

مدل سازی بازشناسی واجی کلمات فارسی

وحید صادقی

گروه زبان انگلیسی، دانشگاه بین‌المللی امام خمینی (ره)

چکیده

در این مقاله مدلی برای درک و بازشناسی کلمات زبان فارسی ارائه شده است. مدل پیشنهادی مبتنی بر نظریه‌ی صوت‌شناختی درک گفتار است. اساس این نظریه، در بازنمود آوایی هر واحد زبانی، مجموعه‌ای از الگوهای صوت‌شناختی به‌عنوان الگوهای آوایی مرجع، ذخیره می‌شوند و درک آن واحد از طریق انطباق الگوهای صوت‌شناختی درون‌داد با الگوهای آوایی مرجع در بازنمایی آوایی صورت می‌گیرد. انگاره‌ی پیشنهادی برای درک و بازشناسی بازنمود آوایی و واجی کلمات از ترکیب مدل المان و مک کلند [۱] تحت عنوان "اثر" و مدل بورسما [۲] تحت عنوان "دستور ادراکی و دستور بازشناسی" است.

واژه‌های کلیدی: بهینگی، دستور ادراکی و بازشناسی، الگوهای آوایی مرجع، محدودیت‌های صوت‌شناختی و واجی

۱- مقدمه

در این مقاله، به مدل‌سازی درک و بازشناسی خودکار کلمات زبان فارسی می‌پردازیم. مدلی که در این مقاله برای بازشناسی کلمات ارائه می‌شود، تلفیقی از مدل المان و مک کلند [۱] تحت عنوان "اثر" و مدل بورسما [۲] تحت عنوان "دستور ادراکی و دستور بازشناسی" است. در این مدل، برای بازشناسی کلمات، رشته علائم صوت‌شناختی گفتار به m قالب زمانی به طول n میلی ثانیه تقسیم شده است، به طوری که بین هر دو قالب زمانی به اندازه‌ی s میلی ثانیه هم‌پوشی زمانی وجود دارد. از نظر آواشناسی، هر قالب زمانی تظاهر صوت‌شناختی یکی از سه مرحله‌ی آغاز، اوج و پایانه‌ی یک بستر تولیدی است. این مدل شامل سه لایه‌ی واحدیاب است:

(۱) لایه‌ی مشخصه‌یاب

(۲) لایه‌ی واج‌یاب

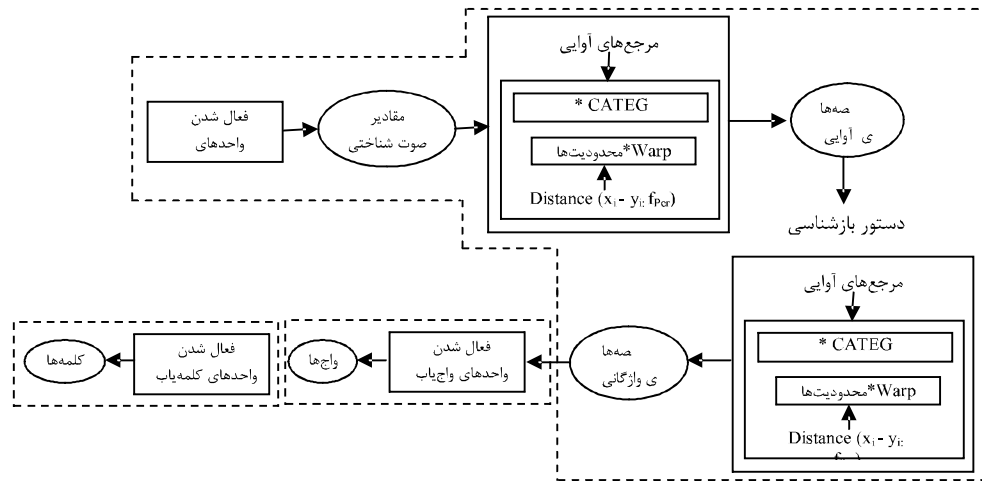
(۳) لایه‌ی کلمه‌یاب

شکل (۱) فرآیند بازشناسی کلمه را در لایه‌های آوایی موردنظر بر اساس مدل پیشنهادی نگارنده نشان می‌دهد.

درون‌داد مدل، امواج صوتی حاصل از فعالیت دستگاه گفتار است. امواج صوتی، مطابق شکل، ابتدا وارد لایه‌ی مشخصه‌یاب می‌شود، لایه‌ی مشخصه‌یاب، امواج صوتی را طی مراحل زیر به مشخصه‌های واجی تبدیل می‌کند:

مرحله‌ی اول: برای هر قالب زمانی در علامت‌گفتاری، مجموعه‌ای از واحدهای مشخصه‌یاب فعال می‌شوند. با فعال شدن هر واحد مشخصه‌یاب، مقدار x_i برای پارامتر صوت‌شناختی متناظر با واحد مشخصه‌یاب به‌دست می‌آید.

مرحله‌ی دوم: در این مرحله، شنونده مقادیر پارامترهای صوت‌شناختی را با استفاده از دستور ادراکی به مشخصه‌های آوایی (ادراکی) دوارزشی تبدیل می‌کند. برای این منظور، وی از دو دسته محدودیت صوت‌شناختی استفاده می‌کند: محدودیت‌های CATEG* و محدودیت‌های WARP*. محدودیت‌های CATEG* به‌طور کلی، درک مقادیر مختلف یک پارامتر صوت‌شناختی را به‌صورت مشخصه‌های آوایی متناظر با آن‌ها غیرمجاز می‌دانند.



شکل (۱) - فرآیند بازشناسی کلمه در انگاره‌ی پیشنهادی نگارنده

به‌ناچار به‌صورت یک مرجع آوایی درک شود. این درک تابع محدودیت‌های *WARP است. صورت‌بندی این محدودیت‌ها به‌صورت زیر است:

$$*WARP(f;d) \Leftrightarrow \exists x_i \in f_{ac} \exists y_i \in f_{perc} \rightarrow |x_i - y_i| < d$$

صوت شناختی و perc ادراکی بر اساس این محدودیت‌ها، الگوی آوایی مرجع y از پارامتر ادراکی f نباید به اندازه‌ی هر مقداری از d متفاوت از مقدار x در مابه‌ازای صوت‌شناختی آن پارامتر باشد. به عبارت دیگر، محدودیت‌های *WARP وجود هر گونه فاصله‌ی ادراکی را بین الگوهای آوایی مرجع یک مشخصه‌ی آوایی و مقادیر صوت‌شناختی آن مشخصه غیرمجاز می‌دانند. آنچه مهم است، آرایش این محدودیت‌ها نسبت به یکدیگر است؛ به این صورت که هر قدر این فاصله‌ی ادراکی بیشتر باشد، رتبه‌ی محدودیت مربوطه در ساخت سلسله‌مراتبی بالاتر می‌باشد و نقض آن، جریمه‌ی سنگین‌تری به‌همراه خواهد داشت. بنابراین نقش محدودیت‌های *WARP، آن است که مقادیر صوت‌شناختی را با حداقل فاصله‌ی ادراکی به الگوهای آوایی مرجع، یعنی مشخصه‌های دوارزشی در بازنمود آوایی، بنگارند.

مرحله‌ی سوم: در این مرحله، شنونده با استفاده از دستور بازشناسی، مشخصه‌های آوایی را به مشخصه‌های واجی (واژگانی) تبدیل می‌کند. برای این منظور، از محدودیت‌های *CATEG و *WARP واجی استفاده می‌شود. محدودیت‌های *CATEG واجی با توجه به مرجع‌های واجی، الگوهای آوایی را در بازنمود آوایی کلمات به تعداد کمتری الگوهای واجی در بازنمود واجی (واژگانی) کلمات مقوله‌بندی می‌کنند.

به‌عبارت دیگر، محدودیت‌های *CATEG اجازه نمی‌دهند مقادیر صوت‌شناختی به‌صورت الگوهای آوایی ناپیوسته درک شوند. از طرف دیگر، مقادیر صوت‌شناختی بر اساس محدودیت Perc باید به یک الگوی آوایی در بازنمود ادراکی نگاشته شوند. حال، سؤال این است که اگر برای درک مشخصه‌ی آوایی f درک x_i به‌صورت y_i ، به‌صورت x_z و y_z بر اساس محدودیت‌های *CATEG غیرمجاز باشد، در آن صورت، مشخصه‌ی f در نهایت چگونه درک می‌شود. برای این منظور از دستور واج‌شناسی به‌پهینگی استفاده می‌کنیم. در این دستور، نحوه‌ی آرایش محدودیت‌ها با یکدیگر متفاوت است. اگر y_i که x_i به آن نگاشته می‌شود، یک الگوی آوایی مرجع باشد؛ یعنی الگویی که به‌صورت یک مقوله‌ی ادراکی در بازنمود آوایی f افزاز شده است، در آن صورت محدودیت *CATEG ناظر بر آن یعنی محدودیت *CATEG ($y_i : x_i$) کمترین رتبه را در ساخت سلسله‌مراتبی آرایش محدودیت‌ها دارد و نقض آن جریمه‌ی کمتری را به‌دنبال خواهد داشت؛ درحالی‌که اگر y_i یک الگوی آوایی مرجع نباشد، محدودیت ناظر بر درک x_i به‌صورت y_i رتبه‌ی بالایی داشته و نقض آن جریمه‌ی سنگینی به‌همراه خواهد داشت. چون فضای صوت‌شناختی یک فضای پیوسته با مقادیر نامحدود و فضای ادراکی، یک فضای ناپیوسته با مقادیر محدود است؛ بنابراین باید گفت، نقش محدودیت‌های *CATEG مقوله‌بندی مجموعه‌ای از مقادیر پیوسته‌ی صوت‌شناختی به‌صورت تعداد محدودی الگوی آوایی است.

محدودیت‌های *WARP زمانی اعمال می‌شوند که مقدار x از پارامتر صوت‌شناختی متناظر با مشخصه آوایی f منطبق بر الگوی آوایی مرجع y نیست. در چنین حالتی x باید

منظور، هر الگوی مرجع را می‌توان به صورت میانگین و پراش صادقی، [۴] یا تابع توزیع امکان یک پارامتر صوت‌شناختی بی‌جن‌خان، [۳] به دست داد. فرض بر آن است که این مرجع آوایی به نوعی در ذهن اهل زبان بازنمایی می‌شود. همپوشی زمانی- مکانی الگوهای مرجع مشخصه‌های واجی در بازنمود آوایی، الگوی مرجع واج‌ها و همپوشی زمانی الگوهای مرجع واج‌ها، الگوی مرجع کلمه را به دست می‌دهد.

ب) مرحله‌ی بازشناسی:

در این مرحله، ساخت آوایی کلمه در فضای صوت‌شناختی به ساخت آوایی کلمه در حوزه‌ی درک شنونده نگاشته شده است و بازنمود واجی کلمه به دست می‌آید؛ یعنی ساخت آوایی کلمه در فضای صوت‌شناختی که به صورت همپوشی زمانی- مکانی الگوهای آوایی پیوسته است با ساخت آوایی کلمه در ذهن شنونده که به صورت همپوشی زمانی- مکانی الگوهای آوایی مرجع است، منطبق شده و الگوی واجی کلمه بازشناسی می‌شود.

نگارنده، الگوی آوایی مرجع برای مشخصه‌های طبقه‌ی واجی، واگذاری- بی‌واکی، محل تولید بستار همخوانی و محل تولید و میزان گرفتگی بستار واکه‌ای را قبلاً با استخراج میانگین و پراش پارامتر یا پارامترهای متناظر با این مشخصه‌ها به دست آورده است. در این مقاله سعی شده است با انتخاب برخی از این پارامترها روش درک و بازشناسی مشخصه‌های واجی (و سپس واج‌ها و کلمه‌ها) مدل‌سازی شود.

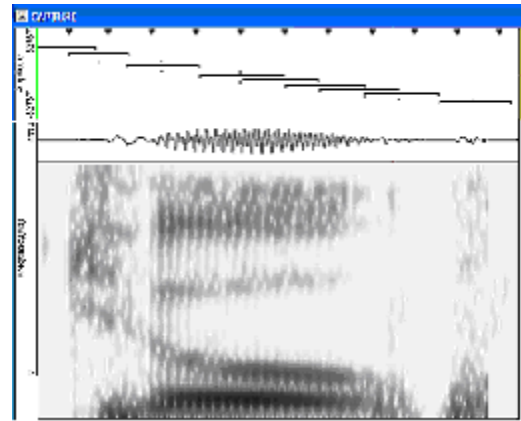
در این بخش، مراحل مختلف بازشناسی واجی کلمه را در مدل پیشنهادی نگارنده برای کلمه‌ی "توپ" نشان می‌دهیم.

۲-۱- لایه‌ی مشخصه‌یاب

گفتیم که بازشناسی مشخصه‌های واجی از علامت صوت‌شناختی گفتار در لایه‌ی مشخصه‌یاب، طی سه مرحله انجام می‌شود:

مرحله‌ی اول:

مقادیر پارامترهای صوت‌شناختی مربوط به هر قالب زمانی تعیین می‌شود. فرض بر آن است که برای هر قالب زمانی، مجموعه‌ای از یک تا چند واحد مشخصه‌یاب فعال می‌شوند. شکل (۲) علامت گفتاری (قسمت بالا) و طیف‌نگاشت (قسمت پایین) رشته‌ی آوایی [tup] را نشان می‌دهد. به طوری که مشاهده می‌شود، این رشته‌ی آوایی به ۱۲ قالب زمانی تقسیم شده است؛ طول هر قالب زمانی ۲۵



شکل (۲) سیگنال زمانی (بالا) و طیف‌نگاشت (پایین) کلمه‌ی "توپ"

محدودیت‌های WARP* واجی نیز در صورت وجود هر گونه فاصله‌ی ادراکی بین درون‌داد آوایی و مرجع‌های واجی، مشخصه‌های آوایی را به عنوان درون‌داد به الگوی واجی که بیشترین شباهت آوایی ممکن (یا کمترین فاصله‌ی آوایی ممکن) را با آن دارد، می‌نگارند.

پس از استخراج مشخصه‌های واجی در لایه‌ی مشخصه‌یاب، این مشخصه‌ها وارد لایه‌ی واج‌یاب می‌شوند؛ در این لایه، با فعال شدن هر واحد واج‌یاب برای هر سه تا شش قالب زمانی، درخت واجی هر واج با ترکیب مشخصه‌ها به دست می‌آید. در مرحله‌ی آخر، واج‌ها به عنوان درون‌داد، وارد لایه‌ی کلمه‌یاب شده و با ترکیب با یکدیگر، درخت واجی کلمه را به دست می‌دهند.

۲- داده‌های زبانی

در طراحی مدل‌های بازشناسی گفتار، دو مرحله‌ی اساسی وجود دارد:

الف) مرحله‌ی تعیین الگوی آوایی مرجع برای کلمات:

چنان‌که می‌دانیم، بازشناسی واجی کلمه، تابعی از بازشناسی کوچک‌ترین واحدهای واجی، یعنی مشخصه‌های واجی است؛ بنابراین الگوی آوایی مرجع برای هر کلمه تابعی از الگوی آوایی مرجع مشخصه‌های واجی سازنده‌ی آن است. الگوی آوایی مرجع هر مشخصه بر حسب الگوهای صوت‌شناختی و با محاسبه‌ی مقادیر پارامترهای صوت‌شناختی تعیین می‌شود. مرجع آوایی باید به گونه‌ای تعیین شوند که تمامی تظاهرات آوایی مشخصه موردنظر را شامل شود. برای این

۱- این مرحله در سامانه‌های بازشناسی گفتار، آموزش نام دارد.

میلی ثانیه است. بین هر دو قالب زمانی به اندازه‌ی ۱۲/۵ میلی ثانیه همپوشی زمانی وجود دارد. این همپوشی زمانی ناظر بر همپوشی بستارهای تولیدی هنگام فعالیت دستگاه گفتار است. حال برای هر قالب زمانی، مجموعه‌ای از واحدهای مشخصه‌یاب را جهت بازشناسی مشخصه‌های واجی فعال می‌کنیم. در این تحقیق از پارامتر کشش نوفه‌ی رهش به‌عنوان واحد مشخصه‌یاب واکداری-بی‌واکی استفاده شده است؛ این پارامتر ناظر بر مدت زمان بازماندن چاکنای از لحظه‌ی باز شدن بست تا لحظه‌ی شروع تناوب چاکنایی برای واکه‌ی بعد می‌باشد و مرجعی مناسب برای مقایسه‌ی تقابل واکداری-بی‌واکی همخوان‌های انسدادی بی‌واک نامیده، بی‌واک دمیده و واکدار است/لده‌فوغد و مدیسون [۵]، (۱۹۹۶: ۴۵). همچنین از همین پارامتر به‌عنوان واحد مشخصه‌یاب محل تولید استفاده شده است. تقابل محل تولید همخوان‌های انسدادی از نظر پارامتر آوایی کشش نوفه‌ی رهش تظاهر آوایی دارد، یعنی فاصله‌ی زمانی میان لحظه‌ی باز شدن بست انفجاری و لحظه‌ی شروع ارتعاش تارآواها برای واکه در بستارهای تولیدی مختلف به‌طور معنی‌داری با یکدیگر تفاوت دارد؛ به این صورت که متوسط کشش نوفه‌ی رهش در بستارهای لبی کم، در بستارهای تیغه‌ای، در یک سطح میانی و در بستارهای بدنه‌ای زیاد است؛ لده‌فوغد و چو، [۶]چو، جان و لده‌فوغد، [۷]. به علاوه، پارامتر نرخ تقاطع صفر و انرژی به‌هنگار را به‌عنوان واحد مشخصه‌یاب طبقه آوایی (گره‌ی ریشه در هندسه مشخصه‌ها) و پارامترهای F1 و F2 را به ترتیب به‌عنوان واحدهای مشخصه‌یاب سطح مقطع گرفتگی و محل تولید بدنه‌ی زبان مورد استفاده قرار داده‌ایم.

مرحله‌ی دوم:

در مرحله‌ی دوم، نحوه‌ی نگاشت مقادیر پیوسته‌ی صوت‌شناختی بر مشخصه‌های ناپیوسته‌ی (دوارزشی) آوایی مدل‌سازی می‌شود. این مدل‌سازی توسط دستور ادراکی صورت می‌گیرد. قبلاً گفتیم که دستور ادراکی شامل دو نوع محدودیت است: محدودیت‌های CATEG* و محدودیت‌های WARP*. درک مقادیر صوت‌شناختی به‌صورت مشخصه‌های آوایی تابعی از میزان نقض محدودیت‌های CATEG* و WARP* با توجه به نحوه‌ی آرایش این محدودیت‌ها نسبت به یکدیگر است.

در این بخش می‌خواهیم روش درک مقادیر F1، NOT و F2، نرخ تقاطع صفر و انرژی به‌هنگار را به‌صورت

مشخصه‌های آوایی متناظرشان در بازنمود آوایی با استفاده از دستور ادراکی توضیح دهیم. ابتدا این روش را برای مشخصه‌های دمش و واک توضیح می‌دهیم.

۲-۱-۱- مدل‌سازی درک مشخصه‌های دمش و واک

در رشته‌ی آوایی [tup] مقدار کشش نوفه‌ی رهش برای قالب‌های زمانی اول، دوم و سوم، [42ms] است. می‌خواهیم درک شنیداری مقدار [42ms] را به‌صورت مقادیر ناپیوسته (+ یا -) مشخصه‌های آوایی [واک] و [دمیده] در چارچوب دستور ادراکی مدل‌سازی کنیم. برای این منظور گام‌های زیر را انجام می‌دهیم:

گام اول: تعیین محدودیت‌های CATEG*

محدودیت‌های CATEG* بر این واقعیت اساسی استوار هستند که به ازای هر مقدار صوت‌شناختی x از مشخصه‌ی آوایی f یک الگوی آوایی y در بازنمود آوایی وجود دارد به‌طوری که درک x به‌صورت y بر اساس محدودیت CATEG*(f: y, x) غیرمجاز است.

حال فرض می‌کنیم برای مقدار [42ms] الگوی آوایی /42ms/ در سطح بازنمود آوایی وجود دارد. محدودیت ناظر بر درک این مقدار به‌صورت الگوی آوایی /42ms/ عبارت است از ([42ms]: /42ms/) CATEG*صادق [۴]. فضای آوایی، واکداری-بی‌واکی و دمش را برای هر طبقه از همخوان‌های انسدادی به سه ناحیه افراز می‌کند؛ به‌طوری که هر ناحیه یکی از گونه‌های [واک] و [دمیده] را نشان می‌دهد. میانگین به‌دست آمده، کشش نوفه‌ی رهش برای انسدادی‌های تیغه‌ای بی‌واک دمیده، بی‌واک نامیده و واکدار به ترتیب 42ms، 18ms و 9ms است. بر این اساس می‌توان فرض کرد که انسدادی‌های تیغه‌ای بی‌واک دمیده، بی‌واک نامیده و واکدار به ترتیب به‌صورت الگوهای ادراکی /47ms/، /18ms/، /9ms/ در حوزه‌ی درک شنونده افراز شده‌اند. این الگوها، الگوهای آوایی مرجع برای درک مقادیر مختلف پارامتر کشش نوفه‌ی رهش به‌حساب می‌آیند. حال فرض کنیم مقدار x [47ms]، [18ms] و [9ms] باشد، در آن صورت محدودیت‌های ناظر بر درک این مقادیر به‌صورت الگوهای آوایی /47ms/، /18ms/، /9ms/ عبارتند از:

$$[18] : [47ms] /47ms/ : \text{CATEG}^* , [18ms] /18ms/ : \text{CATEG}^*$$

به عبارت دیگر، محدودیت‌های مورد نظر به ترتیب درک یک انسدادی تیغه‌ای با کشش نوفه‌ی رهش [42ms] را به صورت بی‌واک دمیده، بی‌واک نادمیده و واکدار غیرمجاز می‌دانند. اما آنچه بسیار اهمیت دارد آرایش این محدودیت‌ها نسبت به یکدیگر است. این آرایش با توجه به فاصله‌ی ادراکی بین [42ms] و الگوهای مرجع موردنظر به صورت زیر است:

$$*Warp([42ms]:/9ms/) \geq *Warp([42ms]:/18ms/) \geq *Warp([42ms]:/47ms/)$$

جدول (۱) نحوه‌ی درک کشش نوفه‌ی رهش [42ms] را به صورت الگوی آوایی مرجع /47ms/ نشان می‌دهد:

[42ms]	Per	*CATEG Others	*Warp (/42ms:/9ms/)	*Warp (/42ms:/18ms/)	*Warp (/42ms:/47ms/)	*C(47ms) *C(18ms) *C(9ms)
/42ms/		*				*
/47ms/					*	*
/18ms/				*		
/9ms/			*			*
nothing	*					

محدودیت Per در این جدول نشان می‌دهد که مقدار x از پارامتر کشش نوفه‌ی رهش باید به صورت یک الگوی آوایی درک شود، در غیر این صورت با حداکثر جریمه‌ی ممکن مواجه خواهد شد.

در این جدول، $(/42ms/ : [42ms])$ *CATEG با توجه به این که کشش نوفه‌ی رهش [42ms]، یک الگوی آوایی مرجع در نظام واجی زبان فارسی نیست، رتبه‌ی بسیار بالایی دارد، بنابراین گزینه‌ی /42ms/ با نقض این محدودیت، پس از nothing بیشترین جریمه را گرفته است. گزینه‌های /47ms/، /18ms/ و /9ms/ با نقض محدودیت‌های *CATEG مربوطه اگرچه جریمه شده‌اند، اما جریمه‌شان به علت پایین بودن رتبه‌ی این محدودیت‌ها که ناشی از مرجع بودن الگوهای آوایی موردنظر است، بسیار کم است. از میان این سه گزینه، /47ms/ به‌عنوان گزینه‌ی بهینه، انتخاب شده است. علت آن، این است که رتبه‌ی $(/47ms/)$ *WARP در این جدول به دلیل فاصله‌ی ادراکی بسیار کم، بین درون‌داد [42ms] و مرجع /47ms/، پایین‌تر از $(/18ms/)$ *WARP و $(/9ms/)$ *WARP است. بنابراین نقض آن برای گزینه‌ی /47ms/ جریمه‌ی کمتری را در مقایسه با /18ms/ و /9ms/

و $(/18ms/ : [18ms])$ *CATEG گفتیم که اگر x کاملاً منطبق بر یک الگوی آوایی مرجع باشد، در آن صورت محدودیت مربوطه، پایین‌ترین رتبه را در میان سایر محدودیت‌ها دارد و نقض آن‌ها جریمه‌ی کمتری را به دنبال خواهد داشت. بر این اساس سه محدودیت اخیر نسبت به محدودیت اول در جدول بهینگی رتبه‌ی پایین‌تری خواهند داشت.

گام دوم: تعیین محدودیت‌های *WARP

در مثال داده شده، مقدار x از پارامتر کشش نوفه‌ی رهش ۴۲ میلی ثانیه است. چون برای مقدار [42ms] یک الگوی آوایی مرجع در حوزه‌ی درک گفتار شنونده وجود ندارد، بنابراین، این مقدار باید به صورت یکی از الگوهای آوایی مرجع /47ms/، /18ms/، /9ms/ درک شود. اگر مبنا را تنها محدودیت‌های *CATEG قرار دهیم در آن صورت هر سه الگوی آوایی مورد نظر به یک اندازه شانس انتخاب شدن به‌عنوان الگوی آوایی مرجع برای مقدار [42ms] را دارند؛ زیرا محدودیت‌های ناظر بر الگوهای آوایی مرجع در انتهای ساخت سلسله‌مراتبی محدودیت‌ها قرار می‌گیرند و آرایش آن‌ها نسبت به یکدیگر آزاد است. این در حالی است که هر مقدار از یک پارامتر صوت‌شناختی، هر بار تنها می‌تواند به صورت یک الگوی آوایی مرجع درک شود و درک آن به‌طور همزمان به صورت سه الگوی آوایی مجاز نیست. بنابراین برای اجتناب از این مشکل اساسی به محدودیت‌های دیگری نیاز داریم که محدودیت‌های *WARP نام دارند.

محدودیت‌های *WARP: محدودیت‌های *WARP مبتنی بر این واقعیت اساسی هستند که اگر مقدار x از یک پارامتر صوت‌شناختی بر یک الگوی آوایی مرجع در حوزه‌ی درک شنونده منطبق نگردد، در آن صورت این مقدار با حداقل فاصله‌ی ادراکی جذب نزدیک‌ترین الگوی آوایی مرجع می‌شود. در مثال داده شده مقدار x، [42ms] است. این مقدار با هیچ یک از الگوهای آوایی مرجع /47ms/، /18ms/، /9ms/ مطابقت ندارد، بنابراین باید جذب یکی از آن‌ها شود. این جذب تابع محدودیت‌های *WARP است. برای مقدار [42ms] با توجه به سه مرجع /47ms/، /18ms/، /9ms/ سه محدودیت *WARP وجود دارد. این محدودیت‌ها عبارتند از:

$$*Warp([42ms]:/47ms/)$$

$$*Warp([42ms]:/18ms/)$$

$$*Warp([42ms]:/9ms/)$$

این محدودیت‌ها به ترتیب نگاشت [42ms] را به الگوهای مرجع /47ms/، /18ms/، /9ms/ غیرمجاز می‌دانند.

۲-۱-۲- مدل‌سازی درک مشخصه محل تولید

بر اساس پارامتر NOT

در این بخش می‌خواهیم توضیح دهیم که چگونه مقادیر پیوسته‌ی پارامتر کشش نوفه‌ی رهش در حوزه درک گفتار به‌صورت مشخصه‌های آوایی [لبی، تیغه‌ای، آبدنه‌ای] درک می‌شوند.

در رشته‌ی آوایی [tup] مقدار کشش نوفه‌ی رهش برای [t] در قالب‌های زمانی اول، دوم و سوم در مجموع 42ms است. می‌خواهیم درک شنیداری مقدار [42ms] را به‌صورت الگوهای آوایی لبی، تیغه‌ای یا بدنه‌ای در چارچوب دستور ادراکی مدل‌سازی می‌کنیم. برای انجام این کار گام‌های زیر را انجام می‌دهیم:

گام اول: تعیین محدودیت‌های CATEG*

برای مقدار [42ms] یک الگوی آوایی /42ms/ در سطح بازنمود آوایی وجود دارد. محدودیتی که درک [42ms] را به‌صورت /42ms/ غیرمجاز می‌داند، محدودیت‌های آوایی [42ms:] /42ms/ CATEG* است. حال فرض می‌کنیم که انسدادی‌های بی‌واک دمیده لبی، تیغه‌ای و بدنه‌ای به ترتیب به‌صورت الگوهای ادراکی /34ms/، /47ms/، /65ms/ در حوزه‌ی درک شنونده افزایش شده‌اند [۴]، این الگوهای آوایی مرجع برای درک مقادیر مختلف پارامتر کشش نوفه‌ی رهش هستند. بنابراین اگر مقدار x منطبق بر این الگوها باشد در آن صورت محدودیت مربوطه پایین‌ترین رتبه را در میان سایر محدودیت‌ها داشته و نقض آن جریمه‌ی کمتری را به دنبال خواهد داشت:

$$*CATEG /34 ms/: [34ms])$$

$$*CATEG /47 ms/: [47ms]) \geq *CATEG /42 ms/: 42ms)$$

$$*CATEG /65 ms/: [65ms])$$

اگر این محدودیت‌ها را با آرایش فوق در جدول بهینگی قرار دهیم و /4ms/، /47ms/، /65ms/ را به‌عنوان گزینه‌های احتمالی برای درون‌داد در نظر بگیریم، در آن صورت هر سه گزینه به یک اندازه شانس انتخاب شدن به‌عنوان برون‌داد بهینه برای درون‌داد موردنظر را دارند؛ زیرا محدودیت‌های ناظر بر آن‌ها نسبت به یکدیگر آرایش آزاد دارند. این در حالی است که دستور باید تنها یکی را به‌عنوان برون‌داد بهینه انتخاب کند.

گام دوم: تعیین محدودیت‌های WARP*

محدودیت‌های WARP* ناظر بر این واقعیت هستند که مقدار x از یک پارامتر صوت‌شناختی باید جذب الگویی شود

که محدودیت‌های رتبه‌ی بالاتر (/18ms/) WARP* و (/9ms/) WARP* را نقض کرده‌اند، به‌همراه دارد.

حال همین مراحل را برای قالب‌های زمانی یازدهم و دوازدهم انجام می‌دهیم. مقدار کشش نوفه‌ی رهش برای این قالب‌های زمانی ۱۳ میلی ثانیه است. برای مقدار [13ms] یک محدودیت CATEG* وجود دارد که درک آن را به‌صورت الگوی ادراکی /13ms/ غیرمجاز می‌داند:

$$*CATEG([42ms]: /42ms/)$$

حال با استناد به نتایج آزمون آماری صادقی [۴] در مورد واگذاری-بی‌واکی همخوان‌های انسدادی، فرض می‌کنیم، نظام آوایی زبان فارسی، مقولات آوایی بی‌واک دمیده، بی‌واک نادمیده و واگذار را در انسدادی‌های لبی، به‌ترتیب به‌صورت الگوهای آوایی مرجع /34ms/، /15ms/، /7ms/ در حوزه‌ی درک گفتار شنونده منطبق می‌سازد. بر این اساس محدودیت‌های ناظر بر درک مقادیر [34ms]، [15ms] و [7ms] به‌صورت الگوهای مرجع /34ms/، /15ms/، /7ms/ را به‌صورت:

$$*CATEG (/34ms/: [34ms]), *CATEG (/15ms/: [15ms])$$

و $*CATEG (/7ms/: [7ms])$ صورت‌بندی می‌کنیم. این سه محدودیت نسبت به محدودیت اول، رتبه‌ی پایین‌تری دارند، زیرا مقادیر درون‌داد را بر الگوهای آوایی مرجع می‌نگارند.

چون برای مقدار [13ms] یک الگوی آوایی مرجع در حوزه‌ی درک گفتار شنونده وجود ندارد، بنابراین، این مقدار باید بر اساس محدودیت‌های WARP* به‌صورت یکی از الگوهای آوایی مرجع /34ms/، /15ms/، /7ms/ درک شود. برای مقدار [13ms] با توجه به سه مرجع /34ms/، /15ms/، /7ms/ سه محدودیت WARP* وجود دارد. این محدودیت‌ها به ترتیب نگاشت کشش نوفه‌ی رهش [13ms] را به الگوهای مرجع /34ms/، /15ms/، /7ms/ غیرمجاز می‌دانند. آرایش این محدودیت‌ها با توجه به فاصله‌ی ادراکی میان مقدار [13ms] و الگوهای مرجع موردنظر به‌صورت زیر است:

$$*Warp([13ms]: /7ms/) \geq *Warp([13ms]: /34ms/) \geq *Warp([13ms]: /15ms/)$$

با توجه به این آرایش، وقتی کشش نوفه‌ی رهش یک همخوان انسدادی لبی 13ms باشد، این همخوان احتمالاً به‌صورت گونه‌ی بی‌واک نادمیده، شنیده می‌شود.

بین واکه‌ها در فضای آواشناختی به دست می‌دهند. در این قسمت با استفاده از پارامترهای F1 و F2 درباره‌ی نحوه‌ی درک مشخصه‌های میزان گرفتگی و محل تولید بستارهای واکه‌ای در فضای آواشناختی واکه‌ها بحث می‌کنیم.

پارامتر F1 ناظر بر متغیر مسیر ارتفاع یا سطح مقطع گرفتگی بدنه‌ی زبان و پارامتر F2 ناظر بر متغیر مسیر محل تولید بدنه‌ی زبان است. با بررسی میانگین تقریبی F1 و F2 برای تعداد زیادی از واکه‌ها در کلمات فارسی، نگارنده چنین فرض می‌کند که پارامتر ارتفاع زبان در حوزه‌ی درک شنونده به سه ناحیه ادراکی /۲۵۰ هرتز/، /۴۵۰ هرتز/ و /۷۵۰ هرتز/ افزاز شده است به طوری که /۲۵۰ هرتز/ الگوی آوایی مرجع واکه‌های افراشته؛ /۴۵۰ هرتز/ الگوی آوایی مرجع واکه‌های نیمه‌افراشته و /۷۵۰ هرتز/ الگوی آوایی مرجع واکه‌های افتاده است. همچنین، فرض می‌کنیم که فضای ادراکی پارامتر F2 به دو ناحیه‌ی ادراکی /۱۸۵۰ هرتز/ و /۸۵۰ هرتز/ افزاز شده است که ناحیه‌ی ادراکی /۱۸۵۰ هرتز/ متناظر با واکه‌های پیشین و ناحیه‌ی ادراکی /۸۵۰ هرتز/ متناظر با واکه‌های پسین است.

برای [u] در [tup] مقدار F1 و F2 در قالب زمانی ششم و هفتم به ترتیب ۲۲۰ و ۷۰۰ هرتز است. ابتدا نحوه‌ی نگاشت مقدار [۲۲۰ هرتز] را به ناحیه‌ی ادراکی /۲۵۰ هرتز/ و درک مشخصه‌ی آوایی [+افراشته] را از قالب زمانی ششم و هفتم توضیح می‌دهیم. برای این منظور گام‌های زیر را انجام می‌دهیم:

گام اول: تعیین محدودیت‌های CATEG*

برای مقدار [۲۲۰ هرتز] یک محدودیت CATEG* وجود دارد که درک این مقدار را به صورت الگوی آوایی متناظر با آن در باز نمود آوایی یعنی /۲۲۰ هرتز/ غیرمجاز می‌داند: ([۲۲۰ هرتز]: [۲۲۰ هرتز]) CATEG*. همچنین، محدودیت‌های ([۲۵۰ هرتز]: [۴۵۰ هرتز]) CATEG* و ([۷۵۰ هرتز]: [۴۵۰ هرتز]) CATEG* به ترتیب ناظر بر درک مقادیر [۲۵۰ هرتز]، [۴۵۰ هرتز] و [۷۵۰ هرتز] به صورت الگوهای آوایی مرجع /۲۵۰ هرتز/، /۴۵۰ هرتز/ و /۷۵۰ هرتز/ هستند. چون مقادیر [۲۵۰ هرتز]، [۴۵۰ هرتز] و [۷۵۰ هرتز] ناظر بر الگوهای آوایی مرجع و مقدار [۲۲۰ هرتز] ناظر بر الگوی آوایی غیرمرجع است، بنابراین سه محدودیت CATEG* اخیر نسبت به محدودیت اول رتبه‌ی پایین‌تری دارند.

که بیشترین شباهت آوایی ممکن را با آن دارد. برای مقدار [42ms] با توجه به الگوهای مرجع /34ms/، /47ms/، /65ms/ سه محدودیت WARP* وجود دارد:

$$*WARP (/42 \text{ ms}/; [34 \text{ ms}])$$

$$*WARP (/42 \text{ ms}/; [47 \text{ ms}])$$

$$*WARP (/42 \text{ ms}/; [65 \text{ ms}])$$

این محدودیت‌ها به ترتیب درک کشش نوفه‌ی رهش [42ms] را به صورت الگوهای ادراکی /34ms/، /47ms/، /65ms/ غیرمجاز می‌دانند. چون شباهت آوایی بین [42ms] و /47ms/ بیشتر از شباهت آوایی بین [42ms] و /34ms/ و مورد اخیر بیشتر از شباهت آوایی بین [42ms] و /65ms/ است، بنابراین آرایش محدودیت‌های مربوطه به صورت زیر خواهد بود:

$$*Warp([42 \text{ ms}]; /47 \text{ ms}/) \geq *Warp([42 \text{ ms}]; /34 \text{ ms}/) \geq *Warp([42 \text{ ms}]; /65 \text{ ms}/)$$

با توجه به این آرایش، کشش نوفه‌ی رهش [42ms] در جدول بهینگی به صورت انسدادی بی‌واک دمیده‌ی تیغه‌ای (/47 ms/) درک خواهد شد.

حال همین مراحل را برای قالب‌های زمانی یازدهم و دوازدهم انجام می‌دهیم. مقدار کشش نوفه‌ی رهش در مجموع این دو قالب زمانی ۱۳ میلی‌ثانیه است. با استناد به نتایج آماری صادقی [۴] فرض می‌کنیم انسدادی‌های بی‌واک نادمیده‌ی لبی، تیغه‌ای و بدنه‌ای به ترتیب به صورت الگوهای ادراکی /15ms/، /18ms/، /24ms/ در حوزه‌ی درک شنونده افزاز شده‌اند. مقدار [13ms] بر هیچ یک از این الگوهای آوایی مرجع منطبق نیست، بنابراین باید بر اساس محدودیت‌های WARP* بر یکی از آن‌ها نگاشته شود. با توجه به فاصله‌ی ادراکی میان [13ms] و الگوی مرجع /15ms/، محدودیت (*Warp ([13ms]: /15ms/)) در مقایسه با سایر محدودیت‌های WARP* رتبه‌ی کمتری داشته و نقض آن جریمه‌ی کمتری خواهد داشت.

$$*Warp([12 \text{ ms}]; /24 \text{ ms}/) \geq *Warp([13 \text{ ms}]; /18 \text{ ms}/) \geq *Warp([13 \text{ ms}]; /15 \text{ ms}/)$$

این آرایش نشان می‌دهد که اگر مقدار کشش نوفه‌ی رهش یک همخوان انسدادی لبی نادمیده‌ی 13ms است، این همخوان احتمالاً به صورت یک همخوان لبی درک می‌شود.

۲-۱-۳- مدل‌سازی درک مشخصه‌ی میزان گرفتگی و محل تولید بستارهای واکه‌ای

فاصله‌ی ادراکی بین، واکه‌ها یک معیار آوایی است که میزان تقابل واجی بین آن‌ها را نشان می‌دهد. فرض بر آن است که پارامترهای F1 و F2 تخمین خوبی از میزان فاصله‌ی ادراکی

گام دوم: تعیین محدودیت‌های *WARP

برای مقدار [۲۲۰ هرتز] با توجه به الگوهای آوایی مرجع ۲۵۰/ هرتز، ۴۵۰/ هرتز و ۷۵۰/ هرتز سه محدودیت *WARP وجود دارد. این محدودیت‌ها به ترتیب درک سازه‌ی اول ۲۲۰ هرتزی را به صورت یک واکه‌ی افراشته (۲۵۰/ هرتز)، نیمه‌افراشته (۴۵۰/ هرتز) و افتاده (۷۵۰/ هرتز) غیرمجاز می‌دانند. این محدودیت‌ها با توجه به میزان فاصله‌ی ادراکی بین مقدار درون داد و الگوهای آوایی مرجع مرتب می‌شوند. به این ترتیب چون فاصله‌ی مقدار ۲۲۰ هرتز به الگوی مرجع ۲۵۰/ هرتز نزدیک، به ۷۵۰/ هرتز دور و به ۴۵۰/ هرتز در یک حد میانی است، بنابراین آرایش محدودیت‌های مربوطه به صورت زیر خواهد بود:

$$*Warp([220Hz]: /250 Hz /) \geq *Warp([220Hz]: /450Hz /) \geq *Warp([220Hz]: /750 Hz /)$$

با توجه به این آرایش و همچنین آرایش محدودیت‌های *CATEG، F1 [۲۲۰ هرتز] به صورت یک واکه‌ی افراشته درک خواهد شد.

حال می‌خواهیم نحوه‌ی درک مشخصه‌ی محل تولید ستاره‌های واکه‌ای را در قالب دستور ادراکی مدل‌سازی کنیم.

گام اول: تعیین محدودیت‌های *CATEG

گفتیم که فضای ادراکی پارامتر F2 به دو ناحیه‌ی ادراکی ۱۸۵۰/ هرتز برای واکه‌های پیشین و ۸۵۰/ هرتز برای واکه‌های پسین، افراز شده است. برای [u] در [tup] مقدار F2 در قالب زمانی ششم ۷۰۰ هرتز است. به‌ازای مقدار [۷۰۰ هرتز] در دستور ادراکی یک محدودیت *CATEG وجود دارد که درک این مقدار را به صورت الگوی آوایی ۷۰۰/ هرتز در باز نمود آوایی غیرمجاز می‌داند. علاوه بر این، به‌ازای هر یک از مقادیر [۱۸۵۰ هرتز]، [۸۵۰ هرتز] یک محدودیت *CATEG وجود دارد که درک مقادیر مربوطه را به صورت الگوهای آوایی مرجع ۱۸۵۰/ هرتز و ۸۵۰/ هرتز غیرمجاز می‌داند. چون مقادیر [۱۸۵۰ هرتز]، [۸۵۰ هرتز] ناظر بر الگوهای آوایی مرجع و مقدار [۷۰۰ هرتز] ناظر بر الگوی آوایی غیرمرجع است بنابراین رتبه‌ی دو محدودیت اخیر نسبت به محدودیت اول در جدول بهینگی پایین‌تر خواهد بود.

با توجه به این که برای مقدار [۷۰۰ هرتز] به‌عنوان درون داد صوت‌شناختی یک ناحیه‌ی ادراکی وجود ندارد، بنابراین شنونده باید این مقدار را به ناحیه‌ی ادراکی که

بیشترین شباهت آوایی را با آن دارد نگاشته و مشخصه محل تولید واکه را درک کند. این فرآیند توسط محدودیت‌های *WARP انجام می‌شود.

گام دوم: تعیین محدودیت‌های *WARP

برای مقدار [۷۰۰ هرتز] با توجه به الگوهای آوایی مرجع ۱۸۵۰/ هرتز و ۸۵۰/ هرتز دو محدودیت *WARP وجود دارد که یکی درک F2 ۷۰۰ هرتزی را به صورت یک واکه‌ی پیشین (۱۸۵۰/ هرتز) و دیگری درک همین سازه را به صورت یک واکه‌ی پسین (۸۵۰/ هرتز) غیرمجاز می‌دانند. آنچه اهمیت دارد آرایش این دو محدودیت نسبت به یکدیگر است؛ چون شباهت آوایی F2 ۷۰۰ هرتزی به الگوی مرجع ۸۵۰/ هرتز بیشتر از الگوی مرجع ۱۸۵۰/ هرتز است بنابراین رتبه‌ی محدودیت *WARP دوم، نسبت به محدودیت *WARP اول، پایین‌تر است؛ که این نشان می‌دهد هر واکه‌ای با F2 ۷۰۰ هرتز به صورت یک واکه‌ی پسین درک می‌شود.

$$*Warp([700Hz]: /1850Hz /) \geq *Warp([700Hz]: /850Hz /)$$

۲-۱-۴- مدل‌سازی درک مشخصه‌ی ریشه

گره‌ی ریشه در هندسه مشخصه‌های واجی ناظر بر طبقات واجی زبان است. طبقات عمده‌ی واجی عبارتند از: انسدادی‌ها، سایشی‌ها و رساها. این طبقات واجی بر اساس وضعیت‌های سه‌گانه‌ی آیرودینامیکی در سطوح حفره‌های بازخوانی مرکب تعیین می‌شوند. بر این اساس، طبقه‌ی انسدادی ناظر بر وضعیت انسداد جهاز صوتی، طبقه‌ی سایشی ناظر بر وضعیت نوفه و طبقه‌ی رسا ناظر بر وضعیت تشدید جهاز صوتی است. طبقه‌ی انسدادی را می‌توان با مشخصه‌های آوایی [+همخوانی، -پیوسته]؛ طبقه سایشی با مشخصه‌های [+همخوانی، +پیوسته] و رسا را با مشخصه [+رسا] نشان داد.

بی‌جن‌خان [۳] برای مقایسه‌ی کمی طبقات واجی زبان فارسی، فضای آوایی را بر حسب دو پارامتر نرخ تقاطع صفر و انرژی به‌هنگار به‌گونه‌ای افراز می‌کند که مرز میان طبقات عمده‌ی واجی در سطح جهاز صوتی تا حد قابل‌قبولی از یکدیگر مشخص باشند. نرخ تقاطع صفر پارامتر آواشناختی برای وضعیت تشدید جهاز صوتی و انرژی به‌هنگار پارامتر آواشناختی برای وضعیت نوفه‌ی جهاز صوتی است. وی [۳] پراکندگی واج‌های زبان فارسی را بر حسب میانگین میانگین‌های نرخ تقاطع صفر و میانگین میانگین‌های انرژی

جدول (۴)

[2815]	Per	*CATEG Others	*Warp (12209); (3689)	*Warp (12209);(1127)	*Warp (12209);(2621)	*C(2621) *C(1127) *C(3689)
/1127/			*			*
/1127/					*	*
/2621/				*		*
/2815/		*				
nothing	*					

جدول (۵)

[852]	Per	*CATEG Others	*Warp (12209); (3689)	*Warp (12209);(2621)	*Warp (12209);(1127)	*C(2621) *C(1127) *C(3689)
/3689/			*			*
/1127/					*	*
/2621/				*		*
/852/		*				
nothing	*					

گزینه‌های /۲۶۲۱/، /۳۶۸۹/ و /۱۱۲۷/ با نقض محدودیت‌های CATEG* مربوطه اگرچه جریمه شده‌اند، اما جریمه‌ی‌شان به علت پایین بودن رتبه‌ی این محدودیت‌ها بسیار کم است. از میان این سه گزینه، /۲۶۲۱/ به عنوان گزینه‌ی بهینه انتخاب شده است. چون شباهت آوایی نرخ تقاطع صفر [۲۲۰۹] به الگوی مرجع /۲۶۲۱/ بیشتر از الگوهای مرجع /۱۱۲۷/ و /۳۶۸۹/ است، بنابراین رتبه‌ی /۲۶۲۱/ (۲۲۰۹): WARP* در این جدول پایین‌تر از محدودیت‌های WARP* دیگر است. چون رتبه‌ی این محدودیت پایین است، نقض آن جریمه‌ی کمی را برای /۲۶۲۱/ به همراه داشته است. این در حالی است که گزینه‌های /۳۶۸۹/ و /۱۱۲۷/ به ترتیب با نقض محدودیت‌های رتبه‌ی بالاتر (/۳۶۸۹/ : [۲۲۰۹] WARP* و /۱۱۲۷/ : [۲۲۰۹] WARP* جریمه‌ی بیشتری گرفته‌اند. در جدول ۴ نیز /۲۶۲۱/ با رعایت همین موارد به عنوان بهترین گزینه برای درک نرخ تقاطع صفر ۲۸۱۵ معرفی شده است.

در جدول (۵) گزینه‌ی /۱۱۲۷/ با توجه به فاصله‌ی ادراکی کم آن با مقدار درون‌داد ۸۵۲ و نحوه‌ی آرایش محدودیت‌های CATEG* و WARP* مربوطه به عنوان گزینه‌ی بهینه برای جذب نرخ تقاطع صفر ۸۵۲ انتخاب شده است.

حال همین مراحل را برای پارامتر انرژی به‌هنگار انجام می‌دهیم. برای [p, [u], t] مقدار انرژی به‌هنگار به ترتیب ۰/۴۳dB، ۰/۹۱dB و ۰/۳۸dB است. چنانچه این مقادیر را به ترتیب به عنوان درون‌داد دستور ادراکی در جداول بهینگی

به‌هنگار برای چهل گویشور فارسی‌زبان به دست می‌آورد و به این ترتیب فضای واجی را به سه ناحیه افزایش می‌کند، به طوری که هر ناحیه یک طبقه از طبقات عمده‌ی واجی را در سطح جهاز صوتی نشان می‌دهد. میانگین این دو پارامتر برای طبقات واجی انسدادی، سایشی و رسا به صورت زیر است:

جدول (۲): میانگین نرخ تقاطع صفر و انرژی به‌هنگار برای طبقات عمده واجی از دادگان فارس دات (بی جن خان، [۴])

طبقات واجی	نرخ تقاطع صفر	انرژی به‌هنگار
انسدادی	۲۶۲۱	۳۴۰ dB
سایشی	۳۶۸۹	۲۹۰ dB
رسا	۱۱۲۷	۱۰۱۵ dB

این میانگین‌ها را به عنوان الگوی آوایی مرجع برای طبقات واجی موردنظر در نظر می‌گیریم. حال مقادیر این پارامترها را برای قالب‌های زمانی مختلف تعیین می‌کنیم. برای [t] در [tup] مقدار نرخ تقاطع صفر و انرژی به‌هنگار در قالب‌های زمانی اول تا سوم به ترتیب ۲۲۰۹ و ۰/۴۳dB است. برای [u] در همین رشته‌ی آوایی، مقدار نرخ تقاطع صفر و انرژی به‌هنگار در قالب زمانی ششم و هفتم ۸۵۲ و ۰/۹۱dB است. برای [p] نیز در رشته‌ی آوایی موردنظر مقدار نرخ تقاطع صفر و انرژی به‌هنگار در قالب‌های زمانی یازدهم و دوازدهم ۲۸۱۵ و ۰/۳۸dB است. می‌خواهیم امکان تعلق قالب‌های زمانی مختلف را به وضعیت‌های سه‌گانه‌ی انسداد، نوبه و تشدید بر اساس دستور ادراکی مدل‌سازی کنیم.

ابتدا پارامتر نرخ تقاطع صفر را در نظر می‌گیریم. برای [p, [u], t] مقدار نرخ تقاطع صفر به ترتیب ۲۲۰۹، ۸۵۲ و ۲۸۱۵ است. جداول ۳ و ۴ نحوه‌ی درک مقادیر [۲۲۰۹] و [۲۸۱۵] را به صورت طبقه‌ی انسدادی و جدول ۵ نحوه‌ی درک مقدار [۸۵۲] را به صورت طبقه‌ی رسا نشان می‌دهد. در جدول ۳ گزینه‌ی /۲۲۰۹/ با نقض محدودیت (۲۲۰۹): CATEG* که به علت مرجع نبودن نرخ تقاطع صفر ۲۲۰۹ رتبه‌ی بسیار بالایی دارد، بیشترین جریمه را گرفته است.

جدول (۳)

[2209]	Per	*CATEG Others	*Warp (12209); (3689)	*Warp (12209);(1127)	*Warp (12209);(2621)	*C(2621) *C(1127) *C(3689)
/3689/			*			*
/1127/				*		*
/2621/					*	*
/2209/		*				
nothing	*					

جداگانه قرار دهیم، انرژی ۰/۴۳dB و ۰/۳۸ dB بر مرجع ۰/۳۴/ یا طبقه‌ی انسدادی و انرژی ۰/۹۱ dB بر مرجع ۰/۱۱۵/ یا طبقه‌ی رسا منطبق می‌گردند.

مرحله‌ی سوم: در مرحله‌ی سوم، نحوه‌ی نگاشت بازنمود آوایی بر بازنمود واجی (واژگانی) مدل‌سازی می‌شود. این مدل‌سازی توسط دستور بازشناسی صورت می‌گیرد. دستور بازشناسی نیز همچون دستور ادراکی شامل دو نوع محدودیت *CATEG* و *WARP* است. تفاوت این محدودیت‌ها با محدودیت‌های دستور ادراکی این است که محدودیت‌های بازشناسی از نوع واجی هستند در حالی که محدودیت‌های ادراکی، صوت‌شناختی هستند. محدودیت‌های *CATEG* بازشناسی به‌طور کلی بازشناسی یک مشخصه‌ی آوایی را در بازنمود آوایی به‌صورت مشخصه‌ی واجی متناظر با آن در بازنمود واژگانی غیرمجاز می‌دانند و محدودیت‌های *WARP* وجود هرگونه فاصله‌ی ادراکی را بین الگوهای واژگانی مرجع یک مشخصه‌ی آوایی و الگوهای آوایی آن مشخصه غیرمجاز می‌دانند. درک الگوهای آوایی به‌صورت الگوهای واژگانی در دستور بازشناسی تابعی از میزان نقض محدودیت‌های *CATEG* و *WARP* با توجه به نحوه‌ی آرایش این محدودیت‌ها نسبت به یکدیگر است.

در بخش قبل از پارامتر NOT برای استخراج مشخصه‌های آوایی واکداری و محل تولید همخوان انسدادی؛ از پارامترهای F1 و F2 برای استخراج مشخصه‌های آوایی ارتفاع زبان و محل تولید بستار واکه‌ای و از پارامترهای نرخ تقاطع صفر و انرژی به‌هنگار برای استخراج مشخصه‌ی ریشه استفاده کردیم. در این مرحله می‌خواهیم روش بازشناسی مشخصه‌های واژگانی را از مشخصه‌های آوایی با استفاده از دستور بازشناسی توضیح دهیم. ابتدا این روش را برای مشخصه‌های آوایی دمش و واک توضیح می‌دهیم.

۲-۱-۵- مدل‌سازی بازشناسی الگوی واک

در مرحله‌ی قبل نشان دادیم که چگونه همخوان انسدادی تیغه‌ای با کشش نوفه‌ی رهش ۴۲ میلی ثانیه به‌صورت یک همخوان بی‌واک دمیده و همخوان انسدادی لبی با کشش نوفه‌ی رهش ۱۳ میلی ثانیه به‌صورت یک همخوان بی‌واک نادمیده درک می‌شود. حال سؤالی که مطرح می‌شود این که چگونه مشخصه‌های [-واک] و [+گسترده چاکنای] برای انسدادی تیغه‌ای [t^h] در نهایت به مشخصه [-واک] کاهش

یافته و به‌صورت این الگوی واژگانی بازشناسی می‌شوند. همچنین چگونگی مشخصه مشخصه‌های [-واک] و [-گسترده چاکنای] برای انسدادی لبی [p] در نهایت به‌صورت الگوی واژگانی [-واک] بازشناسی می‌شوند. ابتدا روش بازشناسی الگوی واژگانی [-واک] را از مشخصه‌های [-واک] و [+گسترده چاکنای] برای انسدادی تیغه‌ای [t^h] شرح می‌دهیم. برای این منظور گام‌های زیر را انجام می‌دهیم:

گام اول: تعیین محدودیت‌های *CATEG*

محدودیت‌های *CATEG* در دستور بازشناسی مبتنی بر این واقعیت هستند که به‌ازای هر الگوی آوایی مرجع در بازنمود آوایی، یک الگوی واجی در بازنمود واژگانی وجود دارد که بازشناسی آن الگوی آوایی بر اساس الگوی واجی موردنظر غیرمجاز است:

$CATEG(x) = x f(:)$ الگوی آوایی در بازنمود آوایی، $f =$ الگوی واجی در بازنمود واژگانی. برای الگوی [-واک]، [+گسترده چاکنای] یک الگوی واجی در سطح بازنمود واژگانی وجود دارد که عبارت است از /-واک/، /+گسترده چاکنای/ محدودیت ناظر بر بازشناسی الگوی آوایی [-واک]، [+گسترده چاکنای] به‌صورت الگوی واجی /-واک/، /+گسترده چاکنای/ عبارت است از:

([-واک]، [+گسترده چاکنای]؛ /-واک/، /+گسترده چاکنای/) *CATEG

در واج‌شناسی جهانی، دمیدگی یک ویژگی نشان‌دار برای انفجاری‌های بی‌واک است. یعنی انفجاری‌های بی‌واک معمولاً نادمیده هستند. به عبارت دیگر، به‌ندرت می‌توان زبانی را یافت که در آن دمیدگی برای انفجاری‌های بی‌واک تمایزدهنده باشد، ولی نادمیدگی نقش تقابلی نداشته باشد کنستویچ، [۸]. این در حالی است که بیشتر زبان‌ها از امکانات واجی مشخصه‌ی واکداری-بی‌واکی برای ایجاد تقابل واژگانی استفاده می‌کنند. زبان فارسی نیز از این الگو پیروی می‌کند. یعنی در نظام آوایی زبان فارسی، برای الگوهای آوایی بی‌واک نادمیده، بی‌واک دمیده و واکدار با توجه به نقش تمایزدهندگی مشخصه [واک] و عدم تمایزدهندگی [گسترده چاکنای] دو الگوی واجی مرجع یا الگوی واژگانی در سطح بازنمود واجی وجود دارد که عبارتند از: /-واک/ و /+واک/، محدودیت‌های ناظر بر بازشناسی الگوی آوایی [-واک] به‌صورت الگوی واجی /-واک/ و [+واک] واک [به‌صورت الگوی واجی /+واک/ عبارتند از: (-واک/، /-واک/

[- واک]، + [گسترده چاکنای] به‌طور کامل با / - واک / و + واک / تناظر آوایی ندارد؛ بنابراین باید جذب یکی از آن‌ها شود. این جذب تابع محدودیت‌های *WARP است. برای [- واک]، + [گسترده چاکنای] با توجه به الگوهای مرجع / - واک / و + واک / دو محدودیت *WARP وجود دارد که چون تناظر آوایی بین [- واک]، + [گسترده چاکنای] و / - واک / بیشتر از تناظر آوایی بین [- واک]، + [گسترده چاکنای] و + واک / است، بنابراین رتبه‌ی محدودیت *WARP دوم نسبت به محدودیت اول پایین‌تر بوده و نقض آن جریمه‌ی کمتری را به‌همراه خواهد داشت.

و [- واک]: + [گسترده چاکنای] *Warp

/ + واک /: + [گسترده چاکنای] *Warp

جدول (۶) نحوه بازشناسی الگوی آوایی [- واک]، + [گسترده چاکنای] را به‌صورت الگوی واجی / + واک / نشان می‌دهد.

جدول (۶)

[+گسترده] [واک]	Per	*CATEG Others	*Warp (+ واک)	*Warp (- واک)	*C (واک) *C (واک)
+گسترده/ /واک/		*			
/ - واک /				*	*
/ + واک /			*		*
nothing	*				

چنان‌چه همین مراحل را برای الگوی آوایی [- واک]، [- گسترده چاکنای] انجام دهیم، بر اساس معیار فاصله‌ی آوایی، / - واک / در سطح بازنمود واژگانی به‌عنوان مرجع واجی برگزیده می‌شود.

۲-۱-۶- مدل‌سازی بازشناسی مشخصه‌ی

محل تولید

در بخش قبل نشان دادیم که وقتی مقدار کشش نوفه‌ی رهش یک همخوان انسدادی بی‌واک دمیده، ۴۲ میلی‌ثانیه باشد، این همخوان به احتمال فراوان به‌صورت یک همخوان تیغه‌ای درک می‌شود. همچنین دیدیم که وقتی کشش نوفه‌ی رهش یک همخوان انسدادی بی‌واک نامیده 13ms باشد، این همخوان احتمالاً به‌صورت یک همخوان لبی درک می‌شود. حال می‌خواهیم توضیح دهیم که چگونه مشخصه‌های [تیغه‌ای] و [لبی] در بازنمود ادراکی واج‌گونه‌های [t^h] و [p] در نهایت به‌صورت مشخصه‌های [تیغه‌ای] و [لبی] در بازنمود واژگانی واج‌های /t/ و /p/ بازشناسی می‌شوند. جدول زیر نحوه‌ی انطباق بین الگوی

[واک] / *CATEG و [+ واک /: + واک /] *CATEG. اگر فرض کنیم محدودیت‌هایی که الگوهای آوایی مرجع را به الگوهای واجی مرجع می‌نگارند در انتهای ساخت سلسله مراتبی و محدودیت‌هایی که این الگوها را به الگوهای واجی غیرمرجع می‌نگارند در ابتدای ساخت سلسله مراتبی آرایش محدودیت‌ها قرار می‌گیرند، در آن صورت آرایش محدودیت‌های مورد نظر به‌صورت زیر خواهد بود:

$$[- واک /: - واک /] *CATEG \geq [+ واک /: + واک /] *CATEG$$

حال اگر درون داد دستور بازشناسی [- واک]، + [گسترده چاکنای] باشد، در آن صورت طبق محدودیت‌های *CATEG هر دو الگوی واجی / - واک / و + واک / می‌توانند مرجع احتمالی برای بازشناسی الگوی آوایی مورد نظر باشند زیرا آرایش آن‌ها نسبت به یکدیگر آزاد است. حال آن‌که دستور بازشناسی باید تنها یکی را به‌عنوان گزینه‌ی بهینه انتخاب کند. برای رفع این مشکل، دستور بازشناسی از محدودیت‌های *WARP استفاده می‌کند.

گام دوم: تعیین محدودیت‌های *WARP

محدودیت‌های *WARP در دستور بازشناسی بر این واقعیت استوار هستند که در صورت عدم تطابق الگوهای آوایی با الگوهای واجی مرجع، این الگوها جذب نزدیکترین الگوی واجی می‌شوند. این محدودیت‌ها به‌صورت زیر صورت‌بندی می‌شوند:

$$*Warp(f; d) \Leftrightarrow \exists x_i \ \varepsilon f_{perc} \ \wedge \ \exists y_i \ \varepsilon f_{Rec} \ \rightarrow \ |x_i - y_i| < d$$

(ادراکی perc واجی; Rec)

بر اساس این محدودیت، الگوی واجی y از ویژگی f نباید به اندازه‌ی هر مقداری از d (خطا در پردازش بازشناسی) متفاوت از الگوی آوایی مرجع x در بازنمود آوایی آن ویژگی باشد. به‌عبارت دیگر، بین الگوهای آوایی و الگوهای واجی باید یک تناظر کامل وجود داشته باشد. هرقدر عدم تناظر آوایی بین الگوهای آوایی و الگوهای واجی مرجع بیشتر باشد، محدودیت *WARP ناظر بر آن‌ها رتبه‌ی بالاتری داشته و نقض آن جریمه‌ی سنگین‌تری را به‌همراه خواهد داشت.

برای ترکیب‌های مختلف مشخصه‌های [واک] و [چاکنای گسترده] در سطح بازنمود آوایی، دو الگوی واجی مرجع در سطح بازنمود واژگانی وجود دارد: / - واک / و + واک /، بنابراین برای الگوی [- واک]، + [گسترده چاکنای] دو مرجع / - واک / و + واک / در ناحیه‌ی واجی افزاز شده است. چون

آوایی [تیغه‌ای] با الگوی واجی /تیغه‌ای/ را در بازنمود واژگانی نشان می‌دهد. در این جدول، به ازای هر یک از مشخصه‌های محل تولید [لبی]، [تیغه‌ای]، [بدنه‌ای] و [املازی] یک محدودیت *CATEG وجود دارد. هر چهار محدودیت، الگوهای آوایی درون‌داد را به الگوهای واجی مرجع (فضای واجی محل تولید همخوان‌های انسدادی به چهار ناحیه لبی، /تیغه‌ای/، /بدنه‌ای/ و /املازی/ افزاز می‌شود) نگاشته‌اند. اما با توجه به این که الگوی درون‌داد، [تیغه‌ای] است بنابراین محدودیت (تیغه‌ای) : [تیغه‌ای/] *CATEG نسبت به سه محدودیت دیگر رتبه‌ی پایین‌تری دارد بوسما، [۲]. همچنین با توجه به این که متناظر با الگوی [تیغه‌ای] یک الگوی واجی مرجع، یعنی /تیغه‌ای/ در فضای واجی افزاز شده است؛ بنابراین برای بازشناسی این الگو نیازی به استفاده از محدودیت‌های *WARP نیست.

جدول (۸)

[+افراشته]	*C (افراشته-افراشته) *C (افراشته)	*CATEG (افراشته)
/+افراشته/	*	
/-افراشته/	*	
/+افراشته=>/		*

جدول (۹)

[+پسین]	*CATEG (-پسین)	*CATEG (+پسین)
/+پسین/		*
/-پسین/	*	

۲-۱-۸- مدل‌سازی بازشناسی مشخصه‌ی

ریشه

در بخش قبل درباره‌ی نحوه‌ی درک مقادیر پارامترهای نرخ تقاطع صفر و انرژی به‌هنگار به‌صورت طبقات آوایی انسدادی [+همخوانی، - پیوسته] و رسا [+رسا] بحث کردیم. حال می‌خواهیم نحوه‌ی بازشناسی این مشخصه‌ها را به‌صورت الگوهای واژگانی /همخوانی، - پیوسته/ و /+رسا/ توضیح دهیم. طبقات آوایی [+همخوانی، - پیوسته] و [+رسا] در جداول ۱۰ و ۱۱ به‌ترتیب به‌عنوان درون‌داد دستور بازشناسی قرار گرفته‌اند. در هر جدول محدودیت‌های *CATEG و *WARP مربوطه مشخص شده‌اند. جدول ۱۰ نحوه‌ی نحوه‌ی انطباق [+همخوانی، - پیوسته] و جدول ۱۱ نحوه‌ی انطباق [+رسا] با طبقه‌ی واجی رسا را نشان می‌دهد:

جدول (۷)

[تیغه‌ای]	*C (لبی) *C (بدنه‌ای) *C (املازی)	*CATEG (تیغه‌ای)
/بدنه‌ای/	*	
/لبی/	*	
/تیغه‌ای=>/		*
/املازی/	*	

حال اگر درون‌داد، [لبی] باشد، با توجه به پایین‌تر بودن رتبه‌ی محدودیت (لبی) : [لبی/] *CATEG نسبت به سایر محدودیت‌های *CATEG، مشخصه لبی/ به‌عنوان مشخصه‌ی واجی بهینه انتخاب خواهد شد.

۲-۱-۷- مدل‌سازی بازشناسی مشخصه‌ی

ارتفاع زبان و محل تولید بستارهای

واکه‌ای

در بخش قبل نشان دادیم اگر واکه‌ای تولید شود که F1 آن ۲۲۰ هرتز و F2 آن ۷۰۰ هرتز باشد، شنونده آن واکه را به‌صورت یک واکه [+افراشته] و [+پسین] درک می‌کند. حال برای بازشناسی این مشخصه‌ها به‌صورت مشخصه‌های واژگانی /+افراشته/ و /+پسین/ از دستور بازشناسی استفاده می‌کنیم.

جداول زیر نحوه‌ی آرایش محدودیت‌های *CATEG و چگونگی بازشناسی مشخصه‌های واژگانی /+افراشته/ و /+پسین/ در چارچوب دستور بازشناسی نشان می‌دهد:

جدول (۱۰)

+همخوانی - پیوسته	*CATEG +رسا	*CATEG +همخوانی - پیوسته	*CATEG +همخوانی - پیوسته
/+رسا/	*		
+همخوانی - پیوسته		*	
+همخوانی - پیوسته			*

قالب زمانی دسترسی دارد. تعداد قالب‌های زمانی که هر واحد کلمه‌یاب می‌تواند به آن‌ها دسترسی داشته باشد، با توجه به طول کلمه‌ی متغیر است.

برای قالب زمانی اول تا دوازدهم واحد کلمه‌یاب /tup/ را فعال می‌کنیم. چنان‌که ملاحظه شد، برای قالب‌های زمانی اول تا پنجم واج /t/؛ ششم و هفتم واج /u/ و هشتم تا دوازدهم واج /p/ به‌دست آمد. حال واحد کلمه‌یاب /tup/ با ترکیب این واج‌ها درخت واجی کلمه‌ی "توپ" را استخراج می‌کند.

۳- نتیجه‌گیری

در این مقاله مدلی برای درک و بازشناسی کلمات زبان فارسی ارائه شد. بر اساس این مدل، در بازنمود آوایی هر واحد زبانی، مجموعه‌ای از الگوهای صوت‌شناختی، به‌عنوان الگوهای آوایی مرجع ذخیره می‌شوند و درک آن واحد از طریق انطباق الگوهای صوت‌شناختی درون‌داد با الگوهای آوایی مرجع در بازنمایی آوایی صورت می‌گیرد. مدل پیشنهادی ترکیبی از مدل المان و مک کلند [۱] تحت عنوان "اثر" و مدل بورسما [۲] تحت عنوان "دستور ادراکی و دستور بازشناسی" است. این مدل شامل سه لایه‌ی واحدیاب است: (۱) لایه‌ی مشخصه‌یاب، (۲) لایه‌ی واج‌یاب و (۳) لایه‌ی کلمه‌یاب. درون‌داد مدل، امواج صوتی است. امواج صوتی، ابتدا وارد لایه‌ی مشخصه‌یاب می‌شود. لایه‌ی مشخصه‌یاب با استفاده از دو دستور ادراکی و بازشناسی، امواج صوتی را به مشخصه‌های واجی تبدیل می‌کند. سپس، مشخصه‌ها وارد لایه‌ی واج‌یاب می‌شوند و این لایه، واحد واجی را از طریق ترکیب مشخصه‌های واجی در قالب‌های زمانی مربوطه استخراج می‌کند. در مرحله پایانی لایه‌ی کلمه‌یاب با ترکیب واج‌ها در چندین قالب زمانی، کلمه را استخراج می‌کند.

دستور ادراکی و بازشناسی پیشنهادی نگارنده، در لایه‌ی مشخصه‌یاب مبتنی بر واج‌شناسی بهینگی است. مدل واج‌شناسی بهینگی، واج‌شناسی کلمات را درک انسانی مدل‌سازی می‌کند و می‌توان از آن برای بازشناسی ماشینی کلمات مجزا از هم یا پیوسته استفاده کرد.

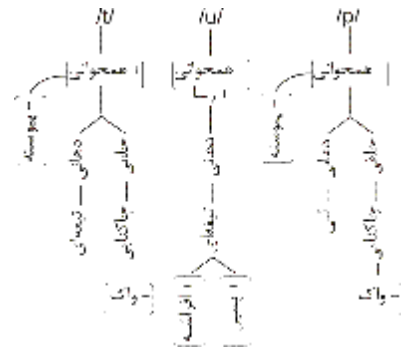
۴- منابع

- [1]- Elman and McClelland (1988). **The Trace Model**. Psycholinguistic Approaches To Spoken Word Recognition. Tutorial 2B
- [2]- Boersma, Paul (1998). Functional phonology: Formalizing the interactions between articulatory and perceptual devices. ph.D dissertation, University of Amsterdam.

جدول (۱۱)

/+رسا/	*CATEG +همخوانی -پیوسته	*CATEG +همخوانی +پیوسته	*CATEG +رسا
/+رسا/	*		*
+همخوانی +پیوسته		*	
+همخوانی -پیوسته			

حال به مثال داده شده بازمی‌گردیم. برای سه قالب زمانی اول و همچنین قالب‌های زمانی چهارم و پنجم که ناظر بر همپوشی بستار همخوانی /t/ با بستار واکه‌ای /u/ است، واحد واج‌یاب /t/ را فعال می‌کنیم. برای این قالب‌های زمانی مشخصه‌های واجی -/واک/، اتیغه‌ای، /+همخوانی/، - پیوسته/ به‌دست آمد. واحد واج‌یاب /t/ با ترکیب این مشخصه‌ها درخت واجی ۳ را برای واج /t/ استخراج می‌کند:



شکل (۳) شکل (۴) شکل (۵)

برای قالب‌های زمانی ششم و هفتم واحد واج‌یاب /u/ را فعال می‌کنیم. برای این قالب‌های زمانی مشخصه‌های واجی /+افراشته/، /+پسین/، /+رسا/ به‌دست آمد. واحد واج‌یاب /u/ با ترکیب این مشخصه‌ها درخت واجی ۴ را برای واج /u/ استخراج می‌کند.

برای قالب‌های زمانی دهم، یازدهم و دوازدهم و همچنین قالب‌های زمانی هشتم و نهم که ناظر بر همپوشی آغاز /p/ با پایانه /u/ است، واحد واج‌یاب /p/ را فعال می‌کنیم. برای این قالب‌های زمانی مشخصه‌های واجی -/واک/، /البی/، /+همخوانی/، -پیوسته/ به‌دست آمد. واحد واج‌یاب /p/ با ترکیب این مشخصه‌ها درخت واجی ۵ را برای واج /p/ استخراج می‌کند.

۲-۳- لایه‌ی کلمه‌یاب: در آخرین لایه، کلمه‌ها از رشته علائم صوت‌شناختی گفتار استخراج می‌شوند. برای این منظور برای هر سه قالب زمانی مجموعه‌ای از واحدهای کلمه‌یاب فعال می‌شود. با توجه به این که واحدهای کلمه‌یاب با یکدیگر همپوشی دارند، هر واحد به بیش از سه

[۳]- بی‌جن‌خان، محمود (۱۳۷۴). بازنمایی واجی و آوایی زبان فارسی و کاربرد آن در بازشناسی خودکار گفتار، پایان‌نامه دکتری، دانشگاه تهران، دانشکده ادبیات و علوم انسانی، گروه زبان‌شناسی.

[۴]- صادقی، وحید (۱۳۸۵). بازشناسی واجی کلمات فارسی: رویکردی مبتنی بر نظریه بهینگی، پایان‌نامه دکتری، دانشگاه تهران، دانشکده ادبیات و علوم انسانی، گروه زبان‌شناسی.

- [5]- Ladefoged, P. and Maddieson, I. (1996). **The Sounds of the World's languages**. Cambridge M, blackwell **Modelling Persian Spoken Word Recognition In this paper, A model**.
- [6]- Ladefoged, P. and Cho, T. (1999). **Variation and Universals in Vot: Evidence from 18 languages**. Journal of phonetics, 27, 207-229.
- [7]- Cho, T., S.- A. Jun, S. Jung and P. Ladefoged (2000). **An acoustic and Aerodynamic study of consonants in Cheju**. Speech Sciences, 7, 109-142
- [8]- Kenstowicz, M. (1994). **Phonology in Generative Grammar**. Cambridge, MA, and Oxford: Blackwell Publishers.



وحید صادقی تحصیلات خود را در مقطع کارشناسی رشته‌ی آموزش زبان انگلیسی در سال ۱۳۷۶ به پایان رساند. وی در سال ۱۳۷۹ در رشته‌ی زبان‌شناسی در مقطع کارشناسی ارشد از دانشگاه تهران فارغ التحصیل شد و در سال ۱۳۸۵ نیز در همان رشته مدرک دکترای خود را اخذ نمود. حوزه‌ی پژوهشی مورد علاقه‌ی وی پردازش گفتار و درک و بازشناسی گفتار است.

نشانی رایانامک (پست الکترونیکی) ایشان عبارت است از:
va_sadeghi2000@yahoo.com: