

فصل اول
کلیات تحقیق

مقدمه:

این فصل به معرفی نظام آوایی واکه‌های زبان فارسی و بیان دلایل انجام تحقیق اختصاص دارد و سپس به معرفی اهمیت واکه‌ها و سه فرمنت اول با هدف دست یافتن به شاخص‌های آماری سه فرمنت اول، طراحی فیلتر مناسب و نهایتاً دست یافتن به مدل کلی برای هر واکه برای تلفظ درست واکه‌ها پرداخته می‌شود. این فصل شامل قسمت‌هایی چون مسأله استخراج سه فرکانس فرمنت، ضرورت پرداختن به واکه‌ها و سه فرمنت اول و اهمیت پرداختن به این موضوع، تعریف مفاهیم و اهداف استخراج و شبیه‌سازی سه فرمنت اول زبان فارسی است.

۱-۱ طرح مسأله

کوچکترین واحد در نظام آوایی یک زبان واج نامیده می‌شود که به دو دسته واکه (مصوت) و همخوان (صامت) تقسیم‌بندی می‌شود. در واقع اگر صدایی که از حنجره، خارج می‌شود حنجره در تولید آن صدا باز و بسته شود و باعث لرزش تارهای صوتی بشود و در حلق و دهان هیچ مانع جدی در برابر جریان هوا نباشد، آن صدا واکه نامیده می‌شود که در زبان فارسی شامل ای (/i/)، آ (/a/)، او (/u/)، (/e/)، آ (/æ/) و (/o/) می‌باشد. همخوان (صامت) آوایی است که به هنگام تولید آن در برابر جریان هوا نوعی گرفتگی بوجود می‌آید. در دانش آواشناسی صامت به صدایی گفته می‌شود که با بسته شدن کامل و یا جزئی مجرای صوتی فوقانی ایجاد می‌گردد. مجرای صوتی فوقانی به قسمتی از مجرای صوتی گفته می‌شود که بالاتر از حنجره قرار دارد. زبان فارسی ۲۳ صامت دارد. همخوان‌ها به سه دسته انفجاری شامل /p/، /b/، /t/، /d/، /k/، /g/ و... همچنین سایشی مانند /s/، /f/، /v/، /ʃ/ و... بالآخره ناسوده مانند /r/، /m/، /n/ و... تقسیم می‌شوند. فرکانس فرمنت‌ها نشان‌دهنده فرکانسهای تشدید دستگاه‌گفتار می‌باشد. فرمنت‌ها پارامترهای مربوط به لوله صوتی می‌باشند که برای هر آوایی مقدار فرکانس فرمنت متفاوتی در مقایسه با آواهای دیگر وجود دارد که مشخصه همان آوا به شمار می‌رود.

پایین‌ترین قله در طیف فرکانسی به عنوان اولین فرمنت و دومین قله به عنوان دومین فرمنت و بقیه‌ی فرمنت‌ها به این ترتیب مشخص می‌شوند. F با شماره مناسب به عنوان نشانه‌ای برای نشان دادن مقدار فرکانس فرمنت‌ها اختصاص یافته است. سه فرمنت اول در تشخیص واکه توسط شنونده نقش مهمی دارند و در واکه‌ها F_1 با ارتفاع زبان و F_2 با وضعیت زبان (جلو یا عقب بودن زبان) و F_3 با شکل لبها (گرد یا گسترده بودن) در ارتباط است.

حال با در نظر گرفتن این مفاهیم می‌توان طرح کلی مسأله در این تحقیق را به این ترتیب بیان کرد. در این تحقیق از تعدادی زن و مرد خواسته می‌شود ۲۴ کلمه را که شامل واکه‌های زبان فارسی می‌شوند، تلفظ کنند. ابتدا سعی شده با استفاده از نرم افزار Adobe Audition هر یک از این واکه‌ها را در کلمه مربوطه‌اش ضبط نمود و پس از ذخیری فایل‌های صوتی در این نرم افزار، با نرم افزار Praat مقدار فرکانس فرمنت‌ها برای هر یک از واکه‌ها را تعیین نمود. پس از استخراج فرکانس فرمنت‌های هر واکه روش نرمال‌سازی مناسب با در نظر گرفتن عامل جنسیت برای نرمال‌سازی واکه‌ها در نظر گرفته می‌شود. در مرحله بعدی پس از نرمال‌سازی فرمنت‌های هر واکه، در نرم افزار Matlab با در نظر گرفتن مقادیر بدست‌آمده برای هر واکه با برنامه‌نویسی در نرم افزار مطلب، فیلتر مناسب را طراحی نمود. به این ترتیب معیاری متناسب با هر واکه خواهیم داشت و با توجه به نتیجه حاصل می‌توان میزانی تلفظ درست آن واکه توسط فرد مورد نظر را شناسایی کرد. قابل ذکر است که نمونه‌های مرجع ما نمونه‌های واکه‌های زبان فارسی گفته‌شده از دو گروه زن و مرد می‌باشد.

۱-۲ ضرورت و اهمیت تحقیق

واکه‌ها هسته اصلی کلمات را تشکیل می‌دهند. خطا در تشخیص صحیح واکه می‌تواند باعث سوءتفاهم در درک همخوان مجاور شود و سه فرمنت اول از اجزای اصلی تشخیص واکه توسط شنونده شناخته می‌شوند. اهمیت این موضوع زمانی بیشتر می‌شود که فرد در تولید واکه آسیب خفیفی داشته باشد. در این حالت تنها با ارزیابی ادراکی نمی‌توان به اختلال و کیفیت آن پی‌برد. تجزیه و تحلیل آکوستیکی هر واکه می‌تواند به ما در تشخیص این اختلال کمک کند و از آنجایی که نرمال‌سازی واکه‌ها و شبیه‌سازی برای زبانهای مختلف با استفاده از نرم‌افزارهای مختلف مدتهاست شروع شده است، انجام این کار برای زبان فارسی می‌تواند معیار مناسبی برای تشخیص تلفظ صحیح واکه‌های زبان فراهم کند.

از طرف دیگر شبیه‌سازی واکه‌ها با استفاده از نرم‌افزار Matlab در راستای پردازش سیگنال گفتار می‌تواند مقدمه‌ای برای شبیه‌سازی همخوانها قرار گیرد تا نهایتاً بتوان گفتار را در قالب کلمه و سپس جمله شبیه‌سازی کرد. این پایان‌نامه گوشه کوچکی از این هدف را محقق می‌سازد و می‌تواند در ادامه برای دانشجویان دیگر در این مقطع به عنوان پایان‌نامه تعریف شود و تکمیل گردد تا نهایت این امر یعنی شبیه‌سازی گفتار در قالب جمله تحقق یابد.

این پایان‌نامه در حوزه صدا و تصویر بخصوص پردازش صدا در سازمان صدا و سیما کاربردهای زیادی دارد. اما می‌توان این نکته را نیز در نظر گرفت که با توجه به اینکه شبیه‌سازی واکه‌ها بر اساس زبان معیار برای زبان فارسی انجام نشده این پایان‌نامه می‌تواند به تلفظ درست افرادی که سعی دارند زبان خود را به زبان معیار نزدیک کنند از جمله گویندگان و مجریان برنامه‌های تلویزیونی و رادیویی و همچنین افراد خارجی که سعی در یادگیری زبان فارسی دارند، کمک بسزایی بنماید.

هنوز در سازمان صدا و سیما آزمون جامعی برای گفتار فارسی در خصوص تلفظ کردن درست واکه‌ها که نقش اصلی در تلفظ صحیح کلمه دارند، تبیین نشده است تا افراد را از نظر درست ادا کردن واکه‌ها ارزیابی کند. با در نظر گرفتن این واقعیت که یکی از اهداف صدا و سیما صیانت از فرهنگ و زبان فارسی است این پایان‌نامه می‌تواند از نظر آکوستیکی این موضوع را بررسی کند و با بررسی تحقیق‌هایی که در دنیا در این خصوص صورت گرفته است، آزمون‌های مناسب را در راستای تامین این هدف فراهم کند تا اینکه یک معیار علمی در کنار معیارهای کیفی در راستای انتخاب مناسب افراد جهت گویندگی باشد.

۳-۱ اهداف تحقیق

در این پایان‌نامه سعی می‌شود که اهداف زیر تحقق یابد:

۱- دست یافتن به شاخص‌های آماری سه فرمنت اول برای هر واکه که در واقع اساس نرمال‌سازی واکه‌های زبان فارسی است.

۲- طراحی فیلتر مناسب با فرکانس فرمنت‌های هر واکه ی زبان فارسی با استفاده از نرم افزار Matlab.

۳- نهایتاً دست یافتن به مدل کلی برای هر واکه تا بتواند معیاری برای تلفظ درست واکه برای گویندگان اعم از زن و مرد قرار گیرد. واضح است که با این معیار می‌توان به یک الگوی آکوستیکی برای تلفظ واکه‌ها دست یافت.

۴-۱ سؤال تحقیق

کوچکترین واحد در نظام آوایی یک زبان واج نامیده می‌شود که به دو دسته همخوان‌ها (صامت‌ها) و واکه‌ها (مصوت‌ها) تقسیم می‌شوند. در این میان واکه‌ها از اهمیت بیشتری برخوردار هستند چرا که خطا در تشخیص صحیح واکه می‌تواند باعث سوء تفاهم در درک همخوان مجاور شود.

مسئله‌ای که در این تحقیق مطرح می‌باشد، پاسخ به این سؤال است که با توجه به منابع مختلفی که باعث تغییر در سیگنال گفتار می‌شود، دیده می‌شود که شنوندگان سالم در تشخیص گفتار دچار اشتباه نمی‌شوند. با این وجود آیا می‌توان برای واکه‌ها الگویی در نظر گرفت و یا می‌توان هر یک از واکه‌ها را با یک فیلتر شبیه‌سازی کرد؟

برای توضیح بیشتر می‌توان به منابع تغییر در سیگنال گفتار اشاره کرد. سه عامل اصلی گوینده، داده‌های زبانی و نویز محیط باعث تغییر در سیگنال گفتار می‌شود. عامل گوینده که شامل جنس، سن، لهجه، میزان تحصیلات، طبقه اجتماعی و حالات روحی می‌باشد. داده‌های زبانی که شامل بافت آوایی، آهنگ و تکیه جمله می‌باشد و همچنین عامل سوم، نویز محیطی است که گوینده در آن محیط صحبت می‌کند و این عامل بر روی سیگنال تولید شده تأثیر می‌گذارد.

در این تحقیق سعی خواهد شد با پیاده کردن روش‌های مختلف نرمال‌سازی تأثیر عوامل تغییردهنده سیگنال گفتار را به حداقل رساند و با انتخاب مناسب‌تر روش نرمال‌سازی واکه‌ها را با یک فیلتر شبیه‌سازی نمود.

بنابراین با وجود تفاوت‌های موجود رسیدن به الگویی که بتواند بیانگر یک واکه باشد نیاز به تحقیق، دقت و بکارگیری نرم افزارهای مناسب خواهد داشت.

۱-۵ فرضیات تحقیق

هرچند در این تحقیق از مردان و زنانی استفاده می‌شود که در بازی سنی ۲۰ تا ۳۰ سال قرار دارند اما در اینجا عامل سن ثابت فرض می‌شود چرا که عامل سن زمانی متغیر می‌باشد که افراد در دوره‌های سنی متفاوتی یعنی دوری سنی بچه‌گی، نوجوانی، جوانی، میان‌سالی و پیری قرار گرفته باشند. همچنین در این تحقیق عامل لهجه و آهنگ گفتار را ثابت فرض می‌کنیم چرا که از افرادی در این تحقیق استفاده می‌شود که گویشی نزدیک به معیار داشته باشند و بیشتر از گویندگان و گزارشگرانی استفاده می‌شود که عامل لهجه در آنان کم‌رنگ می‌باشد. از طرفی آهنگ را نیز ثابت فرض می‌کنیم چرا که کلمات در جملاتی نشانده شده‌اند که الگوی آهنگ جملات خبری را دارند بطوری که هسته تکیه جمله بر روی کلمه مورد نظر تحقیق قرار دارد و همچنین قبل از ضبط به صورت اختصاصی با هریک از افراد تمرین می‌شود تا به آهنگ مورد نظر برای این تحقیق دست یابند.

در این تحقیق به واسطه‌ی اینکه از تمامی افراد در یک استودیو تست گرفته شده است عامل نویز محیط نیز ثابت فرض می‌شود.

۱-۶ تعریف مفاهیم

واکه‌های زبان فارسی شامل ای با نشانی /i/ و آ با نشانی /a/ و او با نشانه /u/ و اِ با نشانی /e/ و آ با نشانی /æ/ و اُ نیمه‌بسته با نشانی /o/ می‌باشد. واکه‌های ای /i/ و او /u/ بسته و واکه‌های آ /æ/ و آ /a/ باز هستند. واکه‌های اِ /e/ و اُ /o/ بین این دو گروه قرار دارند. ای /i/ و اِ /e/ و آ /æ/ پیشین و واکه‌های آ /a/ و اُ بسته /o/ و او نیمه‌باز /u/ به ترتیب پسین هستند. واکه‌های ای /i/ و آ /æ/ گسترده و واکه اِ /e/ نیمه‌گسترده می‌باشند و واکه آ /a/ و او /u/ گرد و اُ /o/ همگی گردد می‌باشد. فرمنت: فرکانس شدید دستگاه گفتار می‌باشد.

فصل دوم

مبانی نظری و پیشینه تحقیق

۲-۱ مقدمه

در این فصل ابتدا به بحث و بررسی در خصوص بررسی زبان‌شناختی تولید واکه‌ها در زبان فارسی، بررسی روش‌های نرمال‌سازی، منابع تغییردهندی گفتار و معیار برای انتخاب مناسب‌ترین روش نرمال‌سازی موضوع شبیه‌سازی و استخراج سه فرمت اول واکه‌های زبان فارسی پرداخته می‌شود و سپس به مطالعات و پژوهش‌های قبلی که در خصوص موضوع تحقیق صورت گرفته است، پرداخته می‌شود.

۲-۲ مبانی نظری

مبانی نظری در اینجا شامل بررسی زبان‌شناختی تولید واکه‌ها و بررسی روش‌های نرمال‌سازی می‌باشد که به آنها پرداخته می‌شود.

۲-۲-۱ بررسی زبان‌شناختی تولید واکه‌ها در زبان فارسی

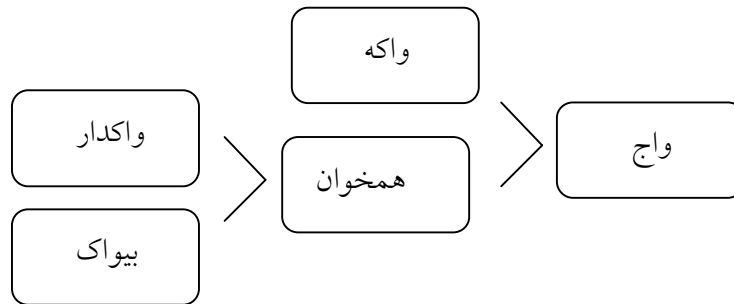
با توجه به اینکه گفتار در هر زبانی با کنار هم قرار گرفتن آواها در آن زبان ایجاد می‌شود، در بررسی واکه‌ها که یکی از اجزاء تشکیل‌دهنده ی آوا در زبانهای مختلف می‌باشد، به‌نظر می‌رسد نقطه شروع و آغازگر در خصوص شبیه‌سازی فرمتهای واکه‌های زبان فارسی بررسی زبان‌شناختی تولید واکه‌ها در زبان فارسی باشد. بدین منظور در قدم اول به تعریف و دسته‌بندی آواهای زبان فارسی و سپس به نقش دستگاه تنفس در چگونگی شکل‌گیری صداهای گفتار پرداخته می‌شود.

آواها^۱ صداهای واضح و مجزایی هستند که سنگ بنای گفتار را در زبان تشکیل می‌دهند. مطالعه و دسته‌بندی صداها در یک زبان به عنوان آواشناسی^۲ شناخته می‌شود (Goldberg, ۲۰۰۰, chap.۲).

۱. Phone

۲. Phonetics

کوچکترین واحد در نظام آوایی یک زبان که جایگزین کردن آن با دیگری می‌تواند سبب تغییر در معنی شود واج^۱ نامیده می‌شود. واج‌ها در زبان فارسی به دو دسته همخوان^۲ و واکه^۳ تقسیم می‌شوند.



همخوان‌ها یا صامت‌ها آوایی هستند که به هنگام تولید، بازدم در گذار از اندام‌های گویایی به مانع برخورد می‌کنند و به عبارت دیگر در دانش آواشناسی صامت‌ها صدای گفتاری هستند که با بسته شدن کامل یا جزئی مجرای صوتی فوقانی ایجاد می‌شوند. مجرای صوتی فوقانی به قسمتی از مجرای صوتی گفته می‌شود که بالاتر از حنجره قرار دارد و از آنجایی که رابطی یک به یکی بین شمار صدای صامت‌ها در زبان‌های دنیا با تعداد حروف صامت در هر الفبایی وجود ندارد، زبان‌شناسان سامانه‌ای ایجاد کرده‌اند به نام الفبای بین‌المللی آواشناسی (IPA) International Phonetic Alphabet تا برای هر صامت مشخص یک نماد تعیین کنند.

زبان فارسی ۲۳ همخوان دارد که در IPA برای صامت‌ها نماد زیر در نظر گرفته شده است.

y, l, n, m, h, f, v, ʒ, x, ʃ, r, ʔ

G, tʃ, k, g, dʒ, z, s, d, tɸ, b

در جدول ۱-۲ همخوان‌های زبان فارسی بر اساس شیوه‌ی تولید، محل تولید و واگذاری^۴ آن‌ها نشان داده شده‌اند. همه همخوان‌های زبان فارسی به وسیله جریان برون‌سوی هوای ششی یا بازدم تولید می‌گردند. هنگام تولید همخوان‌ها در نقطه‌ای از دستگاه گفتار مانع یا تنگی‌ای ایجاد می‌گردد. برحسب این که این انسداد یا تنگی کامل یا ناقص باشد شیوه‌ی تولید همخوان شکل می‌گیرد. چگونگی حالت لب‌ها بر کیفیت آوایی همخوان‌ها کم و بیش اثر می‌گذارد. این اثر از آن جایی ناشی می‌شود که شکل لب‌ها حجم حفری

۳. Phoneme
 ۴. Consonant
 ۳. Vowel
 ۴. Voicing

دهان را که به منزله‌ی یک تشدیدگر است تغییر می‌دهد. محل تولید به محلی از دستگاه گفتار گفته می‌شود که در آن همخوان مورد نظر تولید می‌شود و علامت واک‌دار^۱ و بی‌واک^۲ بودن در هر یک از سلولهای جدول به این ترتیب مشخص می‌شود که علامت سمت چپ نشانگر بی‌واک بودن و علامت سمت راست نشانگر واک‌دار بودن است.

جدول ۱-۲. همخوان‌های زبان فارسی بر اساس محل تولید، نحوه تولید و واکداری آنها

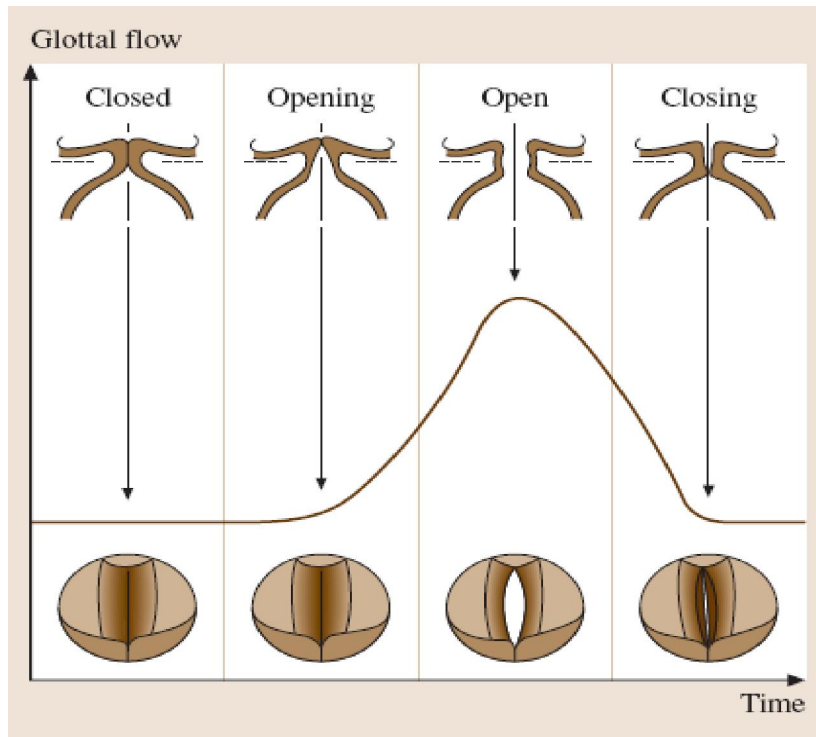
محل تولید نحوه تولید	دولبی	لب و دندانی	دندانی	لثوی	لثوی کامی	کامی	ملازی	چاکنایی
انفجاری	p, b		t, d			k, g	G	ʔ
سایشی		f, v		s, z	ʃ, ʒ		x	h
انفجاری - سایشی					tʃ, dʒ			
لرزشی				r				
خیشومی	m			n				
روان				l		y		

زبان‌شناسان درباره وجود شش واکه‌ی ای (/i/)، آ (/a/)، او بسته (/u/)، ا (/e/)، آ (/æ/)، ا نیمه بسته (/o/) در نظام آوایی زبان فارسی اتفاق نظر دارند (ثمره، ۱۳۶۴).

واکه‌ها آواهای واک‌دار هستند که در اثر باز و بسته شدن حنجره و لرزش تارهای صوتی بدون بروز هیچ مانعی در مجرای گفتار از قبیل گرفتگی و یا تنگی مجرا که موجب ایجاد سایش گردد، تولید می‌شوند. به عبارت دیگر، واکه‌ها بدون حرکت معنی‌دار اندامهای گویایی در هنگام تولید ادا می‌شوند و فقط موقعیت زبان، مقدار انقباض آن و شکل لب بطور موثر در دسته‌بندی صدای واکه‌ها نقش دارد. انقباض در اینجا اشاره به کیفیت نزدیکی زبان به سقف دهان دارد (Goldberg, ۲۰۰۰, chap.۲).

^۱. Voiced
^۲. Unvoiced

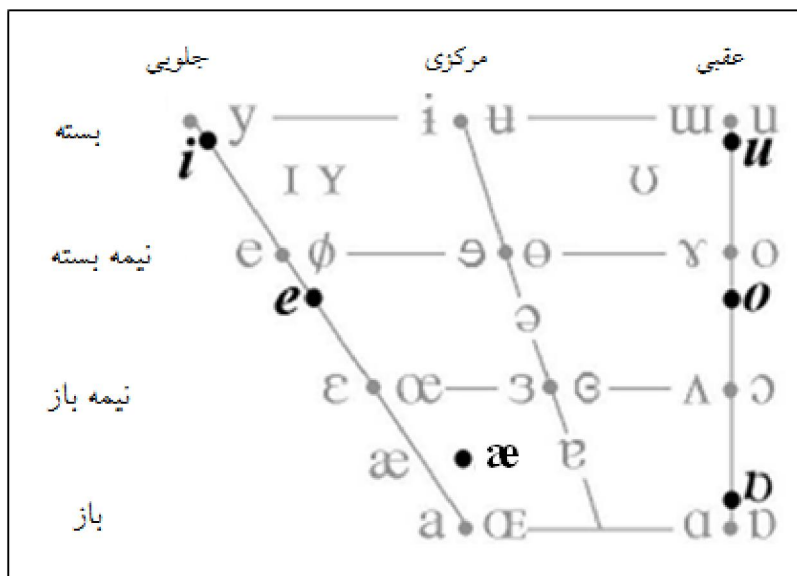
همان طوری که اشاره شد در اثر باز و بسته شدن حنجره و لرزش تارهای صوتی واکه‌ها ایجاد می‌شود. برای اینکه تصویر روشنی از باز و بسته شدن حنجره در ذهن ایجاد شود شکل ۱-۲ باز و بسته شدن حنجره در تولید واکه‌ها را به تصویر کشیده است.



شکل ۱-۲. باز و بسته شدن حنجره و شکل موج جریان هوا پس از عبور از حنجره را نشان می‌دهد (D.o'shoughnessy, ۲۰۰۰, chap ۱۶).

واکه‌های زبان فارسی همه دهانی هستند، یعنی هنگام تولید آن‌ها مسیر بینی، به وسیله نرم‌کام، مسدود است.

در شکل ۲-۲ نحوی تولید واکه‌های زبان فارسی در مقایسه با سایر زبان‌ها نشان داده شده‌اند. در این شکل ستون عمودی میزان بازشدگی دهان را در حین تولید واکه‌ها نشان می‌دهد. محور افقی بیانگر این است که کدام قسمت از زبان برای تولید واکه‌ی مورد نظر درگیر می‌شود.

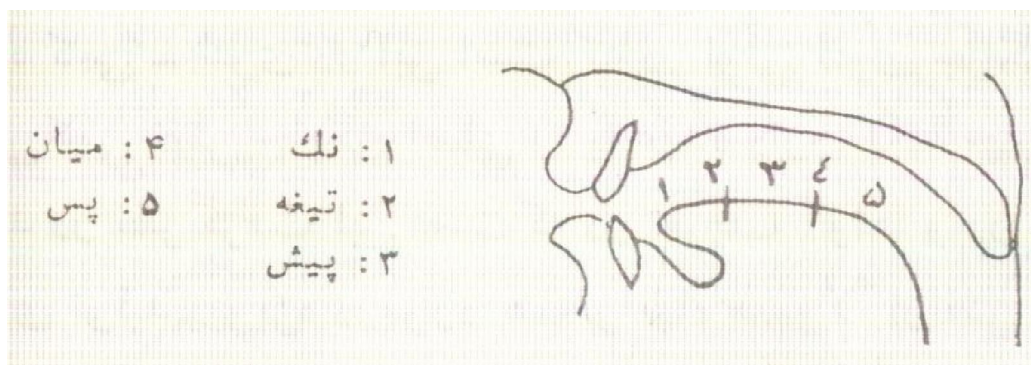


شکل ۲-۲. واکه‌های زبان فارسی در مقایسه با سایر زبان‌ها: محور عمودی میزان بازشدگی دهان و محور افقی محل از زبان را که در حین بیان واکه درگیر است نشان می‌دهد (فاضل، ۱۳۸۸).

همان‌طوری که در شکل ۲-۲ مشاهده می‌شود بر روی محور افقی اصطلاحاتی چون عقبی (پیشین)، مرکزی (میانی) و جلویی (پسین) دیده می‌شود و به همین ترتیب بر روی محور عمودی اصطلاحاتی چون بسته، نیمه بسته، نیمه باز و باز دیده می‌شود در اینجا به توضیح مختصری در خصوص این اصطلاحات پرداخته می‌شود.

در تولید واکه‌ها زبان در تماس با سقف دهان قرار نمی‌گیرد. تنها در واکه‌های بسته ممکن است به آن نزدیک شود. در نتیجه در تعریف واکه‌ها از جایگاههای تولید^۱ نمی‌توان استفاده کرد بلکه واکه‌ها را با توجه به آن قسمتهایی از زبان که در تولیدشان نقش فعال دارند، تعریف می‌کنند. تقسیم‌های زبان در شکل ۱-۳ نشان داده شده است و اصطلاحاتی هم که برای قسمتهای مختلف زبان به کار می‌روند همراه با شکل نشان داده شده است.

^۱.Place of costigaltion



شکل ۳-۲. قسمت‌های مختلف زبان (ثمره، ۱۳۸۶).

نوک زبان آن بخش از زبان است که در حالت استراحت در تماس با دندان‌های پیشین و پایین قرار دارد. تیغی زبان آن قسمت از زبان است که در حالت استراحت در تماس با لثه بالائی قرار می‌گیرد. قسمت پیشین زبان در حالت استراحت در تماس با پیشکام، کام و پسکام قرار می‌گیرد. قسمت پسین در حالت استراحت در تماس با نرمکام و ملاز قرار دارد. قسمت میانی مرز بین قسمت پیشین و قسمت پسین است.

همان‌طوری که قبلاً گفته شد محور عمودی بیانگر ارتفاع زبان یا میزان باز شدگی دهان می‌باشد. فک پایین می‌تواند بر روی محور خود به سوی پایین حرکت کند. با حرکت فک به سوی پائین، دندان‌های زیرین و زبان نیز به پایین کشیده می‌شوند. در نتیجه این حرکت، دهان باز می‌شود و زبان از سقف دهان و دندان‌های زیرین از دندان‌های بالا فاصله می‌گیرند. در این حالت فاصلی بین زبان و سقف دهان را به اصطلاح ارتفاع زبان می‌گویند. درجه‌ی ارتفاع زبان بسته به میزان افراستگی متفاوت است. درآشناسی این ارتفاع را گاه به سه درجه و گاه به چهار درجه تقسیم می‌کنند. ما در اینجا به پیروی از اصول خطِ آوانگاری بین‌المللی این ارتفاع را به چهار درجه به صورت زیر تقسیم می‌کنیم.

۱- باز: وقتی که فک، زبان و دندان‌های پایین تا آنجا که میسر است از سقف دهان فاصله گرفته

باشند. آن‌گونه که در تلفظ /a/ و /æ/ در زبان فارسی اتفاق می‌افتد.

۲- بسته: وقتی که فک، زبان و دندان‌های پایین تا آنجا که میسر است به سقف دهان نزدیک شده

باشند. آن‌گونه که هنگام تلفظ /i/ و /u/ در زبان فارسی اتفاق می‌افتد.

۳- نیمه باز: وقتی که اندام‌های فوق حدود یک چهارم ارتفاع باز را داشته باشند. آن‌گونه که در تلفظ

/ε/ و /œ/ در انگلیسی مشاهده می‌شود.

۴- نیمه بسته: وقتی که اندامهای فوق حدود سه چهارم ارتفاع باز را داشته باشند. آن گونه که در تلفظ /o/ و /e/ در زبان انگلیسی مشاهده می‌شود.

حالا پس از دسته‌بندی آواها و بررسی واژه‌ها به نقش دستگاه تنفس در تولید آوا و چگونگی شکل‌گیری صدای گفتار پرداخته می‌شود.

دستگاه گفتاری از سیستم تنفسی، حنجره و لوله صوتی تشکیل شده است. سیستم تنفسی، منبع تولید انرژی است که شامل پرده‌ی دیافراگم، شش‌ها و نای می‌باشد. حنجره به واسطه‌ی تارهای صوتی منبع تولید صوت است. لوله صوتی، نیز نقش تشدیدگر را بازی می‌کند. همان‌طور که در شکل ۴-۲ نشان داده شده است، این لوله شامل گلوگاه، حفری بینی و حفری دهان می‌شود.

این قسمت به این دلیل به نام لوله صوتی اطلاق می‌شود چون در دستگاه تنفس اگر از انحنای محل تلاقی حلق و حفری دهانی چشم‌پوشی شود می‌توان این مسیر را به یک لوله تشبیه نمود که در یک انتها توسط حنجره‌ای که تارهای صوتی در آن قرار دارد بسته می‌شود و در انتهای دیگر که لبها قرار دارد باز می‌باشد (Hayward, ۲۰۰۰, chap ۴).

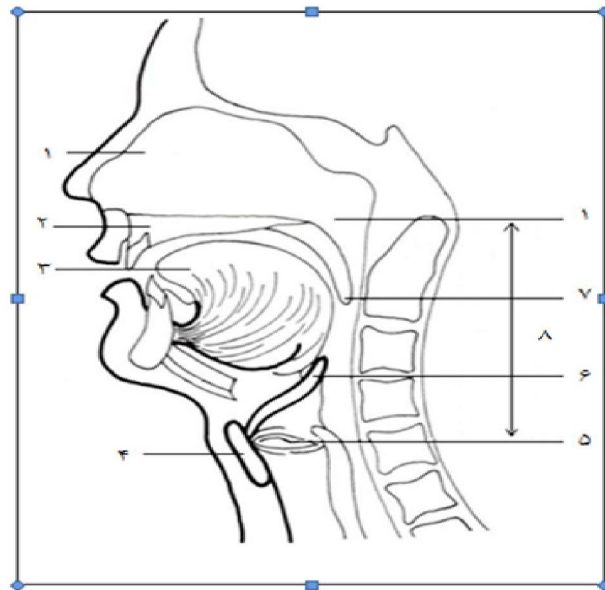
با این توصیف به مسیر عبور هوای بالای حنجره لوله صوتی گفته می‌شود. شکل لوله صوتی در تولید گفتار بسیار حائز اهمیت است. لوله صوتی به سه قسمت اصلی حفره‌های حلق، دهان و بینی تقسیم می‌شود. که این حفرات نقش تشدیدگر برای صدای خروجی از تارهای صوتی ایفا می‌کنند.

فرکانسهای رزونانس این لوله فرکانسهای تشدید یا فرمتهای نامیده می‌شود. فرکانس فرمنت دومین عامل اصلی است که تعیین می‌کند کدام آوا بوسیله دستگاه گفتار تلفظ شده است (Goldberg, ۲۰۰۰, chap. ۲).

صداسازی با خروج با فشار هوای ریه‌ها که تولید یک نوفه می‌کند شروع می‌شود. این نوفه اولیه با اعمال لب‌ها و دندان‌ها و سایر اندام‌های گویایی تغییر کرده و واج‌ها را در هر زبانی می‌سازد. انقباض و حرکت ماهیچه‌های شکمی بر دیافراگم فشار آورده و موجب خروج هوا از شش‌ها می‌شود. همان‌طوری که قبلاً اشاره شد اگر در حین عبور جریان هوا، تارهای صوتی به ارتعاش در آیند و تولید صدا کنند، به آن صدا، واک‌دار و اگر بدون ارتعاش تارهای صوتی هوا عبور کند به آن صدا بی‌واک گفته می‌شود.

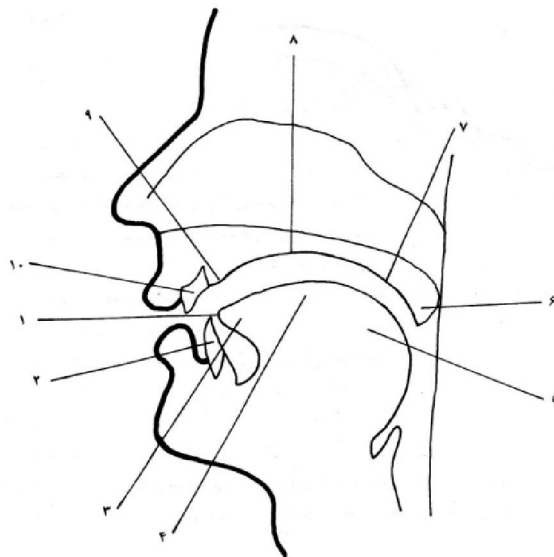
اگر هوا وارد حفره بینی شده و از آن‌جا خارج شود به صدای مورد نظر خیشومی گفته می‌شود و اگر هوا از دهان خارج شود، آن صدا را دهانی می‌نامند. در حفره دهان اندام‌های گویایی که شامل نوک زبان،

دندان‌های پایین، تیغی زبان، جلوی زبان، عقب و ریشی زبان می‌شوند، تغییراتی را در موج صدا ایجاد کرده و از آن جا هوای مرتعش شده توسط لب‌ها پراکنده می‌شود (Ladefoged، ۱۹۶۸).



شکل ۴-۲. برش گلوگاه و دهان

- ۱- حفره‌ی بینی: ۲- سخت‌کام: ۳- زبان: ۴- غضروف تیروئید:
 ۵- تارهای صوتی: ۶- دریچه‌ی نای: ۷- نرم‌کام: ۸- گلوگاه (ثمره، ۱۳۸۶).

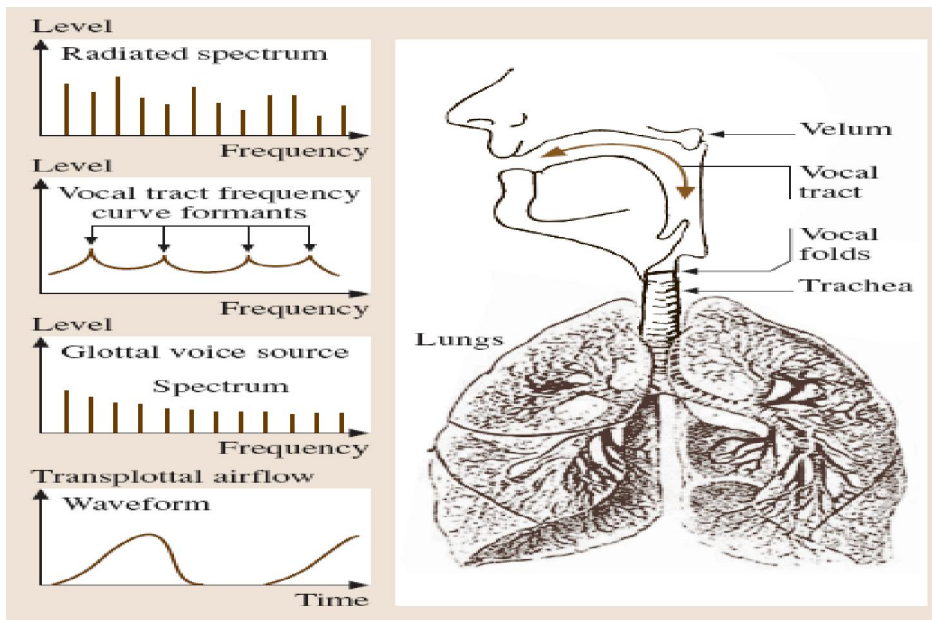


شکل ۵-۲. برش دهان

- ۱- نوک زبان: ۲- دندان‌های پایین: ۳- تیغی زبان: ۴- جلوی زبان: ۵- عقب زبان: (ثمره، ۱۳۸۶).

بر اساس نظریه‌ی منبع- فیلتر که در شکل ۶-۱ نشان داده شده است. صدای واکه در نتیجه سه فرآیند تولید می‌شود.

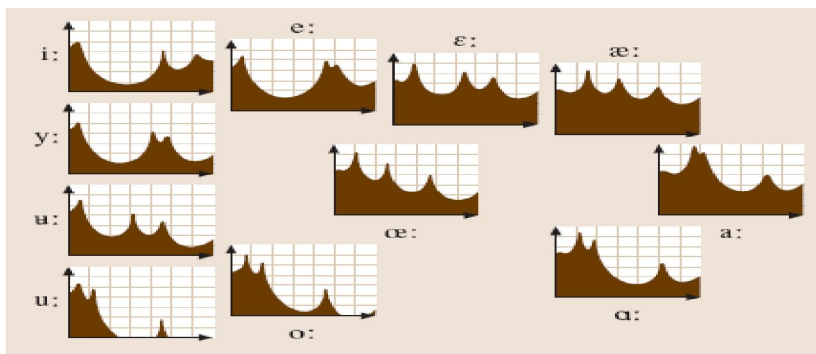
- ۱- تولید جریان یکنواختی از هوا در ششها که به این مرحله آغازش Initiation گفته می‌شود.
- ۲- تبدیل این جریان به یک تکانه متناوب با استفاده از تارهای حنجره که سیگنال تحریک خوانده می‌شود.
- ۳- پاسخ لوله صوتی به این سیگنال تحریک که بوسیله تابع انتقال دستگاه تنفس انجام می‌شود (D.o'shoughnessy, ۲۰۰۰, chap ۱۶).



شکل ۶-۲. قسمتهایی از دستگاه گفتار و روند تولید صدا (D.o'shoughnessy, ۲۰۰۰, chap ۱۶).

در شکل ۶-۲ علاوه بر اینکه قسمتهای مختلفی را از دستگاه گفتار شامل ششها، نای، تارهای صوتی و نرمکام نمایش می‌دهد، روند تولید صدا را نیز با استفاده از نمودارهای ترسیم‌شده در زمان وفرکانس از پایین به بالا نشان می‌دهد. نمودار اول شکل موج جریان هوا در ششها را در حوزی زمان نشان می‌دهد که پس از ارتعاش تارهای صوتی، تکانه‌های متناوب صدا را متناظر با تعدادی از همسازهای

قوی که دامنه‌اش با فرکانس کاهش می‌یابد، ایجاد می‌کند. این طیف فرکانسی مطابق با خصوصیات انتقال صدای دستگاه تنفس شامل قله‌ها و دره‌های بین آنها فیلتر می‌شود در طیفی از لبهای باز متشعشع می‌شود. همان‌طوری که در شکل ۷-۲ دیده می‌شود فرمتهای در نقش قله‌هایی می‌باشند که دامنه بیشتری از فرکانس‌های مجاور دارند و هر واکه در مقایسه با واکه دیگر فرکانس فرمنت متفاوتی دارد که از ویژگیهای آن واکه به شمار می‌رود.



شکل ۷-۲. سطوح دامنه و فرمنت‌های واکه‌های زبان سوئدی نشان می‌دهد (D.o'shoughnessy, ۲۰۰۰, chap ۱۶).

در واقع در شکل‌گیری آوا دو فاکتور عامل و فیلتر نقش دارد. تحریک به وسیله هوای بیرون رانده شده از ششها ایجاد می‌شود و همان‌طوری که قبلاً گفته شد در صورتیکه این جریان هوا باعث ارتعاش تارهای صوتی شود تکانه شبه متناوبی را ایجاد می‌کند که این تکانه آواهای واک‌دار (voice) را شکل می‌دهد. سپس لوله صوتی تحت تاثیر خصوصیات تشدید زبان، حلق و حفری دهانی در نقش فیلتر بر روی تحریک ایجاد شده تغییر ایجاد می‌کند.

دومین عامل در تولید آوا بوسیله نويز بوجود می‌آید. به این ترتیب جریان یکنواخت ایجاد شده نويزی را خلق می‌کند و طیفی که از انقباض قسمتهای فوقانی دستگاه تنفس ایجاد می‌شود اصوات بی‌واک (unvoiced) را شکل می‌دهد اما با توجه به اینکه در اینجا تنها تولید واکه‌ها مورد بررسی قرار می‌گیرد، صرفاً به بحث درباره سیگنالهایی با تکانه‌های متناوب پرداخته می‌شود.

در هنگام تولید واکه تعدادی از اندامهای گویایی در شکل‌گیری مقدار پارامتری فرمتهای نقش دارند. به این ترتیب می‌توان گفت :

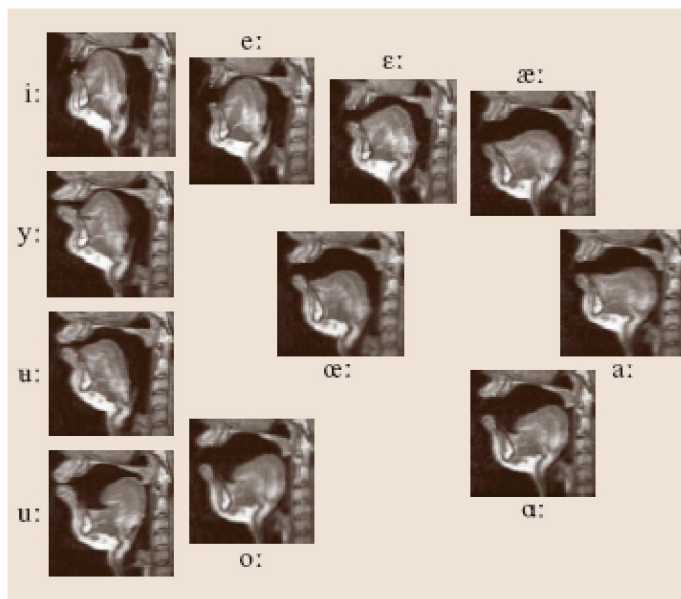
F₁ بوسیله ارتفاع زبان کنترل می‌شود.

F₂ با حرکت جلو- عقب زبان (پسین- پیشین) بطور معنی‌داری تغییر می‌کند.

F₃ با شکل لبها، گرد یا گسترده بودن در ارتباط است (D.o'shoughnessy, ۲۰۰۰, chap ۱۶).

منظور از شکل لبها، گرد یا گسترده که در بالا به آن اشاره شد، شکلی است که لبها هنگام تولید آوا به خود می‌گیرند که در طبیعت آوا تأثیر می‌گذارد. در جریان تولید آوا ممکن است لبها به پیش رانده شوند و گرد شوند. در این صورت حالت لبها را گرد می‌گویند. آن مختصه‌ی آوایی که به کمک گرد شدن لبها در آوا ظاهر می‌شود در رابطه با واکه به نام گرد خوانده می‌شود. از طرف دیگر آن مختصه‌ی آوایی نیز که با گسترده شدن لبها در آوا ظاهر می‌شود به همین نام یعنی گسترده مرسوم است.

به این ترتیب اندام‌های گویایی در تولید هر واکه شکل‌های مختلفی می‌گیرند. در شکل ۸-۲ شکل‌گیری اندام‌های گویایی در تولید واکه‌های زبان سوئدی نمایش داده شده است که تعدادی از آنها برای شش واکه زبان فارسی نیز مصداق دارد.



شکل ۸-۲. شکل‌گیری اندام‌های گویایی در تولید واکه‌های زبان سوئدی (D.o'shoughnessy, ۲۰۰۰, chap ۱۶).

۲-۲-۲ بررسی روش‌های نرمال‌سازی

زبان‌شناسان اهمیت اول، دوم و سومین پیک یا فرمنت را در تعیین کیفیت واکه تشخیص داده‌اند اما مقدار فرمنت‌ها در افراد مختلف به علت تفاوت در اندازه لوله صوتی متفاوت می‌باشد. هر چند این تفاوتها برای برآورده کردن برخی اهداف همچون استخراج تفاوت‌های گویندگان به سبب خصوصیات ذاتی افراد و ارائه یک مدل آکوستیکی از واکه‌های یک زبان مفروض مطلوب به نظر می‌رسد اما برای اینکه

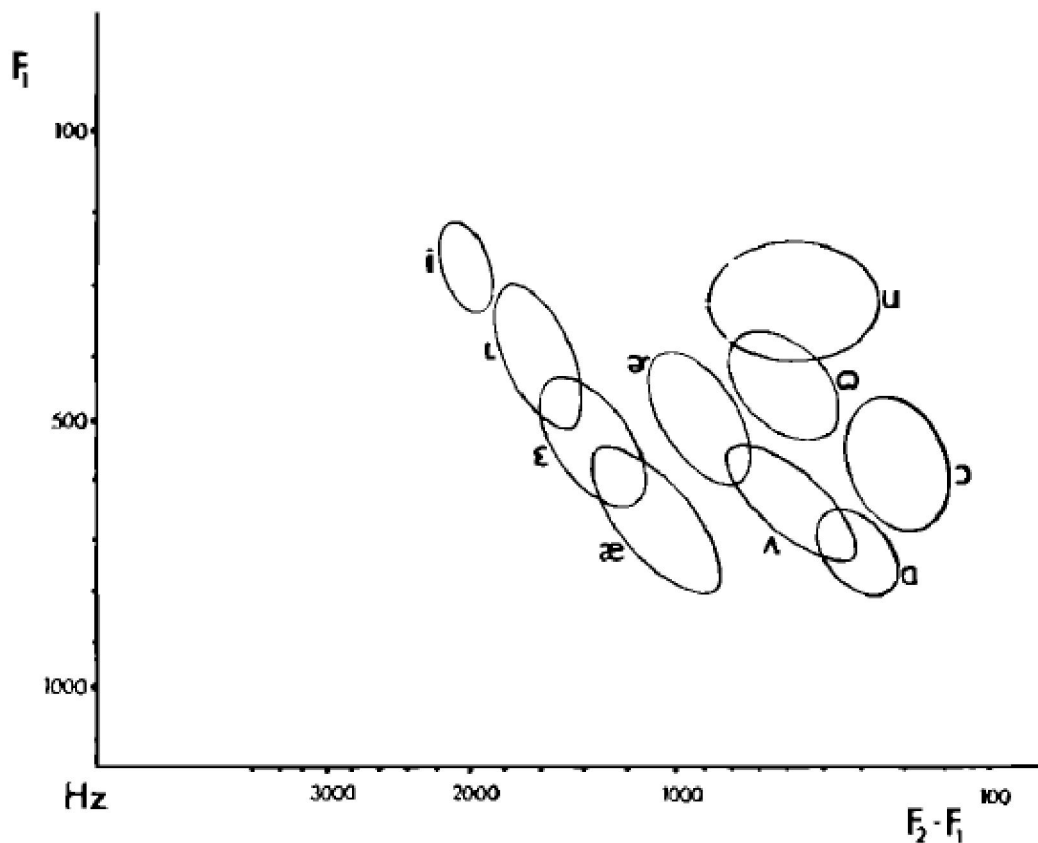
بتوان بطور مؤثر اثر ویژگی گوینده را بر روی داده‌ها حذف کرد از نرمال‌سازی استفاده می‌شود. چرا که نرمال‌سازی یک وسیله با ارزش برای دسته‌بندی واژه‌ها می‌باشد (Disner, ۱۹۷۹).

ماکزیمم کاهش در واریانس درون گروه واژه بعنوان معیاری برای درجه‌بندی الگوریتم‌ها برای نرمال‌سازی واژه‌ها در نظر گرفته می‌شود. اما این معیار به‌تنهایی کافی نیست. بنابراین برای تضمین اینکه روندی که در نرمال‌سازی واژه‌ها پیش می‌گیریم درست باشد، باید دقت نمود چون نمی‌توان تنها به خروجی تأکید کرد. زیرا روند نرمال‌سازی باید ارائه درستی از ارزشهای وابسته به زبان‌شناسی داشته باشد. منظور از ارزشهای وابسته به زبان‌شناسی حفظ اطلاعات آوایی در داده‌های حاصل می‌باشد.

دیسنر (۱۹۷۵) واژه‌ها در زبانهای مختلف را در فضای فرمتهی با F_1 در امتداد عرض و اختلاف بین F_1 و F_2 در امتداد طول ترسیم کرد. لُبانو^۱ فضای اشغال‌شده به وسیله هر واژه را با یک بیضی متوسط مشخص کرد. داویس^۲ (۱۹۷۶) برای هر دسته از نقاط واژه یک بیضی با شعاع انحراف استاندارد ۲ در امتداد محورها متمایل به اجزاء اصلی در نظر گرفت. این در حالی است که لادفوغد^۳ (۱۹۷۹) مسافت در امتداد محورها را متناسب با مقیاس میل (MEL) ارائه داد. وی معتقد بود که این طرح نوعی نمایش خوب برای درک فاصله در فضای آوایی می‌باشد.

در شکل ۹-۲ نمونه‌ای از این نمودار به همراه بیضی‌هایی که برای هر واژه ترسیم شده نمایش داده شده است.

۲. Lobanov
۳. Davis
۱. Ladefoged



شکل ۹-۲. هر بیضی ترسیم شده دسته‌ای از نقاط واکه را نشان می‌دهد (Disner, ۱۹۷۹).

روش بیضی برای مقایسه کاهش پراکندگی در روش نرمال‌سازی با دقت قابل توجهی عمل می‌کند. به این ترتیب این روش کاربرد عملی برای مقایسه بین روش‌های نرمال‌سازی دارد و در واقع در این روش بهترین روش آنی است که اندازه بیضی‌ها را بیشتر کاهش می‌دهد. البته این نکته باید ذکر شود که شکل واقعی فضای بیضی از یک زبان به زبان دیگر خیلی متفاوت می‌باشد که این ممکن است به سبب اثر گویندگان متفاوت، تعداد واکه‌ها و توزیع واکه‌ها در فضای فرمنت برای زبانهای مختلف باشد (Disner, ۱۹۷۹).

بطور سنتی روش نرمال‌سازی واکه توسط اینسورث^۱ (۱۹۷۵) و نیری^۲ (۱۹۸۹) به دو دسته واکه-درون‌گرا^۳ و واکه-برون‌گرا^۴ تقسیم می‌شوند. این تقسیم‌بندی مطابق با نوع داده‌ای که به کار

۱. Ainsworth
 ۲. Neary
 ۳. Vowel-intrinsic
 ۴. Vowel-extrinsic