

بهنجارسازی فضای واکه‌ای زبان فارسی

نسیم اسفندیاری*

بتول علی‌نژاد**

چکیده

ترسیم واکه‌ها در فضای $F1 * F2$ نه تنها از نظر صوت‌شناختی تولیدی بلکه از نظر رویکردهای ادراکی نیز از اصلی‌ترین ابعاد در بیان کیفیت واکه‌ای محسوب می‌شوند. در ترسیم فضای واکه‌ای مشکل اصلی‌ای که اکثر محققان با آن روبرو هستند این است که ارزش‌های سازه‌های اندازه‌گیری شده برای یک واکه یکسان تولید شده توسط افراد مختلف، متفاوت است. جهت حذف تنوعات سازه‌ای ناشی از منابع ایجاد تغییر در سازه‌ها همچون تفاوت‌های فیزیولوژیکی میان گویندگان، موضوع بهنجارسازی واکه‌ای مطرح می‌شود. انتخاب میان تکنیک‌های بهنجارسازی اگرچه بسته به طبیعت و هدف بررسی می‌تواند متفاوت باشد ولی همان‌طور که همگی زبان‌شناسان بر آن توافق دارند فرایند بهنجارسازی و عملکرد آن یک امر وابسته به زبان است و یک فرایند بهنجارسازی نمی‌تواند موثرترین فرایند در همه‌ی زبان‌ها محسوب شود. از این‌رو در این بررسی جهت بهنجار کردن داده‌های زبان فارسی به ارزیابی شیوه‌های مختلف بهنجارسازی در زبان فارسی بر اساس دو معیار الف) تکنیک محاسبه میزان همپوشانی و ب) تکنیک تحلیل تشخیصی پرداخته می‌شود. شاخصه *الف* با هدف شناسایی مقیاسی که نسبت به سایرین بهتر بتواند اثر شخص را حذف کند اعمال می‌شود و معیار ب در جهت حفظ اطلاعات واجی و به منظور ارزیابی قدرت تفکیکی یک تکنیک در جهت جداسازی بسامدهای سازه‌ای بهنجارسازی شده به گروه‌های متمایز آوایی تعریف می‌شود. این تحقیق تکنیک نری را به عنوان مطلوب‌ترین شیوه بهنجارسازی جهت حذف و یا کمینه کردن تاثیرات فیزیولوژیکی با حفظ تمایزات واجی در زبان فارسی معرفی می‌کند. همچنین فضای واکه‌ای زنان چه در حالت تکیه‌دار و چه بی تکیه بزرگتر از مردان نمایش داده شده است.

کلید واژه‌ها: فضای واکه‌ای، بهنجارسازی واکه‌ای، تکنیک بهنجارسازی واکه-درون‌گرا و تکنیک بهنجارسازی واکه-برون‌گرا

۱. مقدمه

مطالعه صوت‌شناختی واکه‌ها که منجر به تعیین مشخصه‌های نمودهای آوایی واکه‌ها می‌شود با نمودار فضای واکه‌ای نشان داده می‌شود. نمودار فضای واکه‌ای یک روش گرافیکی است که نشان می‌دهد یک واکه در چه ناحیه‌ای از فضای تولیدی و فیزیکی و شنیداری واقع شده است. این کار اولین بار توسط اسنر^۱ (۱۹۴۷) و جوس^۲ (۱۹۴۸) (به نقل از هرینگتون^۳، ۲۰۱۰) انجام شد و از آن زمان به بعد طرح F1* F2 به‌عنوان یکی از شیوه‌های استاندارد در مقایسه کیفیت واکه‌ای در کل مطالعات آواشناختی شناخته شد. منتهی همان‌طور که فلین^۴ (۲۰۱۱) بیان می‌کند یکی از مشکلات اصلی در بررسی فضای واکه‌ای این است که گویندگان دارای مجرای صوتی با ابعاد و مشخصات یکسان نمی‌باشند و این منجر به این امر می‌شود که نتوان تفاوت موجود در ارزش‌های سازه‌ای را در اثر تغییرات واکه‌ای در نظام واکه‌ای و یا در اثر تفاوت‌های آنا‌تومیکی و فیزیولوژیکی در میان گویندگان شناسایی کرد. در این رابطه پیترسون و بارنی^۵ (۱۹۵۲) و پوتر و استاینبرگ^۶ (۱۹۵۰) با نمایش طیفی از واکه‌های یکسان تولیدشده توسط اشخاص متفاوت بیان داشتند که تفاوت‌های قابل‌ملاحظه‌ای در سازه‌های واکه‌ای بالاخص در میان گروه‌های جنسیتی و سنی وجود دارد. هرینگتون (۲۰۱۰) علت این نوع تنوعات آکوستیکی را تفاوت در شکل و اندازه مجرای صوتی در افراد مختلف می‌داند. لدفوگد و برادبنت^۷ (۱۹۵۷) نیز با پرداختن به موضوع مذکور، این نوع تنوعات را به عنوان تنوعات وابسته به گوینده نامگذاری و آنها را به دو گروه تنوعات شخصی و تنوعات اجتماعی-زبانی تقسیم‌بندی می‌کنند. منشا تنوعات شخصی در سیگنال مربوط به تفاوت میان گویندگان در شکل و اندازه مجرای صوتی و حنجره و منشا تنوعات اجتماعی-زبانی مربوط به تفاوت‌های شخصیت‌های اجتماعی گویندگان (همچون سطح تحصیلاتی، جنسیت، قومیت و غیره) می‌شود. پولس^۸ و همکاران (۱۹۷۳) نیز با بررسی تنوعات موجود در بسامد سازه‌ها، علت تنوعات را وابسته به سه منبع واکه، گوینده و خطای اندازه‌گیری می‌دانند. لدفوگد (۱۹۷۶)، پیترسون و بارنی (۱۹۵۲) پوتر و استاینبرگ (۱۹۵۰) نیز با کنترل متغیر واکه و خطای اندازه‌گیری متذکر شدند که بسامد سازه‌های واکه‌ای تفاوت قابل‌ملاحظه‌ای در یک واکه‌ی یکسان برای گویندگان مختلف نشان می‌دهد. به عبارت دیگر، واکه‌های تولیدشده توسط افراد مختلف، در نقاط متفاوتی بر روی نمودار قرار می‌گیرند. به‌عنوان مثال، از آنجایی که طول مجرای صوتی زنان حدود ۱۵٪ کوچک‌تر از طول مجرای صوتی مردان است، بنابراین مقادیر سازه‌ای آنها متفاوت خواهد بود. بدین صورت یکی از مهمترین مسائل در تحلیل آکوستیکی واکه به امر حذف و یا کمینه‌کردن تاثیرات تفاوت مجرای صوتی در بسامد نسبی سازه‌های پایین در گویندگان مختلف اختصاص پیدا می‌کند.

تکنیکی که در مطالعات صوت‌شناختی جهت کمینه نمودن نتایج صوت‌شناختی حاصل از منابع ایجادکننده تغییر به کاربرده می‌شود بهنجارسازی واکه‌ای^۹ نام دارد. می‌توان بهنجارسازی واکه‌ای را ایجاد گشتار بر روی نمودهای آکوستیکی نمونه‌های واکه‌ای دانست که هدف آن، کمینه کردن نتایج آکوستیکی منابع مختلف در نمود آکوستیکی

¹ C. Essner

² M. Joos

³ J. Harrington

⁴ N. Flynn

⁵ G.E. Peterson & H.L. Barney

⁶ R.K. Potter & J.C. Steinberg

⁷ P. Ladefoged & D.E. Broadbent

⁸ L.C. Pols et al

⁹ vowel normalization

واکه است. به عبارت دیگر، هدف بهنجارسازی واکه‌ای حذف تنوعات سازه‌ای ناشی از منابع ایجاد تغییر در سازه‌ها همچون تفاوت‌های فیزیولوژیکی میان گویندگان است. به نظر می‌رسد چنین فرایندی معرف رفتار شنوندگان به هنگام تعیین کیفیت واکه شنیده شده و انتساب آن واکه به گروه واکه‌ای خاص باشد؛ چرا که پیش‌فرض این رفتار، حذف مشخصه‌های شخصی گویندگان به جهت دست‌یابی به نقطه هدف واکه‌ای به منظور ادراک واکه مورد نظر است. بنابراین یکی از مسائل اصلی در تحلیل آکوستیکی واکه، تلاش در جهت حذف تأثیرات فیزیولوژیکی گویندگان بر بسامد سازه‌های پایین واکه‌ها می‌باشد. از این رو، یکی از اهداف اصلی در بررسی آکوستیکی واکه‌ها تنظیم فرمول‌های قابل اجرا در جهت حذف تأثیرات تفاوت‌های گویندگان در رخدادهای آوایی یکسان از طریق 'اصلاح' ابعاد سازه‌ای است (نری^۱)، چراکه نمی‌توان بسامدهای سازه‌ای خام در مقیاس هرتر را در گویندگان مختلف مستقیماً مورد بررسی قرار داد.

تکنیک‌های بهنجارسازی واکه‌ای توسط آینسورث^۲ (۱۹۷۵) و نری (۱۹۸۹) بنا به نوع داده‌ای که در فرایند به کار می‌رود به دو دسته واکه-درون‌گرا^۳ و واکه-برون‌گرا^۴ تقسیم‌بندی می‌شوند. فرایند بهنجارسازی واکه-درون‌گرا، به هدف هدف ارائه مدل‌سازی در واکه‌های ادراکی انسان، تنها از اطلاعات آکوستیکی درون یک واکه استفاده می‌کند و شامل استفاده از مقیاس‌های هرتر، بارک، مل و ارب در بهنجارسازی می‌شود. این در حالی است که فرایند بهنجارسازی واکه-برون‌گرا، به هدف ارائه الگوریتم تشخیص خودکار گفتار، به اطلاعاتی بیش از یک واکه‌ی گفته شده توسط گوینده نیاز دارد و در برگیرنده تکنیک‌های مطرح شده توسط گریستمان، لبانوی، نری، وات و فابریشوز و غیره می‌شود.

در حوزه ادراک گفتار، مطالعات در تأیید یکی از انواع بهنجارسازی درون‌گرا یا برون‌گرا دلالت بر عدم توافق عمومی زبانشناسان دارد. جوس (۱۹۴۸)، لدفوگد و برادبنت (۱۹۵۷) بر این اعتقاد هستند که بهنجارسازی می‌بایست برون‌گرا باشد چراکه واکه‌های یک گوینده مفروض در رابطه با واکه‌های ممیز همان گوینده ادراک می‌شود. از طرف دیگر آسمن^۵ و همکاران (۱۹۸۲)، وربراگ و استرانگ^۶ (۱۹۷۶) بیان می‌کنند که برای شنوندگان هیچ تفاوتی در شناسایی واکه‌ها در محیط‌های مختلف همچون تغییر گوینده به هنگام تولید یک واکه نسبت به واکه قبل وجود ندارد. بعلاوه مسأله اختلاف نظر تنها محدود به انتخاب دو رویکرد درون‌گرا یا برون‌گرا نمی‌شود بلکه در انتخاب فرایندهای مطرح در هر رویکرد نیز اتفاق نظر وجود ندارد.

بخشی از این عدم اجماع در انتخاب و معرفی یک رویکرد و یا فرایند بهنجارسازی خاص به این علت است که فرایند بهنجارسازی یک امر مطلق نیست بلکه بسته به زبان مورد بررسی، فرایندی ممکن است در حذف تنوعات بهتر از فرایندهای دیگر عمل کند. بنابراین فرایند بهنجارسازی یک امر وابسته به زبان است و یک فرایند بهنجارسازی نمی‌تواند برای همه زبانها موثرترین فرایند محسوب شود (فراری-دیسنر^۷، ۱۹۸۰). در اینجا شایان ذکر است که علیرغم نبود اتفاق نظر در میان زبان‌شناسان در ارائه الگوریتم‌های متفاوت بهنجارسازی، همگی یک مجموعه اهداف کلی را در انجام

¹ T. M. Neary

² W. A. Ainsworth

³ vowel-intrinsic

⁴ vowel-extrinsic

⁵ P.F. Assmann

⁶ R.R. Verbrugge & W. Strange

⁷ S. Ferrari- Disner

این فرایند مشخص می‌کنند که به صورت زیر بیان می‌شود (هیندل^۱، ۱۹۷۸؛ فراری-دیسنر، ۱۹۸۰؛ توماس^۲، ۲۰۰۲؛ لانگ استروف^۳، ۲۰۰۶؛ توماس و کندال^۴، ۲۰۰۷؛ فابریشوز و همکاران^۵، ۲۰۰۹؛ کلاپر^۶، ۲۰۰۹ و فلین، ۲۰۱۱):

الف) حذف و یا کمینه کردن تنوعات میان گویندگان که ناشی از تفاوت‌های ذاتی-فیزیولوژیکی باشد؛

ب) حفظ تنوعات میان گویندگان که ناشی از تفاوت‌های میان طبقات اجتماعی همچون سن، جنسیت، گویش و یا تغییرات آوایی باشد؛

ج) حفظ طبقات واکه‌ای و تفاوت‌های واجی

د) الگوبرداری از فرایندهای شناختی که به شنوندگان این مجال را می‌دهد که واکه‌های تولید شده توسط گویندگان مختلف را بهنجارسازی کنند.

البته، رعایت همه این اهداف تا حدی آرمان‌گرایانه می‌نماید چراکه هیچ الگوریتم بهنجارسازی نمی‌تواند به‌طور همزمان همه اهداف بالا را برآورده سازد (فلین، ۲۰۱۱؛ توماس، ۲۰۰۲؛ آدانک^۷ و همکاران، ۲۰۰۴؛ بیگهام^۸، ۲۰۰۸؛ توماس و کندال^۹، ۲۰۰۷). بنابراین زبان‌شناسان بسته به طبیعت و هدف بررسی، اهمیت بیشتری را برای یک هدف نسبت به اهداف دیگر قائل می‌شوند. به‌عنوان مثال در مطالعات آوایی-اجتماعی به حفظ تنوعات اجتماعی بیشتر تأکید می‌شود. یا در مطالعات برپایه ادراک، هدف بهنجارسازی نزدیک‌شدن به فرایند ادراکی انسانی تعریف می‌شود (فلین، ۲۰۱۱؛ روسنر و پیکرینگ^{۱۰}، ۱۹۹۴ و سیردال و گوپال^{۱۱}، ۱۹۸۶).

در این بررسی، پیش‌فرض مطالعه این است که شنوندگان جهت طبقه‌بندی واکه‌های شنیده شده از گویندگان مختلف تنها تنوعات واجی را در سیگنال صوت انتخاب می‌کنند و به تنوعات آناتومیکی و فیزیولوژیکی توجهی ندارند. در این رویکرد انتخاب شده در حوزه ادراک، بهنجارسازی واکه‌ای یک فرایند مجزا در ادراک واکه محسوب می‌شود که به‌وسیله آن شنوندگان قادر می‌شوند که اطلاعات ثابت واجی را از سیگنال صوتی جهت طبقه‌بندی واکه‌ای انتخاب کنند (آدانک، ۲۰۰۳). بنابراین، این بررسی علاوه بر اهمیت و الگوبرداری از فرایندهای شناختی، حذف تفاوت‌های فیزیولوژیکی را در اولویت قرار می‌دهد و با توجه به مطلق نبودن فرایندهای بهنجارسازی در زبان‌های مختلف در این بررسی در ابتدا جهت بهنجارسازی داده‌های زبان فارسی در جهت حذف اثر فیزیولوژیکی به ارزیابی شیوه‌های مختلف بهنجارسازی در زبان فارسی پرداخته می‌شود. هدف بررسی تعیین مطلوب‌ترین تکنیک بهنجارسازی در زبان فارسی در جهت حذف تأثیر جنسیت بر مقادیر سازه‌ای می‌باشد. در اینجا لازم به ذکر است که از آنجایی که حفظ تفاوت‌های واجی از جمله ضوابط اساسی در فرایند بهنجارسازی محسوب می‌شود، این امر نباید تحت فرایند بهنجارسازی دچار خدشه گردد. به‌عبارت‌دیگر بهنجارسازی‌ای که در آن حذف تنوعات فیزیولوژیکی منجر به خسارت سنگین در حوزه

¹ D. Hindle

² E. Thomas

³ C. Longstrof

⁴ E. Thomas & T. Kendall

⁵ A.H. Fabricius & et al

⁶ C. Clopper

⁷ P. Adank

⁸ D. Bigham

⁹ E. Thomas & T. Kendall

¹⁰ B. Rosner & J. Pickering

¹¹ A.K. Syrdal & H.S. Gopal

حفظ تنوعات واجی گردد پذیرفته نمی شود. از این رو پس از ارزیابی تکنیک‌ها براساس معیار اول که به حذف تنوعات فیزیولوژیکی می پردازد تکنیک اعلام شده مجدداً با معیار دوم، معیار تحلیل تشخیصی، ارزیابی می شود. معیار تعریف شده در ارزیابی فرایندهای بهنجارسازی در جهت کاهش تنوعات آناتومیکی و فیزیولوژیکی به میزان ترفیع انطباق فضاهای مختلف واکه‌ای اختصاص داده شده و معیار انتخابی در ارزیابی تکنیک‌های بهنجارسازی در حفظ اطلاعات واجی یک تحلیل آماری است که در آن از ماتریس کوواریانس متغیرهای اکوستیکی جهت طبقه‌بندی نمونه‌ها استفاده می شود. در نهایت براساس نتیجه‌ی به دست آمده از ارزیابی تکنیک‌های بدست آمده با دو معیار بیان شده داده‌های زبان فارسی بهنجارسازی می شود و فضای واکه‌ای در فضای $F1 * F2$ به تفکیک دو گروه واکه تکیه‌بر و بی تکیه در زبان فارسی ارایه می گردد.

۲. روش تحقیق

داده‌های زبان فارسی براساس فارسی معیار و از طریق ضبط مستقیم اخبار پخش شده توسط ۱۰ گوینده اخبار شبکه سراسری سیمای جمهوری اسلامی ایران که از کاربران حرفه‌ای زبان معیار محسوب می شوند در آبان ماه ۱۳۹۱ جمع آوری و در چارچوب آواشناسی آزمایشگاهی مورد تجزیه و تحلیل قرار می گیرند. ضبط مستقیم از طریق اتصال USB به درگاه خروجی تلویزیون در جهت جلوگیری از آلوده شدن داده‌های تحقیق توسط منابع تولیدکننده خش در محیط صورت گرفته است. در اینجا لازم به ذکر است که از آنجا که داده‌های گردآوری شده براساس گفتار پیوسته صورت گرفته و امکان تفکیک بافتی در آن‌ها میسر نبوده است با پیروی از مک‌نی لاگ (۱۹۷۰)، دانیلوف و هامبرگ (۱۹۷۳)، وینی فرد استرانگ (۱۹۸۹) قطعه‌ایستای واکه، قطعه هدف در نظر گرفته می شود. جهت شناسایی نقطه هدف واکه‌ای با استفاده از تکنیک هرینگتون (۲۰۱۰) که دلالت بر بیشینه بودن میزان بسامد در نقطه هدف واکه‌ای و همچنین مشاهدات چشمی حدود ۴۰۰۰ نمونه متغیر سازه اول و دوم در ۲۰۰۰ طیف صوتی با استفاده از نرم افزار پرات ویرایش 5.2.1.2 اندازه گیری و مورد تجزیه و تحلیل قرار گرفته‌اند. تحلیل‌های صوت‌شناختی فضاهای واکه‌ای به تفکیک دو متغیر جنسیت (مرد و زن) و بافت آوایی همجا (تکیه‌بر و بی تکیه) در چهار گروه مردان تکیه‌بر، مردان بی تکیه، زن تکیه‌بر و زنان بی تکیه انجام می گیرند. با توجه به کنترل متغیرهایی همچون سن، گویش، تحصیلات، سرعت گفتار و غیره منشا تفاوت‌های موجود در ارزش‌های سازه‌ای را می توان در اثر تفاوت فیزیولوژیکی در میان گویندگان دانست.

۳. مبانی نظری: بررسی شیوه‌های بهنجارسازی

همان‌طور که بیان شد به‌طور سنتی شیوه‌های بهنجارسازی واکه‌ای توسط آینسورث (۱۹۷۵) و نری (۱۹۸۹) به دو دسته واکه-درون گرا و واکه-برون گرا تقسیم‌بندی می شوند. در تکنیک‌های بهنجارسازی واکه-درون گرا همه اطلاعات لازم جهت این نوع بهنجارسازی تنها در یک نمونه واکه‌ای نهفته است و از ترکیب ارزش‌های سازه‌ای $F1$ و $F2$ معمولاً $F3$ و بندرت $F4$ و گاهی پهنای باند سازه‌ای به دست می آید. این در حالی است که در فرایندهای بهنجارسازی واکه-برون گرا از اطلاعات صوتی آن واکه و واکه‌های دیگر زبان استفاده می شود. به عبارت دیگر، متغیرهای بهنجارسازی واکه-برون گرا به مقایسه ارزش‌های سازه‌ای واکه‌های مختلف تولید شده توسط گوینده مفروض می پردازند. جهت آشنایی بیشتر با این دو شیوه بهنجارسازی، اهداف، مقیاس‌ها و معادلات مطرح در آنها ارائه می شود.

۳-۱. شیوه‌های بهنجارسازی واکه-درون‌گرا

انگیزه اصلی در بهنجارسازی درون‌گرا نزدیک شدن به شیوه ادراک انسان از واکه است که این هدف توسط سیردال و گوپال (۱۹۸۶) از طریق ایجاد گشتار در مقیاس هرتز پایه‌ریزی شده است. مقیاس هرتز (HZ) روشی خطی در ارائه بسامد سازه‌ها است، در حالی که مقیاس‌های ادراکی، غیرخطی هستند و می‌توانند از طریق نزدیک‌تر کردن بازنمایی بسامدی گفتار به عملکرد اندام حلزونی گوش انسان جنبه‌های ادراکی آواهای زبان را بهتر از مقیاس هرتز مشخص کنند (ترانمولر^۱، ۱۹۸۱). بنابراین فرایند بهنجارسازی درون‌گرا شامل تبدیل مقیاس هرتز به مقیاس‌های غیرخطی همچون بارک، بارک، مل، ارب می‌شوند که به شرح آن‌ها می‌پردازیم.

مقیاس بارک یک مقیاس آوایی-ذهنی است که توسط ابرهارد زویبکر در سال ۱۹۶۱ در مدل بلندی رسائی صوت مطرح شده است. در این مدل تأثیرات گوش میانی و بیرونی در انتقال صوت مورد بررسی قرار می‌گیرد. سیستم شنیداری انسان را می‌توان متشکل از مجموعه‌ای از فیلترها و به عبارتی دیگر باندهای بحرانی دانست. اصطلاح باند بحرانی به دامنه‌ی موثر بسامدی اطلاق می‌شود که به آن همه نقاط غشاء پایه گوش پاسخ می‌دهند (فلتچر^۲، ۱۹۴۰). بدین صورت اندازه‌گیری پهنای باند بحرانی معمولاً شامل آزمایش در حوزه بلندی صوت است. زویبکر در تقسیم محدوده بسامدی شنیداری به باند بحرانی متذکر می‌شود که این باندها از طریق اندازه‌گیری‌های دقیق در آزمایش بر روی آستانه آواهای مرکب در ادارک گفتار و بلندی آواهای مرکب انجام شده است (ترانمولر، ۱۹۸۱).

معادله ترانمولر (۱۹۹۰) بیشترین انطباق را با جدول باند بحرانی زویبکر (۱۹۶۱) برقرار می‌کند که به قرار زیر است:

معادله ۳-۱

$$F_i^Z = 26.81 * \left(\frac{f1}{1960 + F_i} \right) - 0.53.$$

مقیاس مل، یک مقیاس ادراکی در زیروبمی است که در سال ۱۹۴۰ توسط استیونز و همکاران نامگذاری شده است. نام مل از کلمه ملودی گرفته شده است که حاکی از آن است که این مقیاس براساس مقایسات زیروبمی استوار است. این مقیاس حاصل آزمایش بر اصوات ساده سینوسی می‌باشد که در آن از شنوندگان خواسته می‌شود که اندازه فاصله زیروبمی موجود میان یک جفت نواختی را به چهار فاصله مساوی ادراکی و در نتیجه سه نواخت تقسیم کنند. بنابراین بسامد جفت نواختی در طی آزمایش مرتب تغییر می‌کند. معادله تبدیل هرتز به مل در ۳-۲ بیان شده است (استیونز^۳ و همکاران، ۱۹۴۰):

معادله ۳-۲

$$. F_i^M = 2595 * \log_{10} \left(1 + \frac{f}{700} \right).$$

امروزه استفاده از مقیاس مل، تنها به علت اولویت تاریخی آن است چرا که این مقیاس بسیار نزدیک به مقیاس درجه باند بحرانی ارزیابی شده است.

¹ H. Tranmuller

² H. Flechter

³ S. Stevens & et al.

مقیاس ارب یا باند بحرانی مستطیلی متساوی الاضلاع، مقیاسی است که توسط مور و گلسبرگ^۱ (۱۹۸۳) با ویرایش مدل بلندی زوییکر بدست آمده است. هر دو مقیاس از عملکرد غشاء پایه در اندام حلزونی گوش داخلی برای تجزیه امواج صوتی به بسامدهای سازنده بدست آمده‌اند. تنها تفاوت گشتار ارب با گشتار بارک را می‌توان در دو نکته خلاصه کرد. اول اینکه گشتار ارب خیلی بیشتر از گشتار بارک به مقیاس لگاریتمی محض نزدیک است که این مسئله بالاخص برای بسامدهای پایین تر صدق می‌کند. ثانیاً، مقیاس ارب در پهنای باند بحرانی باریک‌تر و کوچکتر از مقیاس بارک نمایش داده می‌شود (آدانک، ۲۰۰۳).

بررسی‌های مور و گلسبرگ حاکی از آن است که انتخاب بسامدهای شنیداری بر اساس باند بحرانی مستطیلی متساوی الاضلاع به عنوان تابعی از بسامد مرکزی صورت می‌گیرد. تابع تبدیل بسامد به ارب توسط مور و گلسبرگ (۱۹۸۳) در معادله ۳-۳ مطرح می‌شود:

معادله ۳-۳

$$. F_i^E = 29.4 \times (\ln(0,00437 * Fi + 1)).$$

در اینجا ذکر این مسئله لازم است که اگرچه در سالهای اخیر بهنجارسازی غیرخطی بسامد به یک رویکرد شناخته شده جهت کاهش تنوعات نظام مند در آناتومی مجرای صوتی تبدیل شده است ولیکن شواهد موجود حاکی از آن هستند که تکنیک‌های بهنجارسازی برون‌گرا در کاهش تفاوت میان گویندگان دارای عملکردی بهتری می‌باشند (فراری-دیسنر ۱۹۸۰).

۲-۳. شیوه‌های بهنجارسازی واکه-برون‌گرا

انگیزه اصلی در بهنجارسازی برون‌گرا رسیدن به طبقه‌بندی صحیح واکه‌ها جهت به کار بستن در حوزه تشخیص خودکار گفتار (ASR) است که توسط گرستمان^۲ (۱۹۶۸) پایه‌ریزی شده است. سیستم تشخیص خودکار گفتار، رایانه را قادر به شناسایی کلمات گفته شده توسط افراد می‌کند. بنابراین تنها عامل اصلی در عملکرد این سیستم تنوعات میان گویندگان می‌باشد و آنچه مورد نیاز است تکنیک یا گشتاری است که بتواند تفاوت‌های میان گویندگان را مجزا نماید. بدین ترتیب نمود آکوستیکی واکه را تعیین می‌نمایند.

فرایند بهنجارسازی گرستمان (۱۹۶۸) از مجموعه‌های مختلفی از پارامترهای آوایی به عنوان درونداد داده‌ای استفاده می‌کند. در نخستین ویرایش این فرایند، گرستمان ارزش‌های کمینه و بیشینه سازه‌های اول و دوم را برای هر یک از نظام‌های واکه‌ای گویندگان در سطوح ثابت ۰ تا ۹۹۹ هرتز قرار می‌دهد. در ویرایش‌های بعدی ارزش‌های سازه‌ای را برای سازه‌های اول در کمینه ۲۵۰ هرتز و بیشینه ۷۵۰ هرتز و برای سازه دوم در کمینه ۸۵۰ و بیشینه ۲۲۵۰ تعریف می‌کند که این مقادیر معرف نوع بهتری از گفتار می‌باشند. وی مقادیر اولین و دومین سازه در هر واکه برای هر گوینده را در معادله ۳-۴ قرار می‌دهد تا مقادیر بهنجارشده آن حاصل شود.

معادله ۳-۴

^۱ B.C.J. Moore & B.R. Glasberg

^۲ L. Gerstmann

$$F_{ti}^{Gerstman} = 999 \times \frac{F_{ti} - F_{ti}^{min}}{F_{ti} - F_{ti}^{min}}$$

در فرمول از ارزش کمینه F_i حاصل از همه واژه‌ها برای گوینده t و از بیشینه ارزش همه واژه‌ها در همان گوینده حاصل می‌شود.

روش بهنجارسازی لبانو^۱ (۱۹۷۱) اگرچه یکی از قدیمی‌ترین تکنیک‌های بهنجارسازی واژه-برون گرامحسوب می‌شود هنوز در میان بهترین و موثرترین معادلات قرار می‌گیرد. تکنیک لبانو بر گشتار داده‌ها به Z-score استوار است که از طریق تفریق میانگین بسامدی گوینده و تقسیم آن بر انحراف از معیار برای هر پارامتر به طور جداگانه محاسبه می‌شود (معادله ۳-۵) به عبارت دیگر این گشتار از مرکز فضای واژه‌ای و همچنین پراکندگی میانگین از مرکز جهت کمینه کردن تفاوت میان گویندگان بهره می‌جوید.

معادله ۳-۵

$$F_i^{lobanov} = (F_i - \bar{F}_i) / SD_i.$$

F_i در فرمول مقدار میانگین بسامد سازه‌ای است که با در نظر گرفتن همه واژه‌های یک گوینده برای یک سازه تعریف می‌شود و SD_i به انحراف معیار از مقدار میانگین اشاره می‌کند. بدین طریق تأثیر این گشتار بر فضای واژه‌ای هر گوینده این است که آن را در مرکزیت صفر میانه قرار می‌دهد و بنابراین محورها بر اساس میزان انحراف از معیار از میانگین گوینده تعریف می‌شوند. در فرایند بهنجارسازی لبانو سیستم واژه‌ای بر مرکز آن ثابت نگه داشته می‌شود در حالی که در فرایند بهنجارسازی گرستمان سیستم واژه‌ها در نقاط انتهایی ثابت فرض می‌شود. از آنجایی که تمایزات واجی نسبتاً کمی در این منطقه حاصل می‌شود بنابراین انتظار می‌رود که این فرایند مداخله کمتری در مکان واژه‌ها در فضای آوایی نسبت به دیگر تکنیک‌هایی مانند گرستمان ایجاد کند (فراری-دیسنر، ۱۹۸۰).

نری (۱۹۷۸) از معادله لگاریتمی در جهت ایجاد گشتار در سازه واژه‌ها استفاده می‌کند. از مقیاس لگاریتمی در اصل به‌عنوان مقیاس در زیربومی آهنگین استفاده می‌شده است و از آنجا که بسامد پایه دو نت معرف فاصله آهنگین خاص بوده‌اند، تفاوت میان بسامدهای لگاریتمی در این فاصله ثابت در نظر گرفته شده است (نری، ۱۹۷۸ و آدانک، ۲۰۰۳). در معادله نری، فضاهای واژه‌ای گویندگان از طریق هم‌راستا شدن در بسامد سازه‌ای میانگین گویندگان با هم مقایسه می‌شوند و بدین طریق، لگاریتم طبیعی ارزش سازه‌ای یک گوینده محاسبه می‌شود و در نهایت، میانگین لگاریتم بسامد سازه‌ای همه واژه‌های گوینده از آن کسر می‌شود (معادله ۳-۶) (فلین، ۲۰۱۱).

معادله ۳-۶

$$F_i^N = Ln. (F_i) - \mu_{in} (F_i).$$

وات و فابریشوز^۲ (۲۰۰۲) برای هر گوینده در فضای واژه‌ای $F_1 * F_2$ یک مرکز ثقل (S) تعریف می‌کنند. مرکز ثقل بر اساس اندازه‌گیری بسامد سازه‌های واژه‌های ممیز [i u a] محاسبه می‌شود چراکه پیش‌فرض در این انتخاب این است که این واژه‌ها معرف ارزش‌های کمینه و بیشینه گویندگان برای سازه‌های F_1 و F_2 هستند، به طوری که [i] دارای ارزش

¹ B. M. Lobanov

² D. Watt & A. Fabricius

کمینه در F1 و ارزش بیشینه در F2 است در حالی که [a] دارای ارزش کمینه در F1 و F2 می‌باشد). بدین طریق مرکز ثقل (S) از طریق محاسبه میانگین Fr برای واکه‌های [u, a, i] محاسبه می‌شود و سپس همه اندازه‌گیری‌های انجام شده در F بر ارزش S تقسیم می‌گردد.

معادله ۳-۷

$$S(Fi) = \begin{cases} \frac{Fi(i)+Fi(a)+Fi(u)}{3} & i = 1. \\ \frac{Fi[i]+Fi(u)}{2} & i = 2. \\ Fi^N = \frac{Fi}{S(Fi)}. \end{cases}$$

در این پژوهش، انتخاب فرایندهای بهنجارسازی چه درون‌گرا و چه برون‌گرا کاملاً بر اساس نتیجه کارایی فرایندها در بررسی‌های پیشین و همچنین نوع داده و اندازه‌گیری‌های مورد نیاز در فرایندها بوده است. در نهایت به‌طور خلاصه می‌توان روش‌های بهنجارسازی بررسی شده را در جدول ۳-۱ خلاصه نمود:

جدول ۳-۱) انواع روش‌های بهنجارسازی استفاده شده در این تحقیق

فرایندهای بهنجارسازی واکه-درون‌گرا	
هرتز	وضعیت مرجع، بسامد سازه در Hz
بارک	گشتار بارک بر روی مقیاس بسامد
مل	گشتار مل بر روی مقیاس بسامد
ارب	گشتار ارب بر روی مقیاس بسامد
فرایندهای بهنجارسازی واکه-برون‌گرا	
گرستان	فرایند بهنجارسازی گرستان (۱۹۶۸)
لبانو	فرایند z-score لبانو (۱۹۷۱)
نری	فرایند میانگین لگاریتم نری (۱۹۷۸)
وات و فابریشوز	فرایند محاسبه مرکز ثقل وات و فابریشوز (۲۰۰۲)

۴- معیارهای ارزیابی شیوه‌های بهنجارسازی

۴-۱. معیار همپوشانی فضاهای واکه‌ای

در بررسی حاضر جهت ارزیابی تکنیک‌های بهنجارسازی به جهت حذف و یا کمینه کردن تاثیرات آناتومیکی از معیار همپوشانی فضاهای واکه‌ای بهره می‌بریم. از این معیار زبان‌شناسان بسیار در بررسی‌های خود استفاده کرده‌اند که ما در اینجا به لحاظ روشن شدن مباحث نظری به ذکر بررسی وات و فابریشوز (۲۰۰۲) و فلین (۲۰۱۱) اکتفا می‌کنیم. وات و فابریشوز (۲۰۰۲) از این معیار در ارزیابی تکنیک بهنجارسازی ارائه شده توسط خود به جهت ترفیع در انطباق فضاهای مختلف واکه‌ای گویندگان در فضای F1*F2 با دیگر تکنیک‌های بهنجارسازی و فلین (۲۰۱۱) در ارزیابی شیوه‌های بهنجارسازی استفاده می‌کنند.

در روش وات و فابریشوز فضای واکه‌ای بر اساس مثلث واکه‌ای و بر اساس اندازه‌گیری‌های بسامد سازه‌ای در واکه‌های ممیز [a], [u], [i] محاسبه می‌شود. در این روش میزان مطابقت مثلث‌های واکه‌ای بر اساس دو پارامتر سنجیده می‌شود. در ابتدا نسبت مساحت مثلث واکه‌ای گویندگان زن نسبت به مساحت مثلث واکه‌ای گویندگان مرد و سپس درجه همپوشانی میان دو مثلث واکه‌ای که بر اساس درصد همپوشانی مثلث واکه‌ای مردان با زنان و یا بالعکس محاسبه می‌گردد. در این معیار، وجود شباهت بیشتر در فضای واکه‌ای گویندگان و همچنین همپوشانی بیشتر در فضاها حاکی از حذف تنوعات ناشی از تفاوت‌های فیزیولوژیکی میان گویندگان می‌باشد.

فلین (۲۰۱۱) به ارزیابی ۲۰ تکنیک بهنجارسازی واکه‌ای با تکنیکی مشابه با تکنیک ذکر شده و البته با الهام گرفتن از تکنیک فابریشوز (۲۰۰۹) و وات و فابریشوز (۲۰۰۲) می‌پردازد. وی ارزیابی فرایندها را براساس توانایی هر فرایند در مساوی و هم‌راستا نمودن فضاهای واکه‌ای گویندگان مختلف در فضای واکه‌ای چهار ضلعی انجام می‌دهد. در روش فلین، فضای واکه چهار ضلعی متشکل از چهار زاویه است که مختصات آن عبارت از نقاط بیشینه F2- کمینه F1، کمینه F2- کمینه F1، کمینه F2- بیشینه F1 و بیشینه F1 می‌باشد.

فرمول کلی برای سنجش میزان یکسان نمودن فضاهای واکه‌ای در هر فرایند، از طریق معادله مساحت چهار ضلعی حاصل می‌شود. شایان ذکر است از آنجا که مساحت‌های ارزیابی شده در هر فرایند بهنجارسازی بر اساس واحدهای آن فرایند تعریف می‌شود، امکان مقایسه مستقیم فضاهای واکه‌ای وجود ندارد. به همین علت فابریشوز و همکاران (۲۰۰۹) و به پیروزی از آنها فلین (۲۰۱۱) از مربع ضریب همبستگی واریانس (SCV) استفاده می‌کنند. تفسیر مربع ضریب همبستگی واریانس این گونه است که فرایند بهنجارسازی که در آن SCV پایین‌تر باشد، آن فرایند در کاهش واریانس مساحت فضای واکه‌ای میان گویندگان و نهایتاً یکسان نمودن مساحت فضاها موفق‌تر عمل کرده است.

هم‌راستا بودن فضاهای واکه‌ای گویندگان نیز امری بسیار مهم می‌باشد که باید در محاسبات مورد توجه واقع شود چرا که ممکن است دو فضای واکه‌ای با اشکال مختلف و حتی همپوشانی ضعیف مساحت یکسان داشته باشند. جهت محاسبه هم‌راستا بودن فضاهای واکه‌ای از معیار میزان همپوشانی فضاها استفاده می‌شود. با توجه به وجود مقیاس‌های مختلف در فرایندهای بهنجارسازی، همپوشانی کلی فضاهای واکه‌ای با میزان درصد محاسبه می‌شود. بنابراین درصد بالای همپوشانی در یک فرایند به این معنا است که آن فرایند در هم‌راستا نمودن فضاهای واکه‌ای عملکرد مطلوبی داشته است.

حال با توضیحات ارائه شده در این معیار، تنها مسئله باقیمانده به موضوع مثلث و یا چهار ضلعی واکه‌ای و در نظر گرفتن کل اضلاع فضای واکه‌ای اختصاص داده می‌شود. جوس (۱۹۴۸)، لیبرمن^۱ و همکاران (۱۹۷۲) بیان می‌کنند که واکه‌های ممیز [i, u, a] را می‌توان به‌عنوان متعادل‌کننده‌های اصلی در نظام واکه‌ای در نظر گرفت. به‌علاوه محققان در حوزه ادراک بیان می‌کنند که یک شنونده با شنیدن واکه‌های ممیز می‌تواند طبقه‌بندی نمونه‌های واکه‌ای تولید شده توسط یک گوینده جدید را بهتر انجام دهند (آدانک، ۲۰۰۳). همچنین بیان شده است که واکه‌های ممیز، محتمل‌ترین کاندیدها جهت ارائه نظام واکه‌ای هستند زیرا این واکه‌ها جایگاه‌های کرانه‌ای را در فضای واکه‌ای تولیدی گوینده اشغال می‌کنند و به‌صورت بسامدهای سازه‌ای نهایی در فضای واکه‌ای آکوستیکی ظاهر می‌شوند. لیلینکرانتس و لیندبلاد^۲

¹ P. Lieberman

² J. Liljencrants & B. Lindbom

(۱۹۷۲) نیز در باب ارجحیت واکه‌های ممیز متذکر می‌شوند که از آنجا که واکه‌های ممیز در نهایت فضای واکه‌ای ممکن قرار می‌گیرند، از نظر صوت‌شناختی کاملاً از هم متمایز هستند و احتمال اشتباه شنونده در ادراک این سه واکه وجود ندارد.

بنابر مستندات ارائه شده در نظر گرفتن واکه‌های ممیز در مثلث واکه‌ای جهت ارزیابی مساحت، کافی به نظر می‌رسد منتهی یکی از انتقادات مطرح در انتخاب مثلث واکه‌ای این است که در این روش با مبنای قرار دادن سه واکه، تغییرات حوزه‌ای دیگر واکه‌ها نادیده گرفته می‌شود و در نهایت دقت لازم در بررسی اعمال نمی‌شود. با توجه به منطقی بودن انتقاد مذکور، در این بررسی مبنای مساحتی به منظور افزایش دقت محاسباتی بر پایه شش ضلع ایجاد شده در فضای واکه‌ای صورت می‌پذیرد.

۴-۲. معیار تحلیل تشخیصی

تحلیل تشخیصی^۱ یک تکنیک تشخیصی استاندارد است که در آن از ماتریس کوواریانس متغیرهای اکوستیکی جهت طبقه‌بندی نمونه‌ها استفاده می‌شود. از این معیار نری (۱۹۷۸)، سیردال (۱۹۸۴)، آدانک و همکاران (۲۰۰۴) جهت ارزیابی تکنیک‌های بهنجارسازی در حفظ اطلاعات واجی بهره گرفته‌اند. معیار تحلیل تشخیصی یک تحلیل آماری است که جهت پیش‌بینی متغیر وابسته از طریق متغیرهای غیروابسته عمل می‌کند. در این روش از متغیرهای اکوستیکی به‌عنوان متغیر غیر وابسته و یا عامل پیش‌بینی‌کننده^۲ و از گروه واکه‌ای مورد نظر به‌عنوان متغیر وابسته^۳ استفاده می‌شود. در صورتی که تکنیک بهنجارسازی مورد نظر درصد بالایی را در طبقه‌بندی صحیح واکه‌ای دارا باشد این بدان معنا است که آن تکنیک در حفظ تنوعات واجی موفق عمل کرده است.

به طور کلی می‌توان اهداف استفاده از تکنیک تحلیل تشخیصی را به‌صورت زیر بیان نمود:

الف) طبقه‌بندی نمونه‌ها در گروه‌ها

ب) ارزیابی یک نظریه از نظر طبقه‌بندی نمونه‌ها طبق پیش‌بینی

ج) تعیین اقتصادی‌ترین روش در تمایز میان گروهی

د) بررسی مشخصه‌ای که بر اساس آن گروه‌بندی به بهترین نحو انجام شود.

اگرچه همگی اهداف بیان شده در یک جهت می‌باشند و آن عبارت از کمینه کردن احتمال طبقه‌بندی اشتباه در نمونه‌ها است ولیکن هدف الف یعنی طبقه‌بندی نمونه‌ها در گروه‌ها در این بررسی مورد استفاده قرار گرفته است.

۵- تجزیه و تحلیل شیوه‌های بهنجارسازی در زبان فارسی

۵-۱. تجزیه و تحلیل شیوه‌های بهنجارسازی با معیار میزان همپوشانی فضاهای واکه‌ای

در این قسمت، شاخصه نسبت همپوشانی فضاهای واکه با هدف شناسایی مقیاسی که نسبت به سایرین بهتر بتواند اثر شخص را حذف کند ارائه می‌گردد. منظور از حذف اثر شخص در این شاخصه انطباق فضاهای واکه‌ای ۱۰ گوینده بر

^۱ discriminant analysis

^۲ discriminant analysis

^۳ dependant variable

یکدیگر می‌باشد. منظور از فضای واکه‌ای ناحیه محصور میان چندضلعی‌ای است که از اتصال ۶ واکه بدست می‌آید. مقیاس یا روش بهنجارسازی مناسب قادر خواهد بود فضای واکه‌ای این ۱۰ شخص را حتی‌الامکان بر یکدیگر منطبق نماید. انطباق بیشتر فضاهای واکه‌ای ۱۰ شخص به معنای این است که مساحت بیشتری از فضای دو بعدی تشکیل شده از مقیاس مورد بررسی در فضای واکه‌ای هر ۱۰ نفر وجود داشته باشد. این مساحت با عنوان مساحت ناحیه مشترک در فضاهای واکه‌ای نامگذاری می‌شود. برای قابل مقایسه بودن این مقدار در روش‌های مختلف بهنجارسازی کافی است که آن را بر مساحت اجتماع فضاهای واکه‌ای تقسیم کنیم. منظور از ناحیه اجتماع فضاهای واکه‌ای، ناحیه‌ای از فضای دو بعدی است که حداقل در یکی از فضاهای واکه‌ای قرار گرفته باشد. نسبت مساحت ناحیه مشترک در فضاهای واکه‌ای بر مساحت اجتماع فضاهای واکه‌ای همان شاخصه مورد نظر است که با عنوان تناسب مساحتی نامگذاری می‌شود. در دو جدول ۱-۵ و ۲-۵ مساحت فضای واکه‌ای زبان فارسی برای هر شخص همراه با مساحت ناحیه مشترک در فضاهای واکه‌ای و مساحت اجتماع فضاهای واکه‌ای در مقیاس‌های مختلف نشان داده شده است.

جدول ۱-۵) مساحت فضای واکه، مساحت مشترک، مساحت اجتماع و تناسب برای هر شخص در مقیاس‌های بهنجارسازی

درون‌گرا

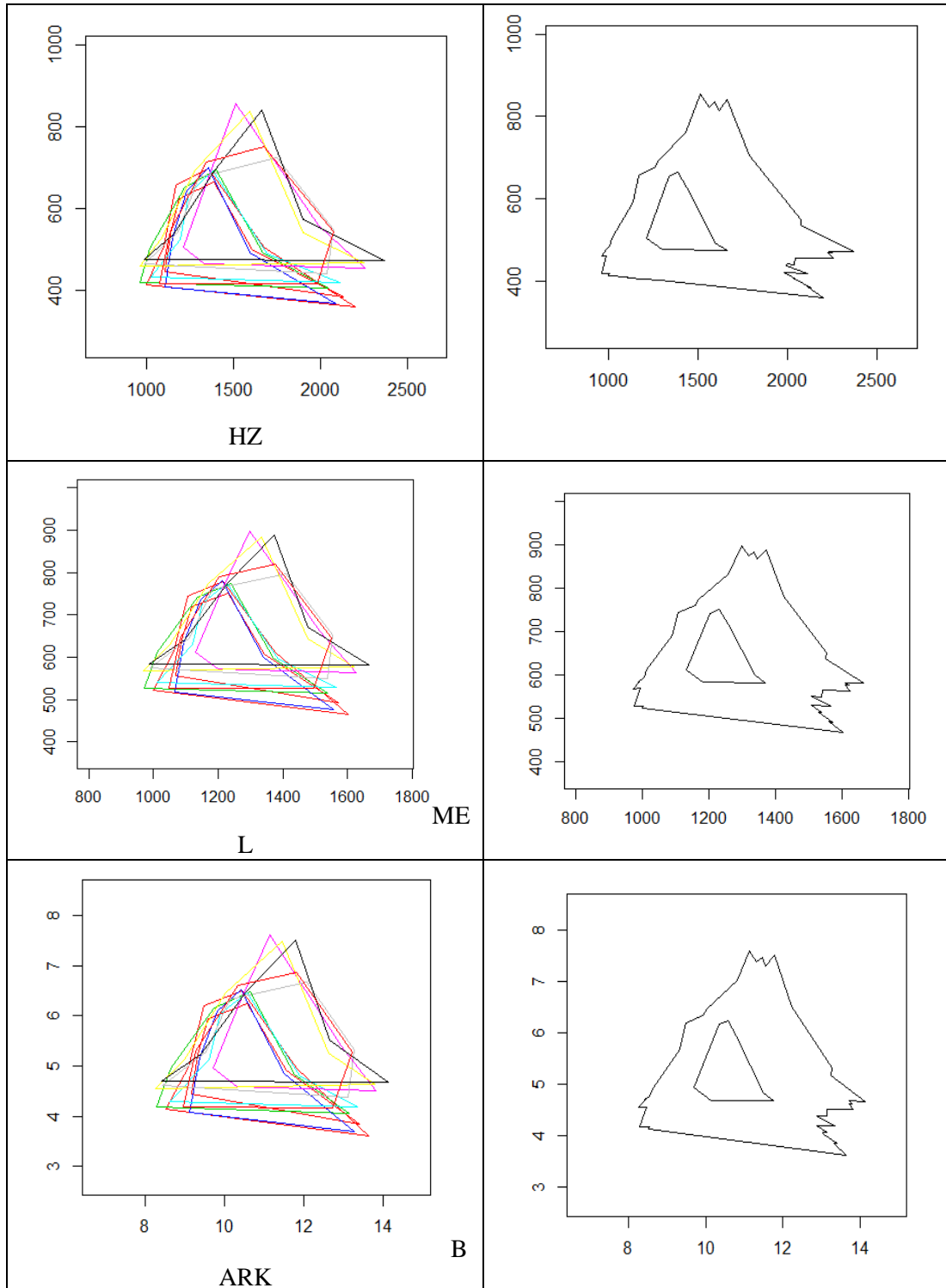
Person	Scale			
	HZ	Mel	Bark	Erb
1	168844.9	84422.5	6.740630	12.30055
2	107421.8	53413.52	4.268828	7.726098
3	145321.2	73416.74	5.898785	10.65933
4	126136.4	63466.65	5.069716	9.2628
5	112842.2	55340.59	4.416973	7.959246
6	195060.6	80867.82	6.279091	10.87848
7	208994.0	92918.06	7.336637	12.8074
8	197601.9	88243.26	6.931605	12.25601
9	195894.8	85020.19	6.669921	11.63034
10	237832.0	105602.1	8.275764	14.6711
Intersection area	43965.77	21116.63	1.692824	2.969823
Union area	340568	154797.8	1.218005	2.177845
Ratio	0.129	0.136	0.139	0.136
Rank	5	4	3	4

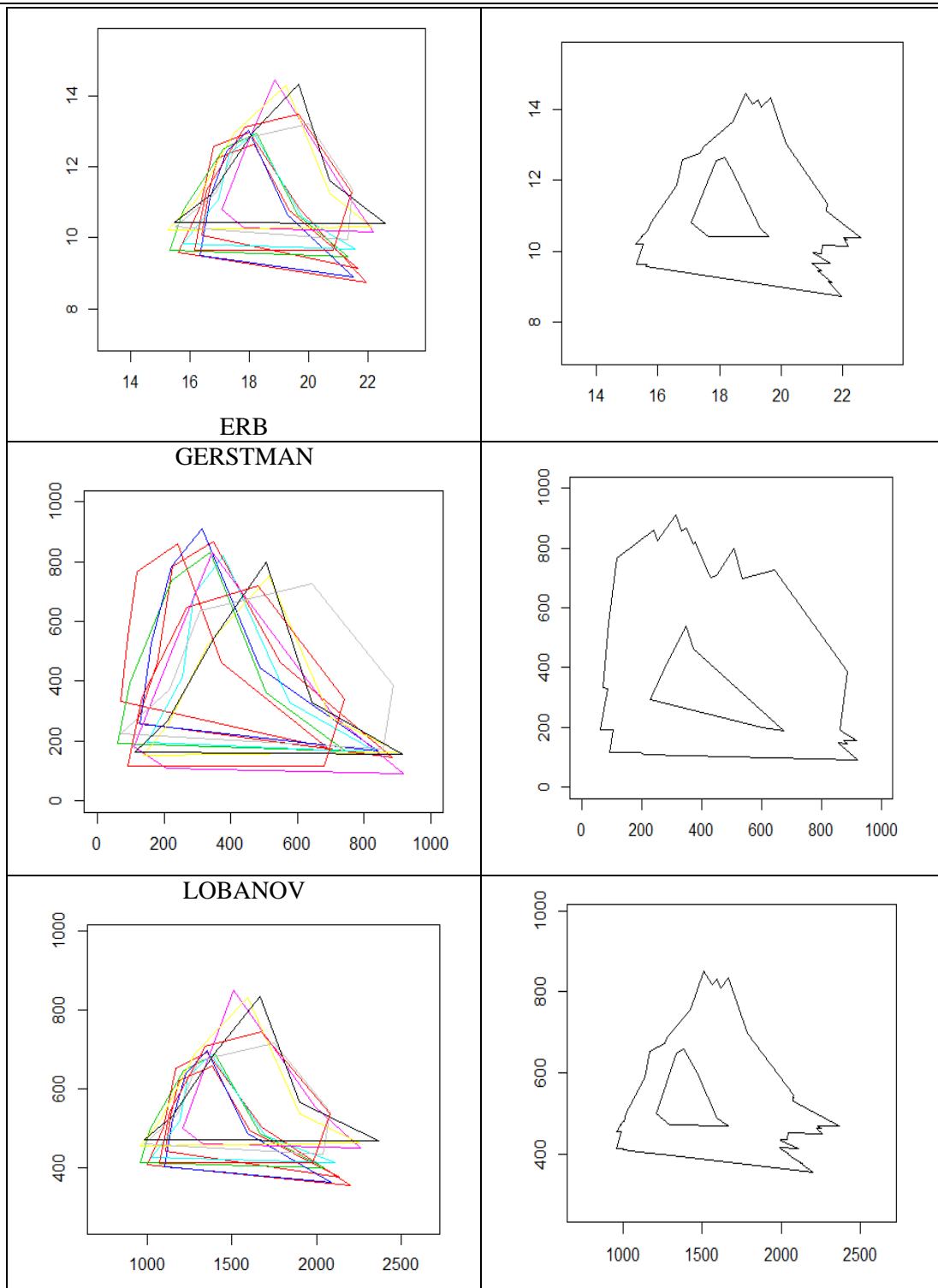
جدول ۲-۵) مساحت فضای واکه، مساحت مشترک، مساحت اجتماع و تناسب برای هر شخص در مقیاس‌های بهنجارسازی

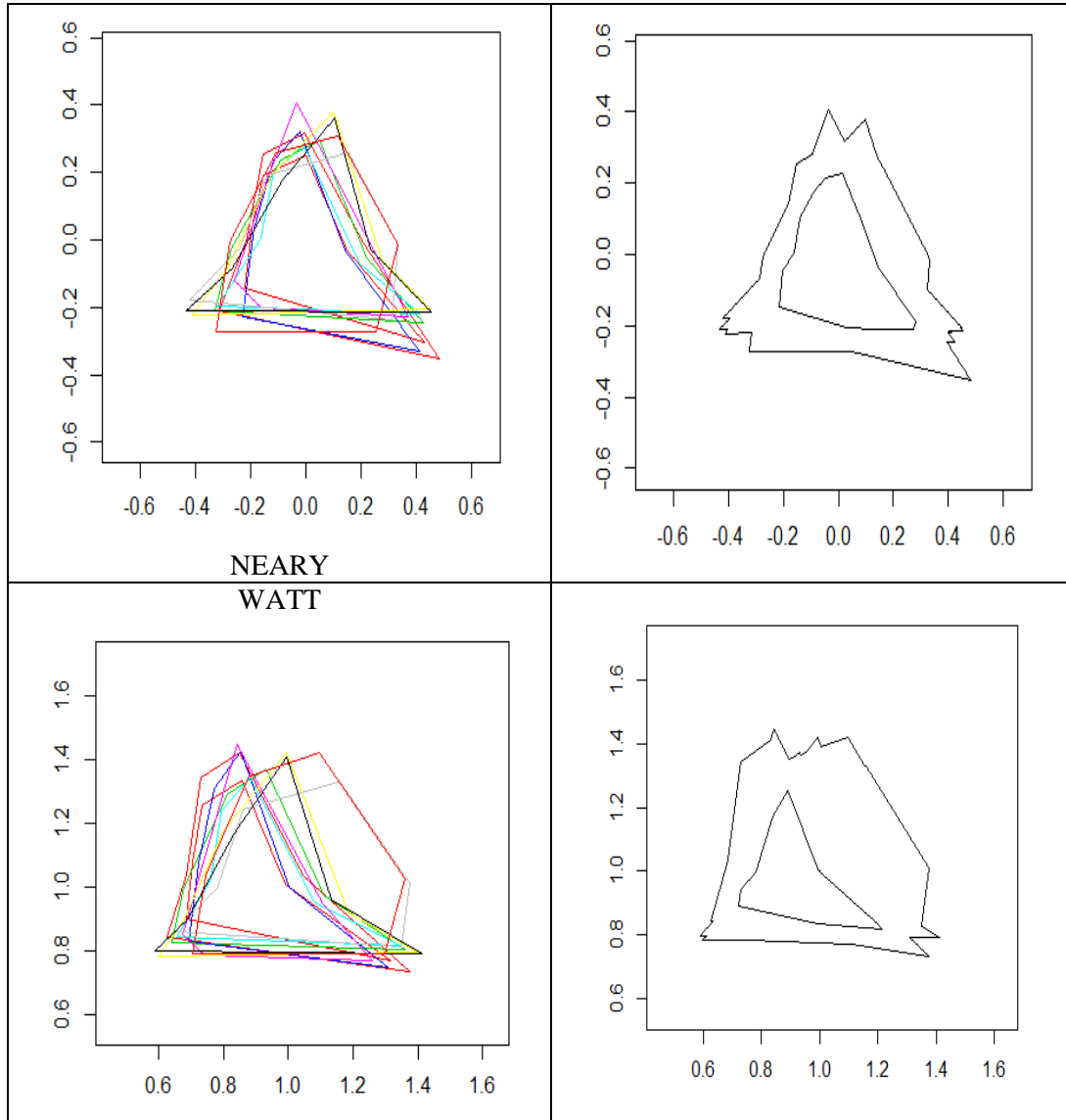
برون‌گرا

Person	Scale			
	Gerstman	Lobanov	Neary1	Watt & Fabricious
1	231553.6	168844.9	0.2457217	0.2152019
2	157808.5	107421.8	0.1527363	0.1329540
3	214101.9	145321.2	0.2110864	0.1913277
4	204814.7	126136.4	0.1853474	0.1606064
5	172711.7	112842.2	0.1563961	0.1399790
6	278465.8	195060.6	0.2006872	0.1835661
7	198563.1	208994.0	0.241374	0.2212098
8	289999.9	197601.9	0.2332502	0.2405694
9	200227.8	195894.8	0.2176764	0.1956755
10	279687.3	237832.0	0.2795693	0.2945615
Intersection area	55494.86	43925.81	0.1207648	0.08380059
Union area	488303	340124.4	0.3432165	0.3763751
Area Ratio	0.114	0.129	0.352	0.223
Rank	6	5	1	2

همان طور که ذکر شد با در نظر گرفتن بیشتر بودن میزان تناسب مساحتی به معنای مطلوبتر بودن فرایند، مقیاس نری به عنوان شیوه برتر معرفی می گردد زیرا در این مقیاس فضاهای واکه ای ۱۰ نفر تطابق بیشتری می یابند. در شکل ۱-۵ فضاهای واکه ای بهنجارسازی شده ۱۰ شخص فارسی زبان همراه با ناحیه اشتراک و اجتماع در فضاهای واکه ای در روش های مختلف بهنجارسازی نشان داده شده است.







شکل ۵-۱) نمود دیداری فضاهای واکه ای بهنجارسازی شده ۱۰ شخص (چپ) همراه با ناحیه اشتراک و اجتماع در فضاهای واکه ای (راست) در روش های مختلف بهنجارسازی

همان طور که در شکل ۵-۱ نیز مشهود است بیشترین میزان همپوشانی در فضای واکه ای اشخاص مختلف در جهت حذف و یا کمینه کردن تاثیرات فیزیولوژیکی با در نظر گرفتن شش ضلعی واکه ای در تکنیک نری ایجاد می شود. واضح است که تکنیک نری چه از نظر محاسباتی و چه از نظر شهودی بهترین تکنیک در حذف اثر جنسیت در نظام واکه ای زبان فارسی محسوب می گردد. تکنیک لبانو نیز پس از تکنیک نری در یکسان نمودن تناسب مساحتی و میزان همپوشانی فضاها موفق تر از دیگر تکنیک ها عمل نموده است. بنابراین می توان گفت که نتیجه به دست آمده از ارزیابی تکنیک های مختلف بهنجارسازی در حذف اثر فیزیولوژیکی در اشخاص مختلف بر پایه شش ضلعی واکه ای، شیوه نری را بعنوان موثرترین روش در نظام واکه ای زبان فارسی معرفی می کند.

۵-۲. تجزیه و تحلیل شیوه‌های بهنجارسازی با معیار تحلیل تشخیصی

در بررسی حاضر، از معیار تحلیل تشخیصی در دو بخش متفاوت استفاده می‌شود. در بخش اول قدرت انواع روش‌های بهنجارسازی در حفظ اطلاعات واجی و یا به عبارت فنی‌تر تفکیک کردن واژه‌ها سنجیده می‌شود که در نتیجه آن مطلوب‌ترین تکنیک در تفکیک واژه‌ای تعیین می‌گردد و در بررسی‌های که لازمه آن ایجاد تفکیک واژه‌ای بهتر است قابل استفاده و استناد قرار می‌گیرد. در بخش دوم، قدرت مقیاس نری به‌عنوان موثرترین تکنیک در حذف تاثیر فیزیولوژیکی، در حفظ اطلاعات واجی و تفکیک کردن واژه‌ها بررسی می‌گردد زیرا همانطور که قبلاً نیز به آن اشاره گردید اگرچه روش بهنجارسازی مطلوب در این بررسی در جهت حذف و یا کمینه کردن اثر شخص و جنس تعیین گردیده ولی انتظار ما از این روش آن است که بتواند اثر واژه‌ها را حفظ و آن‌ها را به خوبی تفکیک نماید و با ایجاد گشتار در داده‌ها لطمه جدی به امر تفکیک واژه‌ای وارد نسازد. از این‌رو در این بخش قدرت تفکیک واژه‌ای نری نسبت به تکنیک برتر ارزیابی می‌شود.

تحلیل تشخیصی برای مدل‌هایی با متغیر وابسته با مقادیر رسته‌ای (categorical variable) که مرتبط با یک یا چند متغیر پیشگو هستند کاربرد دارد. در مدل واژه‌ای زبان فارسی طبقات واژه‌ای که شامل ۶ گروه واژه است به‌عنوان متغیر وابسته و متغیرهای F1 و F2 به‌عنوان متغیرهای پیشگو تعریف می‌شوند. این تحلیل به تفکیک اشخاص مختلف انجام می‌گردد و در نهایت با محاسبه میانگین آن‌ها میزان قدرت تفکیک واژه‌ای اعلام می‌شود. شیوه‌ای که دارای قدرت تفکیک بالاتر برای واژه‌ها باشد در این رویکرد مطلوب ارزیابی می‌شود.

میزان قدرت تفکیک با توجه به شاخصی بنام درصد رده‌بندی درست (classified percentages correctly) بررسی می‌گردد. این شاخص درصد مواردی را که براساس مدل در نظر گرفته شده به درستی تشخیص و تفکیک می‌شوند را اعلام می‌کند. در حالت بررسی قدرت تفکیک واژه بزرگ‌تر بودن درصد رده‌بندی درست، هدف تعریف می‌گردد. در جدول ۳-۵ میانگین درصد رده‌بندی درست در اشخاص مختلف در ستون average نشان داده شده است. در این تحلیل از آنجا که مقادیر بزرگ‌تر مطلوب‌تر می‌باشند بر این اساس انواع روش‌های بهنجارسازی رتبه‌بندی شده و رتبه آن‌ها در ستون rank مشخص شده است.

جدول ۳-۵) میانگین درصد رده‌بندی درست در اشخاص مختلف در مقیاس‌های مختلف بهنجارسازی جهت بررسی قدرت

تفکیک واژه‌ای

Scale	Person										Average	Rank	
	1	2	3	4	5	6	7	8	9	10			
M1=HZ	89.5	79.6	91.0	86.2	84.9	77.6	86.7	79.5	91.5	80.5	84.70	۵,۱۲	4
M2=Mel	90.4	81.5	89.9	87.5	87.7	80.6	88.0	80.8	91.5	82.9	86.08	۴,۲۱	1
M3=bark	90.4	81.5	89.9	87.5	87.7	80.6	85.3	80.8	89.8	82.9	85.64	۳,۹۳	2
M4=Erb	89.5	80.6	88.8	87.5	87.7	79.1	85.3	80.8	89.8	82.9	85.20	۴,۰۴	3
M5=Gerstman	89.5	79.6	91.0	86.2	84.9	77.6	86.7	79.5	91.5	80.5	84.70	۵,۱۲	4
M6=lobanov	89.5	79.6	91.0	86.2	84.9	77.6	86.7	79.5	91.5	80.5	84.70	۵,۱۲	4
M7=Neary1	88.6	81.5	87.6	85.0	87.7	79.1	85.3	80.8	88.1	82.9	84.66	۳,۴۱	5
M8=Watt & Fabricious	89.5	79.6	91.0	86.2	84.9	77.6	86.7	79.5	91.5	80.5	84.70	۵,۱۲	4

بر اساس جدول ۳-۵ دو روشی که بالاترین رتبه‌ها را در تفکیک واکه‌ای به خود اختصاص می‌دهند مقیاس‌های مل و بارک هستند. البته در روش نری هر چند میانگین نسبت به سایر روش‌ها بیشتر بوده و در نتیجه رتبه بالاتری دارد ولی مقدار انحراف استاندارد در این روش نسبت به سایر روش‌ها کمتر می‌باشد که از محاسن آن به‌شمار می‌رود. در مرحله دوم که هدف ارزیابی قدرت تفکیک واکه‌ای نری نسبت به تکنیک مل است، به‌دنبال روش یا روش‌هایی هستیم که با مقیاس مل در تفکیک واکه‌ای تفاوت معناداری نداشته باشند. از این رو روش مل را با هر یک از روش‌های دیگر به‌طور مجزا از طریق آزمون پارامتری t-زوجی مقایسه می‌کنیم. فرض صفر در این آزمون‌ها برابری قدرت تفکیک واکه در دو روش مورد بررسی می‌باشد. در اینجا متذکر می‌شویم که می‌توان از ۷ آزمون، t-زوجی را به‌عنوان پس‌آزمونی برای آزمون اندازه‌های تکرار شده استفاده نمود. از آنجاکه این آزمون‌ها با علم به عدم برابری هشت روش و رد شدن فرض صفر در آزمون اندازه‌های تکرار شده انجام می‌شود، لذا سطح معناداری را به‌جای ۰/۰۵ با حاصل ۰/۰۵ تقسیم بر ۷ مقایسه می‌کنیم که ۷ تعداد آزمون‌های t-زوجی است. نتایج آزمون در جدول ۴-۵ نشان داده شده است.

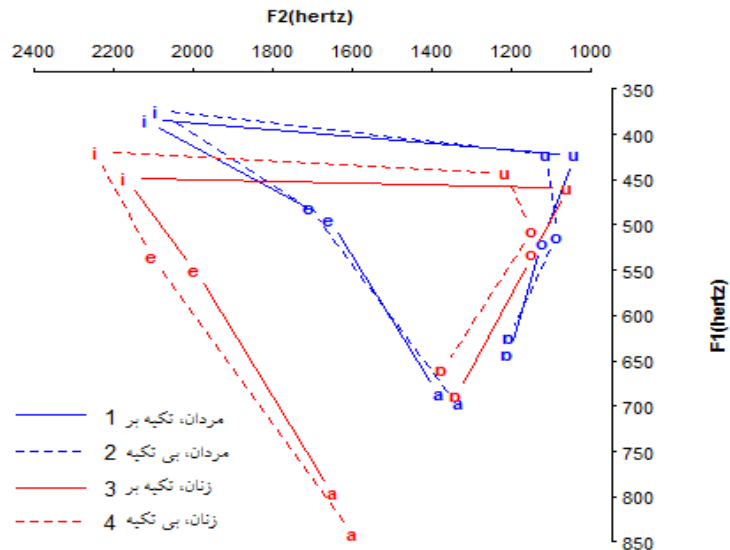
جدول ۴-۵) پس‌آزمون جهت مقایسه مقیاس برتر با سایر مقیاس‌ها در قدرت تفکیک واکه‌ای

		Mean	Std. Deviation	Std. Error Mean			Sig=.05/7=.00714
Pair 1	m2 - m1	1.380	1.259	.398	3.467	9	.00708
Pair 2	m2 - m3	.440	.957	.303	1.454	9	.17997
Pair 3	m2 - m4	.880	.914	.289	3.045	9	.01390
Pair 4	m2 - m5	1.380	1.259	.398	3.467	9	.00708
Pair 5	m2 - m6	1.380	1.259	.398	3.467	9	.00708
Pair 6	m2 - m7	1.420	1.321	.418	3.398	9	.00790
Pair 7	m2 - m8	1.380	1.259	.398	3.467	9	.00708

در جدول ۴-۵ سطح معناداری مربوط به مقایسه روش دوم (مل) با روش‌های ۳، ۴ و ۷ که به‌ترتیب مربوط به روش‌های بارک، ارب و نری است از ۰/۰۰۷۱۴ بیشتر می‌باشد. پس فرض صفر در این موارد رد نمی‌شود. بدین معنا که فرض برابری قدرت تفکیک واکه در روش مل با روش‌های بارک، ارب و نری رد نمی‌شود و می‌توان این روش‌ها را به اندازه روش مل مناسب دانست.

بنابراین به‌عنوان نتیجه‌گیری کلی از تجزیه و تحلیل روش‌های بهنجارسازی روش نری را به‌عنوان بهترین روش در کمینه کردن اثر فیزیولوژیکی شخص با حفظ تمایزات واجی و قدرت بالا در تفکیک واکه‌ای پیشنهاد می‌کنیم. حال پس از مشخص شدن بهترین روش جهت بهنجارسازی فضای واکه‌ای زبان فارسی، در ابتدا نمودار فضای واکه‌ای بر اساس تجزیه و تحلیل متغیرهای صوت‌شناختی سازه‌های اول و دوم واکه‌ها در دو هزار نمونه واکه‌ای با در نظر گرفتن دو متغیر جنسیت (مرد و زن) و بافت آوایی هجا (تکیه‌بر و بی‌تکیه) در چهار گروه مردان تکیه‌بر، مردان بی‌تکیه، زنان تکیه‌بر و زنان بی‌تکیه در شکل ۵-۲ نشان داده شده است.

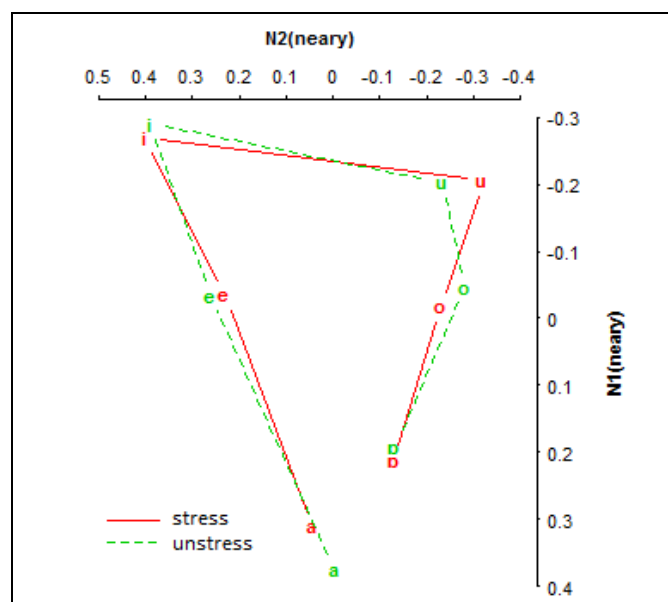
vowel polygon split by gender, context



شکل ۲-۵: نمودار فضای واکه‌ای به تفکیک چهار گروه پس از حذف داده‌های پرت

همان‌طور که در شکل ۲-۵ مشخص است فضای واکه‌ای در گروه زنان چه در حالت تکیه‌بر، چه در حالت بی‌تکیه بزرگتر از فضای واکه‌ای در گروه مردان است.

ارزیابی شیوه‌های بهنجارسازی (هرتز، مل، بارک، ارب، گرستان، لبانو، نری، وات و فابریشوز) بر اساس معیار میزان همپوشانی و تناسب فضاهای واکه‌ای دو گروه زنان و مردان صورت گرفته شده است. در میان شیوه‌های بررسی شده شیوه بهنجارسازی نری با داشتن رتبه یک در انطباق فضاهای واکه‌ای مردان و زنان به عنوان مطلوب‌ترین شیوه در زبان فارسی در جهت حذف و یا کمینه کردن اثر فیزیولوژیکی اشخاص با مزیت حفظ اطلاعات واجی معرفی شد. بر این اساس فضاهای واکه‌ای بهنجارسازی شده در زبان فارسی به صورت شکل ۳-۵ می‌باشد.



شکل ۳-۵ (فضاهای واکه‌ای بهنجارسازی شده در زبان فارسی)

لازم به ذکر است که در شکل ۳-۵ با توجه به این که با انجام فرایند بهنجارسازی، تفاوت‌های میان دو گروه جنسیتی زنان و مردان به کمینه مقدار خود رسیده است نمودار فضای واکه‌ای تنها به تفکیک تکیه‌بر و بی تکیه ترسیم شده است. این شکل تفاوت واکه‌ها را در بافت تکیه‌دار و بی تکیه نشان می‌دهد. همان گونه که مشخص است واکهٔ پسین افزوده در بافت تکیه‌دار پسین تر از بافت بدون تکیه است و بیشترین تفاوت را نسبت به سایر واکه‌ها نشان می‌دهد. در ضمن واکه افتاده پیشین در بافت بی تکیه افتاده تر از حالت تکیه‌دار است.

۶- نتیجه گیری

در این بررسی بهنجارسازی واکه‌ای در زبان فارسی با هدف حذف و یا کمینه کردن تنوعات سازه‌ای ناشی از تفاوت‌های فیزیولوژیکی میان گویندگان صورت گرفته شده است. نتیجه کلی تجزیه و تحلیل بهنجارسازی، روش نری را به عنوان بهترین روش در کمینه کردن اثر فیزیولوژیکی شخص با حفظ تمایزات واجی و قدرت بالا در تفکیک واکه‌ای پیشنهاد می‌کند. پس از بهنجارسازی فضای واکه‌ای، فضای واکه‌ای زبان فارسی نشان داده شد و مشخص گردید که فضای واکه‌ای در گروه زنان چه در حالت تکیه‌بر، چه در حالت بی تکیه بزرگتر از فضای واکه‌ای در گروه مردان نشان داده شده است.

منابع

- Adank, P. (2003). Vowel Normalization: a Perceptual-Acoustic Study of Dutch Vowels. PhD thesis. University of Nijmegen.
- Adank, P., Van Hout, R. & Smiths, R. (2004). "A comparison of vowel normalization procedures for language variation research". *Journal of the Acoustical Society of America* 116(5), 3099-3107.
- Ainsworth, W. A. (1975). "Intinsic and extrinsic factors in vowel judgement". in Auditory analysis and Perception of Speech, edited by G. Fant and M. A. A. Tatham . Pp. 103-113. London: Academic Press.
- Assman, P. F., Neary, T. M. ,& Hogan, J.T. (1982). "Vowel identification: Orthographic, perceptual, and acoustic aspects". *Journal of the Acoustical Society of America*, 71, 975-989.
- Bigham, D. (2008). Dialect Contact and Accommodation among Emerging Adults in a University Setting. Ph.D. thesis, The University of Texas at Austin.
- Clopper, C. (2009) "Computational methods for normalizing acoustic vowel data for talker differences". In *Language & Linguistic Compass* 3(6), 1430-1420.
- Daniloff, R. & Hammarberg, R. (1973). "On defining coarticulation". *Journal of Phonetics* 1, 239-348.
- Essner, G. (1947). "Recherche sur la structure des voyelles orales". *Archives Neerlandaises de Phonetique Experimentale* 20, 40-77.
- Fabricius, A. H., Watt. D. & Johnson. D. E. (2009). "A comparison of the three speaker-intrinsic vowel formant frequency normalization algorithms for sociophonetics". *Language Variation and Change* 21, 413-434
- Ferrari-Disner, S. (1980). "Evaluation of vowel normalization procedures". *Journal of the Acoustical Society of America* 67. Pp. 253-261.
- Ferrari-Disner, S. (1984). "Insights on vowel spacing. In Maddieson". *Patterns of sounds*, Pp. 136-155. Cambridge: Cambridge University Press.
- Fletcher, H. (1940). "Auditory patterns". *Review of modern physics* 12, 47-65.
- Flynn, N. (2011) "Comparing vowel formant normalization procedures". *York Papers in Linguistic Series* 2, 1-13.
- Gerstmann, L. (1968). "Classification of self-normalized vowels". *IEEE Transactions of Audio Electroacoustics* 16(1), 78-80.
- Gimson, A. C. (1989). *An Introduction to the Pronunciation of English*. 4th ed. London: Edward Arnold.

- Harrington, J. (2010). *Phonetic Analysis of Speech Corpora*. London: Wiley-Blackwell publication.
- Hindle, D. (1978). "Approaches to formant normalization in the study of natural speech". Pp. 161-72. In *Linguistic Variation: Models and Methods*". New York: Academic Press.
- Joos, M. (1948). *Acoustic Phonetics*. Language Monograph, 23. Baltimore: The Linguistic Society of America.
- Klein, W., Plomp, R. & Pols, L. W. S. (1970). "Vowel spectra, vowel spaces and vowel identification". *The Journal of Acoustical Society of America* 48, 999-1009.
- Ladefoged, P. (1976). *Three Areas of Experimental Phonetics*. Oxford: Oxford University Press.
- Ladefoged, P., and Broadbent, D. E. (1957). "Information conveyed by vowels". *Journal of the Acoustical Society of America* 29, 88-104.
- Langstrof, Ch. (2006). *Vowel Change in New Zealand English-Patterns and Implications*. PhD Dissertation. Christchurch, New Zealand: University of Canterbury.
- Lieberman, P., Crelin, E. S., & Klatt, D. H. (1972). "Phonetic ability and related anatomy of the newborn and adult man, neanderthal man and the chimpanzee". *American Anthropologist* 74, 287-307.
- Liljencrants, J. & Lindblom, B. (1972). "Numerical simulation of vowel quality systems: the role of perceptual contrast". *Language* 48, 839-862.
- Lobanov, B. M. (1971) "Classification of Russian vowels spoken by different speakers". *Journal of the Acoustical Society of America* 49, 606-608.
- Moore, B. C. J. & Glasbery, B. R. (1983). "Suggested formulae for calculating auditory filter band width and excitation patterns". *The Journal of Acoustical Society of America* 74, 750-753.
- Neary, T. M. (1978). *Phonetic Feature System for Vowels*. Indian university: Linguistic Club.
- Neary, T. M. (1989). "Static, dynamic and relational properties in speech perception". *Journal of the Acoustical Society of America* 85, 2088-2113.
- Peterson, G. E. & Barney, H. L. (1952). "Control methods used in a study of the vowels". *Journal of the Acoustical Society of America* 24, 175-84.
- Pols, L. C. W., Tromp, H. R. C. & Plomp, R. (1973). "Frequency analysis of Dutch vowels from 50 male speakers". *Journal of the Acoustical Society of America* 53, 1093-1101.
- Potter, R. K. & Steinberg, J. C. (1950). "Toward the specification of speech". *Journal of Acoustical Society of America* 22, 807-820.
- Rosner, B. S. & Pickering, J. B. (1994). *Vowel Perception and Production*. Oxford: Oxford University Press.
- Stevens, S. & Volkman, S. J. (1940). "The relation of pitch to frequency: a revised scale". *Journal of the Acoustical Society of America* 53, 329-353.
- Strange, W. (1989). "Evolving theories of vowel perception". *Journal of the Acoustical Society of America* 85, 2081-2087.
- Syrdal, A. K. (1984). "Aspects of a model of the auditory representation of American English vowels". *Speech Communication* 4, 121-135.
- Syrdal, A. K., Gopal, H.S. (1986). "A perceptual model of vowel recognition based on auditory representation of American English vowels". *Journal of the Acoustical Society of America* 79(4), 1086-1100.
- Thomas, E. & Kendall, T. (2007). NORM: The Vowel Normalization and Plotting Suite. Online resource. URL:<http://ncsl.p.lib.ncsu.edu/tools/norm>
- Thomas, E. (2002). *Instrumental Phonetics*. In *The Handbook of Language Variation and Change*. Oxford: Blackwell.
- Tranmuller, H. (1981). "Perceptual dimension of openness in vowels". *Journal of the Acoustical Society of America* 69, 1465-1475.
- Tranmuller, H. (1990). "Analytical expression for the tonotopic sensory scale". *Journal of the Acoustical Society of America* 88, 97-100.
- Watt, D. & Fabricius, A. (2002). "Evaluation of a technique for improving the mapping of multiple speaker s vowel spaces in the F1~ F2 plane". *Leeds Working Papers in Linguistics and Phonetics* 9, 159-173.
- Verbrugge, R. R. & strange, W. (1976). "What information enables a listener to map a talkers vowel space?" *Journal of the Acoustical Society of America* 60, 198-212
- Zwicker, E. (1961). "Subdivision of the audible frequency range into critical bands as a function of frequency". *Journal of the Acoustical Society of America* 33, 248.