بررسی ویژگی‌های واکه‌های زبان فارسی در حالت‌های واضح و وُکودشده و مقایسهٔ مقدماتی آن با واکه‌های زبان انگلیسی

مسعود متصدی زرندی[1]
دانشگاه علوم پزشکی تهران
شهره جلایی[3]
دانشگاه علوم پزشکی تهران
زهرا شیرژیان[5]
دانشجوی دکتری مهندسی پزشکی، دانشگاه علوم پزشکی تهران
سیده مهناز احمدی[7]
دکترای شنوایی شناسی، دانشگاه کالیفرنیای جنوبی

فرنوش جارالله‌ی[2]
دانشگاه علوم پزشکی ایران
امیرهمایون جعفری[4]
دانشگاه علوم پزشکی تهران
نینا حسینی کیونانی[6]
کارشناس ارشد زبان شناسی همگانی، دانشگاه الزهرا

امیرسالار جعفرپیشه[8]
دانشگاه علوم بهزیستی و توانبخشی و دانشگاه علوم پزشکی تهران

چکیده

درک واکه‌ها به واسطهٔ انتقال اطلاعات طیفی و زمانی صورت می‌گیرد. وجود اطلاعات طیفی برای درک واکه‌ها حائز اهمیت است، اما با تضعیف آن در سیگنال‌های ساختگی (مانند گفتار وُکودشده)، وجود اطلاعات زمانی در پوش، هنوز می‌تواند سیگنال را قابل فهم گرداند. هدف این مطالعه تعیین ویژگی‌های آکوستیکی واکه‌های واضح و وُکود شده در گویشوران فارسی‌زبان و مقایسهٔ مقدماتی آن با واکه‌های گویشور انگلیسی‌زبان، برای نخستین بار بود. ده گویشور

1. motasaddi@yahoo.com
2. jarollahi.f@iums.ac.ir
3. jalaeish@tums.ac.ir
4. amir_j73@yahoo.com
5. z.shirjiyan@gmail.com
6. nina.hosseini@student.alzahra.ac.ir
7. mahnazahmadi@gmail.com
8. jafarpisheh@gmail.com

فارسی‌زبان (۵ زن و ۵ مرد) و یک گویشور مرد انگلیسی‌زبان، واکه‌های سادهٔ زبان خود را سه بار تولید کردند. سپس واکه‌ها با استفاده از نرم‌افزار متلب (MATLAB) وُکود شدند. شکل موج، طیف‌نگاشت و نمودار LPC واکه‌های واضح و وُکودشده، توسط نرم‌افزارPraat اندازه‌گیری و سپس مقایسه و تحلیل شد. براساس نتایج این تحقیق، در گویشوران فارسی‌زبان F0 و F1 فضای واکه‌ای و دیرشِ واکه‌های واضح در زنان بیشتر از مردان بود، F2 و F3 تفاوتی نداشت، ولی شدت کمتر بود. همین‌طور، فرکانس سازه‌های اول تا سومِ فارسی با انگلیسی تفاوت داشت. فضای واکه‌ای فارسی از انگلیسی بزرگتر بود. دیرش واکه‌های انگلیسی از فارسی بیشتر، اما شدت آنها کمتر بود. در واکه‌های وُکودشدهٔ هردو زبان، F0 و فرکانس سازه‌ها قابل شناسایی نبوده است و نمودار LPC آنها برخلاف واکه‌های واضح، قله‌های گردشده با دامنه‌های کمتر داشتند. دیرش واکه‌های وُکودشده با واضح تفاوتی نداشت، اما شدت آنها در فارسی کاهش و در انگلیسی افزایش یافت. نتایج وُکودکردنِ واکه‌های زبان‌های فارسی و انگلیسی نشانگر تفاوت‌های طیفی واکه‌های وُکودشده نسبت به واکه‌های واضح بود.

کلیدواژه‌ها: واکه، ویژگی‌های آکوستیکی، زبان فارسی، زبان انگلیسی، محرک‌های وُکودشده.

۱- مقدمه

بیش از شصت و پنج سالِ پیش عقیدهٔ پژوهشگران بر آن بود که وجود نشانه‌های طیف فرکانسی[۱] برای بازشناسی[۲] انواع سیگنال‌های گفتاری، اعم از واکه، همخوان، واژه و جمله، ضروری است (لیک‌لیدر[۳] و پولاک[۴]، ۱۹۴۸؛ شانون[۵] و همکاران، ۱۹۹۵). گرچه، براساس پژوهش‌های اولیه، از میان سه ویژگی آکوستیکی گفتار واضح[۶] (ساختار ظریف[۷]، پریودیسیتی[۸] و

1. frequency spectrum cues
2. recognition
3. J. C. R. Licklider
4. I. Pollack
5. R. V. Shannon
6. clear speech

۷. ساختار ظریف (fine structure): تغییراتی که در شکل موج در داخل دوره‌های مجزا در اصوات دوره‌ای، و یا در فاصله‌های زمانی کوتاه در اصوات غیردوره‌ای رخ می‌دهد به عنوان اطلاعات ساختار ظریف در نظر گرفته می‌شود. از نظر آکوستیکی، ساختار ظریف اطلاعاتی را درباره طیف صدا می‌دهد، همچنین حاوی الگوهای سازه است (رُزن، ۱۹۹۲).

۸. پریودیسیتی (periodicity): ویژگی‌های سیگنال گفتاری که به تفکیک محرک‌های دوره‌ای از غیردوره‌ای مربوط می‌شود (رُزن، ۱۹۹۲).

پوش[۱])، ساختار سازه‌ای[۲](سازه‌ها و فرکانس پایه[۳]) را مهم‌ترین مشخصه‌های فیزیکی و ادراکی برای واکه‌ها می‌دانستند (لدِفوگد[۴] و دیسنر[۵]، ۲۰۰۵؛ ریکی‌مور[۶]، ۲۰۰۶ و محمدی و همکاران، ۱۳۹۰)، اما دیگر ویژگی‌ها، به ویژه پوش سیگنال گفتاری، نیز از اهمیت بسیاری برخوردارند (شانون و همکاران، ۱۹۹۵). با ایجاد تغییرات نامطلوب در اجزای مختلف سیگنال‌های گفتاری، از جمله تضعیف اطلاعات طیفی، به اهمیت وجود پوش سیگنال پی برده شده است. در واقع، پرکاربردترینِ این نوع محرک‌های گفتاری، سیگنال وُکودشده[۷] (VS) است. گفتار وُکودشده، صوت گفتاریِ ساخته‌شده‌ای است که در آن یک یا چند باند حامل جایگزین سیگنال گفتاری می‌شود، به‌طوری‌که پوش دامنهٔ سیگنال اصلی حفظ می‌شود اما اطلاعات طیفی[۸] آن از بین می‌رود (شانون و همکاران، ۱۹۹۵؛ ریکی‌مور، ۲۰۰۶). درک اصوات گفتاریِ وکودشده شبیه به درک صدا توسط افرادی است که عمل کاشت حلزون را انجام داده‌اند، یعنی صداها شبیه به پچ‌پچ خشن و تیز است. بنابراین، مشخصهٔ اصلیِ گفتارِ وُکودشده، کاهش شدید اطلاعات فرکانسیِ سیگنالِ گفتاریِ اصلی است. تعداد باندها در بازشناسی سیگنال ساخته‌شده، متغیرِ بسیار حائز اهمیتی است، به‌طوری‌که افزایش باندهای فرکانسی[۹] از یک تا چهار، قابلیتِ فهم VS را افزایش می‌دهد. بنابراین VS تک باند بسیارکم اما VS سه و چهار باند نسبتاً قابل فهم است (شانون و همکاران، ۱۹۹۵).

از سال‌های ۱۹۹۰، براساس مطالعات سایکوآکوستیک، مدل‌های تئوریکی که از شواهد فیزیولوژیکی ساخته شده و بالاخره براساسِ شواهد موفقیت‌آمیز بسیاری از بیماران کاشت حلزون شده، این فرضیه تقویت شد که اطلاعات زمانی اصوات گفتاری (نسبت به اطلاعات طیفی این صداها) به‌تنهایی می‌تواند در درک گفتار مؤثر باشد (رُزن، ۱۹۹۲). این مطالعات

---

۱. پوش (envelope): پوش سیگنال اشاره به تغییراتی در دامنهٔ کلی موج گفتاری دارد (رُزن، ۱۹۹۲).

2. formant structure
3. fundamental frequency
4. P. Ladefoged
5. S. F. Disner
6. H. Riquimaroux
7. vocoded speech
8. spectral information

۹. باند فرکانسی (frequency band): محدوده‌ای از فرکانس‌ها که بین دو حد خاص قرار دارد (عادل‌قهرمان و همکاران، ۱۳۸۵).

اطلاعات بسیار ارزشمندی را درباره نحوهٔ پردازش گفتار و درک آن در اختیار گذاشته‌اند (ریکی‌مور، ۲۰۰۶)، و کماکان علاقه به پژوهش بر روی این ویژگی‌ها ادامه دارد.

امروزه به خوبی روشن شده است که به‌رغم تغییرات زیادی که در رسیدن اصوات گفتاری به گوش رخ می‌دهد، می‌توان گفتار را به خوبی دریافت و درک کرد. در آزمایشگاه نیز، با ایجاد تغییرات شدید و تجزیهٔ سیگنال، درک گفتار بسیار قوی است. وجود نشانه‌های آکوستیکی برای شناخت و درک گفتار، ضروری نیست زیرا در سیگنال گفتاری حشو فراوانی وجود دارد و بنابراین نشانه‌های آکوستیکی متعددی برای هویت آوایی وجود دارد (اسکات[۱]، ۲۰۰۵). افراد برای تشخیص سیگنال آکوستیکی که یکی از نشانه‌هایش حذف شده باشد، بر دیگر علائم موجود در آن سیگنال اتکا می‌کنند. هرچه انتخاب فرکانسی در سیستم شنوایی ضعیف‌تر شده باشد، نقش فاکتورهای زمانی به‌طور واضحی افزایش می‌یابد (رُزن، ۱۹۹۲). بنابراین سیگنال گفتاری با اطلاعات طیفی تضعیف‌شده، هنوز می‌تواند قابل فهم باشد (مایلدنر[۲]، ۲۰۰۸). این موضوع مربوط به این است که درک ما از گفتار، قوی و انعطاف‌پذیر است، و توسط مسیرهای پردازشی چندگانه و موازی حمایت می شود (اسکات، ۲۰۰۵؛ هرویس‌آدلمن[۳] و همکاران، ۲۰۱۱).

تاکنون پژوهش‌های متعددی بر روی ویژگی‌های آکوستیکی واکه‌های زبان انگلیسی (لی[۴]، ۱۹۷۳؛ رُزن[۵]، ۱۹۹۲؛ هلینبرند[۶] و همکاران، ۱۹۹۵؛ چو[۷] و لده‌فوگد، ۱۹۹۹؛ بی‌کن[۸] و اورلیکوف[۹]، ۲۰۰۰؛ لده‌فوگد و دیسنر، ۲۰۰۵؛ برمن[۱۰]، ۲۰۰۷) و وُکودکردن آنها (شانون و همکاران، ۱۹۹۵ و ۱۹۹۸؛ اویدا[۱۱] و ناکاجیما[۱۲]، ۲۰۰۸؛ چن[۱۳] و لوئیزو[۱۴]، ۲۰۱۱) انجام شده است. بیشتر پژوهش‌های صورت‌گرفته برروی محرک‌های گفتاری وُکودشده در زبان‌های

1. S. K. Scott
2. V. Mildner
3. A. G. Hervais-Adelman
4. W. A. Lea
5 . S. Rosen
6. J. Hillenbrand
7. T. Cho
8. R. J. Baken
9. R. F. Orlikoff
10. A. Behrman
11. K. Ueda
12. Y. Nakajima
13. F. Chen
14. P. C. Loizou

انگلیسی و ژاپنی بوده است (شانون و همکاران، ۱۹۹۸؛ فو[1] و شانون، ۱۹۹۹؛ لیو[2] و همکاران، ۲۰۱۲؛ شوارتز[3] و کَترجی[4]، ۲۰۱۲).

در زبان فارسی پژوهش‌های چندی به بررسی فرکانس سازه‌های اول و دوم (F1، F2)، و فضای واکه‌ای واکه‌های واضح پرداخته‌اند (سپنتا، ۱۳۷۷؛ بی‌جن‌خان، ۲۰۰۰؛ انصارین، ۲۰۰۴؛ صالحی و همکاران، ۱۳۸۷ و محمدی و همکاران، ۱۳۹۰)، اما در واقع می‌توان بیان داشت که برای اولین بار وُکودکردن واکه‌های فارسی در این پژوهش انجام شده است. به دلیل فقدان محرک‌های گفتاری فارسی وُکودشده، تاکنون هیچ پژوهشی در زبان فارسی با این نوع محرک‌ها انجام نشده است. بالطبع هیچ پژوهش مقایسه‌ای بین شباهت‌ها و تفاوت‌های آکوستیکی واکه‌ها در زبان‌های فارسی و انگلیسی، در حالت‌های واضح و وُکودشدهٔ آنها صورت نگرفته است.

ازآنجاکه ویژگی‌های آکوستیکی، واجی و واژگانی زبان فارسی با زبان انگلیسی و همچنین دیگر زبان‌ها متفاوت است (محمدی و همکاران، ۱۳۹۰؛ رُزن، ۱۹۹۲؛ اویدا و ناکاجیما، ۲۰۰۸؛ اِلرمایر[5] و همکاران، ۲۰۱۲؛ چن[6]، ۲۰۱۲)، نیاز به پژوهش‌های آکوستیکی بنیادی در واج‌های زبان فارسی برای اهداف گوناگون نمایان است. ازجمله لازم است انطباق پروتزهای کاشت که با ویژگی‌های زبان انگلیسی طراحی‌شده است در کاربران فارسی‌زبان به نحو شایسته انجام شود، و یا نحوهٔ پردازش و درک شنوایی گفتار فارسی مورد بررسی قرار گیرد. بر این اساس، هدف پژوهش حاضر تعیین ویژگی‌های واکه‌های زبان فارسیِ واضح و وُکودشده و همچنین مقایسهٔ توصیفی مقدماتی با واکه‌های مشابه در زبان انگلیسی است.

1. Q. J. Fu
2. C. Liu
3. K. C. Schvartz
4. M. Chatterjee
5. W. Ellermeier
6. F. Chen

۲- روش بررسی

پژوهش حاضر به بررسی و مقایسهٔ مقدماتیِ شش واکهٔ ساده در زبان فارسی شامل /u/، /o/، /ɑ/، /a/، /e/، /i/[۱] و هشت واکهٔ زبان انگلیسی امریکایی شامل /u/، /ʊ/، /ʌ/، /ɑ/، /a/، /ɛ/، /ɪ/، /i/[۲] می‌پردازد. واکه‌های انگلیسی مشابه واکه‌های فارسی عبارت از /u/، /ʊ/، /ɑ/، /a/، /ɪ/، /i/ هستند که به‌ترتیب در واژه‌های heed-hid-had-hod-hood-who'd به‌کار رفته‌اند. برای تعیین ویژگی‌های آکوستیکی واکه‌های واضح در زبان فارسی از پنج گویشور زن (میانگین سنی: ۲۵/۶±۱/۶۷ سال) و پنج گویشور مرد (میانگین سنی: ۲۸±۲/۴۵ سال) بومی زبان فارسی با گونهٔ معیار که دارای شنوایی طبیعی و حنجرهٔ کاملاً طبیعی بودند، که به‌طور تصادفی از بین دانشجویان کارشناسی ارشد انتخاب شده بودند استفاده شد. گویشوران واکه‌ها را در قالب /hVd/ به‌صورت هجاهای بی‌معنا با سه بار تکرار، خارج از جمله و بدون آهنگ و تکیه تولید کردند و از طریق ضبط صدای آنها در اتاق اکوستیک کلینیک شنوایی‌شناسی دانشگاه علوم پزشکی ایران و با برنامه Praat واکه‌های واضح تولید شد، و در مجموع ۱۸۰ نمونهٔ آوایی فارسی به‌دست آمد: {مجموع ۱۸۰ نمونه (۶ واکه * ۳ بار تکرار * ۱۰ گویشور)}.

برای تعیین ویژگی‌های آکوستیکیِ واکه‌های واضح در زبان انگلیسی، صدای یک گویشور مرد ۲۵ سالهٔ انگلیسی‌زبان با لهجهٔ غرب آمریکا و بومی ایالت یوتا، که در دسترس بود، در استودیوی صدابرداری با سرعت نمونه ۴۸[۳] کیلوهرتز و با سه بار تکرار که ۸ واکهٔ زبان انگلیسی را تولید کرد، ضبط شد و بدین‌ترتیب ۲۴ نمونهٔ آوایی به‌دست آمد: {مجموع ۲۴ نمونه (۸ واکه * ۳ بار تکرار * ۱ گویشور)}. سپس همه واکه‌های تولیدشده در هردو زبان در گویشوران مرد وُکود شد. به دو دلیل فقط صدای گویشوران مرد برای وُکودکردن واکه‌ها انتخاب شد: اولاً، گویشور زن انگلیسی‌زبان برای مقایسه با گویشوران زن فارسی‌زبان در دسترس نبود. ثانیاً، در پژوهش‌های مختلف بر روی محرک‌های وُکودشده از گویشوران زن یا مرد استفاده شده است، چون جنس صدا تأثیری بر وُکودکردن ندارد. بدین‌ترتیب ۲۷۰ نمونهٔ آواییِ واکه‌های فارسی وُکودشده {مجموع ۲۷۰ نمونه (۶ واکه * ۳ بار تکرار * ۵ گویشور مرد * ۳ باند)} و ۵۴ نمونهٔ

---

۱. به‌ترتیب در واژه‌های hid-hed-had-hɑd-hod-hud

۲. به‌ترتیب در واژه‌های heed-hid-head-had-hud-hod-hood-who'd

3. sampling rate

آواییِ واکه‌های انگلیسی وُکودشده: {مجموع ۵۴ نمونه (۸ واکه * ۳ بار تکرار * ۱ گویشور مرد * ۳ باند)} به‌دست آمد.

۲-۱- نحوهٔ انجام کار

با به‌کارگیری نرم‌افزار «متلب» سیگنال‌ها به چهار، ده و بیست باند تقسیم شدند. علت انتخاب این تعداد باند براساس مطالعات قبل بود که چهار باند حداقلِ تعداد باندی است که سیگنال‌های گفتاری تاحدودی قابل فهم است (ریکی‌مور، ۲۰۰۶) و بالای هشت باند بسیار زیاد قابل فهم است و به همین دلیل ده باند انتخاب شد و بیست باند نیز بسیار مشابه سیگنال‌های واضح است. نحوهٔ تقسیم فرکانسی به‌صورت لگاریتمی براساس فرمول گرین وود[1] (۱۹۹۰) انجام شد.

پس‌از کسب حالت وُکودشدهٔ واکه‌ها که با استفاده از نرم‌افزار متلب صورت گرفت، کلیهٔ نمونه‌های آوایی فارسی و انگلیسی در هردو حالت واضح و وُکودشده در قالب فایل‌های wave در هردو بخش پژوهش با استفاده از نرم‌افزار Praat ویرایش ۵، ۲، ۳۴ شکل موج و طیف‌نگاشت آنها به‌دست آمد و مورد تجزیه و تحلیل صوت شناختی قرار گرفت. بدین‌صورت که محدودهٔ وضعیت ثبات[2] درحدود ۵۰ میلی‌ثانیهٔ وسط طیف‌نگاشت مشخص شد. در شکل ۱ نمونه‌ای از طیف‌نگاشت و موج صوتی واکهٔ /e/ در واژهٔ /hed/ نشان داده شده است، که در این شکل محدودهٔ اندازه‌گیری‌شده برای اندازه‌گیریِ فرکانسِ سازه‌ایِ واکه مشخص است، به‌طوری‌که مناطق تیرهٔ افقی سازه‌های این واکه‌اند، و از پایین به بالا به‌صورت F1، F2، F3 نام‌گذاری می‌شوند. فرکانس پایه و شدت نیز در همین منطقه مشخص‌شده‌اند که با استفاده از منوی pitch و intensity به‌دست آمدند. از آنجا که سازه‌های بالاتر از F3 تأثیر چندانی در تعیین کیفیت واکه و تشخیص آن ندارند، بنابراین در بررسی‌های صوت‌شناختی، فرکانس پایه و فرکانس سازه‌های اول تا سوم اندازه‌گیری می‌شوند (بورسما[3] و وینینگ[4]، ۲۰۰۷). بنابراین مشخصه‌های F0، F1، F2، F3، فضای واکه‌ای، شدت و دیرش در واکه‌های فارسیِ واضحِ

1. D. D. Greenwood
2. steady-state time
3. P. Boersma
4. D. Weenink

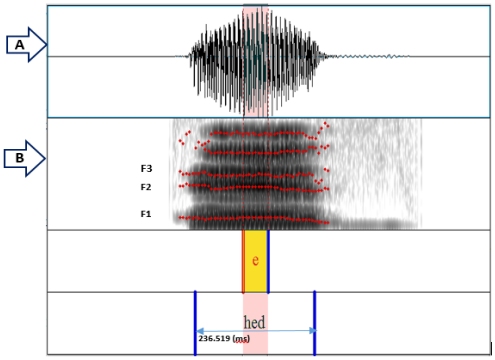
گویشوران زن و مرد و واکه‌های انگلیسیِ واضح در گویشور مرد، و مشخصه‌های شدت، دیرش و نمودار LPC[1] واکه‌های فارسی و انگلیسی وُکودشده در گویشوران مرد فارسی‌زبان و گویشور مرد انگلیسی‌زبان به‌دست آمده و مقایسه شدند. LPC طیفی است یک‌دست با قله‌هایی که نمایندهٔ سازه‌های واکه یا همخوان‌های شبه‌واکه است (زویکر[2] و ترهارت[3]، ۱۹۸۰) و با مقایسهٔ LPC واکه‌های واضح و وُکودشده می‌توان به تفاوت ساختار سازه‌ای این دو نوع محرک پی برد.

مقادیر به‌دست‌آمده از فرکانس سازه‌های اول و دوم در پژوهش حاضر برطبق الگوی پیشنهادیِ زویکر و ترهارت (۱۹۸۰) توسط دی‌تردینگ در اکسل از هرتز به بارک تبدیل شد، زیرا از نظر درک دیداری نمودار بارک نسبت به نمودار هرتز نمایش بهتری دارد (شکل ۲). در این شکل، محور افقی جایگاه زبان و محور عمودی ارتفاع زبان را در تولید واکه نشان داده و مساحت آن برحسب هرتز مربع به‌دست آمده است. با توجه به فرمول پیشنهادی کنت[4] و کیم[5] (۲۰۰۸) مساحت فضای واکه‌ای به‌روش زیر محاسبه شد:

Area: 0. 5 *{(/i/F2 * /e/F1 + /e/F2 * /a/F1+ /a/F2 * /ɑ/F1 + /ɑ/F2 * /o/F1+ /o/F2 * /u/F1 + /u/F2 * /i/F1) – (/i/F1 * /e/F2 + /e/F1 * /a/F2+ /a/F1 * /ɑ/F2 + /ɑ/F1 * /o/F2+ /o/F1 * /u/F2 + /u/F1 */i/F2)}

نهایتاً برای توصیف مقادیر F3,F2,F1,F0 و دیرش و شدت از شاخص‌های آمار توصیفی (میانگین و انحراف معیار)، برای مقایسهٔ ویژگی‌های آکوستیکی واکه‌های واضح در بین دو جنس زن و مرد از تی-مستقل[6]، برای مقایسهٔ این ویژگی‌ها در بین شش واکهٔ فارسی در نمونه‌های آوایی زنان و مردان از روش آماری آنوا[7]، برای مقایسهٔ ویژگی‌های دیرش و شدت واکه‌ها در حالت‌های واضح و وِکودشده در نمونه‌های آوایی فارسی از آزمون تی زوجی[8] و در نمونه‌های آوایی انگلیسی از آزمون من-ویتنی[9] استفاده شد. سطح معناداری P<0. 005 درنظر گرفته شد و نهایتاً داده‌ها با نرم‌افزار SPSS ویرایش ۲۱ توصیف و تحلیل شدند.

1. linear predictive coding
2. E. Zwicker
3. E. Terhardt
4. R. D. Kent
5. Y. Kim
6. t-test
7. ANOVA
8. paired t-test
9. Mann-Whitney



شکل ۱- شیوهٔ اندازه‌گیریِ واکهٔ [*e*] در تلفظ واژهٔ «*hed*» توسط گویشور مرد (*A*: موج صوتی، *B*: طیف‌نگاشت)

۳- نتایج

پنج گویشور زن (میانگین سنی: ۱/۶۷±۲۵/۶ سال) و پنج گویشور مرد (میانگین سنی: ۲/۴۵±۲۸ سال) مشارکت‌کننده در این پژوهش، بومی زبان فارسی با گونهٔ معیار و یک گویشور مرد ۲۵ سالهٔ انگلیسی‌زبان با لهجهٔ غرب امریکا و بومی ایالت یوتا، دارای سطح شنوایی طبیعی و حنجرهٔ کاملاً طبیعی بودند. با تولید واکه‌های واضح توسط گویشوران فارسی و انگلیسی‌زبان، ویژگی‌های آکوستیکی شامل F0, F1, F2, F3، دیرش و شدت به‌دست آمد (جداول ۱ و ۲ به‌ترتیب برای واکه‌های فارسی و انگلیسی). با وُکودکردن واکه هیچ‌یک از ویژگی‌های سازه‌ها وجود نداشته و تکیه بر دیرش و شدت آنهاست. مقایسهٔ نتایج واکه‌های واضح و وُکودشده در گویشوران مردِ فارسی و انگلیسی‌زبان در جدول ۳ گزارش شده است.

۳-۱- ویژگی‌های واکه‌های واضح در گویشوران فارسی‌زبان

در هردو گویشوران مرد و زن F0 در واکهٔ /i/ و سپس در واکهٔ /u/ بیشترین مقدار بوده و فرکانس F0 واکهٔ /ɑ/ کمترین مقدار را داشته است. تحلیل F0 در نمونه‌های آوایی مردان (F (5. 359)=0. 527, P=0. 76) و در نمونه‌های آوایی زنان (F (5. 387)=1/1, P=0. 37) در تولید شش واکه از نظر آماری معنادار نبود، اما مقدار F0 در زنان بیشتر از مردان بود (F(1. (179) = 1062. 38 , P<0. 0001) (جدول ۱).

جدول ۱- میانگین و انحراف معیار فرکانس پایه، سازه‌های اول، دوم، سوم، دیرش و شدت واکه‌های گویشوران زن و مرد فارسی‌زبان

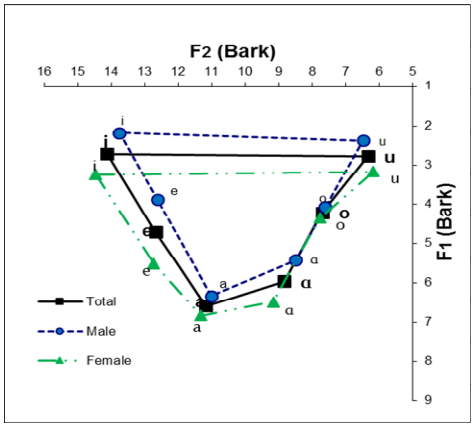
		/hid/	/hed/	/had/	/hɑd/	/hod/	/hud/
F0	مرد	۱۳۸ (۱۹/۵۵)	۱۳۴ (۱۸/۷۱)	۱۳۰ (۱۸/۰۱)	۱۲۹ (۱۹/۲۳)	۱۳۳ (۱۸/۵۲)	۱۳۷ (۱۹/۷۲)
	زن	۲۳۴ (۱۹/۸۷)	۲۲۳ (۱۴/۶۸)	۲۲۶ (۲۴/۸۷)	۲۲۰ (۲۰/۱۴)	۲۲۵ (۱۷/۸۰)	۲۲۳ (۱۸/۱۹)
	مجموع	۱۸۶ (۵۲/۴۹)	۱۷۹ (۴۸/۴۴)	۱۷۸ (۵۳/۴۷)	۱۷۴ (۵۰/۱۶)	۱۷۹ (۵۰/۴۵)	۱۸۵ (۵۲/۱۶)
F1	مرد	۲۲۳ (۴۱/۸۳)	۴۰۴ (۸۰/۳۲)	۶۹۳ (۱۱۹/۷۴)	۵۸۲ (۶۳/۹۲)	۴۲۶ (۶۹/۶۱)	۲۴۳ (۶۲/۲۵)
	زن	۳۳۵ (۷۱/۸۱)	۵۹۳ (۶۸/۲۹)	۷۵۹ (۱۷۳/۹۴)	۷۱۲ (۱۳۸/۲۵)	۴۵۷ (۱۱۸/۷۴)	۳۲۷ (۶۹/۸۷)
	مجموع	۲۷۹ (۸۱/۱۱)	۴۹۸ (۱۲۰/۸۵)	۷۲۶ (۱۵۰/۴۷)	۶۴۷ (۱۲۴/۸۹)	۴۴۱ (۹۶/۸۸)	۲۸۵ (۷۷/۶۱)
F2	مرد	۲۲۱۶ (۷۳/۱۸)	۱۸۶۰ (۲۵/۹۹)	۱۴۶۱ (۱۰۱/۴۲)	۹۹۹ (۸۳/۳۴)	۸۶۷ (۱۰۶/۲۶)	۷۱۳ (۱۱۲/۷۰)
	زن	۲۴۸۲ (۹۹۲/۱۹)	۱۸۸۹ (۴۵۴/۹۰)	۱۵۲۷ (۱۴۲/۲۸)	۱۱۰۶ (۲۱۹/۰۶)	۸۹۰ (۱۹۴/۲۰)	۶۷۴ (۱۱۹/۲۲)
	مجموع	۲۳۴۹ (۷۰۴/۳۴)	۱۸۷۴ (۳۱۶/۹۱)	۱۴۹۴ (۱۲۵/۹۹)	۱۰۵۲ (۱۷۱/۶۶)	۸۷۸ (۱۵۶/۴۲)	۶۹۴ (۱۱۵/۷۰)
F3	مرد	۲۸۳۱ (۳۱۹/۵۹)	۲۴۱۵ (۱۶۱/۸۰)	۲۲۲۶ (۲۷۱/۰۳)	۲۷۱۳ (۳۸۸/۷۴)	۳۰۴۳ (۴۶۸/۳۶)	۳۰۸۱ (۲۵۵/۵۴)
	زن	۳۴۳۱ (۲۸۲/۶۶)	۲۷۳۷ (۳۰۴/۹۲)	۲۰۶۶ (۳۸۹/۱۶)	۲۱۰۰ (۷۸۶/۵۸)	۲۴۶۹ (۹۷۴/۹۴)	۳۲۴۷ (۴۶۳/۰۹)
	مجموع	۳۱۳۱ (۴۲۵/۴۲)	۲۵۷۶ (۲۹۰/۳۴)	۲۱۴۶ (۳۳۹/۴۲)	۲۴۰۷ (۶۸۴/۷۱)	۲۷۵۶ (۸۰۶/۲۵)	۳۱۶۴ (۳۷۷/۲۰)
دیرش	مرد	۱۶۸ (۴۰/۲۸)	۱۷۵ (۴۶/۸۹)	۱۸۶ (۴۵/۱۸)	۱۹۷ (۴۴/۸۳)	۱۸۲ (۴۰/۷۹)	۱۸۵ (۴۴/۶۴)
	زن	۲۲۲ (۵۵/۹۴)	۲۱۱ (۴۶/۲۶)	۲۱۰ (۴۵/۰۹)	۲۲۶ (۴۶/۳۰)	۲۱۷ (۵۱/۳۳)	۲۲۰ (۵۲/۶۵)
	مجموع	۱۹۵ (۵۵/۲۴)	۱۹۳ (۴۹/۱۸)	۱۹۸ (۴۶/۰۰)	۲۱۱ (۴۷/۱۷)	۲۰۰ (۴۸/۸۰)	۲۰۲ (۵۱/۱۸)
شدت	مرد	۸۳/۸۷ (۲/۹۲)	۸۲/۸۷ (۲/۴۷)	۸۲/۲۰ (۳/۳۹)	۸۲/۷۳ (۲/۴۹)	۸۲/۹۳ (۲/۴۹)	۸۳/۰۰ (۳/۴۰)
	زن	۸۱/۰۰ (۳/۰۷)	۷۸/۸۰ (۲/۲۷)	۷۹/۴۷ (۳/۰۰)	۷۹/۹۳ (۱/۹۴)	۷۹/۶۰ (۲/۲۹)	۸۰/۲۷ (۳/۱۳)
	مجموع	۸۲/۴۳ (۳/۲۹)	۸۰/۸۳ (۳/۱۱)	۸۰/۸۳ (۳/۴۴)	۸۱/۳۳ (۲/۶۲)	۸۱/۲۷ (۳/۰۰)	۸۱/۶۳ (۳/۵۰)

نتایج حاصل از اندازه‌گیریِ F1 در نمونه‌های آواییِ گویشوران زن نشان می‌دهد، واکهٔ /u/ کمترین و واکهٔ /a/ بیشترین فرکانس را به خود اختصاص داده‌اند (F (84. 13027)=40. 183, P<0. 0001)، و در نمونه‌های آواییِ گویشوران مرد واکهٔ /i/ کمترین و واکهٔ /a/ بیشترین

فرکانس را داشته‌اند (F (84. 5890)=87. 070, P<0. 0001). در مجموع، مقدار F1 در گروه زنان با درجهٔ اطمینان ۹۵ درصد بیشتر از مردان بود (t (178)=-3. 501, P<0. 001) (جدول ۱).

نتایج F2 در نمونه‌های آوایی هردو گویشوران زن و مرد نشان داد که واکهٔ /u/ کمترین و واکهٔ /i/ بیشترین فرکانس را دارد و این مقادیر از نظر آماری معنادار، و به‌ترتیب در مردان (F (5. 89)=680. 016, P<0. 0001) و در زنان (F (5. 89)= 31. 016, P<0. 0001)، بودند. آزمون تی-مستقل اختلاف معناداری بین F2 گویشوران زن و مرد حتی در تک‌تک واکه‌ها نشان نداد (t (178)=-0. 807, P=0. 42).

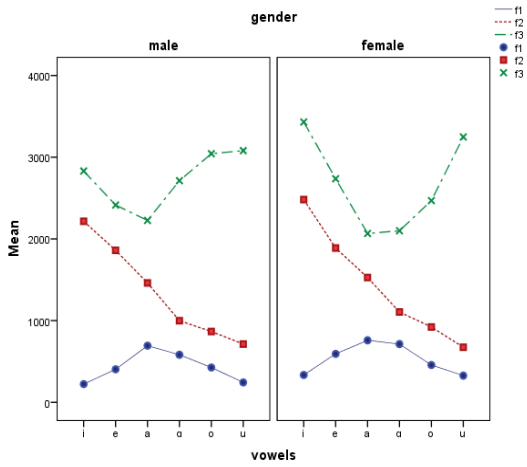
براساس مقادیر به‌دست‌آمده از فرکانس دو سازهٔ اول واکه‌ها و به‌منظور درک دیداری بهتر آنها، فضای واکه‌ای برای هردو گروه مرد و زن و میانگین کلی ترسیم شد (شکل ۲). مساحت فضای واکه‌ای محاسبه شده در این شکل در گویشوران زن بیشترین مقدار (۸۸۳۹۷۱ هرتز مربع) نشان داده است. سپس مساحت مجموع (۸۲۸۱۸۳ هرتز مربع)، و نهایتاً فضای واکه‌ای گویشوران مرد (۵۸۲۵۶۹ هرتز مربع) کمترین مقدار را به خود اختصاص داده‌اند.



شکل ۲- فضای واکه‌ای گویشوران زن و مرد فارسی‌زبان و میانگین کلی برحسب بارک

نتایج به‌دست‌آمده برای F3 نشان می دهند که در نمونه‌های آواییِ گویشوران زن، واکهٔ /a/ کمترین و واکهٔ /i/ بیشترین سطح معناداری آماری را از خود نشان داده است (F (5. 89)= 14. 064, P<0. 0001)، اما در نمونه‌های آوایی گویشوران مرد واکهٔ /a/ کمترین اما واکهٔ /u/

بیشترین فرکانس را داشت که از نظر آماری معنادار بود (F $_{(5, 89)}$= 16. 480, P<0. 0001). آزمون تی-مستقل برای F3 مجموع بین گویشوران زن و مرد اختلاف معناداری نداشت (t $_{(178)}$=0. 455, P=0. 65) (جدول ۱). در بررسی دیداریِ میانگین فرکانس‌های سازه‌های واکه‌های فارسی، نمودار خطی ترسیم شد (شکل ۳). با توجه به این نمودار، به‌طورکلی زنان فرکانس سازه‌ای بالاتری نسبت به مردان داشتند.



شکل ۳- نمودار خطی فرکانس‌های سازه‌های اول، دوم و سوم واکه‌های زبان فارسی

کمترین دیرش واکه‌ها در نمونه‌های آوایی گویشوران زن براساس جدول ۱ به‌ترتیب /a/ و /e/ و بیشترین دیرش متعلق به واکهٔ /ɑ/ بود (F $_{(5, 89)}$= . 239, P=0. 944)، ولی در نمونه‌های آوایی گویشوران مرد /i/ کمترین و /ɑ/ بیشترین دیرش را داشت (F $_{(5, 89)}$= . 770, P=0. 574)، که البته این تفاوت‌ها از نظر آماری معنادار نبود. در مجموع، دیرش واکه‌های زنان بیش از مردان بود (t $_{(178)}$=-5. 137, P<0. 001).

کمترین شدت واکه‌ها در نمونه‌های آوایی گویشوران زن، براساس جدول ۱، متعلق به واکهٔ /e/ و بیشترین شدت متعلق به واکهٔ /i/ بود (F $_{(5, 89)}$= 1. 195, P=0. 319)، ولی در نمونه‌های آوایی گویشوران مرد، واکهٔ /a/ کمترین و واکهٔ /i/ بیشترین شدت را داشتند (F $_{(5, 89)}$=0. 525, P=0. 757)، اما این تفاوت‌ها از نظر آماری معنادار نبود. در حالت کلی، واکه‌های گویشوران

مرد در مقایسه با واکه‌های گویشوران زن شدت بیشتری را نشان می‌دهند $F_{(1, 179)} = 56.302, P<0.0001$).

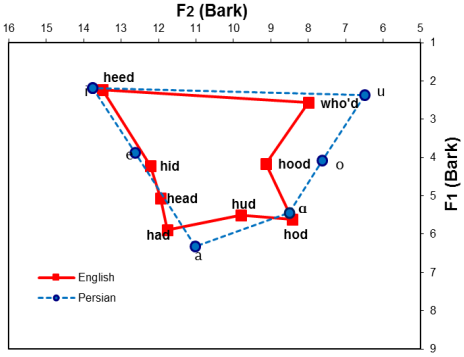
۳-۲- مقایسهٔ ویژگی‌های آکوستیکی واکه‌های واضح گویشوران مرد فارسی و انگلیسی‌زبان

ویژگی‌های آکوستیکی هشت واکهٔ سادهٔ زبان انگلیسی تولیدشده توسط گویشور مرد (جدول ۲)، با مطالعات پژوهشگران دیگر در زبان انگلیسی با نمونه‌های بیشتر مطابقت دارد (لده‌فوگد و دیسنر، ۲۰۰۵؛ هیلنبرند و همکاران، ۱۹۹۵). مقایسهٔ ویژگی‌های آکوستیکی واکه‌های واضح در دو زبان نشان داد که بیشترین فرکانس پایه متعلق به واکهٔ /u/ در واژهٔ Who'd و کمترین مربوط به واکهٔ /ɪ/ در واژهٔ hid بود. بیشترین مقدار F0 در هردو زبان تقریباً مشابه بود اما کمترین F0 در زبان انگلیسی با زبان فارسی تطابق نداشت. ضمناً در مجموعِ همه واکه‌ها، F0 مردان فارسی‌زبان بیشتر از F0 مرد انگلیسی‌زبان بود. بیشترین فرکانس سازهٔ اول متعلق به /a/ در واژهٔ had، و کمترین فرکانس سازهٔ اول متعلق به واکهٔ /i/ در واژهٔ heed بود که هردو با زبان فارسی تطابق داشت.

جدول ۲- میانگین و انحراف معیار فرکانس پایه، سازه‌های اول، دوم، سوم، دیرش و شدت واکه‌های زبان انگلیسی توسط گویشور مرد (n=1)

	Heed	hid	head	had	hud	hod	Hood	who'd
F0	۱۰۴ (۱/۷۳)	۱۰۲ (۲/۵۲)	۱۰۶ (۲/۳۱)	۱۰۵ (۱/۱۶)	۱۱۱ (۳/۰۶)	۱۱۲ (۱/۰۰)	۱۰۶ (۱/۵۳)	۱۲۲ (۱/۰۰)
F1	۲۲۸ (۴/۱۶)	۴۴۱ (۲۲/۹۱)	۵۳۷ (۸/۵۱)	۶۳۷ (۱۳/۳۲)	۵۹۰ (۹/۰۷)	۶۰۴ (۵/۵۱)	۴۳۴ (۷/۰۲)	۲۶۳ (۵/۶۹)
F2	۲۱۱۴ (۲۱/۱۷)	۱۷۴۹ (۱۱/۷۹)	۱۶۷۵ (۳۰/۳۵)	۱۶۳۲ (۱۷/۵۲)	۱۲۱۹ (۳۸/۵۳)	۹۸۶ (۳۰/۰۵)	۱۲۴۱ (۳۴/۹۳)	۹۲۲ (۶۴/۱۷)
F3	۳۰۷۱ (۶۰/۹۲)	۲۵۶۷ (۲۱/۲۲)	۲۵۶۰ (۵/۶۹)	۲۵۱۱ (۴۵/۵۵)	۲۶۸۵ (۴/۵۱)	۲۵۴۷ (۳۸/۵۸)	۲۵۵۶ (۳۳/۹۸)	۲۲۳۹ (۴۷/۴۵)
دیرش	۲۷۱ (۱۹/۸۱)	۱۷۴ (۱۲/۵۰)	۱۷۸ (۴/۹۳)	۲۹۳ (۱۶/۰۴)	۲۰۸ (۱۷/۰۴)	۲۳۹ (۱۳/۳۲)	۱۸۵ (۷/۵۱)	۲۳۴ (۱۰/۵۰)
شدت	۶۴/۶۷ (۰/۵۸)	۶۷/۰۰ (۱/۰۰)	۶۹/۳۳ (۰/۵۸)	۶۷/۶۷ (۰/۵۸)	۶۸/۶۷ (۱/۱۶)	۷۰/۳۳ (۰/۵۸)	۷۰/۳۳ (۱/۱۶)	۷۰/۳۳ (۰/۵۸)

واکهٔ /i/ در واژهٔ heed بیشترین فرکانس سازه‌های دوم و سوم، و /u/ در واژهٔ Who’d کمترین فرکانس سازه‌های دوم و سوم را داشتند، که شبیه F2 فارسی هستند اما با F3 فارسی تفاوت دارند. در مجموع، فرکانس سازه‌های اول تا سوم در زبان انگلیسی بیشتر از زبان فارسی است. این نتایج نشان داد که فرکانس سازه‌های زبان فارسی با زبان انگلیسی تفاوت زیادی دارد. مقایسهٔ فضای واکه‌ای مردان در دو زبان فارسی و انگلیسی در پژوهش حاضر نشان داد که فضای واکه‌ایِ گویشوران مردِ فارسی‌زبان بزرگتر از فضای واکه‌ایِ گویشور مرد انگلیسی‌زبان است (شکل ۴).

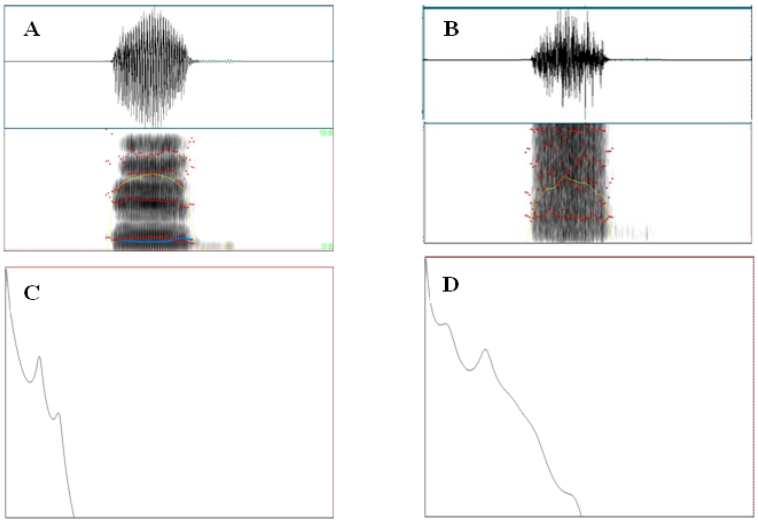


شکل ۴- مقایسهٔ فضای واکه‌ای گویشوران مرد فارسی و انگلیسی‌زبان

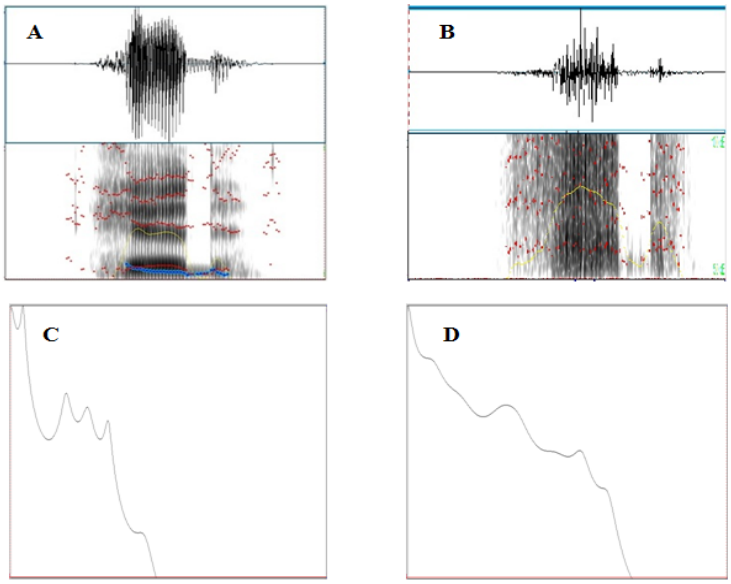
بیشترین دیرش واکه در زبان انگلیسی مربوط به واکهٔ /a/ در واژهٔ had، و کمترین آن مربوط به واکهٔ /ɪ/ در واژهٔ hid، به‌ترتیب برابر با ۲۹۳ و ۱۷۴ میلی‌ثانیه، بود. درحالی‌که این نتیجه در تطابق با دیرش واکه‌های زبان فارسی نیست. در فارسی بیشترین دیرش مربوط به واکهٔ /ɑ/ و کمترین مربوط به واکهٔ /i/ بود. در مجموع، دیرش واکه‌های انگلیسی از واکه‌های فارسی بیشتر، اما شدت آنها کمتر است، به‌طوری‌که طبق جداول ۱ و ۲ بیشترین شدت در فارسی مربوط به واکهٔ /i/ و کمترین شدت مربوط به واکهٔ /a/ بود، درحالی‌که در انگلیسی واکهٔ /i/ در واژهٔ heed کمترین شدت را به خود اختصاص داد که در تضاد با زبان فارسی بود و بیشترین شدت متعلق به واکه‌های پسین در زبان انگلیسی بود.

۳-۳- ویژگی‌های آکوستیکی واکه‌های وُکودشده در گویشوران مرد فارسی و انگلیسی‌زبان

تفسیر شکل موج، طیف‌نگاشت و نمودار LPC واکه‌های واضح و وُکودشده در هردو زبان گویای تفاوت زیاد در بین ویژگی‌های آکوستیکی واکه‌های واضح و وُکود شده است (شکل‌های ۵ و ۶)، به‌طوری‌که در واکهٔ واضح، F0 و سازه‌ها کاملاً قابل شناسایی است و نمودار LPC دارای قله‌های تیز است اما در واکه‌های وُکودشده F0 و ساختار سازه‌ای قابل شناسایی نیست و نمودار LPC آن دارای قله‌های با وضوح کمتر، گردشده و دامنه‌های کمتر است.



شکل ۵- واکهٔ واضح و وُکودشدهٔ /hed/ در زبان فارسی. A: موج صوتی و طیف‌نگاشت در حالت واضح، B: موج صوتی و طیف‌نگاشت در حالت وُکودشده، C: LPC در حالت واضح، D: LPC در حالت وُکودشده.



شکل ۶- واکهٔ واضح و وُکودشدهٔ /hɪd/ در زبان انگلیسی. A: موج صوتی و طیف‌نگاشت در حالت واضح، B: موج صوتی و طیف‌نگاشت در حالت وُکودشده، C: نمودار LPC در حالت واضح، D: نمودار LPC در حالت وُکودشده.

در جدول ۳ مقادیر دیرش و شدت واکه‌های وُکودشدهٔ فارسی و انگلیسی مقایسه شده است.

جدول ۳- مقادیر دیرش و شدت واکه‌های فارسی و انگلیسی گویشوران مرد در حالت‌های واضح و وُکودشده

		فارسی						انگلیسی							
		hid	hed	had	hɑd	Hod	hud	Heed	Hid	Had	Hud	Hod	Hood	Who'd
دیرش	واضح	۱۶۸	۱۷۵	۱۹۷	۱۸۶	۱۸۵	۱۸۲	۲۷۱	۱۷۴	۲۹۳	۲۰۸	۲۳۹	۱۸۵	۲۳۴
	NV-4	۱۶۲	۱۷۰	۱۸۰	۱۸	۱۷۵	۱۸۳	۲۳۲	۱۸۳	۲۰۲	۲۸۶	۱۷۱	۱۷۱	۲۶۹
	NV-	۱۸۰	۱۷۶	۱۸۸	۱۸	۱۶۷	۱۶۰	۲۳۳	۱۸۵	۲۰۴	۲۹۲	۱۷۲	۱۷۲	۲۶۹
	NV-	۱۸۰	۱۷۶	۱۸۷	۱۸	۱۶۸	۱۶۱	۲۳۴	۱۸۵	۲۰۶	۲۹۲	۱۷۳	۱۷۳	۲۶۹
شدت	واضح	۸۳	۸۳	۸۲	۸۳	۸۳	۸۴	۷۰	۷۰	۶۹	۶۸	۶۹	۶۷	۶۵
	NV-4	۷۳	۷۴	۷۵	۷۵	۷۴	۷۳	۷۷	۷۷	۷۷	۷۶	۷۸	۷۸	۷۶
	NV-	۷۴	۷۵	۷۵	۷۶	۷۶	۷۵	۷۷	۷۹	۷۸	۷۶	۷۷	۷۸	۷۸
	NV-	۷۳	۷۵	۷۷	۷۶	۷۶	۷۳	۷۷	۷۹	۷۹	۷۶	۸۰	۸۰	۸۰

NV-4= وُکود چهار باند. NV-10= وُکود ده باند. NV-20= وُکود بیست باند

براساس نتایج جدول ۳، به‌رغم آنکه دیرش واکه‌های وُکودشده در زبان فارسی نسبت به حالت واضح هیچ تغییری نکرد (t (89)=-1. 102, P=0. 273)، اما شدت آنها کاهش یافت که

این تغییر از نظر آماری معنادار بود (t (89)=21. 1, P<0. 001). در زبان انگلیسی نیز دیرش واکه‌ها در حالت وُکودشده نسبت به حالت واضح هیچ تغییری نکرد اما شدت آنها افزایش یافت.

۴- بحث

براساس یافته‌های این پژوهش، F0 واکه‌های زبان فارسی در گروه زنان بیشتر از گروه مردان بود (P<0. 01)، در گویشوران زن و مرد بیشترین F0 به‌ترتیب مربوط به واکهٔ /i/ و سپس با اختلاف بسیار اندکی مربوط به واکهٔ /u/ و کمترین F0 مربوط به واکهٔ /ɑ/ بود. بیشترین F0 واکه‌های زبان انگلیسی مشابه زبان فارسی بوده است اما کمترین F0 در دو زبان تطابق نداشته‌اند. با وُکودکردن واکه‌ها در هردو زبان فارسی و انگلیسی، F0 قابل اندازه‌گیری وجود نداشت. بیشترین فرکانس سازهٔ اول در گویشوران زن و مرد در این پژوهش /a/ و کمترین آن به‌ترتیب /u/ و /i/ بود. در پژوهش حاضر، بیشترین و کمترین F1 در زبان انگلیسی مشابه زبان فارسی بود. با وُکودکردن واکه‌ها در هردو زبان فارسی و انگلیسی، F1 قابل اندازه‌گیری وجود نداشت. نتایج به‌دست‌آمده در این پژوهش برای F2 در هردو گویشوران زن و مرد نشان داد که واکهٔ /i/ بیشترین و واکهٔ /u/ کمترین فرکانس را داراست. بیشترین F2 در زبان انگلیسی مشابه زبان فارسی در پژوهش حاضر بود، اما کمترینF2 متفاوت بود. در زبان فارسی فضای واکه‌ای گویشوران زن بیشتر از گویشوران مرد به‌دست آمد. فضای واکه‌ای انگلیسی در شش واکهٔ انگلیسی مشابه با واکه‌های فارسی در گویشور مرد کمتر از فضای واکه‌ای زبان فارسی در گویشوران مرد به‌دست آمد. با وُکودکردن واکه‌ها در هردو زبان فارسی و انگلیسی، F2 قابل اندازه‌گیری وجود نداشت. F3 در گویشوران زن نشان داد که واکهٔ /a/ کمترین فرکانس و واکهٔ /i/ بیشترین فرکانس را دارد و برای گویشوران مرد کمترین فرکانس مربوط به واکهٔ /a/ و بیشترین فرکانس مربوط به واکهٔ /u/ بود. بیشترین و کمترین F3 در واکه‌های زبان‌های فارسی و انگلیسی هیچ تطابقی نداشت. با وُکودکردن واکه‌ها در هردو زبان فارسی و انگلیسی، F3 قابل اندازه‌گیری وجود نداشت. اندازه‌گیری دیرش واکه‌ها در این پژوهش نشان داد که دیرش واکه‌های /a/ و سپس با اختلاف بسیار اندکی /e/ کمترین، و دیرش واکهٔ /ɑ/ در گویشوران زن بیشترین بود. دیرش واکه‌های /i/ و /ɑ/ به‌ترتیب کمترین و بیشترین برای گویشوران مرد به‌دست آمد. مقایسهٔ دیرش واکه‌های فارسی و مشابه آنها در انگلیسی در پژوهش حاضر نشان

داد که اولاً دیرش واکه‌های زبان انگلیسی بیشتر از زبان فارسی است و ثانیاً کمترین دیرش واکه در دو زبان متفاوت است، به‌رغم آنکه بیشترین دیرش در هردو زبان متعلق به گروه واکه‌های کشیده است. با وُکودشدن واکه‌ها در هردو زبان فارسی و انگلیسی، دیرش هیچ تغییری نکرد. اندازه‌گیری شدت واکه‌های فارسی در پژوهش حاضر نشان داد که در هردو گویشوران زن و مرد، واکهٔ /i/ بیشترین شدت را داشت، اما واکهٔ /e/ در زنان و واکهٔ /a/ در مردان کمترین شدت را دارا بود. در حالت مجموع، واکه‌های گویشوران مرد شدت بیشتری نسبت به واکه‌های گویشوران زن داشت. برخلاف زبان فارسی، در گویشور مردِ انگلیسی‌زبان، واکهٔ /i/ کمترین شدت و واکه‌های پسین بیشترین شدت را داشتند. به‌طور کلی، شدت واکه‌های گویشوران مرد فارسی‌زبان از شدت واکه‌های گویشور مرد انگلیسی‌زبان بیشتر بود. با وُکودشدن، شدت واکه‌های فارسی کاهش، اما شدت واکه‌های انگلیسی افزایش یافت.

F0 با زیر و بمی صوت ارتباط دارد. بدین معنا که صدای زیر با F0 بالا و صدای بم با F0 پایین در ارتباط است. واکه‌های باز، F0 پایین‌تری نسبت به واکه‌های بسته دارند (تیتز[1]، ۱۹۸۹؛ والن[2] و لویت[3]، ۱۹۹۵). با توجه به مقادیر به‌دست‌آمده در این پژوهش، فرکانس پایهٔ واکه‌های ساده نیز از الگویی مشابه پیروی می‌کند. تفاوت بین زیروبمی فرکانس‌ها در گویشوران زن و مرد تحت تأثیر تفاوت فیزیولوژیکی این گویشوران است. تارآواها در مردان طولانی‌تر از زنان است و به همین دلیل آواها در مردان فرکانس پایهٔ پایین‌تری دارند (لهیسته[4] و پیترسون[5]، ۱۹۶۱؛ علی‌نژاد و حسینی‌بالام، ۱۳۹۲). این موضوع در تحلیل آواهای به‌دست‌آمده نیز صادق است. براساس نتایج پژوهش حاضر، مطابق با پژوهش بی‌جن‌خان (۱۳۶۹) که واکه‌ها را در بافت /hVd/ استفاده کرده، و فشندکی (۱۳۹۰) که واکه‌ها را در بافت /cVb/ و فقط گویشوران زن استفاده کرده بود، واکهٔ /u/ بیشترین میزان فرکانس پایه را داشت. همان‌طور که انتظار می‌رفت، در بین افراد یک جنس اختلاف معناداری به‌دست نیامد زیرا تولید F0 فقط وابسته به طول تارآواهاست. نتایج پژوهش حاضر نیز در مورد واکه‌های زبان انگلیسی امریکایی همسو با نتایج هیلنبرند و همکاران (۱۹۹۵) است که بیشترین میزان F0 در واکه‌های زبان انگلیسی‌آمریکایی

1. I. R. Titze
2. D. H. Whalen
3. A. G. Levitt
4. I. Lehiste
5. G. E. Peterson

متعلق به واکهٔ /u/ است. در مجموع، F0 فارسی بیشتر از F0 انگلیسی بود. براساس پژوهش جیانگ پینگ[1] (۲۰۰۲) که به بررسی تأثیر نژاد بر سازهٔ واکه‌های انگلیسی در بافت /hVd/ پرداخته است، نژاد باعث تفاوت اندازهٔ حفره‌های گویایی در افراد و نیز منجر به تغییر سازه‌های واکه‌ها می‌شود.

F1 با ارتفاع زبان رابطه عکس دارد، بنابراین هرچه از ارتفاع زبان کاسته شود، فرکانس سازهٔ اول افزایش خواهد یافت. نتایج پژوهش حاضر ضمن تأیید این گفته، در تطابق با یافته‌های پژوهشی بی‌جن‌خان (۱۳۶۹)، صالحی و همکاران (۱۳۸۷) در پژوهش‌های خود بر روی گویشوران زن و مرد، و انصارین (۲۰۰۴) و فشندکی (۱۳۹۰) بر روی گویشوران زن بود. همچنین یافته‌های این پژوهش درباره F1 با یافته‌های صالحی و همکاران (۱۳۸۷) و بی‌جن‌خان (۱۳۹۲)، مبنی بر اینکه F1 اول زنان بیشتر از مردان است، همسو است.

F2 بالا با مشخصهٔ پیشین و F2 پایین با مشخصهٔ پسین واکه‌ها در ارتباط است (هیوارد[2]، ۲۰۰۰؛ لده‌فوگد، ۲۰۰۶). نتایج پژوهش حاضر درخصوص F2 با مطالعات بی‌جن‌خان (۱۳۶۹)، انصارین (۲۰۰۴) و فشندکی (۱۳۹۰) در بررسی واکه‌های فارسی همسو است. در پژوهش صالحی و همکاران (۱۳۸۷) و بی‌جن‌خان (۱۳۹۲) برروی واکه‌های زبان فارسی، فرکانس سازهٔ دوم گروه مردان نیز پایین‌تر از گروه زنان گزارش شده است، که مطابق با نتایج پژوهش حاضر است. به‌طورکلی، نویسندگان این مقاله نیز، موافق با نتایج مطالعات قبلی، فرکانس سازه‌های اول و دوم واکه‌های مشابه در زبان‌های مختلف را از هم متفاوت می‌دانند (علی‌نژاد و حسینی‌بالام، ۱۳۹۲؛ مدرسی‌قوامی، ۱۳۹۲).

در این پژوهش نشان داده شد که فضای واکه‌ای زنان بیشتر از مردان است. سیمسون[3] (۲۰۰۰) بیان داشت تفاوت فضای واکه‌ای مردان و زنان (فضای واکه‌ای زنان بزرگتر از مردان) دلایل متفاوتی دارد، از جمله: دلایل آناتومی و همچنین آواشناسی‌اجتماعیِ ارتباط بین F0 و فرکانس سازه‌ها. هرچه تعداد واکه‌های یک زبان بیشتر باشد، فضای واکه‌ای آکوستیکی آن زبان بزرگتر است. مقایسهٔ فضای واکه‌ای در کل F1 و F2 شش واکه در دو زبان فارسی و انگلیسی مشابه است و تفاوت بیشتر در واکه‌های پسین و افتاده مشاهده می‌شود (علی‌نژاد و

1. H. Jianping
2. K. Hayward
3. A. P. Simpson

حسینی‌بالام، ۱۳۹۲). در پژوهش حاضر، برخلاف نظر علی‌نژاد و حسینی‌بالام (۱۳۹۲)، فضای واکه‌ایِ انگلیسی در شش واکهٔ مشابه فارسی، کمتر از فضای واکه‌ای زبان فارسی به‌دست آمد، اما با نظر موافق، تفاوت در واکه‌های پسین وجود دارد. شکل فضای واکه‌ای زبان فارسی مطابق با پژوهش محمدی و همکاران (۱۳۹۰) به‌دست آمد.

در پژوهش حاضر، واکه‌های /a/ و /i/ به‌ترتیب کمترین و بیشترین میزان F3 را به خود اختصاص داده‌اند که همسو با یافته‌های فشندکی (۱۳۹۰) است. F3 واکه‌های زبان انگلیسی در این پژوهش مطابقتی با F3 واکه‌های زبان فارسی نداشت و این مطابق با یافته‌های زبان انگلیسی در مطالعات قبلی است. به‌طور کلی زنان فرکانس سازه‌ای بالاتری نسبت به مردان دارند و این نتیجه همسو با نظرات پترسون و بارنی[1] (۱۹۵۲)، هیلنبرند و همکاران (۱۹۹۵) و بی‌جن‌خان (۱۳۹۲) است.

در تحقیقِ پیشِ رو، فاصلهٔ F1 و F2 واکهٔ /i/ بسیار زیاد اما فاصلهٔ F2 و F3 آن کم بود. در واکهٔ /e/ فاصلهٔ F1 و F2 نسبت به واکهٔ /i/ کمتر است چون F1 افزایش و F2 کاهش یافته است، و فاصلهٔ F2 با F3 بیشتر است، که این نتایج همسو با مطالعهٔ علی‌نژاد و حسینی‌بالام (۱۳۹۲) است. درخصوص واکهٔ /a/ انتظار می‌رود که با پایین‌آمدن زبان در این واکهٔ پسینِ افتان، F1 و F3 افزایش و F2 کاهش یابد که در پژوهش حاضر نیز همین نتیجه به‌دست آمده است. F2 واکهٔ /u/ به عنوان یک واکهٔ پسین افراشته نسبت به همتای پیشین خود یعنی /i/، به دلیل گردی لب‌ها افت قابل ملاحظه‌ای داشت. در واکهٔ /o/، مشابه واکهٔ /u/، به دلیل گردیِ لب‌ها افت F2 نسبت به همتای پیشین آن یعنی /e/ به‌دست آمد. F1 و F3 واکهٔ /ɑ/ به عنوان یک واکهٔ پسین افتاده با همتای پیشین خود یعنی /a/ افزایش، اما F2 آن کاهش یافت، که این نتایج با علی‌نژاد و حسینی‌بالام (۱۳۹۲) در تطابق است.

مطالعات محدودی ویژگی‌های دیرش واکه‌ها و همچنین بررسی‌های بسیار محدودتر، ویژگی شدت واکه‌ها را بررسی کرده‌اند. کشیده‌ترین واکهٔ فارسی در گویشوران زن همسو با مطالعهٔ فشندکی (۱۳۹۰) است اما کمترین دیرش در مطالعهٔ فشندکی و همکاران متعلق به واکهٔ /o/ است که با مطالعهٔ حاضر یکسان نیست. بی‌جن‌خان (۱۳۶۹)، در رابطه با بررسی دیرش واکه‌های ساده بیان می‌کند که واکهٔ /ɑ/ کشیده‌ترین واکه و واکه‌های /o/ و /e/ کوتاه‌ترین‌اند. در

1. H. L. Barney

مطالعهٔ حاضر، دیرش /ɑ/= ۲۲۶ میلی‌ثانیه، در پژوهش بی‌جن‌خان (۱۳۶۹) دیرش /ɑ/ = ۳۰۰ میلی‌ثانیه، و در پژوهش فشندکی (۱۳۹۰)، دیرش /ɑ/ = ۲۹۱ میلی‌ثانیه بود، که در مجموع دیرش واکه‌ها در مطالعهٔ حاضر کمتر از مطالعات قبلی به‌دست آمد. اما دیرش واکه‌های زنان بیش از مردان بود. واکه‌های باز در حالت عادی کشیده‌تر از واکه‌های بسته‌اند و این نتیجه همسو با نتایج مطالعات لهیسته و پترسون (۱۹۶۱)، لده‌فوگد (۲۰۰۶) و مدرسی‌قوامی (۱۳۹۳) است. در بسیاری از پژوهش‌ها کوشیده‌اند دلیل تفاوت دیرش در گروه مردان و زنان را با استفاده از آواشناسی اجتماعی توضیح دهند که واکه‌های کشیده‌تر و میزان دیرش بیش‌تر، با گفتار واضح‌تر، یعنی گفتار زنان ارتباط دارند (هنتون[1] و همکاران، ۱۹۹۵؛ سیمسون و اریک‌دوتر[2]، ۲۰۰۳). هیلنبرند و همکاران (۱۹۹۵) نیز در پژوهش خود به این تفاوت دیرشی در مردان و زنان اشاره کرده‌اند. یکی دیگر از دلایل تفاوت در دیرش‌های به‌دست‌آمده می‌تواند مربوط به بافت واکه باشد زیرا دیرش واکه تحت تأثیر همخوان بی‌واک کمتر و در مجاورت با همخوان‌های واک‌دار طولانی‌تر می‌شود (مرمل‌شِتاین[3]، ۱۹۷۸). در مجموع، دیرش واکه‌های انگلیسی از واکه‌های فارسی در پژوهش حاضر بیشتر بود که مربوط به ساختار تولیدی زبان است. با وُکودشدن واکه، در هردو زبان هیچ تغییری در دیرش به وجود نیامد. همان‌طور که در نحوهٔ ساختِ واکه‌های وُکودشده توضیح داده شد، چون دیرش دستکاری نمی‌شود و پوش سیگنال دست نخورده باقی است، بدیهی است که انتظار می‌رود پس‌از وُکودشدن نیز اختلافی بین دیرش واکه‌های واضح و وُکودشده دیده نشود.

در پژوهش حاضر، بیشترین شدت واکه‌های فارسی در گویشوران زن ۸۱ دسی‌بل و کمترین میزان آن ۷۸/۸ دسی‌بل به‌دست آمد. در گویشوران مرد بیشترین شدت واکه‌ها ۸۳/۹ دسی‌بل و کمترین شدت ۸۲/۲ دسی‌بل به‌دست آمد که این یافته همسو با مطالعات قبلی نیست. مقادیر شدت در پژوهش فشندکی (۱۳۹۰) به‌ترتیب ۷۲/۶ دسی‌بل و ۷۵/۲ دسی‌بل بود، که کمترین شدت واکه مطابق با نتیجه پژوهش حاضر است. شدت واکه‌های گویشوران مرد از گویشوران زن بیشتر بود. در پژوهش حاضر، شدت واکه‌های فارسی از واکه‌های انگلیسی بیشتر بود. فربنکس[4]

1. C. Henton
2. C. Ericsdotter
3. P. Mermelstein
4. G. Fairbanks

و همکاران (۱۹۵۰) بیان داشتند که شدت واکه‌های /u/، /e/ و /o/ به نسبت شدت واکه‌های /i/، /a/ و /ɑ/ بیشتر است. علاوه بر تفاوت بافت و گویش، به نظر می‌رسد که دلیل این تفاوت، اختلاف اندک در میزان شدت در بین واکه‌های مختلف در پژوهش حاضر باشد. زیرا اختلاف بیشترین و کمترین شدت در پژوهش حاضر در حدود ۲ دسی‌بل است، درحالی‌که در مطالعات قبلی بین ۱۰-۸ دسی‌بل بوده است. با وُکودشدن واکه در زبان فارسی در هر سه حالت چهار و ده و بیست باند، شدت کاهش یافته است اما در زبان انگلیسی شدت افزایش یافته است. برای وُکودکردنِ سیگنال لازم است فیلترهای مختلفی بر سر راه سیگنال صوتی اِعمال شوند تا ساختار ظریف آن از بین برود، گرچه نهایتاً نرمالیزیشن صورت می‌گیرد اما این امر موجب تضعیف سیگنال می‌شود، همچنان که در زبان فارسی صورت گرفته است. به نظر می‌رسد دلیل احتمالی افزایش شدت در واکه‌های زبان انگلیسی مربوط به تفاوت شدت آنها با زبان فارسی در حالت واضح باشد و چون مطالعات قبلی در این زمینه وجود ندارد نمی‌توان مقایسه‌ای هم انجام داد و پژوهش‌های گسترده‌تری مورد نیاز است. با این حال در مجموع، سیگنال گفتاری وُکودشده دارای پوش دست نخورده اما بدونِ فرکانس پایه و سازه‌هاست؛ همچنان‌که در پژوهش حاضر در نمودارهای طیف‌نگاشت و LPC قابل مشاهده بود.

ارزش ساخت محرک‌های گفتاری وُکودشده در هر زبانی و دانستن ویژگی‌های آکوستیکی آن در مطالعات شبیه‌سازی‌شده در حیطه‌های سایکوآکوستیک، فیزیولوژی و پردازش گفتار، و در مطالعات بالینی بسیار زیاد است. همچنان‌که نتایج پژوهش حاضر نشان داد، محرک‌های گفتاری واضح و وُکودشده، از نظر طیفی متفاوت هستند، اما به دلیل تشابهِ مشخصه‌های زمانی که دارای پوش یکسان هستند، ویژگی‌های مشترکی در بازنمایی‌های عصبی‌شان وجود دارد. گفتار واضح در سیستم عصبی مرکزی، اساساً توسط اطلاعات فرکانسی پردازش می‌شود، و با وجود مدار عصبی، بازشناسی گفتار با کدشکنی سیگنال‌های گفتاری از اطلاعات فرکانسی رمزگذاری‌شده در سیستم عصبی، صورت می‌گیرد. بنابراین برای سیستم عصبی مرکزی بسیار مشکل است که بدون اطلاعات فرکانسی، سیگنال‌های گفتاری کدشکنی شوند. اما واقعیت این است که سیستم شنوایی قادر است حتی زیروبمی محرک‌های گفتاریِ فاقد اطلاعات فرکانسی را دریافت کند. در سیستم عصبی مرکزی، سیگنال‌های گفتاری همچنین از نظر اطلاعات پوش دامنه نیز رمزگذاری می‌شوند، اما استفاده نمی‌شوند. مسئلهٔ مهم این است که یک شبکهٔ عصبی

مورد نیاز است تا با استفاده از اطلاعات پوش دامنه، سیگنال‌های وُکودشده را به اصوات گفتاری قابل‌فهم ترجمه کند. بنابراین نیاز به پلاستیسیتی در سیستم عصبی مرکزی برای پردازش سیگنال‌های وُکود شده است. انتظار می‌رود نواحی که در پردازش اصوات گفتاری واضح دخیل نیستند، در پردازش سیگنال‌های وُکودشده مشارکت داشته باشند. مطالعات اخیر به وسیلهٔ fMRI نشان داده‌اند که علاوه بر قشر گیجگاهی که به‌طور معمول در پردازش اصوات گفتاری واضح دخالت دارد، قشرهای پیشانی و آهیانه‌ای نیز در پردازش سیگنال‌های وُکودشده دخیل می‌شوند (ریکی‌مور، ۲۰۰۶).

از آنجا که ساخت وُکود واکه‌های زبان فارسی برای اولین بار در این پژوهش انجام شده است، بالطبع تمام نتایجی که از آن و همچنین مقایسه با واکه‌های زبان انگلیسی به‌دست آمده، به عنوان اولین پژوهش مطرح است و نتایج آن را نمی‌توان تعمیم داد و اثبات نتایج این مطالعه متکی بر پژوهش‌های گسترده آینده خواهد بود.

۵- نتیجه‌گیری

پژوهش حاضر، ویژگی‌های آکوستیکی واکه‌های زبان فارسی را توصیف کرده و نشان داده است که با زبان انگلیسی متفاوت است. نتایج وُکودکردن واکه‌های زبان فارسی، و همچنین مقایسهٔ مقدماتی آن با واکه‌های زبان انگلیسی، برای نخستین بار مطرح شد. همچنین تفاوت‌های طیفیِ واکه‌های فارسیِ وُکودشده نسبت به واکه‌های واضح آشکار شد.

۶- سپاسگزاری

- مقاله حاضر بخشی از پژوهش پایان‌نامهٔ دکتری در دانشکده علومِ توانبخشیِ دانشگاهِ علومِ پزشکیِ ایران، و دارای تأییدیه از کمیتهٔ اخلاق به شماره ۱۳۰/۸۴۱/ د/ ۹۲ است.
- از کمک‌های پروفسور رابرت شانون و دکتر جان ویگونسکی از دانشگاه کالیفرنیا و پروفسور متیو دیویس از دانشگاه کمبریج برای ساخت سیگنال‌های وُکودشده، کمال سپاسگزاری را می‌نماید.

منابع

بی‌جن‌خان، محمود. (۱۳۶۹). *تجزیه و تحلیل صوتی واکه‌های ساده و مرکب زبان فارسی براساس نظریهٔ فانت*. پایان‌نامه کارشناسی ارشد، دانشگاه تهران.

بی‌جن‌خان، محمود. (۱۳۹۲). *نظام آوایی زبان فارسی*. تهران: سمت.

سپنتا، ساسان. (۱۳۷۷). *آواشناسی فیزیکی زبان فارسی (توصیف مشخصه‌های صوت شناختیِ واج‌های زبان فارسی)*. اصفهان: نشر گل‌ها.

صالحی، فرانک، اکبر بهرامی، فرهاد ترابی‌نژاد، و محمد کمالی. (۱۳۸۷). «ساختار سازه‌ای واکه‌های زبان فارسی در دانش آموزان ۷ تا ۹ سالهٔ عادی و مبتلا به افت شنوایی متوسط و شدید شهر اصفهان». *شنوایی‌شناسی*، ۱۷، ۵۷-۴۲.

عادل‌قهرمان، منصوره. صدایی، مهین. فتاحی، جمیله. محمدخانی، قاسم. (۱۳۸۵). *فرهنگ جامع شنوایی‌شناسی*. تهران: دانش‌فریار.

علی‌نژاد، بتول و فهمیه حسینی‌بالام. (۱۳۹۲). *مبانی آواشناسی آکوستیکی*. اصفهان: دانشگاه اصفهان.

فشندکی، شهره. (۱۳۹۰). *بررسی صوت‌شناختی (آکوستیکی) واکه‌های سادهٔ زبان فارسی معیار*. پایان‌نامهٔ کارشناسی ارشد، دانشگاه الزهرا.

محمدی، هیوا. محمدی، ریحانه. ترابی‌نژاد، فرهاد و منصور رضایی. (۱۳۹۰). «تعیین ساختار سازه‌ای و فضای واکه‌ای در واکه‌های زبان فارسی». *شنوایی‌شناسی*، ۲۰(۲)، ۸۵-۷۹.

مدرسی‌قوامی، گلناز. (۱۳۹۲). *آواشناسی: بررسی علمی گفتار*. تهران: سمت.

مدرسی‌قوامی، گلناز. (۱۳۹۳). «نگاهی دیگر به ویژگی‌های صوت‌شناختی واکه‌های کوتاه و بلند در زبان فارسی». *پژوهش‌های زبان‌شناسی تطبیقی*، ۸، ۹-۱.

Ansarin, Ali Akbar, (2004). "An acoustic analysis of modern Persian vowels". Paper presented at the SPECOM'2004: 9th Conference Speech and Computer September 20-22, St. Petersburg, Russia.

Baken, R. J., Orlikoff, R. F. (2000). *Clinical measurement of speech and voice*. California: Singular Thomson Learning.

Behrman, A. (2007). *Speech and voice science*. San Diego: Plural Pub.

Bijankhan, M. (2000). "Farsi vowel compensatory lengthening: an experimental approach". *The Fifth International Conference on Spoken Language Processing*, Beijing, China.

Boersma, P., Weenink, D. (2007). *Praat: doing phonetics by computer* Retrieved on July 27, 2007, 34.

Chen, F. (2012). "Predicting the intelligibility of cochlear-implant vocoded speech from objective quality measure". *Journal of Medical and Biological Engineering.* 32, 189-94.

Chen, F., Loizou, P. C. (2011). "Predicting the intelligibility of vocoded speech". *Ear and hearing*. 32(3), 331-38.

Cho, T., Ladefoged, P. (1999). "Variation and universals in VOT: evidence from 18 languages". *Journal of Phonetics*. 27(2), 207-29.

Ellermeier, W., Kattner, F., Ueda, K., Doumoto, K., Nakajima, Y. (2012). "Effects of noise-vocoded speech in the irrelevant-sound paradigm". In *38. Jahrestagung für Akustik–DAGA 2012*. Galrev: Berlin, 336

Fairbanks, G., House, A. S., Stevens, E. L. (1950). "An experimental study of vowel intensities". *Journal of the Acoustical Society of America*. 22(4), 457-59.

Fu, Q. J., Shannon, R. V. (1999). "Effects of electrode configuration and frequency allocation on vowel recognition with the Nucleus-22 cochlear implant". *Ear and hearing*. 20(4), 332-44.

Greenwood, D. D. (1990). "A cochlear frequency-position function for several species-29 years later". *Journal of the Acoustical Society of America.* 87(6): 2592-605.

Hayward, K. (2000). *Experimental phonetics*. London: Longman.

Henton, C., Elenius, K., Branderud, P., (1995). "Cross-language variation in the vowels of female and male speakers": *The 13th international congress of phonetic sciences*. Vol.4, 420-423.

Hervais-Adelman, A. G., Davis, M. H., Johnsrude, I. S., Taylor, K. J., Carlyon, R. P. (2011). "Generalization of perceptual learning of vocoded speech". *Journal of experimental psychology: Human perception and performance*. 37(1), 283-95.

Hillenbrand, J., Getty, L. A., Clark, M. J., Wheeler, K. (1995). "Acoustic characteristics of American English vowels". *Journal of the Acoustical Society of America.* 97(5 Pt 1), 3099-111.

Jianping H. (2002). *Cross-Racial Studies of Human Vowel Tract Dimentions and Formant Structures*. Ph.D. dissertation, Ohio State University. www.ohiolink.edu.

Kent, R. D., & Kim, Y. (2008). Acoustic analysis of speech. In M. J. Ball, M. R. Perkins, N. Muller & S. Howard (eds.), *The Handbook of Clinical Linguistics*. Oxford: Blackwell Publishing.

Ladefoged, P and Disner, S. F. (2005). *Vowels and consonants*, 2nd ed., Oxford: Wiley-Blackwell.

Ladefoged, P. (2006). *A course in phonetics*, 5th ed., Boston: Thomson Wadsworth.

Lea, W. A. (1973). "Segmental and suprasegmental influences on fundamental frequency contours, Consonant types and tones". In Hyman, L M., editor. *Consonant types and tones*. Los Angeles: Linguistics Program, University of Southern California. 15-70.

Lehiste, I., Peterson, G. E. (1961). "Transitions, glides, and diphthongs". *Journal of the Acoustical Society of America*, 33(3), 268-77.

Licklider, J. C. R., Pollack, I. (1948). "Effects of differentiation, integration, and infinite peak clipping upon the intelligibility of speech". *Journal of the Acoustical Society of America*, 20(1), 42-51.

Liu, C., Azimi, B., Tahmina, Q., Hu, Y. (2012). "Effects of low harmonics on tone identification in natural and vocoded speech". *Journal of the Acoustical Society of America*, 132(5), 378-84.

Mermelstein, P. (1978). "Difference limens for formant frequencies of steady-state consonant-bound vowels", *J Acoust Soc Am*. 63, 572-80.

Mildner, V. (2008). *The cognitive neuroscience of human communication*. New York: Taylor & Francis.

Peterson, G. E, Barney, H. L. (1952). "Control methods used in a study of the vowels". *Journal of the Acoustical Society of America*. 24(2),175-84.

Riquimaroux, H. (2006). "Perception of noise-vocoded speech sounds: Sentences, words, accents and melodies". *Acoustical Science and Technology*, 27(6), 325-31.

Rosen, S. (1992). "Temporal information in speech: acoustic, auditory and linguistic aspects". *Philosophical transactions of the Royal Society of London Series B, Biological sciences*, 336(1278), 367-73.

Schvartz, K. C., Chatterjee, M. (2012). "Gender identification in younger and older adults: use of spectral and temporal cues in noise-vocoded speech". *Ear and hearing*. 33(3), 411-20.

Scott S. K. (2005). "The neurobiology of speech perception". In: Cutler A, editor. *Twenty-First Century Psycholinguistics: Four Corner Stones*. 1st ed. Mahwah, New Jersey: Lawrence Erlbaum Associates. 141-157.

Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., Ekelid, M. (1995). "Speech recognition with primarily temporal cues". *Science*, 270(5234), 303-4.

Shannon, R. V., Zeng, F. G., Wygonski, J. (1998). "Speech recognition with altered spectral distribution of envelope cues". *Journal of the Acoustical Society of America*. 104(4), 2467-76.

Simpson, A. P. (2000). "Gender-specific differences in the articulatory and acoustic realization of interword vowel sequences in American English". *5th Seminar on Speech Production: Models and Data Kloster Seeon*.

Simpson, A. P., Ericsdotter, C. (2003). "Sex-specific durational differences in English and Swedish". *Proceedings of 15th ICPhS*, (1113-1116) Barcelona.

Titze, I. R. (1989). "Physiologic and acoustic differences between male and female voices". *Journal of the Acoustical Society of America*, 85(4), 1699-1707.

Ueda, K., Nakajima, Y. (2008). "A consistent clustering of power fluctuations in British, English, French, German, and Japanese". *Trans Tech Comm Psychol Physiol Acoust*, 38(8), 771-76.

Whalen, D. H., Levitt, A. G. (1995). "The universality of intrinsic F0 of vowels". *Journal of Phonetics*, 23(3), 349-66.

Zwicker, E., Terhardt, E. (1980). "Analytical expressions for critical-band rate and critical bandwidth as a function of frequency". *Journal of the Acoustical Society of America*, 68(5), 1523-25.