳: یافته‌های مرتبط با مدل خطی اثر مخلوط در پارامترهای فرکانس پایه و

فرکانس‌های اول تا چهارم در گویشوران زن و مرد

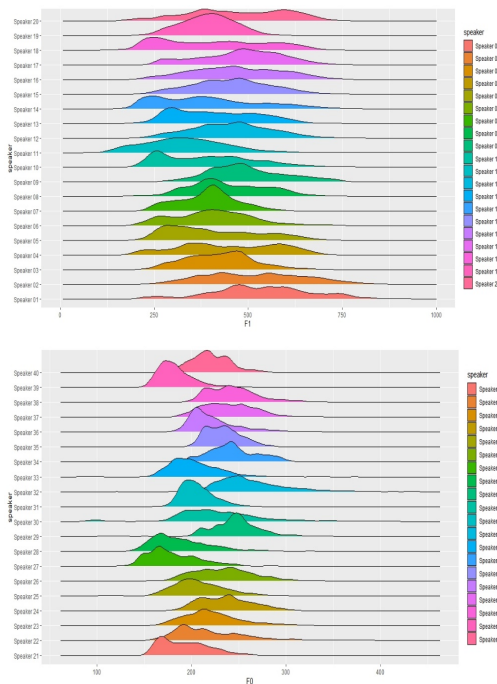
پارامتر	اثر ثابت	اثر تصادفی	نتایج در گروه مردان	نتایج در گروه زنان
F0	گوینده	زبان	$\chi^2(19) < 0.001, p < 0.0001$	$\chi^2(19) < 0.001, p < 0.0001$
		سبک	$\chi^2(19) < 0.001, p < 0.0001$	$\chi^2(19) < 0.001, p < 0.0001$
		زبان*سبک	$\chi^2(19) < 0.001, p < 0.0001$	$\chi^2(19) < 0.001, p < 0.0001$
F1	گوینده	زبان	$\chi^2(19) < 0.001, p < 0.0001$	$\chi^2(19) < 0.001, p < 0.0001$
		سبک	$\chi^2(19) < 0.001, p < 0.0001$	$\chi^2(19) < 0.001, p < 0.0001$
		زبان*سبک	$\chi^2(19) < 0.001, p < 0.0001$	$\chi^2(19) < 0.001, p < 0.0001$
F2	گوینده	زبان	$\chi^2(19) < 0.001, p < 0.0001$	$\chi^2(19) < 0.001, p < 0.0001$
		سبک	$\chi^2(19) < 0.001, p < 0.0001$	$\chi^2(19) < 0.001, p < 0.0001$
		زبان*سبک	$\chi^2(19) < 0.001, p < 0.0001$	$\chi^2(19) < 0.001, p < 0.0001$
F3	گوینده	زبان	$\chi^2(19) < 0.001, p < 0.0001$	$\chi^2(19) < 0.001, p < 0.0001$
		سبک	$\chi^2(19) < 0.001, p < 0.0001$	$\chi^2(19) < 0.001, p < 0.0001$
		زبان*سبک	$\chi^2(19) < 0.001, p < 0.0001$	$\chi^2(19) < 0.001, p < 0.0001$
F4	گوینده	زبان	$\chi^2(19) < 0.001, p < 0.0001$	$\chi^2(19) < 0.001, p < 0.0001$
		سبک	$\chi^2(19) < 0.001, p < 0.0001$	$\chi^2(19) < 0.001, p < 0.0001$
		زبان*سبک	$\chi^2(19) < 0.001, p < 0.0001$	$\chi^2(19) < 0.001, p < 0.0001$

بر مبنای یافته‌های موجود در جدول (۳)، در هر دو گروه زن‌ها و مردها تمایز میان گویشوران در زبان، سبک و تعامل میان سبک و زبان و با در نظر گرفتن پارامترهای فرکانس پایه، فرکانس سازه‌های اول تا چهارم معنادار بوده‌است.

¹ ANOVA

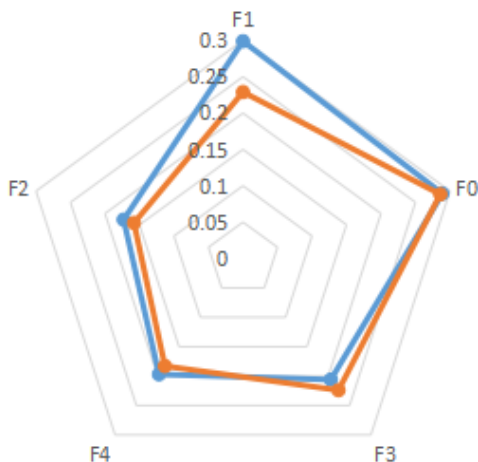
۳.۴. بررسی اهمیت پارامترهای مورد بررسی در دسته‌بندی گوینده

برای بررسی اینکه کدام یک از پارامترهای صوت‌شناختی مورد بررسی از اهمیت بالاتری در دسته‌بندی گوینده برخوردارند، از مدل جنگل‌های تصادفی بهره گرفتیم. در این مدل، گوینده به‌عنوان متغیر وابسته و پارامترهای صوت‌شناختی مورد بررسی به‌عنوان متغیرهای پیش بین وارد مدل جنگل‌های تصادفی شدند. بر مبنای یافته‌های به‌دست آمده برای گروه مردان، فرکانس سازه اول و فرکانس پایه به ترتیب با ۲۹ و ۲۸ درصد بیشترین اهمیت را در دسته‌بندی گوینده داشتند. پس از آن نیز پارامترهای صوت‌شناختی فرکانس سازه سوم با ۲۰ درصد، فرکانس سازه چهارم با ۱۹ درصد و فرکانس سازه دوم با ۱۷ درصد بیشترین تأثیر را در پیش‌بینی دسته‌بندی گوینده داشته‌اند. در گروه زن‌ها نیز فرکانس پایه و فرکانس سازه اول به ترتیب با ۲۸ و ۲۳ درصد بیشترین اهمیت را در دسته‌بندی گویندگان نشان دادند. پس از آن نیز پارامترهای صوت‌شناختی فرکانس سازه سوم با ۲۲ درصد، فرکانس سازه چهارم با ۱۸ درصد و فرکانس سازه دوم با ۱۵ درصد بیشترین نقش را در دسته‌بندی گوینده بازی کرده‌اند. شکل (۱)، گوناگونی فرکانس سازه اول و فرکانس پایه را که به ترتیب در گروه مردها و زن‌ها بیشترین اهمیت را در دسته‌بندی گویندگان داشتند، نشان می‌دهد.



شکل ۲: نمودار «ridgeline» نشان‌دهنده گوناگونی بین-گوینده پارامتر فرکانس سازه اول در گویشوران مرد (بالا) و فرکانس پایه در گویشوران زن (پایین)

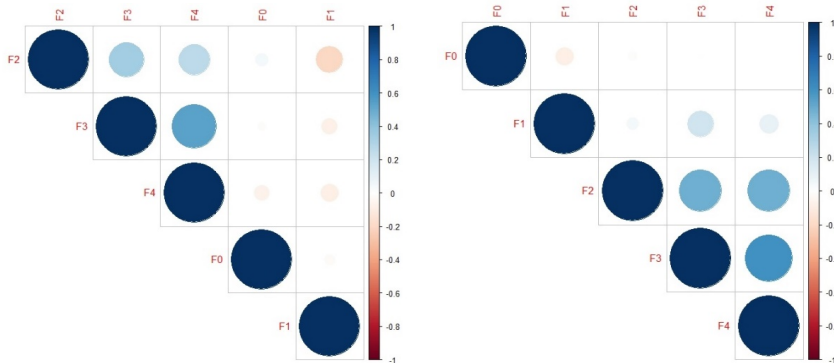
شکل (۳)، اهمیت نسبی هر یک از پارامترها را در دسته‌بندی گویندگان در دو گروه زنان و مردان نشان می‌دهد.



شکل ۳: نمودار رادار نشان‌دهنده اهمیت نسبی پارامترهای صوت‌شناختی فرکانس پایه و فرکانس سازه‌های اول تا چهارم در گروه مردان (رنگ آبی) و گروه زنان (رنگ نارنجی)

۴.۴. ترکیب بهینه پارامترها در نشان دادن گوناگونی‌های بین-گوینده

در بخش‌های پیشین، با توجه به یافته‌های به‌دست آمده از آزمون‌های آماری معلوم شد که هم در گروه مردها و هم در گروه زن‌ها، فرکانس پایه و فرکانس سازه‌های اول تا چهارم میان دو زبان و دو سبک گفتاری مورد بررسی و نیز به هنگام تعامل میان زبان و سبک تفاوت معنادار داشته‌اند. اکنون، برای بررسی اینکه چه ترکیبی از این پارامترها می‌تواند در تشخیص هویت گوینده مؤثرتر واقع شود، همبستگی میان پارامترهای مدنظر را از طریق ضریب همبستگی پیرسون سنجیدیم. یافته‌ها نشان داد که فرکانس پایه همبستگی ضعیفی با فرکانس سازه‌های اول تا چهارم در هر دو گروه زنان و مردان دارد. بنابراین ترکیب فرکانس پایه به عنوان یکی از پارامترهای منبع با فرکانس سازه‌ها به عنوان پارامترهای صافی می‌تواند در نشان دادن گوناگونی‌های بین-گوینده مناسب باشد. شکل (۴)، همبستگی میان پارامترهای گوناگون را نشان می‌دهد. در این نمودار، همبستگی مثبت با رنگ آبی و همبستگی منفی با رنگ قرمز نشان داده می‌شود. تیرگی و یا روشنی رنگ و اندازه دایره با ضریب همبستگی ارتباط مستقیم دارد. هر چه رنگ پررنگ‌تر و دایره بزرگ‌تر باشد، ضریب همبستگی میان پارامترها قوی‌تر است. داده‌های موجود در سمت راست نمودار، ضرایب همبستگی به همراه رنگ‌های متناظر با آن را نشان می‌دهد.



شکل ۴: نمودار دایره‌ای نشان‌دهنده ضریب همبستگی میان پارامترهای مورد بررسی در گویشوران زن (سمت راست) و مرد (سمت چپ)

۵. بحث و بررسی

در این پژوهش به بررسی گوناگونی‌های بین-گوینده در گفتار گویندگان دوزبانه فارسی-انگلیسی پرداختیم. به این منظور، صدای ۲۰ گوینده دوزبانه مرد و ۲۰ گوینده دوزبانه زن را که در دو سبک گفتاری خوانداری و بداهه صحبت می‌کردند، مورد بررسی و تحلیل صوت‌شناختی قرار دادیم. بر مبنای یافته‌های به‌دست آمده، پارامترهای صوت‌شناختی مورد بررسی یعنی فرکانس پایه و فرکانس سازه‌های اول تا چهارم بین دو زبان فارسی و انگلیسی و سبک گفتاری خوانداری و بداهه و نیز در تعامل میان زبان*سبک معنادار بوده اند ($p < 0.001/0$). با این حال، بر اساس مدل جنگل‌های تصادفی آشکار شد که همه پارامترها از اهمیت یکسان در نشان دادن گوناگونی‌های بین-گوینده برخوردار نیستند. از میان پارامترهای مورد بررسی فرکانس سازه اول و فرکانس پایه در گروه مردان و در گروه زنان نیز فرکانس پایه و فرکانس سازه اول به ترتیب از بیشترین اهمیت در طبقه‌بندی گویندگان برخوردار بوده‌اند. پس از این پارامترها، فرکانس سازه سوم در هر دو گروه بیشترین نقش را در ایجاد تمایز و طبقه‌بندی گویندگان بر عهده داشت.

بر مبنای یافته‌های این پژوهش، فرکانس پایه به عنوان یکی از پارامترهایی گزارش شد که از اهمیت بالایی در نشان دادن گوناگونی‌های بین-گوینده و نیز دسته‌بندی گویندگان برخوردار بود. همچنین تحلیل بین-زبانی نشان داد که فرکانس پایه در دو زبان فارسی و انگلیسی با هم تفاوت داشته‌است. اگرچه در پاره‌ای از پژوهش‌های پیشین مانند دورین (Dorreen, 2017) و جانسون و همکاران (Johnson et al., 2020) تفاوتی در الگوی تغییرات فرکانس پایه میان دو زبان گفتاری گویندگان مشاهده نشده بود، با این حال در پژوهش حاضر تفاوت معناداری میان دو زبان

گفتاری گویندگان یعنی فارسی و انگلیسی و نیز دو سبک گفتاری آن‌ها مشاهده شد که این نتیجه هم‌راستا با دستاوردهای پژوهش چنگ (Cheng, 2020) است. با این وجود، به نظر می‌رسد که فرکانس پایه اگرچه تحت تأثیر تغییر زبان و سبک قرار می‌گیرد ولی همچنان ویژگی‌های فردویژه خود را نگه می‌دارد و این توانایی را دارد که گویندگان دوزبانه را در دو سبک گفتاری از هم متمایز و دسته‌بندی نماید.

دیگر پارامتری که هم در گروه زنان و هم در گروه مردان دوزبانه توانست گوناگونی‌های بین-گوینده را به خوبی نشان دهد، فرکانس سازه اول بود. فرکانس سازه اول پارامتری است که با میزان باز بودن دهان در ارتباط است. هر چه مجرای دهان در تولید یک واژه بیشتر باشد، فرکانس سازه اول آن بیشتر است و برعکس (Ladefoged, 2006) با توجه به تفاوت معنادار این پارامتر در گویندگان دوزبانه فارسی-انگلیسی به نظر می‌رسد که گویندگان رفتار متفاوتی در حرکات تولیدی مرتبط با دهان داشته‌اند. این رفتار فردویژه، مشخصاً در میزان باز و بسته کردن دهان نمود داشته‌است که بازتاب آن در مشخصه صوت‌شناختی فرکانس سازه اول پدیدار شده‌است. بیشتر در پژوهش‌هایی که بر روی افراد تک‌زبانه انجام شده‌است، فرکانس سازه‌های بالاتر یعنی فرکانس سازه سوم و چهارم نشانگر گوناگونی‌های بین-گوینده بوده‌اند (Asadi et al., 2018; Gold et al., 2013; McDougall, 2004; Moos, 2010). فرکانس سازه‌های اول و دوم به دلیل ارتباطی که با جایگاه زبان دارند، بیشتر عهده‌دار رمزگذاری محتوای آوایی در نظر گرفته می‌شوند (Ladefoged, 2006). با این حال، در پژوهش‌های اخیر نشان داده‌است که فرکانس سازه اول به‌ویژه زمانی که از قله هجا به سمت پایین حرکت می‌کند و در مسیر کاهشی قرار دارد اطلاعات فردویژه بیشتری در خود حمل می‌کند (He et al., 2019). به بیانی، هنگامی که دهان در حال بسته شدن است، گوناگونی‌های بین-گوینده بیشتر بوده‌است. در پژوهش حاضر، با توجه به اینکه گویندگان به دو زبان متفاوت صحبت می‌کرده‌اند، موضوع انتقال محتوای آوایی نیز بی‌گمان برای گویندگان اهمیت می‌یابد. این احتمال وجود دارد که گویشوران در رمزگذاری زبانی به هنگام تلفظ واژه‌ها متفاوت عمل کرده‌اند که این تفاوت بیشتر در میزان باز و بسته کردن دهان بوده‌است و در نتیجه با تغییر در فرکانس سازه اول همراه شده‌است.

دیگر پارامتری که در نشان دادن گوناگونی‌های بین-گوینده و نیز دسته‌بندی گویندگان دوزبانه عملکرد مناسبی داشته‌است، فرکانس سازه سوم بود. فرکانس سازه سوم پارامتری است که با گردشگی لب‌ها ارتباط دارد؛ به گونه‌ای که گردشگی لب‌ها سبب پایین آمدن این فرکانس می‌شود. بر اساس نتایج، فرکانس سازه سوم در هر دو گروه زنان و مردان دوزبانه در دو سبک

گفتاری خوانداری و بداهه در زبان فارسی نسبت به زبان انگلیسی بیشتر بوده‌است. این امر نشان می‌دهد که گویندگان به هنگام صحبت کردن به زبان انگلیسی بیشتر از گردش‌گی لب‌ها بهره گرفته‌اند. همچنین تفاوت معنادار این پارامتر و عملکرد بهینه آن در بازنمایی گوناگونی‌های بین-گوینده نشان می‌دهد که گویندگان دوزبانۀ این پژوهش در میزان گردش‌گی لب‌ها با هم تفاوت داشته‌اند که این امر سبب شده تا مقدار فرکانس سازه سوم دستخوش این دگرگونی شود و در میان گویندگان متفاوت باشد. این یافته همسو با نتیجه پژوهش آسیائی و همکاران (Asiaee et al., 2019) است که در آن فرکانس سازه‌های اول و سوم طولانی‌مدت به عنوان پارامترهای مناسب در نشان دادن گوناگونی‌های بین-گوینده در گویشوران دو زبانه فارسی-عربی معرفی شده‌اند.

یکی از منبع‌های تغییرات آکوستیکی درون-گوینده در پژوهش حاضر سبک گفتار بوده‌است. همان‌گونه که یافته‌ها نشان می‌دهد پارامترهای آکوستیکی فرکانس پایه و فرکانس سازه‌ها در دو سبک گفتاری خوانداری و بداهه تفاوت چشمگیری داشته‌اند. در این میان، البته نقش زبان گفتاری نیز بسیار اهمیت داشته‌است. بر مبنای یافته‌ها، در گروه مردان فرکانس پایه در سبک بداهه نسبت به سبک خوانداری در هر دو زبان بیشتر بوده‌است. در گروه زنان، این میزان با توجه به زبان گفتاری متغیر است. در این گروه، فرکانس پایه در سبک گفتاری بداهه در زبان فارسی نسبت به سبک خوانداری بیشتر بوده‌است در حالی که در زبان انگلیسی میزان فرکانس پایه در سبک خوانداری بیشتر گزارش شده‌است. آنچه وجه مشترک بین هر دو گروه مردان و زنان است بالاتر بودن میزان فرکانس پایه در سبک بداهه نسبت به سبک خوانداری در زبان فارسی است. این نتیجه نشان می‌دهد که گویندگان رفتار بین-سبکی مشابهی در زبان اول خود یعنی فارسی داشته‌اند. همچنین بر مبنای یافته‌ها، فرکانس سازه اول در ایجاد تغییرات بین-سبکی تأثیر چشمگیری داشته‌است. این احتمال وجود دارد که با توجه به همبستگی میان فرکانس سازه اول و میزان باز بودن دهان، گویندگان به هنگام آدای پاره‌گفتارها در دو سبک خوانداری و بداهه متفاوت عمل کرده‌اند. بر مبنای یافته‌ها، در هر دو گروه مردان و زنان، میزان فرکانس سازه اول در زبان فارسی در سبک خوانداری نسبت به بداهه بیشتر بوده‌است؛ در حالی که در زبان انگلیسی که زبان دوم گویندگان بوده‌است میزان فرکانس سازه اول در سبک بداهه نسبت به خوانداری بیشتر بوده‌است. بر اساس یافته‌ها می‌توان پذیرفت که گویشوران رفتار متفاوت بین-سبکی در هر دو زبان داشته‌اند. به دیگر سخن، گویشوران هنگامی که در سبک خوانداری و به زبان فارسی صحبت می‌کرده‌اند، واکه‌ها را بازتر آدا کرده‌اند؛ در حالی که در زبان انگلیسی یعنی زبان دومشان به هنگام صحبت

کردن در سبک خوانداری واکه‌ها را بسته تر تلفظ کرده‌اند. این یافته‌ها نشان می‌دهد که اگرچه زبان گفتاری در ایجاد تغییرات آکوستیکی نقش مهمی دارد ولی نمی‌توان نقش سبک گفتاری را نیز نادیده گرفت.

نکته‌ای که معمولاً باید در پژوهش‌های آواشناسی قضایی مورد توجه قرار گیرد این است که ترکیب چند پارامتر در قیاس با یک پارامتر بهتر می‌تواند گوناگونی‌های بین-گوینده را نشان دهد. دیگر موضوعی که باید به آن توجه داشت چگونگی ترکیب پارامترهاست. پارامترهایی باید با هم ترکیب شوند که افزون بر نشان دادن مشخصه‌های فردیژه صدا همبستگی بالایی نیز با هم نداشته باشند. اگر دو پارامتر با هم همبسته باشند این به آن معناست که آن دو مشخصه دربردارنده اطلاعات کم و بیش مشابهی هستند و از این روی ترکیب آن‌ها ابعاد تازه‌ای از آوا را نشان نمی‌دهد. یافته‌های پژوهش حاضر نمایانگر همبستگی ضعیف میان پارامترهای صوت شناختی فرکانس پایه و فرکانس سازه‌ها بود. این امر نشان‌دهنده آن است که این پارامترها اطلاعات متفاوتی در مورد صدای گویندگان بازگو می‌کنند و در نتیجه ترکیب آن‌ها می‌تواند در امر تشخیص هویت گوینده سودمند باشد. این یافته‌ها با یافته‌های پژوهش‌های پیشین (Gold et al., 2010; Moos, 2013) نیز هم‌راستا است. این پژوهش‌ها فرکانس سازه را به دلیل عملکرد مستقل از فرکانس پایه و نیز سایر مشخصه‌ها همچون سرعت گفتار معیار مناسبی در مقایسه قضایی گویندگان می‌دانند و بر این باورند فرکانس سازه را می‌توان با سایر پارامترهای آوایی و زبانی ترکیب کرد.

۶. نتیجه‌گیری

در پژوهش حاضر، مجموعه‌ای از پارامترهای صوت شناختی مربوط به منبع یعنی فرکانس پایه و نیز پارامترهای مرتبط با صافی یعنی فرکانس سازه‌ها مورد بررسی و تجزیه و تحلیل صوت شناختی قرار گرفتند تا از این رهگذر بهترین و مناسب‌ترین پارامترهای صوت شناختی تأثیرگذار بر گوناگونی‌های بین-گوینده در گویندگان دوزبانه فارسی-انگلیسی شناسایی شود. یافته‌ها نشان داد که پارامترهای مورد بررسی در زبان فارسی و انگلیسی و نیز میان دو سبک گفتاری خوانداری و بداهه تفاوت معنادار داشته‌اند. فرکانس پایه، فرکانس سازه اول و نیز فرکانس سازه سوم موثرترین پارامترهای صوت شناختی در نشان دادن گوناگونی‌های بین-گوینده میان گویشوران دوزبانه فارسی-انگلیسی گزارش شدند. با توجه به همبستگی اندکی که میان فرکانس پایه و فرکانس سازه‌ها مشاهده شد، روشن شد که ترکیب این دو پارامتر صوت شناختی می‌تواند سرنخ سودمندی در ایجاد تمایز میان گویندگان افراد دوزبانه باشد. همچنین در پژوهش حاضر نشان داده شد که

فرکانس سازه اول دربردارنده اطلاعات فردویژه بسیاری در صدای افراد دوزبانه فارسی-انگلیسی است. بر مبنای یافته‌های این پژوهش، می‌توان این‌گونه نتیجه گرفت که پارامترهای صوت‌شناختی فرکانس پایه و نیز فرکانس سازه پارامترهایی سودمند در تشخیص هویت گوینده هستند و با اینکه پیکره آوایی مورد بررسی از جنبه زبان و سبک گفتاری تغییرات بسیاری داشت ولی این پارامترها همچنان توانسته‌اند گویندگان را از یک‌دیگر بازشناسند.

References

- Abercrombie, D. (1967). *Elements of General Phonetics*. UK: Edinburgh University Press.
- Afshan, A., Kreiman, J., & Alwan, A. (2020). Speaker discrimination in humans and machines: Effects of speaking style variability. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH, 2020-October*(September), 3136–3140. <https://doi.org/10.21437/Interspeech.2020-3004>
- Altenberg, E. P., & Ferrand, C. T. (2006). Fundamental Frequency in Monolingual English, Bilingual English / Russian, and Bilingual English / Cantonese Young Adult Women. *Journal of Voice, 20*(1), 86–96. <https://doi.org/10.1016/j.jvoice.2005.01.005>
- Asadi, H., Nourbakhsh, M., Sasani, F., & Dellwo, V. (2018). Examining long-term formant frequency as a forensic cue for speaker identification: An experiment on Persian. In M. Nourbakhsh, H. Asadi, & M. Asiaee (Eds.), *Proceedings of the First International Conference on Laboratory Phonetics and Phonology* (pp. 21–28). Tehran: Neveesh Parsi Publications.
- Asadi, H., Nourbakhsh, M., & Sassani, F. (2019). Between-speaker variability of voiceless fricatives in Persian. *Language Related Research, 10*(1), 129–147. <http://lrr.modares.ac.ir/article-14-3720-en.html>
- Asiaee, M., Nourbakhsh, M., & Skarnitzl, R. (2019). Can LTF discriminate bilingual speakers? *Proceedings of the 28th Annual Conference of the International Association for Forensic Phonetics and Acoustics (IAFPA)*, 41–42.
- Baldwin, J. R., & French, P. (1990). *Forensic Phonetics*. London: Printer Publishers. <http://hdl.handle.net/1885/91429>
- Boersma, P., & Weenink, D. (2022). *Praat: Doing phonetics by computer* (6.1.39). <http://www.praat.org/>
- Cheng, A. (2020). Cross-linguistic f0 differences in bilingual speakers of English and Korean. *The Journal of the Acoustical Society of America, 147*(2), EL67–EL73. <https://doi.org/10.1121/10.0000498>
- Corrette, R. (2022). *Praat Vocal Toolkit*. <http://www.praatvocaltoolkit.com>
- Debruyne, F., Decoster, W., Van Gijssels, A., & Vercammen, J. (2002). Speaking fundamental frequency in monozygotic and dizygotic twins. *Journal of Voice: Official Journal of the Voice Foundation, 16*(4), 466–471. [https://doi.org/10.1016/s0892-1997\(02\)00121-2](https://doi.org/10.1016/s0892-1997(02)00121-2)
- Dorreen, K. (2017). *FUNDAMENTAL FREQUENCY DISTRIBUTIONS OF BILINGUAL SPEAKERS IN FORENSIC SPEAKER COMPARISON* [The University of Canterbury]. Retrieved from <<https://doi.org/10.1017/CBO9781107415324.004>>

- Foulkes, P., & Barron, A. (2000). Telephone speaker recognition amongst members of a close social network. *International Journal of Speech, Language and the Law*, 7(2), 180–198. <https://doi.org/10.1558/sll.2000.7.2.180>
- Gold, E. (2014). *Calculating likelihood ratios for forensic speaker comparisons using phonetic and linguistic parameters*. (Phd dissertation), The University of York, Toronto, Canada.
- Gold, E., French, P., & Harrison, P. (2013). Examining long-term formant distributions as a discriminant in forensic speaker comparisons under a likelihood ratio framework. *Proceedings of Meetings on Acoustics*, 19. <https://doi.org/10.1121/1.4800285>
- Grosjean, F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language*, 36(1), 3–15. [https://doi.org/https://doi.org/10.1016/0093-934X\(89\)90048-5](https://doi.org/https://doi.org/10.1016/0093-934X(89)90048-5)
- Grosjean, F. (2010). *Bilingual: Life and Reality*. Cambridge: Harvard University Press. <http://www.jstor.org/stable/j.ctt13x0ft8>
- He, L., Zhang, Y., & Dellwo, V. (2019). Between-speaker variability and temporal organization of the first formant. In *The Journal of the Acoustical Society of America*, 145(3), 209–214. <https://doi.org/10.1121/1.5093450>
- Heeren, W., Vloed, D., & Vermeulen, J. (2014). Exploring long-term formants in bilingual speakers. *IAFPA's 2014 Annual Conference Book of Abstracts*, 2014, 3, 39.
- Holm, S. (2003). Individual use of acoustic parameters in read and spontaneous speech. *Phonum* 9, 9, 157–160.
- Järvinen, K., Laukkanen, A., & Aaltonen, O. (2013). Speaking a foreign language and its effect on F0. *Logopedics Phoniatrics Vocology*, 38(2), 47–51. <https://doi.org/10.3109/14015439.2012.687764>
- Johnson, K. A., Babel, M., & Fuhrman, R. A. (2020). Bilingual acoustic voice variation is similarly structured across languages. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH, 2020-October*, 2387–2391. <https://doi.org/10.21437/Interspeech.2020-3095>
- Keshavarz, M. H., & Ingram, D. (2002). The early phonological development of a Farsi-English bilingual child. *International Journal of Bilingualism*, 6(3), 255–269. <https://doi.org/10.1177/13670069020060030301>
- Kinoshita, Y. (2001). *Testing Realistic Forensic Speaker A Likelihood Ratio Based Approach Using Formants* [The Australian National University]. Retrieved from <<https://doi.org/10.1558/ijsl.v9i1.133>>.
- Laan, G. P. M. (1997). The contribution of intonation, segmental durations, and spectral features to the perception of a spontaneous and a read speaking style. *Speech Communication*, 22(1), 43–65. [https://doi.org/https://doi.org/10.1016/S0167-6393\(97\)00012-5](https://doi.org/https://doi.org/10.1016/S0167-6393(97)00012-5)
- Ladefoged, P. (2006). *A course in phonetics* (5nd ed.). USA: Wadsworth Cengage Learning.
- Lo, J. J. H. (2021). *Cross-Linguistic Speaker Individuality of Long-Term Formant Distributions: Phonetic and Forensic Perspectives*. 416–420. <https://doi.org/10.21437/interspeech.2021-1699>
- McDougall, K. (2004). Speaker-specific formant dynamics: An experiment on Australian English /aI/. *International Journal of Speech Language and the Law*, 11(1 SE-Articles), 103–130. <https://doi.org/10.1558/sll.2004.11.1.103>

- Mok, P. P. K., Xu, R. B., & Zuo, D. (2015). Bilingual speaker identification: Chinese and English. *International Journal of Speech, Language and the Law*, 22(1), 57–77. <https://doi.org/10.1558/ijssl.v22i1.18636>
- Moos, A. (2010). Long-term formant distribution as a measure of speaker characteristics in read and spontaneous speech. *The Phonetician*, 101/102, 7–24.
- Nolan, F., & Grigoras, C. (2005). A case for formant analysis in forensic speaker identification. *International Journal of Speech, Language and the Law*, 12(2), 143–173. <https://doi.org/10.1558/sll.2005.12.2.143>
- Paradis, J. (2001). Do bilingual two-year-olds have separate phonological systems? *International Journal of Bilingualism*, 5(1), 19–38. <https://doi.org/10.1177/13670069010050010201>
- Park, S. J., Yeung, G., Vesselinova, N., Kreiman, J., Keating, P. A., & Alwan, A. (2018). Towards understanding speaker discrimination abilities in humans and machines for text-independent short utterances of different speech styles. *The Journal of the Acoustical Society of America*, 144(1), 375–386. <https://doi.org/10.1121/1.5045323>
- R Core Team. (2021). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Retrieved from <<https://www.r-project.org/>>
- Rose, P. (2002). *Forensic Speaker Identification*. London: Taylor and Francis.
- Rose, P., & Clermont, F. (2001). A comparison of two acoustic methods for forensic speaker discrimination. *Acoustics Australia*, 29(1), 31–35.
- Sambur, M. R. (1975). Selection of Acoustic Features for Speaker Identification. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 23(2), 176–182. <https://doi.org/10.1109/TASSP.1975.1162664>
- Vaheb, A., Choobasti, A. J., Sabeti, B., Mortazavi Najafabadi, S. H. E., & Safavi, S. (2018a). Mirasvoice: A bilingual (English-Farsi) speech corpus. *LREC 2018 - 11th International Conference on Language Resources and Evaluation*, 3, 2903–2906.
- Vaheb, A., Choobasti, A. J., Sabeti, B., Mortazavi Najafabadi, S. H. E., & Safavi, S. (2018b). Mirasvoice: A bilingual (English-Farsi) speech corpus. *LREC 2018 - 11th International Conference on Language Resources and Evaluation*, May, 2903–2906.

