



وزارت علوم، تحقیقات و فناوری

پژوهشگاه علوم و فناوری اطلاعات ایران

گزارش طرح

ساخت پیکره گفتاری هجای CV زبان فارسی مبتنی بر الگوی نوایی گفتار برای استفاده در برنامه تبدیل متن به گفتار

مجری: زهرا محمودزاده

همکار طرح: وحید صادقی

۱۳۹۲

به نام خدا

فهرست مطالب

عنوان	شماره صفحه
گزارش طرح	أ
فصل اول معرفی طرح پژوهشی	۱
۱-مقدمه	۲
۱-۱. بیان مسئله	۲
۱-۲. پیشینه پژوهش	۴
۱-۳. مراحل اجرایی پژوهش	۵
فصل دوم: پیکره های گفتاری	۷
۲-مقدمه	۸
۲-۱. روش هم گذاری	۹
۲-۲. پیکره متنی	۱۱
۲-۳. ضبط پیکره زبانی	۱۲
۲-۴. تقطیع و برجسب دهی هجاها	۱۲

۱۳.....	۲-۵. آماده سازی دادگان هجایی (پردازش اولیه هجاها جهت تعیین آغازهای تناوب چاکنایی)
۱۳.....	۲-۶. پیاده سازی الگوهای آهنگ و تنظیم پارامترهای آوایی
۱۵.....	فصل سوم بافت نوایی
۱۶.....	۳. مقدمه
۱۶.....	۳-۱. واجشناسی نوایی
۱۹.....	۳-۱-۱. فرایندهای آوایی حساس به بافت نوایی
۲۲.....	۳-۱-۲. موضع تکیه بر
۲۳.....	۳-۲. پیاده سازی مشخصه های نوایی
۲۵.....	۳-۳. الگوی نوایی زبان فارسی
۲۶.....	۳-۳-۱. سیستم ثبت آهنگ
۲۷.....	۳-۳-۲. تکیه واژگانی
۲۸.....	۳-۳-۳. تکیه نوایی
۲۸.....	۳-۳-۴. گروه تکیه ای (AP)
۳۰.....	۳-۳-۵. ساختار AP
۳۱.....	۳-۳-۶. نواخت مرز نما
۳۲.....	۳-۳-۷. تکیه زیرویمی هسته (NPA)
۳۴.....	۳-۳-۸. گروه آهنگ
۳۵.....	۳-۴. هجای فارسی و الگوی نواختی آن
۳۵.....	۳-۴-۱. هجای CV
۳۵.....	۳-۴-۲. الگوهای نواختی

۴۱.....	۳-۴-۳. بررسی آکوستیکی گروه تکیه‌ای
۴۴.....	فصل چهار پیکره متنی
۴۵.....	۴. مقدمه
۴۵.....	۴-۱. طراحی جملات حامل هجاهای نوایی:
۴۶.....	۴-۲. هجای بی‌تکیه و ازگانی
۴۸.....	۴-۲-۱. همخوان‌های انفجاری
۵۱.....	۴-۲-۲. همخوان‌های سایشی
۵۴.....	۴-۲-۳. همخوان‌های انسایشی
۵۵.....	۴-۲-۴. همخوان‌های رسا
۵۸.....	۴-۳. هجای تکیه‌بر و ازگانی
۵۹.....	۴-۳-۱. همخوان‌های انسدادی
۶۷.....	۴-۳-۲. همخوان‌های سایشی
۷۵.....	۴-۳-۳. همخوان‌های انسایشی
۷۷.....	۴-۳-۴. همخوان‌های رسا
۸۲.....	۴-۴. تشدید
۸۴.....	۴-۴-۱. همخوان‌های انفجاری
۸۷.....	۴-۴-۲. همخوان‌های سایشی
۸۸.....	۴-۴-۳. همخوان‌های انسایشی
۸۹.....	۴-۴-۴. همخوان‌های رسا
۹۲.....	فصل پنج: ضبط و تهیه پیکره گفتاری

۵.مقدمه	۹۳
۵-۱.ضبط و تقطیع هجاها	۹۳
۵-۲.پارامترهای آکوستیکی	۹۵
۵-۲-۱.دیرش واکه	۹۵
۵-۲-۲.میانگین زیروبمی	۹۷
۵-۲-۳.میانگین شدت انرژی	۱۰۱
۵-۳.تحلیل نتایج برای برنامه تبدیل متن به گفتار	۱۰۶
فصل شش	۱۰۹
نتیجه گیری	۱۰۹
فهرست منابع	113

فصل اول

معرفی طرح پژوهشی

۱. مقدمه

از جمله اهداف استراتژیک پژوهشگاه، گسترش فناوری اطلاعات در دو حوزه نرم‌افزاری و سخت‌افزاری است. پیکره زبانی گفتاری یک از بخشهای مهم و اصلی نرم‌افزارهای سنتز گفتار مبتنی بر پیکره است. برای تبدیل متن به گفتار در این برنامه‌ها، پیکره گفتاری آن بخشی است که حاوی صورتهای صوتی آواها، هجاها، کلمات، جملات و مشخصه‌های اکوستیکی و نوایی آنهاست و برنامه برای تولید صورت آوایی یک متن، به آن رجوع می‌کند. هرچه پیکره تهیه شده به لحاظ خصوصیات زبانی و نوایی^۱ گفتار، غنی‌تر و با جزئیات آواشناختی بیشتری ساخته شود، صورت گفتاری تولید شده به وسیله نرم‌افزار بهتر، طبیعی‌تر و به گفتار انسان شبیه‌تر است.

۱-۱. بیان مسئله

اخیرا روشی در بازسازی گفتار مطرح شده است که بر اهمیت جایگاه و بافت نوایی در تهیه برابره‌های آوایی واحدهای بازسازی و تاثیر آن بر کیفیت گفتار بازسازی شده تاکید بسیار دارد. این روش مبتنی بر این فرض است که با انتخاب برابره‌های آوایی یک واحد بازسازی از بافت‌های نوایی متنوع و ذخیره‌سازی آنها در دادگان برنامه، واحدهای بازسازی با حفظ ویژگی‌های طبیعی زنجیره‌ای^۲ و زبرزنجیره‌ای^۳ خود در جایگاه‌های نوایی مختلف وارد زنجیره گفتار بازسازی شده می‌شوند که این مسئله باعث می‌شود کیفیت گفتار بازسازی شده بهبود یابد (Yoon, 2008, 2007, Hirschberg & Prieto, 1996). با وجود واحدهای حساس به بافت نوایی اول آن که واحدها با حفظ ویژگی‌های

^۱ . prosody

^۲ . segmental

^۳ . suprasegmental

نوایی طبیعی و ذاتی خود در جایگاه‌های نوایی مختلف از جمله دیرش^۴، زیروبمی^۵ و شدت انرژی^۶ وارد زنجیره گفتار بازسازی شده می‌شوند و حجم تغییرات و تنظیمات نوایی لازم برای بازسازی آهنگ گفتار پس از هم‌گذاری^۷ واحدها کاهش می‌یابد. دیگر آن که تغییرات زنجیره‌ای پیوسته گفتار طبیعی در گفتار بازسازی شده حفظ می‌شود. با این حال ساخت یک پیکره زبانی بزرگ که دربرگیرنده واحدهای زبانی در جایگاه‌های نوایی مختلف باشد، کار بسیار دشواری است ضمن آن که تبدیل این پیکره بزرگ زبانی به یک پیکره گفتاری پیوسته و تقطیع هجاهای پرتعداد از سطح آن بسیار وقت گیر و طاقت فرسا است. بنابراین برابری آوایی واحدهای بازسازی در چنین مدل‌هایی معمولاً از سطح چند جایگاه نوایی محدود و مهم از نظر ساخت سلسله مراتبی نوایی یا الگوی نوایی گفتار انتخاب می‌شوند و سایر برابری‌های آوایی واحدهای مورد نظر پس از هم‌گذاری با اعمال تغییرات و تنظیمات نوایی حاصل می‌شوند.

برای تهیه پیکره هجایی، هجاهای زبان فارسی در هر یک از چهار جایگاه نوایی (۱) موضع فاقد تکیه (واژگانی ۲) موضع تکیه بر واژگانی، دارای تکیه زیروبمی هسته در سطح جمله و دارای نواخت مرزی افتان (۳) موضع تکیه بر در کلمه، دارای تکیه زیروبمی غیرهسته در سطح جمله و دارای نواخت مرزی خیزان (۴) موضع تکیه بر در کلمه و دارای تکیه زیروبمی غیرهسته و نواخت مرزی خیزان (جملات سؤالی) قرار می‌گیرند. هر یک از این جایگاه‌های نوایی بر تظاهر آوایی هجای مورد نظر تاثیر می‌گذارد و این هجاها بسته به جایگاه‌های نوایی به صورت‌های مختلف تولید می‌شوند.

از آنجا که واجگان فارسی از ۶ واکه و ۲۳ همخوان تشکیل یافته است، با توجه به سه نوع هجای CV، CVC و CVCC شمار بالقوه هجاها در زبان فارسی ۷۶۳۱۴ خواهد بود که از این تعداد کمتر از ۶۰۰۰ فعلیت دارد. برای تهیه پیکره مورد نظر نمونه‌های هجای CV در ۴ بافت نوایی مختلف (در بافت جمله) که در بالا ذکر شد تولید می‌شوند. جملات پس از ضبط، با استفاده از نرم‌افزار تحلیل صوتی Praat، برحسب الگوی نوایی توسط آواشناس تقطیع و برچسب‌دهی می‌شوند. تهیه پیکره هجایی در قالب نظریه خودواحد-عروضی پی‌یرهامبرت^۸ (۱۹۸۰) و لد^۹ (۱۹۹۶) و نیز الگوی آهنگ گفتار فارسی سادات تهرانی (۲۰۰۷) صورت می‌گیرد و برای ثبت نواخت‌ها در منحنی زیروبمی از

4. duration

5. pitch

6. intensity

7. concatenation

8. Pierrehumbert, J.

9. Ladd, D.R.

نظام نشانه گذاری تویی^{۱۰} (ToBI) که مجموعه ای از نشانگرهای نواختی و مرزی است، استفاده می‌شود. در تهیه دادگان هجایی، علاوه بر بافت نوایی، بافت آوایی برابره‌های آوایی هجاها نیز حائز اهمیت فراوان است و بر کیفیت گفتار بازسازی شده تاثیر می‌گذارد. اثر این دو عامل باید به شکل موثری در برنامه اعمال شود تا کیفیت گفتار بازسازی شده تا حد مطلوبی ارتقا یابد.

۲-۱. پیشینه پژوهش

دادگان صوتی یکی از بخش‌های ضروری نرم‌افزارهای تبدیل متن به گفتار است. الگوی آهنگ زبان فارسی از سوی افرادی نظیر توحیدی (۱۹۷۴)، اسلامی (۱۳۸۴)، ماهجانی (۲۰۰۳) مورد مطالعه و بررسی قرار گرفته است. اما تاکنون پیکره زبانی که مبتنی بر ساختار نوایی زبان فارسی باشد، تهیه نشده است. اسلامی و همکاران (۱۳۸۸) در طراحی بخش دادگان صوتی موتور تبدیل متن به گفتار «گویا» از دو نوع دادگان صوتی هجایی و دایفونی استفاده نمودند. آنها بخشی از هجاها را از دادگان گفتاری پژوهشکده پردازش هوشمند علائم، «فارس‌دات بزرگ» و بخشی را با استفاده از فرهنگ فارسی معین استخراج نمودند. علاوه بر آن برخی از هجاها را نیز در کلمات بی‌معنی^{۱۱} قرار داده و پس از تولید استخراج نمودند. به لحاظ ساختار نوایی و برجستگی، آنها هجاها را مورد نظر را در موضع بی‌تکیه آغاز کلمات مجزا قرار داده و پس از تولید، استخراج نمودند. بنابراین تولید این هجاها تنها در موضع بی‌تکیه و آغاز کلمه و خارج از بافت یعنی در کلمات مجزا صورت گرفته و از این رو علاوه بر آن که هجاها فاقد تنوع برجستگی و نوایی هستند، از آنجا که تولید آنها خارج از بافت صورت گرفته، فاقد ویژگی‌های گفتار پیوسته نیز هستند.

پیکره هجایی نوایی نوعی پیکره کاربردی است که در نرم‌افزار تبدیل متن به گفتار فارسی مورد استفاده قرار می‌گیرد و قابل استفاده در تحقیقات آواشناسی و زبانی نیز هست. از جمله پیکره‌های زبانی موجود می‌توان به فارس‌دات و نیز پیکره ارقام تلفنی تهیه شده در مؤسسه پردازش هوشمند علائم اشاره کرد. به نظر نمی‌رسد که در تهیه این پیکره‌ها تنوعات نوایی لحاظ شده باشد.

اخیرا مطالعاتی در زمینه نقش جایگاه و بافت نوایی در بازسازی گفتار صورت گرفته است. یون^{۱۲} (۲۰۰۷) به طراحی و ارزیابی یک سامانه بازسازی گفتار به روش هم‌گذاری واحدهای دایفونی^{۱۳} حساس به ساختار نوایی در زبان کره‌ای می‌پردازد. وی دایفون‌ها را در مواضع متفاوت نوایی مثل

¹⁰ . tones and break indices

¹¹ . nonsense

¹² . Yoon, K.

¹³ . diphone

گروه آهنگ^{۱۴}، گروه تکیه‌ای^{۱۵} و کلمه واجی^{۱۶} و در جایگاه‌های متفاوت کلمه قرار می‌دهد و با انجام یک آزمایش ادراکی به ارزیابی شنیداری واحدهای حاصل از هم‌گذاری دایفون‌های نوایی می‌پردازد. نتایج پژوهش وی نشان می‌دهد که شنوندگان نمونه‌های حاصل از دایفون‌های نوایی را ترجیح می‌دهند و آنها را طبیعی‌تر تشخیص می‌دهند.

صادقی (۱۳۸۹) به طراحی و ارزیابی یک سامانه بازسازی گفتار فارسی به روش هم‌گذاری واحدهای حساس به بافت نوایی می‌پردازد. وی نمونه‌های گفتاری را به سه روش متفاوت بازسازی می‌کند. در روش اول، کلمات از طریق هم‌گذاری هجاهای حساس به بافت نوایی بدون پردازش فرکانس پایه^{۱۷} (F0) و دیرش^{۱۸} هجا ساخته می‌شوند. در روش دوم کلمات با هم‌گذاری هجاهای غیرنوایی ساخته شده و سپس در مرحله بعد پردازش نوایی پارامترهای فرکانس پایه (F0) و دیرش هجا صورت می‌گیرد و در روش آخر کلمات بدون هیچ‌گونه پردازش نوایی از طریق هم‌گذاری هجاهای غیرنوایی ساخته می‌شوند. نمونه‌های حاصل سپس در یک آزمون ادراکی مورد ارزیابی از سوی آزمودنی‌ها قرار گرفتند. نتایج آزمون ادراکی نشان می‌دهد که گفتار بازسازی‌شده حاصل از دو روش اول و دوم به طور معناداری طبیعی‌تر است. علاوه بر آن میزان طبیعی بودن گونه گفتاری حاصل از روش اول بیشتر از روش دوم است.

این آزمایش‌ها نشان می‌دهد که واحدهای بازسازی حساس به بافت نوایی، کیفیت گفتار بازسازی‌شده را تا حد قابل‌قبولی بهبود می‌بخشد.

۳-۱. مراحل اجرایی پژوهش

۱- مطالعه ساختار نوایی زبان فارسی

۲- انجام مطالعات زبانشناختی جهت تعیین الگوی مناسب استخراج هجاها

۳- تهیه پیکره نوشتاری

۴- ضبط دادگان

14 . intonational phrase

15 . accentual phrase

16 . phonological word

17 . fundamental frequency

18 . duration

۵- تقطیع و برجسب دهی هجاها و تولید پیکره گفتاری

روش‌شناسی سه گام اول به صورت توصیفی شامل مطالعه ادبیات و منابع موجود در زمینه ویژگیهای نوایی گفتار فارسی و استخراج بافت نوایی و آوایی دادگان هجایی است.

روش‌شناسی دو گام آخر به صورت آزمایشگاهی و شامل ضبط داده‌ها، تقطیع هجاها و برجسب‌دهی الگوی نوایی آنها با استفاده از نرم‌افزار Praat و براساس پارامترهای اکوستیکی الگوی نوایی گفتار.

خروجی نهایی این طرح پیکره گفتاری هجای CV زبان فارسی است که به لحاظ آکوستیکی و الگوی نوایی فارسی تقطیع و برجسب‌دهی می‌شود. این پیکره می‌تواند در نرم‌افزار تبدیل متن به گفتار فارسی و یا در مطالعات زبانی مورد استفاده قرار گیرد.

فصل دوم

پیکره های گفتاری

۲. مقدمه

بازسازی گفتار می‌تواند به دو روش مختلف هم‌گذاری^{۱۹} و پارامتری^{۲۰} صورت گیرد. در روش هم‌گذاری برابری آوایی متن ورودی برنامه که واحدهای بازسازی محسوب می‌شوند، با قواعد و اصول خاصی از دادگان صوتی برنامه بازسازی کننده گفتار احضار و کنار هم چیده می‌شوند و شرایط برای تبدیل به گفتار فراهم می‌شود. در تهیه واحدهای بازسازی، ابتدا پیکره^{۲۱} زبانی لازم تهیه می‌شود و سپس پیکره زبانی مورد نظر از سوی یک یا چند نفر اهل زبان در یک محیط آزمایشگاهی خوانده شده و ضبط می‌شود. بعد از انجام پردازش‌های لازم واحدهای بازسازی از پیکره استخراج و در قالب یک دادگان صوتی در چارچوب برنامه بازسازی گنجانده می‌شود. در روش پارامتری بازسازی گفتار که به آن روش مبتنی بر قاعده نیز گفته می‌شود، همه پارامترهای آکوستیکی گفتار مانند سایش^{۲۲}، سازه‌ها^{۲۳} و دامنه نوسان^{۲۴} آواها کنار گذاشته می‌شود و برنامه این پارامترها را در چارچوب قواعد خاص بازسازی می‌کند. در این روش (مانند بازسازی کننده Klatt) ابتدا واحد واجی به چند مشخصه صوت‌شناختی تقسیم می‌شود و هر کدام از این واحدهای صوت‌شناختی براساس "توالی سیگنال‌های کنترلی"^{۲۵} بازتولید می‌شوند. در این روش ویژگی‌های آوایی مربوط به هم‌تولیدی و گذر از صدایی به صدای دیگر لحاظ می‌شود.

¹⁹ . concatenation

²⁰ . parametric

²¹ . corpus

²² . frication

²³ . formants

²⁴ . amplitude

²⁵ . sequence of control signals

۱-۲. روش هم‌گذاری

به طور کلی، کیفیت گفتار بازسازی شده در یک سیستم سنتز گفتار مبتنی بر روش هم‌گذاری به کیفیت واحدهای گفتاری بستگی دارد. کیفیت واحدهای یک دادگان گفتاری به عوامل بسیاری بستگی دارد: نوع واحدها (واج^{۲۶}، دایفون، ترایفون^{۲۷}، هجا کلمه، عبارت زبانی و جمله)، روش برچسب‌دهی (دستی یا خودکار)، تعداد نمونه‌ها برای هر واحد بازسازی، غنای نوایی واحدهای بازسازی و غیره. یون (۲۰۰۷، ۳) اظهار می‌کند که کیفیت گفتار بازسازی شده به طول واحدهای بازسازی بستگی دارد، هر قدر طول واحد بازسازی بلندتر باشد، هم‌تولیدی گفتار خروجی طبیعی‌تر خواهد بود.

در برنامه‌های بازسازی گفتار به روش هم‌گذاری واحدهای متفاوتی برای بازسازی گفتار استفاده می‌شود. در بازسازی‌کننده‌های با کاربرد خاص و محدود، از واحدهای بازسازی بزرگ مانند جملات و عبارت‌ها می‌توان استفاده کرد. این گونه برنامه‌ها مناسب محل‌هایی مانند پایانه‌ها، فرودگاه‌ها، بازارهای بورس و اعلام وضع هوا و غیره است که در آن پیام‌های زبانی به صورت متون نوشتاری خاص و تقریباً کلیشه‌ای تولید می‌شوند و ارتباط کاربر با برنامه بازسازی‌کننده بسیار محدود است. کلمه واحد زبانی کوچکتر از جمله و عبارت است که می‌تواند به عنوان واحد بازسازی گفتار مورد استفاده قرار گیرد (Buron 1968, Chapman 1971, Eady et al. 1987). در طراحی برنامه‌ای جامع برای بازسازی گفتار، کلمه واحد مناسبی نیست، چراکه احصا و نیز ذخیره کردن اطلاعات دیجیتالی همه کلمه‌های زبان عملاً غیرممکن است. به همین دلیل پژوهشگران در بازسازی گفتار از واحدهای کوچکتری استفاده می‌کنند که هم اقتصادی است و هم می‌توان به لحاظ آوایی در استفاده از آنها به تعمیم‌هایی دست پیدا کرد و کیفیت گفتار بازسازی شده را بهبود بخشید. این واحدهای کوچکتر عبارتند از: واج، دایاد^{۲۸}، دایفون، نیم‌هجا^{۲۹} و هجا که در ادامه به شرح مختصر هر کدام می‌پردازیم.

واج به عنوان یک واحد انتزاعی ممیز معنا در زبان به خاطر داشتن فهرست محدود ظاهراً مناسبترین واحد بازسازی است، ولی همه تلاش‌ها در هم‌گذاری واج‌ها در بازسازی گفتار با شکست مواجه شده است و علت اصلی این شکست از این واقعیت ناشی می‌شود که مرز و کیفیت آوایی واج‌ها در زنجیره گفتار با توجه به بافت آوایی به لحاظ صوت‌شناختی متغیر است. واقعیت هم‌تولیدی^{۳۰} بین صداهای

²⁶ . phoneme

²⁷ . triphone

²⁸ . dyad

²⁹ . demisyllable

³⁰ . coarticulation

مجاور در گفتار و واقعیت عروضی^{۳۱} بین صداها در فضای آکوستیکی گفتار باعث می‌شود که استفاده از واج به عنوان واحد بازسازی گفتار با موفقیت همراه نباشد. حتی استفاده از واجگونه‌ها^{۳۲} نیز نمی‌تواند به رفع این مشکل کمک کند، چرا که واجگونه‌ها هم مانند واج‌ها فاقد بسیاری از ویژگی‌های هم‌تولیدی هستند. بنابراین گفتار بازسازی‌شده با استفاده از واج‌ها و حتی واجگونه‌ها به‌زحمت قابل درک خواهد بود (Sivertsen, 1961). بنابراین پژوهشگران به واحدهای دیگری روی آوردند و در انتخاب واحدها دو ملاک را مد نظر دادند؛ اول اینکه این واحدها قابل پردازش با رایانه باشند و به گونه‌ای انتخاب شوند که استفاده از آنها در رایانه امکان‌پذیر باشد و دوم اینکه این واحدها حائز شرایط هم‌تولیدی و گذر از صدایی به صدای دیگر باشند. دایفون، نیم‌هجا و هجا حائز شرایط فوق هستند. دایاد از بخش پایانی یک واحد (واج) و از بخش آغازی واحد بعدی تشکیل می‌شود. اصطلاح دایفون معادلی برای اصطلاح دایاد است؛ اگر چه در مفهوم دقیق‌تر این دو متفاوت از هم هستند (Klatt, 1987). در اصل دایفون ناظر بر ناحیه گذر بین صداهای مجاور است که اثرات هم‌تولیدی و اطلاعات مربوط به گذر از یک آوا به آوای مجاور را در بر می‌گیرد، ولی دایاد بخش پایدار^{۳۳} صدا را نیز شامل می‌شود. نیم‌هجا بیشتر شبیه به دایفون است، با این تفاوت که در هجای CVC مرز دایفون وسط واکه است، در حالیکه در نیم‌هجا پایان ناحیه گذر و آغاز بخش پایدار واکه، مرز تلقی می‌شود. ترایفون واحد دیگری در بازسازی گفتار است که در آن دو گذر واکه به همخوان و همخوان به واکه و یا بالعکس دیده می‌شود (مثلاً VCV یا CVC). هجا نیز یکی دیگر از واحدهای مهم و رایج در بازسازی گفتار به روش هم‌گذاری است که علاوه بر آن که همچون دایفون محدوده گذر بین آواها و شرایط هم‌تولیدی آنها را در بر می‌گیرد، تشخیص مرز این واحد نیز در گفتار پیوسته در مقایسه با واجگونه‌ها و دایفون‌ها آسان‌تر است. انتخاب یکی از این واحدها به عنوان واحد بازسازی فارسی تا حد زیادی وابسته به نظام آوایی زبان فارسی از جمله ساختار هجایی، محدودیت‌های همنشینی واجها در درون هجا و غیره است که مطالعات صورت گرفته اولیه در این زمینه نشان می‌دهد با توجه با سادگی ساختار هجایی فارسی و روشن تر بودن مرز هجاها در گفتار پیوسته در قیاس با دایفون‌ها، هجا واحد مناسب تری برای بازسازی گفتار فارسی است.

علاوه بر این با فرض انتخاب هجا یا دایفون به عنوان واحد بازسازی سوال دیگری که مطرح است آن که آیا به ازای هر هجا (یا هر واحد بازسازی دیگر) یک برابر آوایی در دادگان گفتاری برنامه ذخیره شده و سایر برابره‌های آوایی واحد پس از هم‌گذاری آن در کنار دیگر واحدها از طریق پردازش

³¹ . prosody

³² . allophone

³³ . steady state

نوایی و تنظیمات مقادیر پارامترهای نوایی متناسب با بافت نوایی جدید بازسازی شود (مانند برنامه های بازسازی گفتار الن و دیگران^{۳۴} ۱۹۸۷، الن و دیگران^{۳۵} ۱۹۹۹، اسپروت^{۳۶} ۱۹۹۸) یا آن که با توجه به تاثیر و اهمیت جایگاه و بافت نوایی بر نحوه تظاهر هجاها چندین برابر آوایی برای هر واحد هجا از بافت های نوایی متنوع انتخاب و در دادگان برنامه ذخیره شود به طوری که واحدهای بازسازی با حفظ ویژگی های طبیعی زنجیره ای و زبرزنجیره ای خود در جایگاه های نوایی مختلف وارد زنجیره گفتار بازسازی شده شوند. رویکرد دوم روش جدیدتری در بازسازی گفتار است. طرفداران این روش معتقدند که با وجود واحدهای حساس به بافت نوایی واحدها با حفظ ویژگی های نوایی طبیعی و ذاتی خود در جایگاه های نوایی مختلف از جمله دیرش، زیروبمی و شدت انرژی وارد زنجیره گفتار بازسازی شده می شوند و حجم تغییرات و تنظیمات نوایی لازم برای بازسازی آهنگ گفتار پس از هم گذاری واحدها کاهش می یابد. ضمن آن که تغییرات زنجیره ای پیوسته گفتار طبیعی در گفتار بازسازی شده حفظ می شود. استفاده از واحدهای حساس به بافت نوایی به عنوان واحدهای بازسازی کننده گفتار تا حد زیادی باعث افزایش کیفیت گفتار بازسازی شده می شود با این حال ساخت یک پیکره زبانی بزرگ که دربرگیرنده واحدهای زبانی در جایگاه های نوایی مختلف باشد کار بسیار دشواری است ضمن آن که تبدیل این پیکره بزرگ زبانی به یک پیکره گفتاری پیوسته و تقطیع هجاها از تعداد از سطح آن بسیار وقت گیر و طاقت فرسا است. بنابراین برابری آوایی واحدهای بازسازی در چنین مدل هایی معمولاً از سطح چند جایگاه نوایی محدود و مهم از نظر ساخت سلسله مراتبی نوایی یا الگوی نوایی گفتار انتخاب می شوند و سایر برابری های آوایی واحدهای مورد نظر پس از هم گذاری با اعمال تغییرات و تنظیمات نوایی حاصل می شوند.

۲-۲. پیکره متنی

پس از انتخاب واحد بازسازی کننده گفتار، فهرست تمامی واحدها شناسایی و جمع آوری شده و یک پیکره زبانی که دربرگیرنده واحدهای مورد نظر باشد تهیه می گردد. هجا یکی از واحدهای مهم در بازسازی گفتار به شیوه هم گذاری است که در آن فهرست هجاها با فعل زبان جمع آوری می شود و بعد از فراهم کردن پیکره گفتاری لازم، برابری های آوایی هر کدام از هجاها از پیکره استخراج و سپس در دادگان صوتی برنامه قرار می گیرد.

³⁴ . Allen, J. et al.

³⁵ . Alan, W. et al.

³⁶ . Sproat, R.

به طور مثال با فرض انتخاب هجاهای حساس به بافت نوایی به عنوان واحد بازسازی، ابتدا با در نظر گرفتن محدودیت های واج آرایشی زبان فارسی، فهرست هجاهای بالفعل زبان فارسی را جمع آوری می کنیم. سپس هر هجا را با در نظر گرفتن جایگاه های نوایی پیش بینی شده در درون جملات طبیعی یا جملات حامل قرار می دهیم و به این ترتیب یک پیکره زبانی را که در برگیرنده هجاهای مورد نظر در جایگاه های نوایی مختلف باشند تهیه می کنیم.

۲-۳. ضبط پیکره زبانی

پس از تهیه پیکره زبانی، پیکره مورد نظر در شرایط آزمایشگاهی تولید و ضبط خواهد شد. ضبط سیگنال آوایی در اتاقک ضد صدا با استفاده از تجهیزات سخت افزاری و نرم افزاری تخصصی ضبط گفتار از جمله میکروفون با پاسخ بسامدی مناسب و نسبت سیگنال به نوفه بالا، کارت صوتی مناسب (خروجی سیگنال آوایی) و نرم افزار پرت (Praat) انجام خواهد گرفت.

۲-۴. تقطیع و برچسب دهی هجاها

پس از ساخت پیکره گفتاری، برابری آوایی هر هجا از پیکره استخراج و در دادگان گفتاری برنامه با برچسبی مشخص گنجانده می شود.

به لحاظ واجی در گفتار فارسی رسمی، در یک هجا بیش از یک واکه نمی تواند وجود داشته باشد، از اینرو، تعداد هجاها در هر رشته آوایی می تواند با شمارش واکه ها مشخص شود. اما تعیین مرز بین دو هجا بستگی به شمارش همخوان های بین دو واکه دارد. حداقل تعداد همخوان بین دو واکه یک و حداکثر سه همخوان است. به این ترتیب آرایش همخوانی بین هر دو واکه در هر رشته آوایی به یکی از سه شکل زیر خواهد بود:

الف) VCV ب) VCCV ج) VCCCV

مرز هجا در مورد اول بین V اول و C است، چون هجا در فارسی می تواند با واکه پایان پذیرد، اما نمی تواند با واکه آغاز شود. مثال این مورد، رشته آوایی /dava/ می باشد که به ترتیب گفته شده دو هجای /da/ و /va/ به دست می آید. در مورد دوم، محل برش هجایی بین دو C خواهد بود زیرا اولاً واکه در آغاز هجا قرار نمی گیرد و ثانیاً خوشه دوهمخوانی در آغاز هجا مجاز نیست. نمونه این مورد، رشته آوایی /doxtar/ است که در آن دو هجای /dox/ و /tar/ به دست می آید. در مورد سوم، مرز دو هجا با دلایل ذکر شده در بالا، بین C دوم و سوم است. این نقطه برش، رشته آوایی

مذکور را به دو هجای CVCC و CV تقسیم می‌کند. نمونه این مورد، رشته آوایی /dʒangdʒu/ است که در آن سه همخوان در میان دو واکه واقع شده‌اند. در اینجا محل برش هجایی بین /dʒ/ و /g/ می‌باشد و در نتیجه دو هجای /dʒu/ و /dʒang/ به دست می‌آید.

با توجه به اطلاعات فوق می‌توان گفت در هر رشته آوایی، همخوان قبل از هر واکه بیانگر ابتدای هجا است. بر مبنای این قانون می‌توان مرزهای هجایی را در هر رشته آوایی به آسانی و بدون کمترین شک و ابهامی مشخص کرد اما تشخیص مرز هجاها در فضای آوایی گاه به علت همپوشی و هم‌تولیدی آواها با یکدیگر در مرزهای هجایی دشوار است. در انجام این مرحله از پروژه از معیارها و الگوهای تقطیع هجایی ون‌زنتن^{۳۷} (۱۹۹۸، ۲۰۰۴) برای تشخیص مرزهای هجایی استفاده خواهیم کرد. تقطیع هجایی در نرم افزار پرت انجام خواهد شد.

۲-۵. آماده سازی دادگان هجایی (پردازش اولیه هجاها جهت تعیین آغازهای تناوب چاکنایی)

در برنامه‌های بازسازی گفتار پس از هم‌گذاری واحدها، فرکانس پایه و دیرش هجاهای خاص در زنجیره گفتار بازسازی شده به نسبت معین تغییر می‌کند که در نتیجه آن زیروبمی و دیرش هجاها متناسب با پاره گفتار جدید تغییر کرده و کیفیت گفتار بازسازی شده بهبود می‌یابد. یکی از روش‌های مهم و رایج در پردازش فرکانس پایه و دیرش واحدهای بازسازی روش PSOLA (pitch synchronous overlap and add) است. در این روش پالس‌های حنجره به صورت فریم‌های آنالیز (که طول هر فریم برابر با یک دوره تناوب چاکنایی است) یک به یک از سیگنال حنجره جدا میشوند، سپس تغییرات لازم در حوزه زمان یا زیروبمی بر روی فریم‌ها انجام می‌گیرد و در نهایت فریم‌ها برای تولید سیگنال دلخواه با هم ترکیب می‌شوند.

۲-۶. پیاده سازی الگوهای آهنگ و تنظیم پارامترهای آوایی

در این مرحله مطابق با اطلاعات صرفی و نحوی برنامه بازسازی گفتار، زنجیره واجی و ساخت هجایی متن ورودی مشخص می‌شود. سپس برابری صوتی هر کدام از هجاها از دادگان گفتاری برنامه اخذ می‌شود. بعد از هم‌گذاری، گفتار تولید می‌شود که این گفتار تولید شده با توجه به اطلاعات موجود تغییراتی می‌یابد و گفتار طبیعی‌تر می‌شود. یعنی براساس مرز گروه‌های نحوی و آهنگی و ساختار آنها و شناسایی جایگاه‌های تکیه زیروبمی و انجام تغییرات لازم در این جایگاه‌ها به تولید نوای گفتار

³⁷ . Van Zanten, E.

ساخت پیکره گفتاری هجایی زبان فارسی مبتنی بر الگوی نوایی... □ 14

طبیعی روی گفتار هم گذاری شده می انجامد. همچنین اقداماتی جهت طبیعی تر و واضح تر شدن گفتار صورت می گیرد. این اقدامات پس از انجام تحقیقات مستقل روشن می شود.

فصل سوم

بافت نوایی

۳. مقدمه

هنگامی که گویشور، برای تولید یک واحد آوایی، بستاری در دستگاه گفتار ایجاد می‌کند، پیامدهای صوت‌شناختی آن نوعاً متعدد و متنوع هستند و معمولاً در یک فاصله زمانی نسبتاً طولانی بر روی سیگنال صوتی گسترش و توزیع یافته‌اند.

فرایندهای آوایی مثل هم‌تولیدی، کاهش^{۳۸}، حذف و غیره در گفتار سبب می‌شود تا ویژگی‌های آوایی هر واحد زبانی به میزان زیادی تحت تأثیر واحدهای اطراف قرار گیرد. امروزه سعی بر آن است تا با ادغام برخی از این فرایندها در سیستم‌های فناوری زبان مثل تبدیل متن به گفتار یا بازشناسی گفتار، تاحدی ویژگی‌های طبیعی آواها حفظ شود.

مسئله مهم دیگری که معمولاً در سیستم‌های فناوری زبان کمتر مورد توجه قرار می‌گیرد، ویژگی‌های نوایی گفتار است. تولید هر واحد آوایی تابعی است از بافت نوایی آن یعنی جایگاه آوا در ساختار نوایی زبان.

۳-۱. واجشناسی نوایی

پروزودی^{۳۹} یا نوای گفتار یعنی سازماندهی گفتار به واحدها یا حوزه‌هایی که ساخت سلسله‌مراتبی دارند به گونه‌ای که برخی از این واحدها به لحاظ آوایی برجسته‌تر از سایر واحدها هستند. براساس این تعریف، پروزودی دو نقش ایفا می‌کند: از سویی گروه‌بندی^{۴۰} و ایجاد سازه‌های نوایی^{۴۱} و از سوی

³⁸ . reduction

³⁹ . prosody

⁴⁰ . phrasing

⁴¹ . prosodic constituents

سوی دیگر نقش برجسته‌سازی^{۴۲} واحدهای آوایی در زنجیره گفتار. بنابراین پروزودی شامل (۱) گروه‌بندی گفتار، یعنی گروه‌بندی واحدهای کوچکتر به بزرگتر با اندازه‌های مختلف، مثل گروه‌بندی آواها به هجاها، هجاها به کلمات، یا کلمات به گروه‌ها (منظور از گروه‌ها اینجا می‌تواند معادل یک سازه نحوی باشد یا نباشد) (۲) تکیه^{۴۳}، یعنی برجستگی هجا در درون کلمه (۳) تکیه زیروبمی^{۴۴}، یعنی برجستگی در سطح گروه و توزیع نواخت‌ها در هجاها (نواخت واژگانی) و در گروه‌ها (نواخت آهنگ) (Keating, 2003, 119). همه این جنبه‌های مختلف و متفاوت پروزودی، با یکدیگر ساختار نوایی یک پاره گفتار آوایی را می‌سازد.

از جمله نظریه‌های مطرح در زمینه ساختار نوایی گفتار بر اساس نقش اول یعنی گروه‌بندی و ایجاد واحدهای نوایی، نظریه واجشناسی نوایی نسپر و ووگل^{۴۵} (۱۹۸۶) است که خود بر مبنای مدلی از واجشناسی نوایی است که توسط سلکرک^{۴۶} (۱۹۷۸) برای اولین بار در مقاله‌ای تحت عنوان "ساخت نوایی و رابطه آن با ساخت نحوی"^{۴۷} مطرح شد. در نظریه واجشناسی نوایی نسپر و ووگل (۱۹۸۶، ۱-۷)، بازنمایی ذهنی گفتار از واحدها یا سازه‌هایی تشکیل یافته که ساخت سلسله‌مراتبی یا درختی دارند. این سازه‌ها توسط سرنخهای آوایی در موج صوتی گفتار انتقال پیدا می‌کنند و حوزه عملکرد قواعد واجی و فرایندهای آوایی قرار می‌گیرند.

نسپر و ووگل (۱۹۸۶: ۱۱) گفتار را مرکب از هفت سازه^{۴۸} آوایی برمی‌شمارند که به ترتیب از کوچکترین به بزرگترین سازه عبارتند از: هجا (σ)، پا^{۴۹} (Σ)، کلمه واجی (ω)، گروه واژه‌بست^{۵۰} (C)، گروه واجی^{۵۱} (Φ)، گروه آهنگ (I) و پاره گفتار واجی^{۵۲} (U). سازه‌های نوایی ساخت سلسله‌مراتبی دارند و براساس برخی اطلاعات آوایی، صرفی-نحوی و یا معنایی گفتار ساخته می‌شوند. شکل ۱ ساخت سلسله‌مراتبی سازه‌ها را نشان می‌دهد.

42 . prominence

43 . stress

44 . accent

45 . Nespor, M. & Vogel, I.

46 . Selkirk, E.

47 . " On prosodic structure and its relation to syntactic structure"

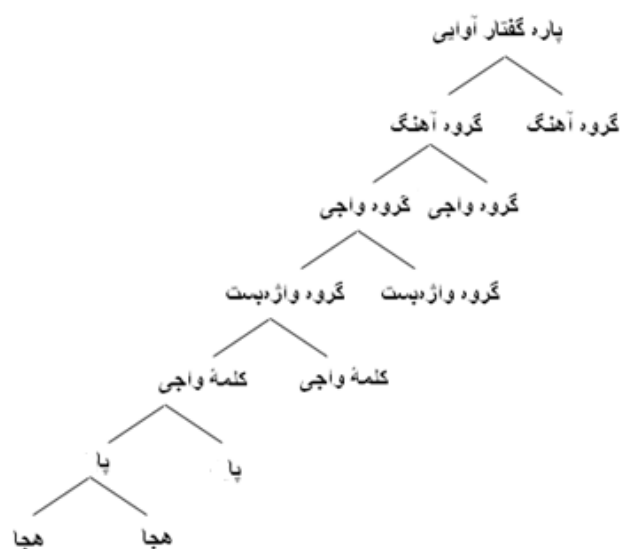
48 . prosodic constituent

49 . foot

50 . clitic group

51 . phonological phrase

52 . phonological utterance



شکل ۳-۱: بازنمایی درختی سازه‌های نوایی

از جمله نظریه‌های مطرح در زمینه نظام آهنگ گفتار بر اساس نقش دوم واجشناسی نوایی، نظریه خودواحد-وزنی آهنگ لد (۱۹۹۶) است که با آثار افرادی نظیر لیبرمن^{۵۳} (۱۹۷۵)، بروس^{۵۴} (۱۹۷۷)، و پیرهامبرت (۱۹۸۰) آغاز شد، آهنگ گفتار از رویدادهای نواختی معنادار به لحاظ واجی مثل تکیه زیرویمی و نواخت‌های مرزی تشکیل یافته است. این رویدادها که شامل زیرویمی خیزان (H) و افتان (L) است در نقاطی در زنجیره گفتار روی می‌دهند. نواخت‌ها به لحاظ آوایی بر محور زمان مترادف دارند و مقادیر آنها بر روی محور عمودی یا فرکانس قابل اندازه‌گیری است. پیکره هجایی نوایی براساس نظریه خودواحد-وزنی لد (۱۹۹۶) و نیز نظام آهنگ زبان فارسی (Sadat Tehrani, 2007) تهیه و استخراج می‌شود که در بخش‌هایی بعدی به آن پرداخته خواهد شد.

در این پژوهش تهیه پیکره هجایی براساس نقش دوم پروزودی یعنی تغییرات برجستگی در سطح جمله صورت می‌گیرد. اما از آنجا که برطبق نقش اول، جایگاه تأثیر عمده‌ای بر مشخصه‌های آوایی واحدهای زبانی، مثل دیرش و تقویت^{۵۵} آکوستیکی دارد، و هجاهای نوایی نیز از موضع آغاز یا پایان کلمه و در جایگاه‌های مختلف جمله استخراج می‌شوند، بنابراین در بخش بعدی به برخی فرایندهای آوایی اشاره می‌شود که حساس به بافت نوایی هستند. این فرایندها نشان می‌دهند که به چه میزان توجه به موضع و ساختار نوایی در تهیه پیکره‌های گفتاری اهمیت دارد.

⁵³ . Liberman, M.

⁵⁴ . Bruce, G.

⁵⁵ . strengthening

۳-۱-۱. فرایندهای آوایی حساس به بافت نوایی

ویژگی‌های نوایی گفتار در پارامترهای آوایی زیروبمی، دیرش و بلندی^{۵۶} ظهور پیدا می‌کند و در سیگنال گفتار منتقل می‌شود و به همین دلیل به این ابعاد آوایی، پارامترهای زبرزنجیری گفته می‌شود. اما پارامترهای آوایی دیگری که معمولاً پارامترهای زنجیری در نظر گرفته می‌شوند و نه زبرزنجیری، نیز تغییرات نوایی سیگنال گفتار را انتقال می‌دهند. به عنوان مثال معمولاً کیفیت واکه نه تنها با تغییر واجی نوع واکه تغییر می‌کند بلکه با عوامل زبرزنجیری مثل تکیه و کشش نیز تغییر می‌کند (Lehiste, 1970). به طور کلی تولید آوایی مشخصه‌های واجی یک واحد گفتاری، واکه یا همخوان، تاحدی به جایگاه آن واحد در کل ساختار نوایی بستگی دارد. می‌توان هر واحد واجی را به صورت مجموعه‌ای از مشخصه‌های آوایی در نظر گرفت که در آخرین گره در پایین درخت بسیار بزرگ ساختار نوایی قرار دارد. در این صورت تولید هر یک از این مشخصه‌ها به تولید سایر مشخصه‌ها در همین واحد، مشخصه‌های واحدهای واجی مجاور و جایگاه مشخصه در کل درخت بستگی دارد. بنابراین ابعاد آوایی زنجیری نیز به همان میزان تحت تأثیر ساختار نوایی هستند که ابعادی که به طور سنتی زبرزنجیری شناخته می‌شوند (Keating, 2003, 119).

چو^{۵۷} (۲۰۰۵) الگوهای آوایی متفاوتی را که ناشی از ساختار نوایی گفتار هستند، تحت عنوان "تقویت نوایی"^{۵۸} قرار می‌دهد و آن را به صورت "گسترش"^{۵۹} زمانی یا مکانی تولید در نتیجه تکیه^{۶۰} و مرز نوایی^{۶۱} "تعریف می‌کند. نسپر و ووگل (۱۹۸۶) معتقدند که فرایندهای سندهی^{۶۲} به وسیله ساختار نوایی محدود می‌شوند: این فرایندها معمولاً در مرزهای نوایی پایتتر رایجتر و قدرتمندترند و کمتر در مرزهای نوایی بالاتر روی می‌دهند.

دو همبسته صوت‌شناختی ساختار نوایی عبارتند از افزایش کشش پایانی^{۶۳} و تقویت آغازین^{۶۴}. افزایش کشش پایانی سبب می‌شود هجاهای پایانی سازه نوایی^{۶۵} طولانیتر از هجاهای میانی آن تولید شوند. همبسته دیگر آن، تقویت آغازین (که کازلا و همکاران^{۶۶} (۲۰۰۷) به صورت افزایش کشش آغازین^{۶۷}

⁵⁶ . loudness

⁵⁷ . Cho, T.

⁵⁸ . prosodic strengthening

⁵⁹ . expansion

⁶⁰ . accent

⁶¹ . sandhi

⁶² . final lengthening

⁶³ . initial strengthening

⁶⁴ . domain-final syllables

⁶⁵ . Kuzla et al.

⁶⁶ . initial lengthening

نیز آن را به کار برده‌اند)، سبب می‌شود همخوانها تماس زبانی-کامی بیشتری نشان دهند، انسدادیها با بستهای بلندتر و زمان شروع واک طولانی‌تری تولید شوند و واکه‌ها اغلب بیشتر چاکنایی و مقاومت بیشتری نسبت به هم‌تولیدی واکه به واکه نشان دهند. به نظر می‌رسد تأثیر هر دو همبسته ساختار نوایی، افزایش کشش پایانی و تقویت آغازین، در بعد عمودی پیوستار نوایی افزایش می‌یابد بدین معنا که در سازه‌های نوایی بالاتر، تأثیر آنها قویتر است (کازلا و همکاران، ۲۰۰۷: ۳۰۲).

کیتینگ (۲۰۰۳، ۱۲۰) اظهار می‌کند که فرایندهایی مثل تقویت آغازین و کشش پایانی نشان می‌دهند که صورت آوایی و واجی کلمات (و واحدهای بزرگتر) در یک زبان به وسیله ساختار نوایی آن زبان تعیین می‌شود. وی با انجام یک آزمایش الکتروپالاتوگرافی (کام‌نگاری) مشاهده نمود که بیشترین میزان تماس بین زبان و کام، زمانی ایجاد می‌شود که همخوان در جایگاه آغازین حوزه‌های نوایی قرار دارد و علاوه بر آن افزایش تقویت تولیدی جنبهٔ تجمعی^{۶۷} نیز دارد یعنی هر چه جایگاه آغازین در ساختار نوایی بالاتر باشد، آن جایگاه قوی‌تر و در نتیجه تولید آن واحد آوایی در آن جایگاه با تماس زبانی-کامی بیشتر همراه خواهد بود.

مرزهای نوایی بر فرایندهای هم‌تولیدی نیز تأثیر می‌گذارند. هم‌تولیدی معمولاً بین واحدهای مجاور روی می‌دهد و مرزها واحدهای آوایی را از یکدیگر جدا می‌کنند. بنابراین هم‌تولیدی در درون کلمات بیشتر روی می‌دهد تا در بین و مرز کلمات. چو^{۶۸} (۲۰۰۲) با استفاده از یک آزمایش EMA^{۶۹} نشان داد که مرزهای نوایی بزرگتر بیشتر مانع هم‌تولیدی می‌شوند تا مرزهای کوچکتر زیرا در آغاز مرزهای نوایی کشش و استحکام آوایی بیشتر است.

مقادیر آوایی واج‌ها برحسب نوع بافت نوایی و نیز جایگاه واحد زبانی در هر یک از بافت‌های نوایی تغییر می‌کند و می‌توان آن را در زبان‌های مختلف مشاهده نمود. مثلاً در زبان کره‌ای بخش آغازین واکه /a/ هنگامی که پس از همخوان سایشی دمیده /s^h/ تولید شود، نفسی‌تر از زمانی است که پس از سایشی سخت /s^{*۷۱}/ تولید شود (Yoon, 2007, 2). میزان نفسی‌شدن واکه اولاً به جایگاه همخوان سایشی، در میان یا آغاز کلمه بستگی دارد. ثانیاً میزان نفسی‌شدن واکه برحسب این که همخوان سایشی در آغاز هریک از گروه‌های نوایی کلمه واجی، گروه تکیه‌ای و یا گروه آهنگ قرار دارد،

^{۶۷} . cumulative

^{۶۸} . Cho, T.

^{۶۹} . Electromagnetic Articulograph

^{۷۰} . tense

^{۷۱} . همخوان‌های سخت در زبان کره‌ای با علامت ستاره نشان داده می‌شوند (Yoon 2007).

مقداری متفاوت دارد و سوم این که مقدار دیرش سایش هر یک از سایشی‌های /s^h/ و /s^{*}/ نیز در هر یک از این حوزه‌های نوایی متفاوت است (Yoon, 2007, 2). بنابراین برای آن که یک سیستم TTS، طبیعی و قابل فهم باشد، لازم است که همه این تنوعات واجگونه‌ای که حساس به بافت نوایی هستند، در برنامه ادغام شود.

بسیاری از تغییرات آوایی حساس به مرز کلمات هستند. در زبان انگلیسی، سایشی /z/ قبل از همخوان انسدادی /t/ در بافتی مثل *has to go* بیواک می‌شود. بیواک شدن این همخوان حتی اگر قبل از یک همخوان بیواک قرار نگیرد، با درجات مختلف در بافتهای مختلف روی می‌دهد و میزان بیواکی آن برحسب این که در مرز هجا، کلمه یا جمله قرار گیرد، متفاوت است (Smith, 1997). بنابراین برای وارد کردن این تنوعات واجگونه‌ای در یک سیستم TTS، باید واحدهای آوایی بزرگتر از کلمه مورد بررسی قرار گیرند و کل بافت گفتار برحسب ساختار نوایی مدلسازی شود و نه منحصرأبافت مجاور هر واحد آوایی.

کیتینگ (۲۰۰۳) افزایش کشش آواها را متأثر از جایگاه آنها در هجا، کلمه یا گروه^{۷۲} می‌داند به نحوی که همخوانها در آغاز یا پایان کلمه نسبت به میان کلمه کشیده‌تر هستند. بنابراین هجاهای نوایی در آغاز یا پایان کلمه قرار داده شدند. کیتینگ و همکاران^{۷۳} (۱۹۹۸: ۱) به بررسی میزان قوت همخوانهای /t, n/ در زبانهای انگلیسی، فرانسه، کره‌ای و تایوانی در آغاز و سازه‌های نوایی هجا، کلمه واجی، گروه تکیه‌ای، گروه آهنگ و پاره گفتار آوایی پرداختند و پارامترهای صوت شناختی، دیرش و کام‌نگاری میزان تماس را بررسی کردند. آنها دریافتند که جایگاه نوایی بر میزان قوت^{۷۴} همخوانها تأثیر دارد به نحوی که با گسترش حوزه سازه نوایی از هجا به پاره گفتار آوایی، میزان قوت همخوان افزایش می‌یابد یعنی میزان تماس افزایش یافته و دیرش همخوان، طولانی‌تر می‌شود.

کازلا و همکاران (۲۰۰۷: ۳۲۰-۳۰۱) تأثیر ساختار نوایی و همگونی واکرفتگی را بر پارامترهای دیرش سایش و میزان ارتعاش تار آواها در همخوانهای سایشی /f, v, z/ در زبان آلمانی بررسی کردند. آنها برای این منظور همخوانهای سایشی را در آغاز سه سازه کلمه واجی، گروه واجی و گروه آهنگ مورد بررسی قرار دادند. یافته‌های آنها نشان داد که دیرش سایش و همگونی واکرفتگی تحت تأثیر ساختار نوایی است بدین صورت که در سازه‌های نوایی بالاتر، دیرش سایش بیشتر (تقویت آغازین) و همگونی (قاعده سندهی) نسبت به سازه‌های پایینتر کمتر است.

⁷² . phrase

⁷³ . Keating, P. A. et al.

⁷⁴ . fortition

بررسی تأثیر سازه‌های نوایی بر همگونی واکرفتگی در همخوانهای سایشی زبان آلمانی از سوی کازلا و همکاران (۲۰۰۷: ۳۱۶) نشان داد که شباهت فرایند همگونی در همه سازه‌های نوایی مورد بررسی (کلمه واجی، گروه واجی، گروه آهنگ) یکسان است (کاهش ارتعاش تار آواها)، اما درجه همگونی در سازه‌ها متفاوت است بدین صورت که در سازه‌های کوچکتر همگونی بیشتر نسبت به سازه‌های بزرگتر روی داده است. کازلا و همکاران (همان: ۳۱۷) معتقدند بدین ترتیب حیطه واجشناسی نوایی که اکنون تنها بر وقوع یا عدم وقوع فرایندهای واجی (تأثیرهای مقوله‌ای) در سازه‌های نوایی تأکید دارد، به سوی تفاوت‌های جزئی‌تر زیرواجی^{۷۵}، تأثیرات مدرج ساختار نوایی گسترش پیدا می‌کند.

در زبان فارسی نیز پژوهش‌های در زمینه تأثیر بافت نوایی بر مشخصه‌های آکوستیکی آواها صورت گرفته است. محمودزاده (۱۳۸۸) تأثیر عامل ساختار نوایی گفتار در پدیده سایشی شدگی همخوان‌های انسایشی^{۷۶} فارسی در چهار سازه نوایی هجا، کلمه واجی، گروه واجی و گروه آهنگ را در گفتار پیوسته بررسی نمود. نتایج پژوهش وی نشان داد که در گفتار پیوسته حذف ناحیه سکوت انسایشی در همه سازه‌های نوایی روی داده است و با افزایش حوزه سازه نوایی، از بسامد حذف دیرش سکوت کاسته شده است.

۳-۱-۲. موضع تکیه‌بر

ریتولد و همکاران^{۷۷} (۲۰۰۴: ۳۷۰) رابطه بین ساختار نوایی با دیرش واکه را در زبان هلندی مورد مطالعه قرار دادند. نتایج پژوهش آنها نشان می‌دهد که دیرش واکه‌ها در جایگاه تکیه‌بر بلندتر از سایر جایگاه‌هاست. علاوه بر آن دیرش واکه‌ها در هجای آغازین کلمه بلندتر از سایر جایگاه‌های کلمه است.

نتیجه پژوهشهای دیل و کاسلمن^{۷۸} (۱۹۹۶: ۱۹۳) نیز نشان داد که تکیه تأثیر معناداری بر پارامتر دیرش سکوت، سایش و زمان افزایش دامنه در همخوان‌های سایشی^{۷۹} و انسایشی دارد به طوری که سبب افزایش مقدار این سه پارامتر می‌شود. تأثیر تکیه تقابلی^{۸۰} بر پارامترهای دیرش سکوت و سایش بدین صورت است که تمایز سایشی-انسایشی در هر دو شرایط تکیه و بی‌تکیه همچنان حفظ می‌شود. آنها نتیجه می‌گیرند که اطلاعات زنجیری و زبرزنجیری به یکدیگر وابسته‌اند و اگرچه تمایز انسایشی-

⁷⁵ . subphonemic

⁷⁶ . affricates

⁷⁷ . Rietveld, T. et al.

⁷⁸ . Diehl, L. R. Castleman, A. W.

⁷⁹ . fricatives

⁸⁰ . contrastive stress

سایشی با تغییر جایگاه تکیه تقابلی همچنان باقی می ماند، اما تکیه تأثیر معناداری بر همه پارامترها دارد و سبب افزایش کمی این پارامترها می شود. در واقع تکیه به لحاظ آوایی سبب تقویت^{۸۱} تمایز واحدهای آوایی می شود.

لاوو^{۸۲} (۲۰۰۱: ۱۰) معتقد است قواعد واجی، همخوانها را در بعضی جایگاهها مورد هدف قرار می دهند. وی قوت همخوان را تابعی از جایگاه نوایی آن توصیف می کند. منظور وی از جایگاه نوایی، جایگاه همخوان در کلمه، آغاز یا میان آن و نیز جایگاه همخوان نسبت به جایگاه تکیه است. لاووا (همان: ۱۱) معتقد است که جایگاه آغاز کلمه قویتر از جایگاه میانی آن است و بنابراین انتظار می رود پیاده سازی مشخصه های همخوانی، بیشتر در آغاز کلمه روی دهد و در نتیجه تضعیف کمتری در این جایگاه نسبت به میان کلمه دیده شود. در مورد جایگاه تکیه، وی پیش بینی می کند که همخوانها در جایگاه تکیه قویترند و بنابراین پیاده سازی مشخصه های همخوانی، بیشتر در جایگاه تکیه بر، نسبت به جایگاه بدون تکیه روی می دهد. با ترکیب این دو عامل قوت همخوانی، می توان گفت که همخوانها در جایگاه تکیه بر آغاز کلمه قویتر و در جایگاه بی تکیه میانی، ضعیفترند.

مهمترین متغیری که بر دیرش تأثیر دارد، تکیه است. جایگاه همخوان در کلمه نیز تأثیر معناداری بر دیرش زمانی آن دارد، اگرچه میزان این تأثیر به اندازه تکیه نیست (لاووا، ۲۰۰۱: ۱۵۹). لاووا (همان: ۱۵۷) چنین نتیجه گیری می کند که تکیه به لحاظ آوایی تأثیر بیشتری بر قوت همخوانی دارد.

انفجار رهش بست نیز در همخوانهای انسدادی، تحت تأثیر متغیرهای بافت و تکیه قرار دارد. نتایج بررسی لاووا (همان: ۱۳۲) نشان می دهد که در زبان انگلیسی، بیشترین میزان حذف انفجار رهش بست، در جایگاه بی تکیه میانی کلمه روی می دهد.

۲-۳. پیاده سازی مشخصه های نوایی

پیاده سازی تنوعات آوایی حساس به بافت نوایی در سیستم های TTS همانند پیاده سازی فرایند هم تولیدی است. هم تولیدی، تغییرات پیوسته واحدهای آوایی در زنجیره گفتار است و به واحدهای مجاور یک واحد آوایی بستگی دارد و نه منحصرأ به بافت نوایی آن. مثلاً تولید /k/ در زبان فارسی در کلمات «کیسه، کر، کار» [cise]، [car]، [kar] تاحدی با یکدیگر فرق دارد زیرا تولید /k/ در

⁸¹ . enhancement

⁸² . Lavoie, L. M.

«کیسه» [cise] و «کر» [car] در جایگاه قبل از واکه‌های پیشین افراشته یا افتاده، تاحدی پیش‌کامی شده است.

پیاده‌سازی هم‌تولیدی در یک سیستم TTS اساساً به نوع روش سنتز بستگی دارد. روش‌های مختلف سنتز گفتار را می‌توان به دو دسته کلی «دانش‌بنیان»^{۸۳} و «ناآگاه‌بنیان»^{۸۴} دسته‌بندی کرد. در یک سیستم دانش‌بنیان، دانش آوایی و آکوستیکی زیادی برای هر فرایند آوایی شامل اطلاعات سازه‌ها، مکان فرکانسی آنها و غیره لازم است اما در یک سیستم «ناآگاه‌بنیان» مثل سیستم‌های هم‌گذاری دایفون (دوواج) یا نیم‌هجا^{۸۵}، دانش کمتری لازم است. در این سیستم‌ها واحد‌ها از بافت‌های مختلف انتخاب می‌شوند و بنابراین تأثیر هم‌تولیدی واج‌های مجاور را با خود به همراه دارند. برای پیاده‌سازی تنوعات پیوسته آوایی حساس به بافت نوایی نیز می‌توان به شیوه‌ای مشابه عمل نمود یعنی می‌توان واحدهای هم‌گذاری را نه تنها بر مبنای بافت زنجیره‌ای بلکه براساس بافت نوایی انتخاب نمود (Yoon 2007, 3).

در سیستم‌های TTS، برای ذخیره فرایندهای آوایی و واجگونه‌ای، به طور سنتی همه تغییرات زنجیره‌ای به غیر از فرایندهای هم‌تولیدی به صورت تغییر مقوله‌ای^{۸۶} مشخصه‌های واجی، به شیوه قواعد واجی SPE با استفاده از قواعد بازنویسی^{۸۷} کدبندی می‌شود مثلاً قواعد تبدیل حرف به آوا و یا قواعد^{۸۸} پساواژگانی در سیستم فوستیوال^{۸۹} (Alan et al. 1999). در سیستم می‌تاک^{۹۰}، با استفاده از شش قاعده بازنویسی، فرایندهای زنجیره‌ای گفتار در مرز کلمات و گروه‌ها کدبندی می‌شود. یکی از این قواعد انسدادی‌های /t, d/ را با همخوان زنجیره‌ای لثوی در مرز کلمات جایگزین می‌کند بدین صورت که اگر /t/ بعد از یک رسای غیرخیشومی و قبل از واکه با تکیه غیراصلی قرار گیرد مثلاً در عبارت sat on a chair در این صورت به صورت زنجیره‌ای تولید می‌شود (Allen et al. 1987, 87). برخی از این قواعد در بخش پساواژگانی سیستم فستیوال پیاده شده‌اند (Alan et al. 2003) اما همان‌طور که پدیدآورندگان نیز پذیرفته‌اند، این بخش به تحقیق و توسعه بیشتری نیاز دارد.

مشکلی که در رابطه با اعمال تغییرات آوایی حساس به بافت نوایی وجود دارد این است که نمی‌توان آنها را به طور کامل و منظم با کمک قواعد بازنویسی تعریف کرد. به عنوان مثال اگرچه می‌توان با

⁸³ . knowledge-based

⁸⁴ . ignorance-based

⁸⁵ . demissyllable

⁸⁶ . categorical

⁸⁷ . rewrite rules

⁸⁸ . postlexical rules

⁸⁹ . Festival

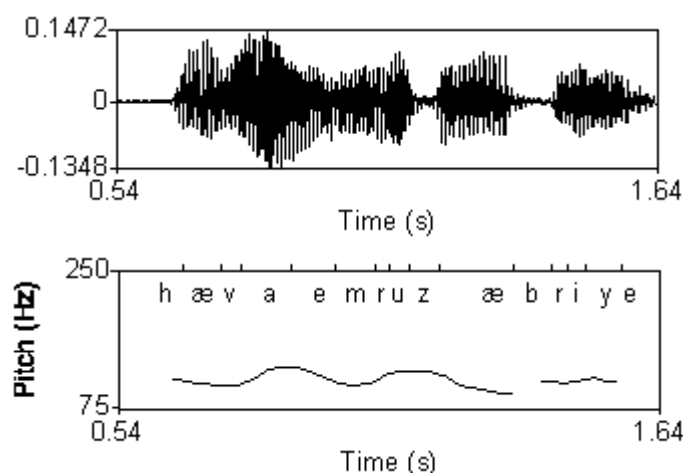
⁹⁰ . MITalk

کمک قواعد پساواژگانی تغییرات مقوله‌ای تناوب زنجی- انسدادی را در sat on the chair و at all و یا تناوب دمیده- نادمیده بین [k^h] دمیده در the car و [k] نادمیده در discover را تعریف کرد اما نمی‌توان قواعد بازنویسی را برای اعمال تغییرات پیوسته دمیدگی واجگونه‌های [k] در گروه‌های نوایی this car, this career, this task به کار برد (Yoon, 2007, 3).

طراحی سیستم هم‌گذاری بدون توجه به تأثیر بافت‌های نوایی مختلف بر تغییرات پیوسته واجگونه‌ای سبب می‌شود تا واحدهای بازسازی ثابتی در همه جایگاه‌های نوایی به کار رود. به عنوان مثال، در یک سیستم سنتز دایفون، هر دایفون به طور مکرر و بدون توجه به جایگاه نوایی آن در جمله، هر گاه که نیاز باشد به کار می‌رود. از سوی دیگر اگر به طبیعت تغییرات آوایی توجه کنیم، ناکارآمدی سیستم‌های مبتنی بر قواعد بازنویسی روشن‌تر می‌شود. به عنوان مثال می‌توان در زبان کره‌ای با استفاده از یک قاعده بازنویسی تولید نفسی واژه /a/ را در جایگاه پس از همخوان /s^h/ در برنامه اعمال کرد اما همان‌طور که گفتیم بسیاری از این تغییرات واجگونه‌ای در واقع پیوسته هستند و مقادیر آنها برحسب بافت نوایی تغییر می‌کند. بهترین راه مدلسازی این تغییرات پیوسته در برنامه این است که واحدهای هم‌گذاری به طور مستقیم از بافت نوایی استخراج و در برنامه وارد شوند.

۳-۳. الگوی نوایی زبان فارسی

آهنگ بخشی از ساختار نوایی زبان فارسی است. برای بررسی آهنگ گفتار، تغییرات زیروبمی که به وسیله منحنی زیروبمی بازنمایی می‌شود، مورد ملاحظه قرار می‌گیرد. منحنی زیروبمی نوسانات زیروبمی در طی زمان را نشان می‌دهد. در مثال زیر (شکل ۳-۲) محور عمودی زیروبمی را برحسب هرتز (Hz) و محور افقی زمان را برحسب ثانیه نشان می‌دهد.



شکل ۳-۲: نمودار آهنگ پاره گفتار « هوا امروز ابری »

تفسیر منحنی‌های آهنگ در چارچوب نظریه خودواحد- وزنی آهنگ (Ladd, 1996) و نظام آهنگ فارسی (Sadat Tehrani, 2007) صورت می‌گیرد. بر اساس این دیدگاه که با آثار افرادی نظیر لیبرمن^{۹۱} (۱۹۷۵)، بروس^{۹۲} (۱۹۷۷)، و پیرهامبرت (۱۹۸۰) آغاز شد، ساختار نواختی از رویدادهای نواختی معنادار به لحاظ واجی مثل تکیه زیروبمی و نواخت‌های مرزی تشکیل یافته است. این رویدادها که شامل زیروبمی خیزان (H) و افتان (L) است در نقاطی در زنجیره گفتار روی می‌دهند. نواخت‌ها به لحاظ آوایی بر محور زمان ترادف دارند و مقادیر آنها بر روی محور عمودی یا فرکانس قابل اندازه‌گیری است.

۳-۱- سیستم ثبت آهنگ

سیستم ثبت آهنگ ToBI (نواخت و فاصله نما) است، سیستمی که در ابتدا برای ثبت آهنگ در زبان انگلیسی آمریکایی بر اساس کارهای پیرهامبرت (۱۹۸۰)، لد (۱۹۸۳) بکمن^{۹۳} و پیرهامبرت (۱۹۸۶)؛ پرایس^{۹۴} و همکاران (۱۹۹۱) و وایتمن^{۹۵} و همکاران (۱۹۹۲) پیشنهاد شد. سیستم ToBI در جلساتی با حضور متخصصینی از علوم مهندسی، روانشناسی، کامپیوتر، و آواشناسی شکل گرفت و هدف آن دستیابی به نظامی مشترک برای نمایش جنبه‌های مختلف نوای گفتار بود. در ToBI اولیه (انگلیسی آمریکایی)، چهار لایه^{۹۶} منحنی فرکانس پایه ی (F0) یک پاره گفتار را همراهی می‌کنند. «لایه املایی»^{۹۷} شامل املای کلمات به کاررفته در پاره گفتار می‌باشد. «لایه فاصله‌نما»^{۹۸} میزان درنگ بین کلمات را نشان می‌دهد و می‌تواند عددی بین ۰ (کمترین درنگ) تا ۴ (بیشترین درنگ) را دارا باشد. «لایه نواخت»^{۹۹} دربرگیرنده نواخت‌های زیروبمی (H و L) است که برای نمایش نواخت‌های کناری و تکیه‌های زیروبمی از آنها استفاده می‌شود. تکیه زیروبمی (مثلاً H*) در ارتباط با هجای تکیه‌بر می‌باشد. نواخت کناری شامل دو نوع نواخت است: نواخت گروه^{۱۰۰} (مثلاً L- که در انتهای گروه میانی^{۱۰۱} واقع می‌شود و عدد لایه فاصله نما برای آن ۳ یا ۴ است) و نواخت مرزنا^{۱۰۲} (مثلاً H% که

^{۹۱} . Liberman, M.

^{۹۲} . Bruce, G.

^{۹۳} . Beckman, M. E.

^{۹۴} . Price, P.

^{۹۵} . Wightman, C.

^{۹۶} . tier

^{۹۷} . orthographic tier

^{۹۸} . break index tier

^{۹۹} . tone tier

^{۱۰۰} . phrase tone

^{۱۰۱} . intermediate phrase

^{۱۰۲} . boundary tone

در انتهای گروه آهنگ^{۱۰۳} با فاصله نمای ۴ قرار می گیرد). لایه متفرقه^{۱۰۴} برای نمایش صداهای غیرزبانشناختی یا اضافه کردن توضیحات به کار می رود مثلاً «خنده» یا «سرفه».

۳-۲. تکیه واژگانی

تکیه واژگانی در کارهای نسبتاً زیادی مورد بحث قرار گرفته است، از آن جمله می توان از لازار (۱۹۹۲)، سامعی (۱۳۷۴)، ماهوتیان (۱۹۹۷)، وحیدیان کامیار (۱۳۷۹)، کهنمویی پور (۲۰۰۳)، و پرمون (۱۳۸۵) نام برد. تاکنون تعریف آواشناختی دقیقی از تکیه در فارسی ارائه نشده است و منظور از هجای تکیه بر عموماً هجایی است که به لحاظ ادراکی برجسته ترین باشد. اهمیت این مفهوم در تحقیق حاضر در این است که تکیه زیروبمی در فارسی بر روی هجای تکیه بر می افتد (اسلامی و بی جن خان، ۱۳۸۱). قوانین اصلی تکیه در فارسی را می توان به صورت زیر جمع بندی کرد.^{۱۰۵} اسامی، صفات، و بیشتر قیود دارای تکیه پایانی هستند. دو مثال از هر مورد در جدول ۳-۱ آمده است (تکیه با علامت بر روی واکه ی هجای تکیه بر نشان داده شده است).

جدول ۳-۱: تکیه واژگانی در اسمها، صفتها و قیدها

اسم	صفت	قید
kabaʔb	kutaʔh	yævaʔʃ
ʃadí	basavaʔd	besaxtí

تکیه افعال بر روی آخرین هجای سازه اصلی است. در مثال xar-íd-am تکیه روی id- است که هجای پایانی ستاک ماضی (xarid) می باشد و شناسه ی æm- نقشی در تکیه ندارد.

بعضی از پیشوندهای فعلی مانند na-/ne- منفی ساز، be-/bo- التزامی ساز، و mi- استمراری ساز تکیه را به خود جذب می کنند. مثال na'-xar-íd-am شامل فعلی با پیشوند منفی ساز است.

برخی کلمات، به لحاظ تکواژشناختی ساده، نیز وجود دارند (بیشتر از ریشه عربی) که دارای تکیه آغازی هستند. به عنوان مثال می توان از واژه های a'mma و ba'le نام برد.

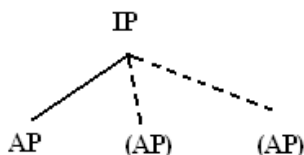
¹⁰³ Intonational Phrase

¹⁰⁴ miscellaneous tier

^{۱۰۵} این جمع بندی شامل تمامی قوانین تکیه فارسی نمی باشد.

۳-۳-۳. تکیه نوایی

ساختار نوایی زبان فارسی شامل دو سازه نوایی است، که عبارتند از گروه تکیه‌ای (AP) و گروه آهنگ (IP). گروه تکیه‌ای کوچکترین واحد نوایی است و یک یا چند گروه تکیه‌ای یک گروه آهنگ را می‌سازد (Sadat Tehrani, 2007, 41). در شکل ۳-۳ دو سازه ساختار نوایی نشان داده شده است. خط چین و پرانتز نشان‌دهنده اختیاری بودن سازه مورد نظر است.



شکل ۳-۳: گروه آهنگ و اجزای آن

دامنه گروه تکیه‌ای به لحاظ نحوی هسته و وابسته‌های آن یعنی XP است که در آن $X=N, V, P$ Adj & Adv است اما دامنه گروه آهنگ خود جمله یعنی inflectional phrase یا IP است.

۳-۳-۴. گروه تکیه‌ای (AP)

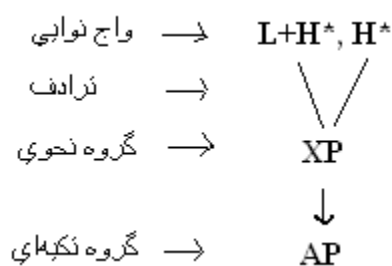
کوچکترین واحد نوایی فارسی گروه تکیه‌ای است. گروه تکیه‌ای خود از واحدهای کوچکتری به نام نواخت تشکیل یافته که در واقع واج‌های نوایی هستند. این نواخت‌ها یا واج‌های نوایی عبارتند از نواخت افتان^{۱۰۶} (L) و نواخت خیزان^{۱۰۷} (H) (Sadat Tehrani, 2007, 41). این دو نواخت تشکیل گروه تکیه‌ای را می‌دهند

نواخت افتان یعنی نرخ ارتعاش کمتر تارهای صوتی و نواخت خیزان یعنی نرخ بالای ارتعاش تارهای صوتی (بدون احتساب سایر پارامترهای آکوستیکی). مثلاً در کلمه دوهجایی «تهران» با صورت آوایی [tehrán]، تکیه واژگانی و نیز زیروبمی (در جمله) بر روی هجای دوم قرار دارد، بنابراین نرخ ارتعاش تارهای صوتی در هجای دوم بیشتر از هجای اول است و بنابراین نواخت هجای اول افتان (L) و نواخت هجای دوم خیزان (H) است. در کلمات تک‌هجایی مثل «کار» [kár] از آنجا که همان هجا تکیه‌بر است بنابراین نواخت کلمات تک‌هجایی فقط شامل نواخت خیزان (H) است.

¹⁰⁶ . low

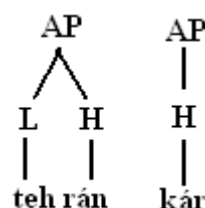
¹⁰⁷ . high

همان‌طور که گفتیم گروه تکیه‌ای برابر است با یک گروه نحوی شامل گروه اسمی، فعلی، حرف اضافه‌ای، صفتی یا قیدی یا به عبارتی هسته با وابسته‌های آن. یک گروه تکیه‌ای با ترادف^{۱۰۸} هر یک از الگوهای واجی نواختی بر یک گروه نحوی شکل می‌گیرد.



شکل ۳-۴: ترادف نواخت بر واحد نحوی

در زبان فارسی نواخت‌ها به دو صورت تشکیل گروه تکیه‌ای می‌دهند و در واقع بازنمایی واجی گروه تکیه به دو صورت است: یکی نواخت بی‌نشان $L+H^*$ بر روی کلمه‌ها، گروه‌ها و عبارات ندایی چندهجایی که تکیه واژگانی بر روی هجای آخر آنها قرار دارد مثل اسم‌ها و صفت‌ها و دیگری H^* بر روی کلمات محتوایی یک‌هجایی و نیز کلمات دارای تکیه آغازین و اغلب افعال (Sadat Tehrani, 2007, 42). نواخت‌های $L+H^*$ و H^* که تشکیل یک گروه تکیه‌ای را می‌دهند، تکیه زیروبمی^{۱۰۹} نیز نامیده می‌شوند.



شکل ۳-۵: گروه تکیه‌ای LH^* و H

قرار گرفتن علامت ستاره (*) بر روی نواخت H نشانه آن است که این نواخت، نواخت اصلی است زیرا تکیه واژگانی در فارسی اغلب بر روی هجای دوم قرار دارد (Sadat Tehrani, 2007, 42).

برای روشن شدن مطلب، پاره گفتار «هوا امروز ابریه» را در نظر بگیرید که ساختار نوایی آن در مثال ۱ نشان داده شده و نمودار آهنگ آن در شکل ۳-۶ آمده است.

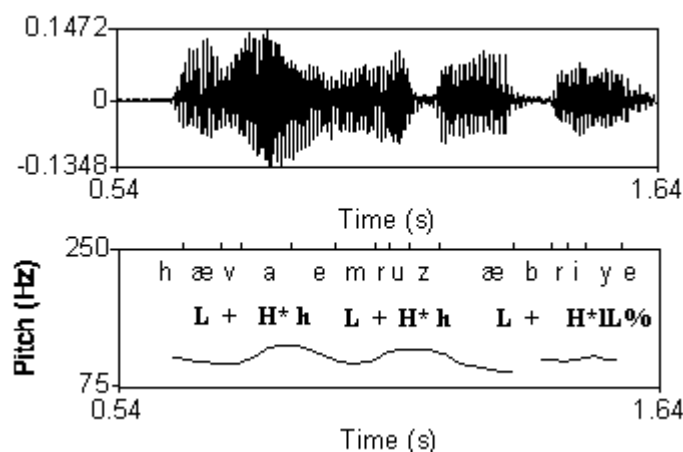
¹⁰⁸ . alignment

¹⁰⁹ . pitch accent

1. hava emruz abri-ye

|| AP1 || AP2 || AP3 ||

|| IP ||



شکل ۳-۶: نمودار آهنگ پاره گفتار «هوا امروز ابری»

همان‌طور که در شکل نشان داده شده، سه گروه تکیه‌ای دارای نواخت واجی $L + H^*$ هستند.

۳-۳-۵. ساختار AP

یک AP معمولاً از یک کلمه محتوایی به همراه واژه‌بستها تشکیل یافته است. مثلاً گروه «کتابم» که از کلمه محتوایی *ketab* به همراه واژه‌بست *-æm* ترکیب یافته است، یک گروه تکیه‌ای یا AP را می‌سازد. کلمات مرکب مثل «کتابخونه» و اسامی مکرر^{۱۱۰} مثل «صندلی مندلی» نیز یک کلمه محسوب می‌شوند و بنابراین یک AP را می‌سازند.

اما گروه تکیه‌ای می‌تواند بیش از یک کلمه را شامل شود. عوامل متعددی مانند سرعت گفتار^{۱۱۱}، تأکید^{۱۱۲}، کشش^{۱۱۳}، موصول^{۱۱۴} و نیز ساختار اطلاعی^{۱۱۵} بر شکل AP تأثیر می‌گذارند و سبب تغییر نحوه شکل‌گیری یک گروه تکیه‌ای می‌شوند (Sadat Tehrani, 2007, 46).

¹¹⁰ . reduplicative nouns

¹¹¹ . speech rate

¹¹² . focus

¹¹³ . length

¹¹⁴ . subordination

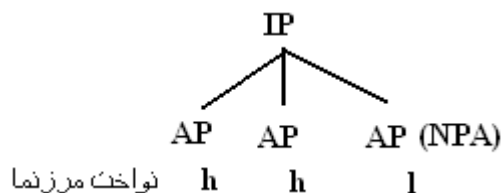
¹¹⁵ . information structure

اگر یک گروه تکیه‌ای AP با واژه‌بست در انتها تولید شود. در صورتی که نواخت مرزنامی واژه‌بست خیزان (h) باشد، معمولاً تغییرات زیروبمی نواخت مرزنامی واژه‌بست (h) بالاتر از نواخت H* تکیه زیروبمی ماقبل واژه‌بست است (Sadat Tehrani, 2007, 47). به طور مثال در جمله ketabi xarid، نواخت H* بر روی هجای دوم ta قرار دارد و نواخت مرزنامی h بر روی واژه‌بست i. بررسی منحنی زیروبمی نشان می‌دهد که نواخت مرزنامی h بالاتر از نواخت واجی H* است.

۶-۳-۳. نواخت مرزناما^{۱۱۶}

بخشی از گروه تکیه که بین تکیه زیروبمی و پایان AP قرار دارد نیز همراه با الگوی نواختی است که نواخت مرزناما نامیده می‌شود. یعنی لبه راست هر AP به وسیله یک نواخت مرزناما مشخص می‌شود که می‌تواند l(low) یا h(high) باشد. در فاصله بین تکیه زیروبمی تا پایان AP گاهی هیچ هجایی وجود ندارد و هجای تکیه‌بر AP همان هجای پایانی است. در این حالت نواخت مرزناما نیز بر روی هجای تکیه‌بر قرار می‌گیرد اما گاهی بین تکیه زیروبمی و پایان AP می‌تواند چندین هجا وجود داشته باشد، در این حالت نواخت مرزناما شامل همه این هجاها تا پایان AP می‌شود (Sadat Tehrani, 2007, 42).

یکی از دلایل وجود نواخت مرزناما، تفاوت بین تکیه زیروبمی هسته (NPA^{۱۱۷}) و غیرهسته است. در بسیاری از جملات ساده بی‌نشان، تکیه زیروبمی هسته بر روی آخرین AP در درون یک گروه آهنگ قرار دارد و نواخت مرزنامی آن افتان یا l است. نواخت مرزنامی سایر گروه‌های تکیه‌ای که هسته نیستند، خیزان یا h است. الگوی نواخت تکیه زیروبمی با حروف بزرگ L یا H و الگوی نواخت مرزناما با حروف کوچک l یا h نشان داده می‌شود. تکیه زیروبمی هسته به لحاظ شنیداری برجسته‌ترین تکیه است و سایر کلمات و هجاها بعد از تکیه زیروبمی هسته، تکیه خود را از دست می‌دهند (سادات تهرانی ۲۰۰۷: ۴۳). شکل زیر جایگاه تکیه زیروبمی هسته را نشان می‌دهد که بر روی آخرین AP قرار دارد و بنابراین نواخت مرزنامی آن l است.

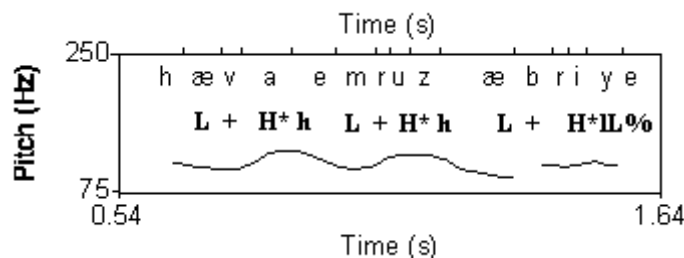


شکل ۷-۳: نواخت مرزناما

¹¹⁶ . boundary tone

¹¹⁷ . nuclear pitch accent

در شکل ۳-۸، هریک از نواخت‌های مرزنامی l (گروه زیروبمی هسته) در گروه تکیه‌ای /abri-ye/ و h در دو AP /hava/ و /emruz/ (تکیه زیروبمی غیرهسته) مشخص شده‌اند.



شکل ۳-۸: نواخت مرزنامی هسته و غیرهسته

هرچ‌برگ^{۱۱۸} (۲۰۰۲: ۳۴) تکیه زیروبمی هسته را «به لحاظ شنیداری برجسته‌ترین تکیه در درون یک گروه نوایی» تعریف می‌کند.

به دلیل کاهش نرخ فرکانس پایه F0، تکیه زیروبمی گروه‌های تکیه‌ای از ابتدا به انتهای پاره گفتار به تدریج کاهش پیدا می‌کند (Gussenhoven, 2004). به همین دلیل تغییرات منحنی زیروبمی آخرین گروه تکیه‌ای که حامل تکیه زیروبمی هسته نیز هست، نسبت به سایر گروه‌های تکیه‌ای پایین‌تر است (Sadat Tehrani, 2007, 44).

۳-۳-۷. تکیه زیروبمی هسته (NPA)

گروه تکیه‌ای هسته که پس از آن تکیه سایر گروه‌ها تا انتهای جمله از بین می‌رود، معمولاً آخرین AP در جمله است با نواخت مرزنامی l. سایر APها که تکیه زیروبمی غیرهسته دارند، همراه با نواخت مرزنامی h هستند. به عنوان مثال در جمله زیر کلمه‌ای که زیر آن خط کشیده شده، همراه با NPA است.

1. mina milanam mimune.

در جملات اسنادی مثبت^{۱۱۹}، تکیه زیروبمی هسته بر روی متمم، عنصر بین فاعل و فعل ربطی قرار می‌گیرد (Sadat Tehrani, 2007, 8-9). در مثال زیر بر روی صفت «عالی» قرار گرفته است.

2. hava ?ali mife.

¹¹⁸ . Hirschberg, J

¹¹⁹ . affirmative copular

در این جملات اگر متمم اول جمله تولید شود و یا جمله فاقد فاعل باشد، در هر دو صورت تکیه زیروبمی هسته بر روی متمم قرار می گیرد:

3. ʔali miʃe hava.

4. ʔali miʃe.

اگر توصیفگر^{۱۲۰} دیگری به متمم اضافه شود، آن توصیفگر تکیه NPA را جذب می کند:

5. hava ʔali-ye ʔali miʃe.

مکان NPA در جملات (s)ov خبری^{۱۲۱} به خاص^{۱۲۲} بودن مفعول بستگی دارد. اگر مفعول خاص باشد، NPA بر روی فعل قرار می گیرد و اگر مفعول خاص نباشد بر روی مفعول قرار می گیرد:

6. ye name-ye mohem umad-e.

7. name mohemm-e' umad.

8. mina film did-e bud.

9. mina film-a ro did-e bud.

در جملات منفی که فعل با نشانگر منفی na منفی شده است، NPA بر روی فعل منفی قرار می گیرد:

10. man be una madjun na'budam.

در جملات امری نیز مانند خبری، اگر مفعول خاص باشد، NPA بر روی فعل قرار می گیرد و اگر مفعول خاص نباشد بر روی مفعول قرار می گیرد:

11. xune bexar!

12. pandʒer' ro beband.

تکیه زیروبمی هسته نمی تواند در مکانی بعد از فعل قرار بگیرد مگر در مورد گروه های قیدی که جابه جا شده اند:

13. mina raft-e xune.

¹²⁰ . post-modifier

¹²¹ . declarative

¹²² . specificity

در جملات سؤالی بله/خیر، با یا بدون ادات سؤالی «آیا»، تکیه زیروبمی هسته بر روی آخرین گروه تکیه‌ای قرار می‌گیرد با این تفاوت که منحنی زیروبمی در سطح بالاتری نسبت به جملات خبری تولید می‌شود. در جملات سؤالی نواخت جمله $H\%$ است:

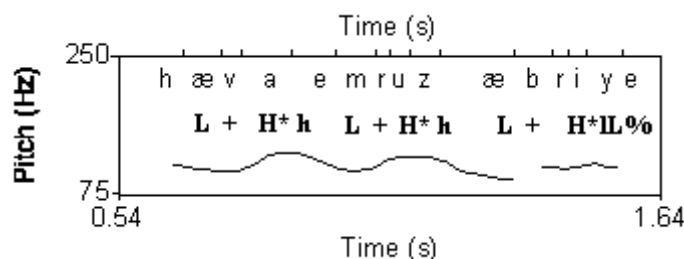
14. *ʃagerda miza ro avordan?*

در جملات پرسشی که با کلمه پرسش ساخته می‌شوند، کلمه پرسشی NPA را جذب می‌کند:

15. *batʃtʃe-ha az koɖʒa ketab xaridan?*

۳-۳-۸. گروه آهنگ

یک یا چند گروه تکیه‌ای (AP) تحت تسلط مستقیم یک گروه آهنگ قرار دارند. گروه آهنگ در جملات ساده برابر با یک پاره گفتار است. به لحاظ واجی، مرز نوایی گروه آهنگ به وسیله نواخت‌های مرزنامی $L\%$ یا $H\%$ بر روی هجای آخر یا بخشی از هجای آخر تعیین می‌شود و به لحاظ آوایی آغاز یک IP معمولاً همراه با آغاز مجدد منحنی زیروبمی و پایان آن معمولاً همراه با مکث یا کشش واکه‌ای است. معمولاً یک تکیه زیروبمی هسته (NPA) در هر گروه آهنگ وجود دارد (Sadat Tehrani, 2007, 50).



شکل ۳-۹: نواخت گروه آهنگ، در این مثال $L\%$

نواخت مرزنامی $L\%$ در انتهای جملات خبری (SOV)، سؤالات بله/خیر، سؤالات استفهامی، جملات امری و ندایی قرار می‌گیرد. نواخت $H\%$ در انتهای جملات با سؤالات بله/خیر، سؤالات کوتاه، سؤالات انعکاسی، ساختارهای همپایه و جملات دارای عبارت موصولی به کار می‌رود (Sadat Tehrani, 2007, 50).

۳-۴. هجای فارسی و الگوی نواختی آن

واجگان زبان فارسی از ۶ واکه و ۲۳ همخوان تشکیل شده است و بدون در نظر گرفتن محدودیت‌های همنشینی واج‌ها در درون هجا، شمار بالقوه هجاها در زبان فارسی ۷۶۳۱۴ خواهد بود. ولی از این تعداد، کمتر از ۶۰۰۰ هجا فعلیت می‌یابند و بقیه ساخت‌های هجایی غیرممکن و یا دارای خلاء هستند.

هجا در زبان فارسی عبارت از یک رشته آوایی پیوسته است که از یک واکه و یک تا سه همخوان تشکیل می‌شود (ثمره ۱۳۷۸). منظور از "رشته آوایی پیوسته" آن است که اجزای سازنده هجا طی یک فرآیند تولید بدون مکث تولید می‌شوند. واکه به منزله مرکز یا هسته یا محور هجاست و همخوان در حکم حاشیه یا دامنه آن است.

در فارسی سه نوع هجا وجود دارد که با قرار دادن C و V به ترتیب به جای همخوان و واکه می‌توان این سه نوع را به صورت CV، CVC و CVCC نشان داد؛ مثلاً در سه کلمه «گو، گوش، گوشت»
/gu/, /guʃ/, /guʃt/

۳-۴-۱. هجای CV

هجای CV متشکل از یک همخوان و یک واکه است. فهرست همخوان‌های فارسی برحسب نحوه تولید در زیر ارائه شده است.

جدول ۳-۲: همخوان‌های زبان فارسی

همخوان	نحوه تولید
b, p, t, d, k, g, q, ʔ	انفجاری
f, v, s, z, ʃ, ʒ, tʃ, dʒ, x, h	سایش و انسایشی
m, n, r, l, j	رسا

واکه‌های فارسی شامل ۶ واکه: a, ɑ, e, o, u, i است. ترکیب همخوان‌ها با واکه‌ها شامل، ۱۳۸ = ۶ × ۲۳ هجا خواهد بود.

۳-۴-۲. الگوهای نواختی

همان‌طور که گفتیم مقادیر آکوستیکی هجاها به بافت زنجیری و نیز زبرزنجیری گفتار حساس است و از این رو اگر این ویژگی‌های آوایی در برنامه بازسازی گفتار لحاظ شود، گفتار حاصل طبیعی‌تر خواهد بود.

الگوی نواختی عبارت است جایگاه و برجستگی هجا در ساختار نوایی. عوامل مؤثر بر الگوی نواختی هجا را می‌توان شامل، جایگاه هجا در ساختار نوایی، تکیه واژگانی، نوع جمله، تأکید، سرعت گفتار، ساختار اطلاعی^{۱۲۳} و غیره دانست. علاوه بر عوامل نوایی، عوامل زنجیری گفتار نیز بر مشخصه‌های آکوستیکی هجاها قابل بررسی است. تأثیر هریک از این موارد بر ویژگی‌های آکوستیکی هجا به تفصیل در ادامه ارائه خواهد شد و برای این منظور هجای /bi/ به عنوان مثال به کار می‌رود.

الف) عامل بافت زنجیری

ویژگی‌های همخوان‌ها و واکه‌های مجاور:

مشخصه‌های آکوستیکی هجای /bi/ تحت تأثیر مشخصه‌های آکوستیکی همخوان و واکه ماقبل آن (/Vbi/، /Cbi/) و نیز مشخصه‌های آکوستیکی همخوان مابعد (/biC/) آن قرار دارد (منظور از C یعنی یک همخوان و V یعنی یک واکه). تغییر مشخصه‌های آوایی در این حالت به دنبال فرایندهای زنجیری گفتار مثل هم‌تولیدی روی می‌دهد.

ب) عامل بافت نوایی

۱- جایگاه هجا در کلمه: آغاز، میان یا پایان کلمه

مشخصه‌های آکوستیکی هجای /bi/ در آغاز کلمه /bidar/ با مشخصه‌های آن در میان کلمه /ʔabidar/ و پایان کلمه /ʔabi/ تفاوت بسیار دارد. همان‌طور که در بخش ۱-۱-۳ گفته شد، دو همبسته ساختار نوایی افزایش دیرش و تقویت نوایی سبب تغییر مشخصه‌های آکوستیکی هجا در هر یک از این مواضع نوایی می‌شود.

۲- تکیه واژگانی: هجای تکیه‌بر واژگانی، هجای بی‌تکیه

مشخصه‌های آکوستیکی هجای /bi/ بسته به آن که در جایگاه تکیه‌بر واژگانی باشد، به عنوان مثال در کلمه /ʔabi/ متفاوت است نسبت به وقتی که در جایگاه بی‌تکیه واژگانی مثلاً در کلمه /bida'r/ باشد. همان‌طور که در بخش ۱-۱-۳ گفته شد، تکیه بر مشخصه‌های آکوستیکی مانند دیرش و قوت همخوانی (تماس زبانی-کامی) تأثیر دارد.

۳- تکیه زیروبمی: هسته، غیرهسته

مشخصه‌های آکوستیکی هجا بسته به این که در کلمه دارای تکیه زیروبمی هسته باشد یا نباشد، متفاوت است. مشخصه‌های آکوستیکی هجا بسته به این که با نواخت‌های زیروبمی H یا L، نواخت‌های مرزنامی h یا l و یا نواخت‌های جمله H% یا L% ترادف داشته باشد، متفاوت است.

در یک AP هسته، نواخت L با همخوان ماقبل واکه تکیه‌بر و نواخت H با همخوان بعد از واکه تکیه‌بر ترادف دارد (Sadat Tehrani, 2007, 5).

• تکیه زیروبمی هسته:

در مثال‌هایی زیر گروه تکیه هسته با خط‌کشی زیرین مشخص شده و نواخت زیروبمی و مرزنامی کلمه حامل هجای bi در زیر آن ارائه شده است.

✓ تکیه واژگانی روی هجای 'bi، ترادف با نواخت زیروبمی H*، نواخت مرزنامی l.

1. abi' bud.

L+H*1

✓ تکیه واژگانی روی هجای 'bi و ترادف با نواخت زیروبمی H*، نواخت مرزنامی l، در انتهای جمله و همراه با نواخت جمله L%.

2. hame raftand to labi'.

L+H*1 L%

✓ فاقد تکیه واژگانی، ترادف با نواخت زیروبمی L و نواخت مرزنامی l.

3. bida'r bud.

L+H*1

✓ فاقد تکیه واژگانی و ترادف با نواخت زیروبمی L، نواخت مرزنامی l و نواخت جمله L%.

4. damha raftand to bi'e'.

L+H*1 L%

• تکیه زیرویمی غیرهسته:

✓ تکیه واژگانی روی هجای 'bi و ترادف با نواخت زیرویمی H* و نواخت مرزنامی h.

5. ?abi'-ye rofae'n bud.

L+H*h

✓ فاقد تکیه واژگانی، ترادف با نواخت زیرویمی L و نواخت مرزنامی h.

6. bidar-e bidar bud.

L+H*h

۴- نوع جمله: مشخصه‌های آکوستیکی هجا تحت تأثیر نوع جمله و مشخصه‌های نوایی جملات متفاوت است.

✓ جملات سؤالی

سه تفاوت آوایی بین جملات خبری و سؤالی بله/خیر عبارتند از: بالا بودن میزان زیرویمی در NPA در جملات سؤالی نسبت به خبری، بالا بودن میزان زیرویمی کلی جملات سؤالی بله/خیر نسبت به جملات خبری و در نهایت افزایش کشش پایانی در جملات سؤالی به ویژه افزایش کشش واکه‌های پایانی (Sadat Tehrani, 2007, 79). به نظر می‌رسد بالا بودن کلی منحنی زیرویمی در جملات سؤالی سبب افزایش زیرویمی نواخت L در این جملات نسبت به جملات خبری شود.

7. damha raftand to un bife'-he?

L+H*1 H%

8. hame raftand to labi'-ye?

L+H*1 H%

در جملات سؤالی بله/خیر مانند مثال‌های زیر که هجای تکیه‌بر NPA هجای پایانی IP نیز هست، نواخت مرزنامی l به لحاظ آوایی ظاهر نمی‌شود (Sadat Tehrani, 2007, 72).

9. damha raftand to bife'?

$L+H^* H\%$

10. hame raftand to labi'?

$L+H^* H\%$

✓ جملات هم پایه

در جملات هم پایه، نواخت مرزنامی IP اول $H\%$ است زیرا جمله ناتمام است در نواخت نواخت $H\%$ در جملات هم پایه به جمله معنای ناقصی می دهد به گونه ای که شنونده در انتظار ادامه گفتار است. نواخت مرزنامی آخرین AP که NPA جمله است، نیز h است (Sadat Tehrani, 2007, 102):

اگر بین NPA و حرف ربط هجاهای دیگری باشد، آنها نیز با زیرویمی بالا تولید می شوند زیرا نواخت مرزنامی تکیه زیرویمی هسته به انتهای IP گسترش پیدا می کند (Sadat Tehrani, 2007, 103):

11. resid bita va behef goftam.

$L+H^*h H\%$

گسترش نواخت مرزنامی بالای AP «رسید» به «بیتا» سبب می شود این گروه اسمی هیچ تکیه زیرویمی نداشته باشد (با میزان زیادی افتادگی^{۱۲۴} یا کاهش فرکانس پایه). این حذف تکیه^{۱۲۵} تا پایان IP اول ادامه پیدا می کند (Sadat Tehrani, 2007, 103).

۵- پی چسب

پی چسب به گروه تکیه ای متصل می شود و نواخت مرزنامی آن AP بر روی آن قرار می گیرد. میزان نواخت مرزنامی پی چسب بالاتر از H^* است. در مثال زیر، واژه /i/ یک پی چسب است که هجای /bi/ را می سازد و نواخت مرزنامی h بر روی آن قرار می گیرد.

12. si'bi az deraxt ofta'd.

¹²⁴ . declination

¹²⁵ . deaccentuation

L+H* h

۶- تأکید

منظور از تأکید، تأکید تقابلی^{۱۲۶} است که به معنای برجسته کردن یک یا چند عنصر در کلام نسبت به سایر عناصر آن (Sadat Tehrani, 2007, 136). مثلاً در جمله "من علی را دیدم"، «علی» در تقابل با افراد دیگر مثلاً «حسن یا مریم» قرار گرفته است: «من علی را دیدم نه مریم را»

واجشناسی تأکید

سازه‌ای که مورد تأکید واقع شده، یک گروه تکیه‌ای (AP) مجزا تشکیل می‌دهد که الگوی نواختی آن تفاوتی با سایر گروه‌های تکیه‌ای معمولی ندارد و به صورت $L+H^*$ است. نواخت مرزنامی آن نیز افتان است. به لحاظ آوایی هم میزان زیروبمی AP های تأکیدی بالاتر و هم دیرش آنها نسبت به گروه‌های تکیه‌ای معمولی بیشتر است. یک AP تأکیدی، تکیه زیروبمی هسته (NPA) یک IP را به وجود می‌آورد و سبب می‌شود سایر سازه‌ها تا انتهای جمله تکیه خود را از دست بدهند (Sadat Tehrani, 2007, 137). مثال:

13. ʔasemun emruz ʔabi'-ye.

14. ʔasemun emru'z ʔabi'-ye.

در جمله اول سه AP وجود دارد که الگوی نواختی هر سه $L+H^*$ است. نواخت مرزنامی دو AP اول خیزان و نواخت مرزنامی گروه آخر که NPA است، افتان می‌باشد. در جمله دوم فقط دو AP دیده می‌شود یکی AP فاعل (ʔasemun) و دیگری AP قید (emru'z). الگوی نواختی هر دو $L+H^*$ است. از آنجا که تأکید بر روی قید قرار دارد، سازه‌های پس از آن تکیه خود را از دست می‌دهند. قید سازه تأکید شده است و همان‌طور که گفتیم سازه تأکیدی، تکیه زیروبمی هسته در جمله را می‌سازد و بنابراین نواخت مرزنامی آن افتان است.

همان‌طور که گفتیم در جملات زیر تکیه زیروبمی هسته بر روی قیده‌های مکان «بیشه و لابی» قرار دارد.

¹²⁶ . contrastive focus

15. damha raftand to bife'.

$L+H*1\ L\%$

16. hame raftand to labi'.

$L+H*1\ L\%$

اگر دو قید «بیشه و لابی» با تأکید تولید شوند، در این صورت الگوی نواختی آنها متفاوت نخواهد بود اما به لحاظ آوایی، با زیروبمی بالاتر و دیرش بیشتر تولید خواهند شد (Sadat Tehrani, 2007, 139).

17. damha raftand to bife'.

$L+H*1\ L\%$

18. hame raftand to labi'.

$L+H*1\ L\%$

۳-۴-۳. بررسی آکوستیکی گروه تکیه‌ای

همه عوامل نوایی که در بالا اشاره شد، بر مشخصه‌های آکوستیکی آواها و بنابراین هجاها تأثیر می‌گذارند. سادات تهرانی (۲۰۰۷: ۱۹۰-۲۰۳) به بررسی آکوستیکی سه AP متفاوت (۱) AP با نواخت مرزنامی h، (۲) AP با نواخت مرزنامی l و (۳) AP تأکیدی می‌پردازد. سه گروه تکیه‌ای که وی مورد مطالعه قرار می‌دهد، در جملات زیر نشان داده شده‌اند:

1) diruz nama'mun umad.

$L+H* h$

2) mal-e nama'mun bud.

$L+H* l$

3) mal-e nama'mun bud.

$L+H* l$

وی پارامترهای دیرش گروه تکیه‌ای، واکه‌ها و همخوان‌های ماقبل و مابعد واکه تکیه‌بر و نیز میزان ترادف L و H را با واکه تکیه و واج‌های اطراف آن را اندازه‌گیری می‌کند. نتایج وی نشان می‌دهد:

(۱) نواخت L همواره با همخوان ماقبل واکه تکیه‌بر ترادف دارد (در مثال بالا با m). در AP‌های تأکیدی با آغازه این همخوان ترادف دارد. در AP با نواخت h با میانه همخوان و در AP با نواخت l اما غیر تأکیدی با بین آغازه و میانه همخوان ترادف دارد

(۲) ترادف نواخت H به نواخت مرزنا بستگی دارد. اگر نواخت مرزنا h باشد، H با میانه واکه پی‌چسب ترادف دارد (در مثال بالا با u در پی‌چسب). اگر نواخت مرزنا l باشد، نواخت H با همخوان بعد از واکه تکیه‌بر ترادف دارد (در مثال بالا با m در پی‌چسب).

(۳) دیرش یک AP تأکیدی به طور معناداری بیشتر از یا AP با نواخت مرزنا h است.

(۴) دیرش یک AP غیر تأکیدی با نواخت مرزنا l بین دو AP دیگر قرار دارد و تفاوت معناداری با آنها ندارد.

(۵) واکه ماقبل واکه تکیه‌بر در AP‌های تأکیدی به طور معناداری بلندتر از AP‌هایی است که نواخت مرزنا h است اما AP‌های غیر تأکیدی با نواخت مرزنا l از این لحاظ تفاوت معناداری با دو AP دیگر ندارند.

(۶) دیرش واکه تکیه‌بر در AP‌های تأکیدی به طور معناداری بلندتر از دو AP دیگر است.

(۷) دیرش واکه پی‌چسب در AP با نواخت مرزنا h به طور معناداری کوتاه‌تر از دو AP دیگر است.

(۸) میزان زیروبمی در دره L بسیار نزدیک به یکدیگر است و تفاوت معناداری بین سه AP نشان نمی‌دهد.

(۹) میزان زیروبمی در قله H بین سه AP متفاوت از یکدیگر است اما تفاوت معناداری نشان نمی‌دهد.

۱۰) میزان زیروبمی نرمال شده (اختلاف H و L تقسیم بر میانگین آنها) در APهای غیر تأکیدی با نواخت مرز نمای l به طور معناداری کمتر از دو AP دیگر (AP تأکیدی و AP با نواخت مرز نمای h) است.

سادات تهرانی (۲۰۰۷) سه گروه تکیه‌ای (AP) را در میان سه جمله خبری قرار می‌دهد. نتایج پژوهش‌های آوایی نشان داده که جایگاه تأثیر مهمی بر مقادیر آکوستیکی پارامترهای آوایی دارد. نرخ گفتار متغیر دیگری است که هم بر عناصر زنجیری و هم زبر زنجیری تأثیر بسزایی دارد. بنابراین می‌توان تفاوت APها را براساس این متغیرها و نیز عوامل دیگری مانند پرسش، ندا، تعجب و غیره نیز بررسی نمود.

فصل چهار

پیکره متنی

۴. مقدمه

در این پروژه سعی بر آن است تا با استخراج هجای حساس به بافت نوایی در برنامه بازسازی گفتار، گفتار حاصل طبیعی و به گفتار انسانی نزدیکتر باشد. اما از آنجا که پرداختن به همه تفاوت‌های نوایی هجاها و لحاظ کردن همه این تفاوت در برنامه بازسازی گفتار پیچیده و دشوار است بنابراین هجاهای مورد نظر تنها از چهار بافت نوایی زیر استخراج می‌شوند:

- ۱- هجای بی‌تکیه واژگانی
- ۲- هجای تکیه‌بر واژگانی، تکیه زیروبمی غیرهسته
- ۳- هجای تکیه‌بر واژگانی، تکیه زیروبمی هسته، نواخت مرزی افتان
- ۴- هجای تکیه‌بر واژگانی، تکیه زیروبمی هسته، نواخت مرزی خیزان

۴-۱. طراحی جملات حامل هجاهای نوایی:

همان‌طور که گفته شد از ترکیب همخوان‌ها با واکه‌ها $138 (6 \times 23)$ هجای CV بدست می‌آید که با احتساب چهار بافت نوایی، تعداد نمونه‌ها به $552 (4 \times 138)$ خواهد رسید. تولید نمونه‌ها در گفتار پیوسته و در درون جمله صورت خواهد گرفت. بافت تولید هجاهای نوایی نیز کنترل خواهد شد بدین صورت که بافت واکه‌ای قبل یا بعد هر هجای نوایی، واکه‌های افتاده /a/ یا /ɑ/ خواهد بود زیرا مشخصه‌های آکوستیکی همخوان‌ها در بافت واکه‌های افتاده کمتر تحت تأثیر

هم تولیدی قرار می گیرد. در بافت همخوانی قبل یا بعد هجای نوایی، همخوان واکندار لبی /b/ تولید خواهد شد. مشخصه های تولیدی همخوان های لبی نیز تأثیر کمتری بر واکه هجای نوایی دارد.

هجاهای نوایی در درون کلمه و جمله تولید می شوند. اما از آنجا که ممکن است کلمه ای در فارسی وجود نداشته باشد که بافت واکه ای و همخوانی شرح داده شده در بالا را داشته باشد، بنابراین برخی از هجاهای نوایی در درون کلمات بی مفهوم و یا جملات بی معنا گنجانده شده اند.

هجاهای نوایی در کلمات دوهجایی تولید می شوند زیرا اگر تعداد هجاها افزایش پیدا کند، به ویژه بر هجای بی تکیه در آغاز کلمه تأثیر گذاشته و سبب تضعیف بیش از حد آن می شود. در کلمات تکیه بر نیز با ایجاد تکیه دومین و یا سومین در کلماتی که بیش از دو هجا دارند، سبب کاهش و تضعیف تکیه در هجای تکیه بر می شود.

۴-۲. هجای بی تکیه واژگانی

در زبان فارسی هجای بی تکیه در آغاز یا میان کلمه قرار دارد. برای تولید پیکره گفتاری، هجاهای بی تکیه در آغاز کلمه و در ابتدای جمله قرار دارند اما برای مشخص شدن قطعه بست همخوان های گرفته و جدا کردن آن از سکوت، حرف اضافه «با» یا «تا» قبل از هجاهای مورد نظر قرار می گیرد تا بدین طریق با جدا شدن قطعه بست همخوان از سکوت، دیرش قطعه بست مشخص شده و امکان تقطیع ناحیه بست نیز میسر باشد. برای این منظور واکه /a/ در کلمات /ba, ta/ «با و تا» قبل از هجاهای نوایی قرار می گیرد. تولید هجاهای بی تکیه در گفتار پیوسته (جمله) و در کلماتی صورت می گیرد که دارای تکیه زیروبمی غیر هسته (H*) هستند.

آغاز هجای پس از هجای بی تکیه، با همخوان لبی واکندار /b/ تولید می شود زیرا همخوان های لبی برخلاف سایر همخوان ها (لثوی یا کامی) گذرهای سازه ای واکه ها را کمتر تحت تأثیر خود قرار می دهند. تولید هر آواها در زنجیره گفتار تحت تأثیر آوای قبلی و بعدی خود است. سازه های واکه ای تحت تأثیر محل تولید و نحوه تولید همخوان بعد و قبل خود قرار می گیرند و معمولاً مقادیر فرکانسی آنها در آغاز^{۱۲۷} یا پایانه^{۱۲۸} افزایش یا کاهش پیدا می کند. استیونس^{۱۲۹} (۲۰۰۰، ۳۴۳، ۳۵۷، ۳۶۷) تفاوت تغییرات گذرهای سازه ای در همخوان های لبی، لثوی و کامی را بدین صورت نشان می دهد که همخوان های لبی سبب کاهش هر سه سازه F1، F2 و F3 می شوند و دیرش قطعه گذر نیز بسیار

¹²⁷ . vowel onset

¹²⁸ . vowel offset

¹²⁹ . Stevens, K. N.

کوتاه در حدود ۵ تا ۱۰ میلی ثانیه است اما در همخوان‌های لثوی و کامی الگوی یکسانی مشاهده نمی‌شود و تغییرات هر سه سازه به یک صورت، افزایش یا کاهش نمی‌یابد و علاوه بر آن دیرش قطعه گذر نیز طولانی‌تر است. برای تهیه پیکره هجایی لازم است که بافت پس از هر هجا به گونه‌ای باشد که کمترین تأثیر را بر هجای تقطیع شده اعمال نماید و بنابراین برای تولید گفتار طبیعی و روان در تبدیل متن به گفتار پردازش هجاها بهتر و در زمان کمتری صورت گیرد.

همخوان لبی واکدار /b/ نیز بر جفت بیواک آن /p/ اولویت دارد زیرا در طول بست تارهای صوتی ارتعاش می‌کنند و نواخت بدون قطع ادامه پیدا می‌کند.

هجای بی‌تکیه در درون جملات منفی با علامت نفی /na, ne/ یا جملات اسنادی تولید می‌شود زیرا بدین طریق تکیه زیروبمی هسته بروی فعل منفی یا مسند قرار گرفته و بنابراین هجای بی‌تکیه در درون کلمه‌ای تولید می‌شود که تکیه زیروبمی غیرهسته دارد.

۴-۲-۱. همخوان‌های انفجاری

/p/

هجاء	کلمه	جملهء حامل
/pi/	/pibi/	با پی پی حرف نزد.
/pe/	/peba/	با پیا حرف نزد.
/pa/	/pabi/	با پپی حرف نزد.
/pa/	/pabe/	فاصله تا پابه زیاده .
/po/	/pobi/	با پپی حرف نزد.
/pu/	/pubi/	با پویی حرف نزد.

/b/

هجاء	کلمه	جمله
/bi/	/bibak/	با بی باک حرف نزد.
/be/	/bebo/	با ببو حرف نزد.
/ba/	/babu/	با ببو حرف نزد.
/ba/	baba	با بابا حرف نزد.
/bo/	/bobo/	با ببو حرف نزد.
/bu/	/bubar/	اون کرم با بو بر گرو نه.

/t/

هجاء	کلمه	جمله
/ti/	/tibag/	با تی بگ گرون می شه.
/te/	/tebar/	با تبار موافق نیست.
/ta/	/tabar/	با تبر شکسته نمی شه.
/ta/	/tabut/taban	با تابوت حمل نشد.
/to/	/tobe/	با توبه درست نمی شه.
/tu/	/tuborg/	با توبرگ هضم نمی شه.

/d/

هجاء	کلمه	جمله
/di/	/diba/	با دیا حرف نزد.
/de/	/debit/	با دبیت نمی شه خرید.
/da/	/dabir/	با دبیر علوم حرف نزد.
/da/	/dabi/	با دابی حرف نزد.
/do/	/dobar/	با دوبار دیدن درست نمی شه.
/du/	/dubi/	با دوبی بازی نکرد.

/k/

هجاء	کلمه	جمله
/ki/	/kibi/	صبحانه با کیبی عالی.
/ke/	/kebek/	اختلاف ساعت با کبک زیاد.
/ka/	/kabab/kabed	با کباب خوشمزه ست.
/ka/	/kabus/	با کابوس بیدار نشد.
/ko/	/kobalt/	مسمومیت با کبالت خطرناکه.
/ku/	/kuba/	فاصله تا کوبا زیاد.

/g/

هجاء	کلمه	جمله
/gi/	/gibon/	با گیون حرف نزد.
/ge/	/geba/	با گبا حرف نزد.
/ga/	/gabbe/gabi	با گبی حرف نزد.
/ga/	/gabon/	فاصله تا گابن زیاد.
/go/	/gobelz/	با گوبلز حرف نزد.
/gu/	/gubi/	با گوبی حرف نزد.

/q/

هجاء	كلمه	جمله
/qi/	/qibo/	با قیو نرفت.
/qe/	/qebal/	با قبال حرف نزد.
/qa/	/qabul/	با قبول موافق نیست.
/qa/	/qabus/	با قابوس نرفت.
/qo/	/qobar/	با غبار كهنه تره.
/qu/	/qubi/	تا غوبی راه زیاده.

/ʔ/

هجاء	كلمه	جمله
/ʔi/	/ʔibal/	با عیبال حرف نزد.
/ʔe/	/ʔebad/	با عباد حرف نزد.
/ʔa/	/ʔabad//ʔabus/	با عبوس حرف نمی زنم.
/ʔa/	/ʔabad/ʔaber	با عابر حرف نمی زنه.
/ʔo/	/ʔobejd/ʔobur	با عبور از اینجا نمی رسی.
/ʔu/	/ʔuberz/	با عوبرز حرف نزد.

۴-۲-۲. همخوان‌های سایشی

/f/

هجاء	کلمه	توضیحات	جمله
/fi/	/fiba/	فدراسیون بستکبال	با فیبا حرف نزد.
/fe/	/febo/		با فبو حرف نزد.
/fa/	/fabi/		با فبی حرف نزد.
/fa/	/fabarm/	کمپانی اسلحه	با فابارم کار نکرد.
/fo/	/foker/		با فوکر پرواز نکرد.
/fu/	/fubus/	بزرگترین ماه مریخ	تا فوبوس فاصله زیاده.

/v/

هجاء	کلمه	جمله
/vi/	/vibi/	با ویبی حرف نزد.
/ve/	/vebal/veber/	با وبر حرف نزد.
/va/	/vaba/	با وبا مبارزه نکرد.
/va/	/vabi/	با وابی کار نکرد.
/vo/	/voban/	با وبان حرف نزد.
/vu/	/vubi/	با ووبی حرف نزد.

/s/

هجاء	کلمه	جمله
/si/	/siba/	با سیبا پول نگرفت.
/se/	/sebil/	با سییل بازی می کرد.
/sa/	/saba/	با صبا حرف نزد.
/sa/	/sabun/	با صابون نشست.
/so/	/sobat/	با ثبات رفتار نکرد.
/su/	/suba/	با سوبا حرف نزد.

/z/

هجاء	کلمه	جمله
/zi/	/ziba/	با زیبا حرف نزد.
/ze/	/zebun//zebel/	با زبل کار می کرد.
/za/	/zaban/	با زبان خودش حرف نزد.
/za/	/zabol/	تا زابل فاصله زیاده.
/zo/	/zoba/	با زوبا تفاوت نداره.
/zu/	/zubin/	با زوین حرف نزد.

/ʃ/

هجاء	کلمه	جمله
/ʃi/	/ʃiba/ ʃibak	با شیبک حرف نزد.
/ʃe/	/ʃeba/	با شبا کار نمی کرد.
/ʃa/	/ʃaban/	با شبان نرفت.
/ʃa/	/ʃabaʃ/	با شاباش جشن نگرفتند.
/ʃo/	/ʃobejr/	با شیر حرف نزد.
/ʃu/	/ʃuberz/	با شوبرز حرف نزد.

/ʒ/

هجاء	کلمه	توضیحات	جملات
/ʒi/	/ʒibo/	شرکت اتصالات ژیبو	با ژیبو کار نکرد.
/ʒe/	/ʒebu/		با ژبو حرف نزد.
/ʒa/	/ʒaba/	نام رودخانه ای در افریقا- دهخدا	تا ژبا فاصله زیاده.
/ʒa/	/ʒabiz/	گروه طبیعتگردی ژابیز گیاه بومادران	با ژابیز سفر نرفت.
/ʒo/	/ʒobo/		با ژبو حرف نزد.
/ʒu/	/ʒubin/	اسم خاص	با ژوین حرف نزد.

/x/

هجاء	کلمه	جمله
/xi/	/xibo/	با خیبو حرف نزد.
/xe/	/xebir/	با خیبر حرف نزد.
/xa/	/xabar/	تا خبر خیلی مونده.
/xa/	/xaban/	با خابان حرف نزد.
/xo/	/xobo/	با خوبو حرف نزد.
/xu/	/xubi/	با خوبی زندگی نکردند.

/h/

هجاء	کلمه	جمله
/hi/	/hiba/	با هیبا حرف نزد.
/he/	/hebo/	با هبو حرف نزد.
/ha/	/habi/	با هبی حرف نزد.
/ha/	/habel/	با هابل کار نکرد.
/ho/	/hobab/	با حباب بازی می کرد.
/hu/	/hubi/	با هوبی بازی نکرد.

۴-۲-۳. همخوان‌های انسایشی

/tʃ/

هجاء	کلمه	جمله
/tʃi/	/tʃibo/	با چیبو کار نکرد.
/tʃe/	/tʃebar/	تا چهار فاصله زیاده.
/tʃa/	/tʃabin/	تا چین فاصله زیاده.
/tʃa/	/tʃabok/	با چابک حرف نزد.
/tʃo/	/tʃobaz/	تا چوباز فاصله زیاده.
/tʃu/	/tʃubin/	با چوبین کار نکرد.

/dʒ/

هجاء	کلمه	جمله
/dʒi/	/dʒiba/	با جیبای پر پول نرفت.
/dʒe/	/dʒebal/	با جبال البرز مقایسه نکن.
/dʒa/	/dʒabin/	با جبین گشاده نپذیرفت.
/dʒa/	/dʒaber/	با جابر حرف نزد.
/dʒo/	/dʒobejr/	با جیر حرف نزد.
/dʒu/	/dʒuba/	تا جوبای آب فاصله زیاده.

۴-۲-۴. همخوان‌های رسا

/m/

هجاء	کلمه	جمله
/mi/	/mibo/	با میبو وصل نشد.
/me/	/mebo/	تا مبو فاصله زیاده.
/ma/	/mabi/	با مبی مشکل نداشت.
/ma/	/mabi/	با مابی حرف نزد.
/mo/	/mobin/	با مبین حرف نزد.
/mu/	/mubiz/	با موبیز الک نکرد.

/n/

هجاء	کلمه	جمله
/ni/	/niba/	تا نیا راه زیاده.
/ne/	/nebat/	با نبات بومی نکاشت.
/na/	/nabat/	با نبات نخورد.
/na/	/nabud/	با نبود آب ممکن نیست.
/no/	/nobel/	با نوبل خوشحال شد.
/nu/	/nuban/	با نوبان حرف نزد.

از آنجا که در همخوان‌های رسا، روان^{۱۳۰} و غلت^{۱۳۱}، تولید قطعه گرفتگی بسیار کوتاه و همراه با الگوی سازه‌ای شبیه به واکه‌هاست، بنابراین به جای تولید واکه قبل از این همخوان‌ها، همخوان انسدادی /b/ در بافت ماقبل تولید می‌شود تا تشخیص سیگنال آنها راحت‌تر صورت گیرد.

/r/

هجاء	کلمه	جمله
/ri/	/riba/	تو خواب ربا حرف می‌زنه.
/re/	/reba/	تو خواب ربا حرف می‌زنه.
/ra/	/rabab/	تو خواب رباب حرف می‌زنه.
/ra/	/rabet/	از رو کتاب رابط فلزی می‌ساخت.
/ro/	/robat/	از رو کتاب ربات درست می‌کنه.
/ru/	/rubah/	با روباه فرار نکرد.

/l/

هجاء	کلمه	جمله
/li/	/libi/	تو این کتاب لیبی بزرگه.
/le/	/lebas/	تو این کتاب لباس نکشیده.
/la/	/labe/	این کتاب لبه‌ی تیزی داره.
/la/	/labe/labi/	تو این کتاب لابی آموزش می‌ده.
/lo/	/lobo/	تو این کتاب لوبو خیلی قشنگه.
/lu/	/lubija/	تو این کتاب لوبیا نکشیده.

130 . liquid

131 . glide

/j/ «ی»

هجاء	کلمه	بافت	جمله
/ji/	/jiba/	/b-b/	تو خواب یی با حرف می زنه.
/je/	/jebar/		تو خواب ییار پرید.
/ja/	/jabab/		توی این کتاب ییاب زیاده.
/ja/	/jabu/		تو این کتاب یابو زیاده.
/jo/	/jobal/		تو خواب یوبال حرف می زنه.
/ju/	/jubil/		تو این کتاب یوبیل قانونی شده.
			سال و قانون یوبیل در مسیحیت

۴-۳. هجای تکیه‌بر واژگانی

در این جملات کنترل بافت با تولید همخوان /b/ بعد از هجاهای نوایی صورت گرفته است. برای تولید هجای تکیه‌بر دارای تکیه زیرویمی غیرهسته، کلمات حامل هجاهای نوایی بیشتر در جملات اسنادی یا منفی تولید می‌شوند. برای تولید هجای تکیه‌بر دارای تکیه زیرویمی هسته، این هجاها بیشتر در جملات اسنادی تولید شده‌اند و یا کلمه حامل هجای نوایی به عنوان قید در جایگاه پس از فعل جمله خبری تولید شده، که در این صورت تکیه زیرویمی هسته بر روی آن قرار می‌گیرد. برای تولید هجای تکیه‌بر دارای تکیه زیرویمی هسته و سؤالی، کلمات حامل در پایان جمله و به صورت سؤالی تولید می‌شوند.

۴-۳-۱. همخوان‌های انسدادی

/p/

هجاء	کلمه	زیروبمی	جمله
/pi/	/tʃapi/	غیر هسته، H*	احزاب <u>چی</u> برنده شدند.
		هسته-افتان، L*	علی <u>چی</u> بود
		هسته-خیزان، H%	علی شد <u>چی</u> ؟
/pe/	/gape/	غیر هسته	روستای گپه بزرگه.
		هسته-افتان	مال گپه بود.
		هسته-خیزان	رفت گپه؟
/pa/	/papa/	غیر هسته	په په برنده شد.
		هسته-افتان	مال په په بود.
		هسته-خیزان	رفت پیش په په؟
/pa/	/papa/	غیر هسته	پاپا <u>برنده</u> شد.
		هسته-افتان	مال <u>پاپا</u> بود.
		هسته-خیزان	رفت پیش <u>پاپا</u> ؟
/po/	/napo/	غیر هسته	نیو برنده شد.
		هسته-افتان	مال نیو بود.
		هسته-خیزان	رفت پیش نیو؟
/pu/	/hapu/, takapu	غیر هسته	هاپو <u>برنده</u> شد.
		هسته-افتان	مال <u>هاپو</u> بود.
		هسته-خیزان	داد به <u>هاپو</u> ؟

/b/

هجاء	کلمه	زیر و بمی	جمله
/bi/	/labi/	غیر هسته	لبی به گوش می رسه.
		هسته-افتان	لبی بود.
		هسته-خیزان	گفتی لبی؟
/be/	/labe/tabe	غیر هسته	رو لبه بلند شد.
		هسته-افتان	رفت رو لبه.
		هسته-خیزان	رفت رو لبه؟
/ba/	/naba/gaba/	غیر هسته	سوره نبا به گوش می رسه.
		هسته-افتان	مال نبا بود.
		هسته-خیزان	گفتی سوره نبا؟
/ba/	/saba/	غیر هسته	صبا برنده شد.
		هسته-افتان	رفت پیش صبا.
		هسته-خیزان	رفت پیش صبا؟
/bo/	/tabo/	غیر هسته	تابو بسیار بده.
		هسته-افتان	تابو بود.
		هسته-خیزان	شد تابو؟
/bu/	/labu/	غیر هسته	لبو با گوشت می پخت.
		هسته-افتان	مال لبو بود.
		هسته-خیزان	خریدی لبو؟

/t/

هجاء	کلمه	زیروبمی	جمله
/ti/	/kati/,	غیر هسته	کتی برنده شد.
		هسته-افتان	رفت پیش <u>کتی</u> .
		هسته-خیزان	رفت پیش <u>کتی</u> ؟
/te/	/kate/	غیر هسته	کته با ماست می خوره.
		هسته-افتان	مال <u>کته</u> بود.
		هسته-خیزان	می پزی کته؟
/ta/	/pata/	غیر هسته	پته برنده می شه.
		هسته-افتان	مال <u>پته</u> بود.
		هسته-خیزان	رفت پیش <u>پته</u> ؟
/ta/	/?ata/xata,	غیر هسته	عطا <u>برنده</u> شد.
		هسته-افتان	مال <u>عطا</u> بود.
		هسته-خیزان	رفت پیش <u>عطا</u> ؟
/to/	/tato/	غیر هسته	تتو به این روش <u>خوبه</u> .
		هسته-افتان	مال <u>تتو</u> بود.
		هسته-خیزان	بلدی <u>تتو</u> ؟
/tu/	/patu/	غیر هسته	یه پتو با بالش می خره.
		هسته-افتان	مال <u>پتو</u> بود.
		هسته-خیزان	خریدی <u>پتو</u> ؟

/d/

هجاء	کلمه	زیر و بمی	جمله
/di/	/badi/yadi/	غیر هسته	یدی بهتره.
		هسته-افتان	یدی بود.
		هسته-خیزان	بلدی کار یدی؟
/de/	/sade/, bade, /dade/	غیر هسته	داده باید خوب باشه.
		هسته-افتان	ساده بود.
		هسته-خیزان	خریدی باده؟
/da/	/dada/	غیر هسته	دد برنده شد.
		هسته-افتان	مال دد بود.
		هسته-خیزان	رفت پیش دد؟
/da/	/?ada//fada/	غیر هسته	فدا برنده شد.
		هسته-افتان	فدا بود.
		هسته-خیزان	بشه فدا؟
/do/	/kado/	غیر هسته	کادو با کاغذ پیچیده شد.
		هسته-افتان	کادو بود.
		هسته-خیزان	می خری کادو؟
/du/	/kadu/	غیر هسته	کدو با ماست خوشمزه ست.
		هسته-افتان	کدو بود.
		هسته-خیزان	می خری کدو؟

/k/

هجا	کلمه	زیروبمی	جمله
/ki/	/saki/	غیر هسته	ساکی برنده شد.
		هسته-افتان	مال ساکی بود.
		هسته-خیزان	رفت پیش ساکی؟
/ke/	/lakke/rake	غیر هسته	رکه بزرگه.
		هسته-افتان	مال رکه بود.
		هسته-خیزان	رفت رکه؟
/ka/	/zaka/	غیر هسته	زکه برنده شد.
		هسته-افتان	مال زکه بود.
		هسته-خیزان	رفت پیش زکه؟
/ka/	/lakka/kaka	غیر هسته	ککا با سم می میرن.
		هسته-افتان	مال ککا بود.
		هسته-خیزان	ریختی رو ککا؟
/ko/	/kako	غیر هسته	کاکو برنده شد.
		هسته-افتان	مال کاکو بود.
		هسته-خیزان	رفت پیش کاکو؟
/ku/	/sakku/maku	غیر هسته	ماکو بزرگه.
		هسته-افتان	مال ماکو بود.
		هسته-خیزان	رفت ماکو؟

/g/

هجاء	کلمه	زیر و بمی	جمله
/gi/	/magi/sagi	غیر هسته	مگی بالاست.
		هسته-افتان	مال مگی بود.
		هسته-خیزان	رفت پیش مگی؟
/ge/	/rage/	غیر هسته	این رگه به اون ور نمی ره.
		هسته-افتان	مال رگه بود.
		هسته-خیزان	خورد تو این رگه؟
/ga/	/laga/	غیر هسته	لگه بزرگه.
		هسته-افتان	مال لگه بود.
		هسته-خیزان	رفت لگه؟
/ga/	/raga/saga	غیر هسته	سگا بازی نمی کردند.
		هسته-افتان	مال سگا بود.
		هسته-خیزان	رفت پیش سگا؟
/go/	/zago/	غیر هسته	زگو برنده شد.
		هسته-افتان	مال زگو بود.
		هسته-خیزان	رفت پیش زگو؟
/gu/	/magu/, /vagu/	غیر هسته	بگو مگو نکرد.
		هسته-افتان	بگو مگو بود.
		هسته-خیزان	بنویسم بگو مگو؟

/q/

هجاء	کلمه	زیروبمی	جمله
/qi/	/naqi/	غیر هسته	نقی برنده شد.
		هسته-افتان	مال نقی بود.
		هسته-خیزان	رفت پیش نقی؟
/qe/	/naqe/	غیر هسته	نقه بزرگه.
		هسته-افتان	مال نقه بود.
		هسته-خیزان	رفت نقه؟
/qa/	/gaqa/	غیر هسته	گقه بزرگه.
		هسته-افتان	مال گقه بود.
		هسته-خیزان	رفت گقه؟
/qa/	/baqa/	غیر هسته	راز بقا بهتره.
		هسته-افتان	مال بقا بود.
		هسته-خیزان	نوشتی بقا؟
/qo/	/laqo/	غیر هسته	لقو بزرگه.
		هسته-افتان	مال لقو بود.
		هسته-خیزان	رفت لقو؟
/qu/	/baqu/ سایشی در این بافت	غیر هسته	بغ بغو به گوش می رسه.
		هسته-افتان	بغ بغو بود.
		هسته-خیزان	بنویسم بغ بغو؟

/ʔ/

هجاء	کلمه	زیروبمی	جمله
/ʔi/	/saʔi/	غیر هسته	ساعی برنده شد.
		هسته-افتان	مال ساعی بود.
		هسته-خیزان	رفت پیش ساعی؟
/ʔe/	/saʔe/	غیر هسته	سعه بزرگه.
		هسته-افتان	مال سعه بود.
		هسته-خیزان	رفت سعه؟
/ʔa/	/raʔa/	غیر هسته	رعه بزرگه.
		هسته-افتان	مال رعه بود.
		هسته-خیزان	رفت رعه؟
/ʔa/	/saʔa/	غیر هسته	سعا برنده شد.
		هسته-افتان	مال سعا بود.
		هسته-خیزان	رفت پیش سعا؟
/ʔo/	/maʔo/	غیر هسته	مائو برنده شد.
		هسته-افتان	مال مائو بود.
		هسته-خیزان	رفت پیش مائو.
/ʔu/	/laʔu/	غیر هسته	لعو برنده شد.
		هسته-افتان	مال لعو بود.
		هسته-خیزان	رفت پیش لعو؟

۴-۳-۲. همخوان‌های سایشی

/f/

هجاء	کلمه	زیروبمی	جمله
/fi/	/safi/	غیر هسته	صفی بهتره.
		هسته-افتان	صفی بود.
		هسته-خیزان	حالا شد صفی؟
/fe/	/kafe/, bafe	غیر هسته	کفه بالا رفت.
		هسته-افتان	مال کفه بود.
		هسته-خیزان	ریختی تو این کفه؟
/fa/	/bafa/	غیر هسته	بفه بزرگه.
		هسته-افتان	مال بفه بود.
		هسته-خیزان	رفت بفه؟
/fa/	/safa/dʒafa	غیر هسته	صفا بالاتره.
		هسته-افتان	با صفا بود.
		هسته-خیزان	رفت صفا؟
/fo/	/?afo/	غیر هسته	آفو برنده شد.
		هسته-افتان	مال آفو بود.
		هسته-خیزان	رفت پیش آفو؟
/fu/	/kafu/	غیر هسته	کافو برنده شد.
		هسته-افتان	مال کافو بود.
		هسته-خیزان	رفت پیش کافو؟

/v/

هجاء	کلمه	زیر وبمی	جمله
/vi/	/ravi/	غیر هسته	راوی برنده شد.
		هسته-افتان	مال راوی بود.
		هسته-خیزان	رفت پیش راوی؟
/ve/	/pave/	غیر هسته	پاوه بزرگه.
		هسته-افتان	مال پاوه بود.
		هسته-خیزان	رفت پاوه؟
/va/	/gava/	غیر هسته	گو بزرگه.
		هسته-افتان	مال گو بود.
		هسته-خیزان	رفت گو؟
/va/	/dava/?ava,sava,nava	غیر هسته	دوا بسیار گرونه.
		هسته-افتان	مال دوا بود.
		هسته-خیزان	خریدی دوا؟
/vo/	/devo/	غیر هسته	یه دوو برنده شد.
		هسته-افتان	مال دوو بود.
		هسته-خیزان	خریدی دوو؟
/vu/	/havu/	غیر هسته	هوو به اون می گن.
		هسته-افتان	مال هوو بود.
		هسته-خیزان	رفت پیش هوو؟

/s/

هجاء	کلمه	زیروبمی	جمله
/si/	/?asi/qasi/	غیر هسته	عاصی برنده شد.
		هسته-افتان	مال عاصی بود.
		هسته-خیزان	رفت پیش عاصی؟
/se/	/lase/	غیر هسته	لته بزرگ شده بود.
		هسته-افتان	مال لته بود.
		هسته-خیزان	خورد تو لته؟
/sa/	/rasa/	غیر هسته	رس بزرگه.
		هسته-افتان	مال رس بود.
		هسته-خیزان	رفت رس؟
/sa/	/?asa/	غیر هسته	عصا بسیار گروه.
		هسته-افتان	مال عصا بود.
		هسته-خیزان	خریدی عصا؟
/so/	/maso/	غیر هسته	مسو برنده شد.
		هسته-افتان	مال مسو بود.
		هسته-خیزان	رفت پیش مسو؟
/su/	/rasu/	غیر هسته	راسو بالای درخته.
		هسته-افتان	مال راسو بود.
		هسته-خیزان	رفت پیش راسو؟

هجاء	کلمه	زیروبمی	جمله
/zi/	/nazi/bazi	غیر هسته	نازی برنده شد.
		هسته-افتان	مال نازی بود.
		هسته-خیزان	رفت پیش نازی؟
/ze/	/mazze/?andaze, taze	غیر هسته	تازه با من حرف نمی‌زنه.
		هسته-افتان	تازه بود.
		هسته-خیزان	خریدی تازه؟
/za/	/zaza/	غیر هسته	زز بزرگه.
		هسته-افتان	مال زز بود.
		هسته-خیزان	رفت زز؟
/za/	/?aza/qaza	غیر هسته	غذا بسیار گروه.
		هسته-افتان	مال غذا بود.
		هسته-خیزان	خوردی غذا؟
/zo/	/gazo/	غیر هسته	گزو برنده شد.
		هسته-افتان	مال گزو بود.
		هسته-خیزان	رفت پیش گزو؟
/zu/	/bazu/	غیر هسته	بازو با دمبل قوی می‌شه.
		هسته-افتان	مال بازو بود.
		هسته-خیزان	بستی به بازو؟

/ʃ/

هجاء	کلمه	زیروبمی	جمله
/ʃi/	/kaʃi/	غیر هسته	کاشی بالا ننداز.
		هسته-افتان	مال کاشی بود.
		هسته-خیزان	گرفتی کاشی؟
/ʃe/	/laʃe/	غیر هسته	لاشه بو گرفته بود.
		هسته-افتان	مال لاشه بود.
		هسته-خیزان	خریدی یه لاشه؟
/ʃa/	/raʃa/	غیر هسته	رشه بزرگه.
		هسته-افتان	مال رشه بود.
		هسته-خیزان	رفت رشه؟
/ʃa/	/?aʃa/	غیر هسته	عُشا به بعد می‌ره.
		هسته-افتان	عشا بود.
		هسته-خیزان	می‌آد موقع عشا؟
/ʃo/	/taʃo/	غیر هسته	تاشو بهتر کار می‌کنه.
		هسته-افتان	مال تاشو بود.
		هسته-خیزان	می‌خری تاشو؟
/ʃu/	/dʒaʃu/	غیر هسته	جاشو برنده شد.
		هسته-افتان	مال جاشو بود.
		هسته-خیزان	رفت پیش جاشو؟

هجاء	کلمه	زیروبمی	جمله
/zi/	/kaʒi/	غیر هسته	با کزی بهش نگو.
		هسته-افتان	مال کزی بود.
		هسته-خیزان	گفتی کزی؟
/ze/	/vaʒe/	غیر هسته	واژه با آوا فرق می کنه.
		هسته-افتان	مال واژه بود.
		هسته-خیزان	نوشتی واژه؟
/za/	/laʒa/	غیر هسته	لژ بزرگه.
		هسته-افتان	مال لژ بود.
		هسته-خیزان	رفت لژ؟
/ʒa/	/vaʒa/	غیر هسته	واژا به چند دسته تقسیم می شن.
		هسته-افتان	مال واژا بود.
		هسته-خیزان	گفتی واژا؟
/zo/	/ʒaʒo/	غیر هسته	ژژو برنده شد.
		هسته-افتان	مال ژژو بود.
		هسته-خیزان	رفت پیش ژژو؟
/zu/	/baʒu/	غیر هسته	بژو برنده شد.
		هسته-افتان	مال بژو بود.
		هسته-خیزان	رفت پیش بژو؟

/x/

هجاء	کلمه	زیروبمی	جمله
/xi/	/saxi/naxi/	غیر هسته	لباس نخى بسیار گرونه.
		هسته-افتان	نخى بود.
		هسته-خیزان	پوشم نخى؟
/xe/	/ɟaxe/	غیر هسته	جخه بزرگه.
		هسته-افتان	مال جخه بود.
		هسته-خیزان	رفت جخه؟
/xa/	/laxa/	غیر هسته	لخه بزرگه.
		هسته-افتان	مال لخه بود.
		هسته-خیزان	رفت لخه؟
/xa/	/saxa/naxa/	غیر هسته	نخا بالای کمده.
		هسته-افتان	مال نخا بود.
		هسته-خیزان	بالای کمدن نخا؟
/xo/	/ɟaxo/	غیر هسته	جخو بزرگه.
		هسته-افتان	مال جخو بود.
		هسته-خیزان	رفت جخو؟
/xu/	/laxu/	غیر هسته	لاخو بزرگه.
		هسته-افتان	مال لاختو بود.
		هسته-خیزان	رفت لاختو؟

/h/

هجاء	کلمه	زیروبمی	جمله
/hi/	/ʃahi/sahi, zahi/	غیر هسته	سرو سهی به اون نمی گن.
		هسته-افتان	مال سهی بود.
		هسته-خیزان	رفت پیش سهی؟
/he/	/lahe/	غیر هسته	لا هه بالای تپه است.
		هسته-افتان	تو لاهه بود.
		هسته-خیزان	می ری لاهه؟
/ha/	/baha/	غیر هسته	بهبه بزرگه.
		هسته-افتان	مال بهبه بود.
		هسته-خیزان	رفت بهبه؟
/ha/	/raha/	غیر هسته	رها برنده شد.
		هسته-افتان	مال رها بود.
		هسته-خیزان	رفت پیش رها؟
/ho/	/vaho/	غیر هسته	وهو برنده نشد.
		هسته-افتان	مال وهو بود.
		هسته-خیزان	رفت پیش وهو؟
/hu/	/yahu/	غیر هسته	ياهو به فارسی راه می افته.
		هسته-افتان	مال ياهو بود.
		هسته-خیزان	رفتی تو ياهو؟

۴-۳-۳. همخوان‌های انسایشی

/tʃ/

هجاء	کلمه	زیروبمی	جمله
/tʃi/	/gatʃi/	غیر هسته	گچی بهتره.
		هسته-افتان	گچی بود.
		هسته-خیزان	خریدی گچی؟
/tʃe/	/batʃtʃe/gatʃe	غیر هسته	گچه بزرگه.
		هسته-افتان	مال گچه بود.
		هسته-خیزان	رفتی گچه؟
/tʃa/	/latʃa/	غیر هسته	لچه بزرگه.
		هسته-افتان	مال لچه بود.
		هسته-خیزان	رفت لچه؟
/tʃa/	/gatʃa/	غیر هسته	گچا بالای خونه است.
		هسته-افتان	مال گچا بود.
		هسته-خیزان	گذاشتی رو گچا؟
/tʃo/	/natʃo/	غیر هسته	ناچو برنده است.
		هسته-افتان	مال ناچو بود.
		هسته-خیزان	رفت پیش ناچو؟
/tʃu/	/matʃu/	غیر هسته	مچو برنده شد.
		هسته-افتان	مال مچو بود.
		هسته-خیزان	رفت پیش مچو؟

/ɖʒ/

هجاء	كلمه	زیروبمی	جمله
/ɖʒi/	/naɖʒi/	غیر هسته	ناجی برنده شد.
		هسته-افتان	مال ناجی بود.
		هسته-خیزان	رفت پیش ناجی؟
/ɖʒe/	/laɖʒe/	غیر هسته	لجه برنده شد.
		هسته-افتان	مال لجه بود.
		هسته-خیزان	رفت پیش لجه؟
/ɖʒa/	/gaɖʒa/	غیر هسته	گجه برنده شد.
		هسته-افتان	مال گجه بود.
		هسته-خیزان	رفت پیش گجه؟
/ɖʒa/	/raɖʒa/	غیر هسته	قطارای رجا بیشتر شدن.
		هسته-افتان	مال رجا بود.
		هسته-خیزان	رفت پیش رجا؟
/ɖʒo/	/baɖʒo/	غیر هسته	باجو برنده شد.
		هسته-افتان	مال باجو بود.
		هسته-خیزان	رفت پیش باجو؟
/ɖʒu/	/maɖʒu/	غیر هسته	مجو برنده شد.
		هسته-افتان	مال مجو بود.
		هسته-خیزان	رفت پیش مجو؟

۴-۳-۴. همخوان‌های رسا

/m/

هجاء	کلمه	زیروبمی	جمله
/mi/	/hami/nami	غیر هسته	لباسای نمی بالای میزند.
		هسته-افتان	لباسا نمی بودن.
		هسته-خیزان	خشک باشه یا نمی؟
/me/	/hame/?ame	غیر هسته	همه برنده هستند.
		هسته-افتان	مال همه بود.
		هسته-خیزان	یکی می‌آد یا همه؟
/ma/	/lama/	غیر هسته	لمه بزرگه
		هسته-افتان	مال لمه بود.
		هسته-خیزان	رفت لمه؟
/ma/	/nama/	غیر هسته	نما بیشتر آجریه.
		هسته-افتان	مال نما بود.
		هسته-خیزان	خریدی نما؟
/mo/	/bamo/	غیر هسته	بمو بزرگه.
		هسته-افتان	مال بمو بود.
		هسته-خیزان	رفت پیش بمو؟
/mu/	/?amu/	غیر هسته	عمو برنده شد.
		هسته-افتان	مال عمو بود.
		هسته-خیزان	رفت پیش عمو؟

/n/

هجاء	کلمه	زیروبمی	جمله
/ni/	/tani/	غیر هسته	برادر تنی برنده شد.
		هسته-افتان	تنی هستند.
		هسته-خیزان	ناتنی هستید یا تنی؟
/ne/	/tane/kane	غیر هسته	تنه با برگها پوشیده.
		هسته-افتان	مال تنه بود.
		هسته-خیزان	رفتی بالای تنه؟
/na/	/?ana/	غیر هسته	انا با این جمله درستره.
		هسته-افتان	مال انا بود.
		هسته-خیزان	گفتی انا؟
/na/	/hana/fana	غیر هسته	حنا برنده شد.
		هسته-افتان	مال حنا بود.
		هسته-خیزان	رفت پیش حنا؟
/no/	/nano/	غیر هسته	نانو به این روش بهتره.
		هسته-افتان	مال نانو بود.
		هسته-خیزان	می شه با نانو؟
/nu/	/nanu/	غیر هسته	ننو به در آویزونه.
		هسته-افتان	مال ننو بود.
		هسته-خیزان	رفت تو ننو؟

از آنجا که در همخوان‌های رسا، روان و غلت، تولید قطعه گرفتگی کوتاه و همراه با الگوی سازه‌ای شبیه به واکه‌هاست، بنابراین به جای تولید واکه قبل از این همخوان‌ها، همخوان انسدادی /b/ در بافت ماقبل تولید می‌شود تا تشخیص سیگنال آنها راحت‌تر صورت گیرد. در بافت پس از هجای نوایی نیز همخوان /b/ تولید می‌شود، بنابراین بافت به صورت /bCV/#b/ است.

/r/

هجای	کلمه	زیروبمی	جمله
/ri/	/?abri/	غیر هسته	ابری بهتره.
		هسته-افتان	ابری بود.
		هسته-خیزان	دوست داری ابری؟
/re/	/xobre/vibre	غیر هسته	خبره برنده شد.
		هسته-افتان	خبره بود.
		هسته-خیزان	رفت پیش خبره؟
/ra/	/libra/	غیر هسته	لیبر برنده شد.
		هسته-افتان	مال لیبر بود.
		هسته-خیزان	رفت پیش لیبر؟
/ra/	/sabra/	غیر هسته	صبرا برنده شد.
		هسته-افتان	مال صبرا بود.
		هسته-خیزان	رفت پیش صبرا؟
/ro/	/zabro/rahro	غیر هسته	زابرو برنده شد.
		هسته-افتان	مال زابرو بود.
		هسته-خیزان	رفت پیش زابرو؟
/ru/	/?abru/	غیر هسته	ابرو بالا ننداز.
		هسته-افتان	مال ابرو بود.
		هسته-خیزان	کشید رو ابرو؟

هجاء	کلمه	زیروبمی	جمله
/li/	/kabli/	غیر هسته	کابلی بهتره.
		هسته-افتان	کابلی بود.
		هسته-خیزان	خریدی کابلی؟
/le/	/moble/	غیر هسته	مبله بهتره.
		هسته-افتان	مبله بود.
		هسته-خیزان	گفتی مبله؟
/la/	/labla/	غیر هسته	لبل برنده شد.
		هسته-افتان	مال لبل بود.
		هسته-خیزان	رفت پیش لبل؟
/la/	/kablə/	غیر هسته	کابلا بالا هستند.
		هسته-افتان	مال کابلا بود.
		هسته-خیزان	گذاشتی کنار کابلا؟
/lo/	/tablo/	غیر هسته	تابلو بالای دیواره.
		هسته-افتان	مال تابلو بود.
		هسته-خیزان	خریدی تابلو؟
/lu/	/diblu/	غیر هسته	دیبلو برنده شد.
		هسته-افتان	مال دیبلو بود.
		هسته-خیزان	رفت پیش دیبلو؟

/j/

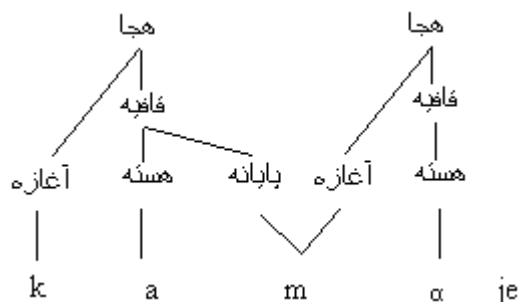
هجاء	کلمه	زیروبمی	جمله
/ji/	/zabji/	غیر هسته	زب یی بزرگه.
		هسته-افتان	مال زب یی بود.
		هسته-خیزان	رفت زب یی؟
/je/	/zabje/	غیر هسته	زب یه باهاش نرفت.
		هسته-افتان	مال زب یه بود.
		هسته-خیزان	رفت پیش زب یه؟
/ja/	/rabja/	غیر هسته	رب یه بزرگه.
		هسته-افتان	مال رب یه بود.
		هسته-خیزان	رفت رب یه؟
/ja/	/labja/fobja	غیر هسته	لب یا بزرگه.
		هسته-افتان	مال لب یا بود.
		هسته-خیزان	رفت لب یا؟
/jo/	/labjo/	غیر هسته	لب یو بزرگه.
		هسته-افتان	مال لب یو بود.
		هسته-خیزان	رفت لب یو؟
/ju/	/labju/	غیر هسته	لب یو بزرگه.
		هسته-افتان	مال لب یو بود.
		هسته-خیزان	رفت لب یو؟

۴-۴. تشدید^{۱۳۲}

علاوه بر هجاهای نوایی فوق، لازم است که در همه این موضع‌ها، هجاهایی با صورت‌های مشدد^{۱۳۳} همخوان‌ها نیز تهیه و ضبط شود زیرا صورت‌های مشدد ویژگی‌های آکوستیکی خاص خود را دارند و از این رو لازم است که در بازسازی گفتار، دادگان شامل صورت‌های مشدد نیز به طور جداگانه باشد. مثلاً در کلمه /tappe/ «تپه» هجای آخر با صورت مشدد همخوان لبی /p/ تولید می‌شود که در این وضعیت دیرش /p/ بلندتر است.

تشدید به دو گروه تشدید حقیقی^{۱۳۴} و عارضی^{۱۳۵} تقسیم می‌شود. اگر تشدید در محل اتصال دو تکواژ و یا مرز دو کلمه مستقل به وجود آید، تشدید عارضی است اما تشدید حقیقی معمولاً در درون کلمه تولید می‌شود. در تشدید عارضی یا در ابتدا هر دو یکسان هستند مثل «کم‌مایه» /kammaje/ یا بر اثر فرایندهای آوایی مثل همگونی در یک تکواژ یا مرز دو تکواژ به وجود می‌آیند. مانند همگونی در «پسته» که به صورت /pesse/ تولید می‌شود (کرد زعفرانلو کامبوزیا و تاج‌آبادی، ۱۳۹۱).

به لحاظ واجی، همخوان مشدد در مرز دو هجا قرار دارد و به دو جایگاه زمانمند متصل است. در مثال زیر همخوان مشدد /m/ در مرز دو هجای اول و دوم قرار دارد (کرد زعفرانلو کامبوزیا و تاج‌آبادی، ۱۳۹۱):



شکل ۴-۱: الگوی هجایی همخوان‌های مشدد

اما به لحاظ آکوستیکی و آوایی، در سیگنال همخوان‌های مشدد، ناحیه بست با دیرش بیشتری نسبت به همخوان‌های غیرمشدد تولید می‌شود. از این رو در تقطیع هجایی، همخوان مشدد به هجای دوم تعلق

¹³² . gemination

¹³³ . geminate

¹³⁴ . true gemination

¹³⁵ . fake gemination

می‌گیرد و از آنجا که در زبان فارسی همخوان‌ها در موضع آغاز کلمه به صورت مشدد تولید نمی‌شوند، بنابراین تهیه هجاهای مشدد فقط در موضع تکیه بر صورت می‌گیرد.

در زبان فارسی، همخوان مشدد همیشه بین دو واکه قرار می‌گیرد. به دلیل تقسیم هجایی، پس از همخوان مشدد و نیز قبل از همخوان مشدد در زبان فارسی همخوان دیگری نمی‌تواند واقع شود (کرد زعفرانلو کامبوزیا و تاج‌آبادی، ۱۳۹۱).

نکته‌ای که در زمینه هجاهای مشدد قابل ذکر است، این است که اولاً لازم است که با انجام پژوهشی آماری بر روی مجموعه بسیار غنی و بزرگی از کلمات فارسی تعیین نمود که همخوان‌های مشدد در کدام هجاهای فارسی تولید می‌شوند زیرا به نظر نمی‌رسد که در زبان فارسی برای همه همخوان‌ها و در هر ۶ بافت واکه‌ای فارسی، کلمات حاوی هجاهای CV مشدد حقیقی موجود باشد هرچند که امکان تولید مشددهای عارضی بیشتر است. کرد زعفرانلو کامبوزیا و تاج‌آبادی (۱۳۹۱) اظهار می‌کنند که معمولاً قبل از همخوان‌های مشدد، واکه‌های کوتاه در زبان فارسی دیده می‌شود و به ندرت واکه‌های کشیده مشاهده می‌شود، به عنوان مثال در کلمه «عمّه» /ʔamme/.

در زبان فارسی در بسیاری موارد، حتی در کلماتی که ریشه عربی دارند، تشدیدزایی صورت می‌گیرد، بنابراین پیکره متنی که در ادامه تهیه شده است تنها هجاهای مشددی (هجای CV مشدد) را شامل می‌شود که احتمال وقوع بیشتری در زبان فارسی دارند.

لازم به ذکر است هجاها در همان بافت‌هایی قرار داده شدند که هجاهای تکیه بر واژگانی (بخش ۴-۳) قرار داده شدند، از این رو جملات حامل ممکن است در برخی موارد بی‌معنا باشند.

۴-۴-۱. همخوان‌های انفجاری

/p/

هجاء	کلمه	زیروبمی	جمله
/pe/	/tappe/	غیر هسته	تپه بزرگه.
		هسته-افتان	رفت رو تپه.
		هسته-خیزان	رفت رو تپه؟
/pa/	/tʃappa/	غیر هسته	چپ پا برنده شد.
		هسته-افتان	مال چپ پا بود.
		هسته-خیزان	رفت پیش چپ پا؟

/b/

هجاء	کلمه	زیروبمی	جمله
/be/	/habbe/	غیر هسته	حبه بزرگه.
		هسته-افتان	مال حبه بود.
		هسته-خیزان	خوردی یه حبه؟
/ba/	/morabba/, /rabba/	غیر هسته	ربا برنده شد.
		هسته-افتان	مال ربا بود.
		هسته-خیزان	رفت پیش ربا؟

/t/

هجاء	كلمه	زیروبمی	جمله
/ti/	/xatti/,	غیر هسته	خطی بهتره.
		هسته-افتان	مال خطی بود.
		هسته-خیزان	گفتی خطی؟
/te/	/latte/, /patte/	غیر هسته	پته بوی بدی می ده.
		هسته-افتان	مال پته بود.
		هسته-خیزان	دوختی پته؟
/ta/	/xatta/,	غیر هسته	خطا با هم تلاقی دارن.
		هسته-افتان	مال خطا بود.
		هسته-خیزان	بنویسه زیر خطا؟

/d/

هجاء	كلمه	زیروبمی	جمله
/de/	/ɖgadde/	غیر هسته	جده بزرگه.
		هسته-افتان	مال جده بود.
		هسته-خیزان	رفت جده؟
/da/	/sadda/	غیر هسته	سدا بزرگن.
		هسته-افتان	مال سدا بود.
		هسته-خیزان	فرو ریختن سدا؟

/k/

هجاء	کلمه	زیروبمی	جمله
/ke/	/lakke/	غیر هسته	لگه بزرگه.
		هسته-افتان	مال لگه بود.
		هسته-خیزان	گفتی لگه؟
/ka/	/lakka/	غیر هسته	لگا با آب می‌رن.
		هسته-افتان	مال لگا بود.
		هسته-خیزان	ریختی رو لگا؟
/ku/	/sakku/	غیر هسته	سگو بزرگه.
		هسته-افتان	مال سگو بود.
		هسته-خیزان	رفت رو سگو؟

/q/

هجاء	کلمه	زیروبمی	جمله
/qi/	/laqqi/	غیر هسته	لقی بهتره.
		هسته-افتان	مال لقی بود.
		هسته-خیزان	گفتی لقی؟
/qe/	/taraqqe/	غیر هسته	رقه بزرگه.
		هسته-افتان	مال رقه بود.
		هسته-خیزان	رفت رقه؟
/qa/	/haqqa/	غیر هسته	حقّا به سادگی بدست نمی‌آن.
		هسته-افتان	مال حقّا بود.
		هسته-خیزان	گفتی حقّا؟

۴-۴-۲. همخوان‌های سایشی

/s/

هجاء	کلمه	زیروبمی	جمله
/se/	/dasse/	غیر هسته	دسه بلند رو بردار.
		هسته-افتان	مال دسه بود.
		هسته-خیزان	خورد تو دسه؟

/z/

هجاء	کلمه	زیروبمی	جمله
/ze/	/mazze/	غیر هسته	مزّه به غذا نداده .
		هسته-افتان	بامزّه بود.
		هسته-خیزان	گفتی بامزّه؟

۴-۴-۳. همخوان‌های انسایشی

/tʃ/

هجاء	کلمه	زیروبمی	جمله
/tʃe/	/batʃtʃe/	غیر هسته	بچه بزرگ شد.
		هسته-افتان	مال بچه بود.
		هسته-خیزان	رفت پیش بچه؟
/tʃa/	/batʃtʃa/	غیر هسته	بچا بالای خونه هستند.
		هسته-افتان	مال بچا بود.
		هسته-خیزان	رفتی پیش بچا؟

/ɖʒ/

هجاء	کلمه	زیروبمی	جمله
/ɖʒi/	/haɖʒɖʒi/	غیر هسته	حجی بزرگ شد.
		هسته-افتان	مال حجی بود.
		هسته-خیزان	گفتی حجی؟
/ɖʒe/	/zaɖʒɖʒe/	غیر هسته	ضجه بلند می‌زد.
		هسته-افتان	مال ضجه بود.
		هسته-خیزان	گفتی ضجه؟
/ɖʒa/	/laɖʒɖʒa/	غیر هسته	لجا بهتره.
		هسته-افتان	مال لجا بود.
		هسته-خیزان	گفتی لجا؟

۴-۴-۴. همخوان‌های رسا

/m/

هجاء	کلمه	زیروبمی	جمله
/mi/	/sammi/	غیر هسته	سمی بهتره.
		هسته-افتان	غذا سمی بود.
		هسته-خیزان	باشه سمی؟
/me/	/?amme/	غیر هسته	عمه برنده هستند.
		هسته-افتان	مال عمه بود.
		هسته-خیزان	رفت پیش عمه؟
/ma/	/?amma/	غیر هسته	اما به من نگو.
		هسته-افتان	مال اما بود.
		هسته-خیزان	گفتی اما؟

/n/

هجاء	کلمه	زیروبمی	جمله
/ni/	/fanni/	غیر هسته	فنی بهتره.
		هسته-افتان	مال فنی بود.
		هسته-خیزان	گفتی فنی؟
/ne/	/banne/	غیر هسته	بنه بزرگه.
		هسته-افتان	مال بنه بود.
		هسته-خیزان	خریدی بنه؟
/na/	/fanna/	غیر هسته	فنا بهتره.
		هسته-افتان	مال فنا بود.
		هسته-خیزان	گفتی همه فنا؟

/r/

هجاء	کلمه	زیروبمی	جمله
/ri/	/serri//farri//barri/	غیر هسته	سرّی بهتره.
		هسته-افتان	مال سرّی بود.
		هسته-خیزان	گفتی سرّی؟
/re/	/darre/	غیر هسته	درّه بزرگه.
		هسته-افتان	مال درّه بود.
		هسته-خیزان	رفت تو درّه؟
/ra/	/barra/	غیر هسته	برّا بزرگن.
		هسته-افتان	مال برّا بود.
		هسته-خیزان	رفت پیش برّا؟
/ro/	/darro/	غیر هسته	دررو بهتره.
		هسته-افتان	مال دررو بود.
		هسته-خیزان	گفتی دررو؟

/l/

هجاء	کلمه	زیروبمی	جمله
/le/	/kalle/	غیر هسته	کله به خوبی پخته.
		هسته-افتان	مال کله بود.
		هسته-خیزان	خریدی کله؟
/la/	/kalla/	غیر هسته	کلا با دست پاک می‌شن.
		هسته-افتان	مال کلا بود.
		هسته-خیزان	گفتی کلا؟

/j/

هجاء	کلمه	زیر و بمی	جمله
/je/	/somajje/	غیر هسته	سمیه بزرگ شد.
		هسته-افتان	مال سیمیه بود.
		هسته-خیزان	رفت پیش سیمیه؟
/ja/	/sijja/, /sajja/	غیر هسته	سیا بزرگ شده.
		هسته-افتان	مال سیا بود.
		هسته-خیزان	رفت پیش سیا؟

فصل پنج

ضبط و تهیه پیکره گفتاری

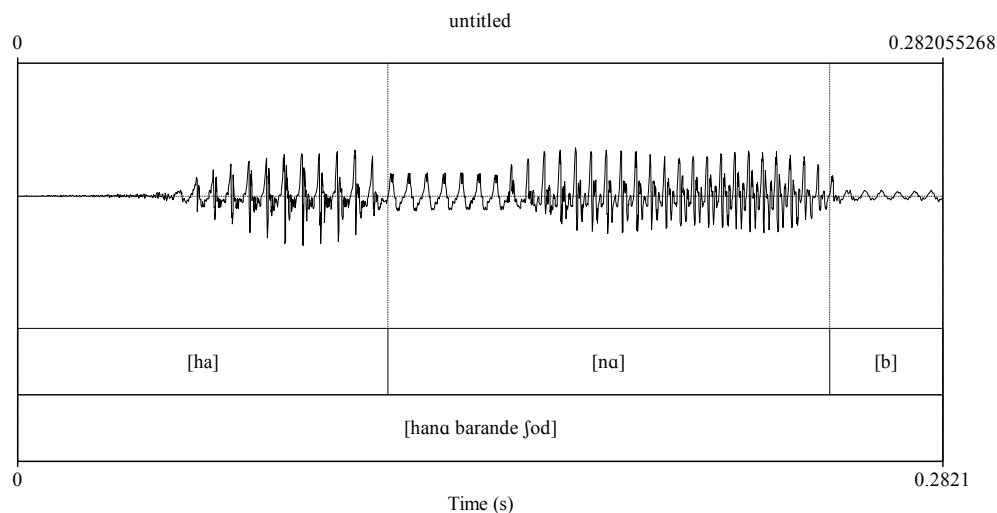
۵. مقدمه

هجاهای نوایی در چهار جایگاه بدون تکیه، تکیه بر هسته، غیرهسته و سؤالی در بافت آوایی ذکر شده در فصل چهار، توسط یک گویشور مرد با لهجه فارسی معیار (که در شهر تهران به آن صحبت می‌شود)، ضبط و ذخیره شد. ضبط داده‌ها با استفاده از یک میکروفون شور مدل SM58 در محیطی آرام و بدون هرگونه صدا صورت گرفت. در این دادگان هجایی نوایی علاوه بر هجاهای ساده فارسی هجاهای مشدد نیز ضبط و به پیکره افزوده شد.

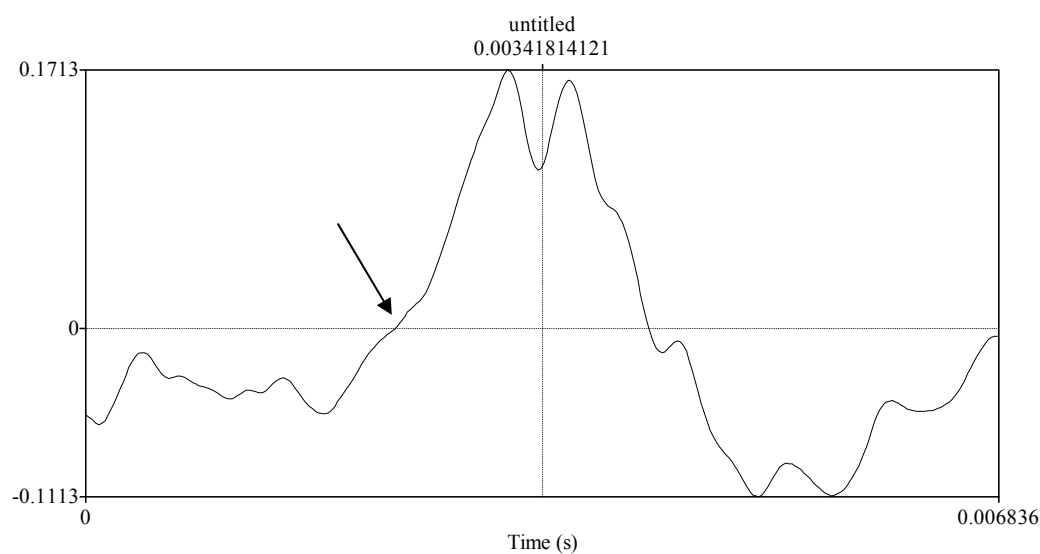
۵-۱. ضبط و تقطیع هجاها

تقطیع جملات و ذخیره هجاها (هم هجاهای ساده و هم هجاهای مشدد) در نرم‌افزار پرت (Praat) صورت گرفت. تقطیع هجاها براساس سیگنال آکوستیکی و طیف‌نگاشت انجام شد (شکل ۵-۱). بدین صورت که مرز هجاها با یافتن مرز آواها براساس تغییرات پالس حنجره مشخص شد و محل تقطیع همخوان و سپس واکه هجا زمانی در نظر گرفته شد که سیگنال آکوستیکی خط صفر را قطع می‌کند یا انرژی دامنه در کمترین مقدار خود نزدیک خط صفر تولید می‌شود زیرا برطبق نظر اسلامی و همکاران (۱۳۸۸) کم بودن انرژی دامنه در این نقطه سبب می‌شود تا اتصال بعدی هجاها بهتر و دقیقتر صورت گیرد. محل قطع به گونه‌ای است که عدم پیوستگی و اعوجاج طیفی بین قطعات متوالی حداقل است. کم بودن دامنه در مرزها تا حدودی، ناپیوستگی در محل‌های اتصال را کاهش می‌دهد. به‌طور خاص در مورد هجاهایی که با واج‌های واکدار آغاز شده و یا به پایان می‌رسند، مرز هجاها در محل عبور سیگنال از خط صفر گذارده شده و پالسهای واک به صورت کامل در محدوده هجا قرار دارد. به عبارتی، مرز در ابتدای اولین و انتهای آخرین پالس واک واقع گشته است. این شیوه تعیین ابتدا و انتهای هجا سبب می‌شود هنگامی که در سنتز گفتار، واج‌های مرزی، واکدار هستند، مسئله عدم تطابق در سیگنال رخ ندهد و مشکلی از این لحاظ در کیفیت پیش نیاید.

در شکل ۱-۵ شیوه تقطیع هجای [na] در بافت جمله "حنا برنده شد" در جایگاه تکیه‌بر غیرهسته نشان داده شده است. همان طور که گفته شد محل تقطیع در محل عبور سیگنال از خط صفر تعیین شد (شکل ۲-۵).



شکل ۱-۵: شیوه تقطیع هجایی نوایی



شکل ۲-۵: محل عبور سیگنال از خط صفر

در مورد هجاهایی که واج انتهایی آنها واکدار است، نقطه‌ای به عنوان پایان هجا در نظر گرفته شده که در آن زمان در سیگنال، ارتعاش تارآواها یعنی حالت واکداری به پایان رسیده است.

هجاهای حاصل در پرونده های جداگانه بر حسب جایگاه نوایی در یک دیسک فشرده ذخیره شدند تا برای تغییرات بعدی وارد برنامه تبدیل متن به گفتار شود. در برنامه تبدیل متن با قرار دادن هجاها در جملات میزان کیفیت آنها سنجیده می شود و در صورت لزوم تغییراتی در پارامترهای آکوستیکی آنها که در ادامه مورد تحلیل قرار می گیرند، صورت خواهد گرفت.

۲-۵. پارامترهای آکوستیکی

تکیه عبارت است از برجستگی یک هجا نسبت به هجای مجاور. به لحاظ آکوستیکی چهار مشخصه آوایی سبب برجستگی یک هجا در مقابل سایر هجا می شوند که عبارتند از زیروبمی، شدت انرژی یا بلندی هجا، دیرش یا طول هجا و کیفیت واکه^{۱۳۶} هجا (Roach, 2000, 94-95).

در این پژوهش سه پارامتر آکوستیکی دیرش واکه هجا، میانگین زیروبمی و میانگین شدت انرژی واکه هجا مورد تحلیل قرار گرفت تا روشن شود که آیا جایگاه های نوایی سبب تولید هجاهایی با مقادیر آکوستیکی و در نتیجه شنیداری و درکی متفاوت می شوند یا خیر. برای این منظور مقادیر آکوستیکی ذکر شده در نرم افزار پرت اندازه گیری و به لحاظ آماری بررسی شد. تعداد نمونه های مورد مطالعه ۱۶۵۶ (۶ واکه X ۲۳ همخوان X ۴ جایگاه نوایی X ۳ پارامتر آکوستیکی) است که به طور دستی و براساس معیارهای ذکر شده ابتدا مورد تقطیع هجایی و سپس مورد اندازه گیری و تحلیل آماری (تحلیل واریانس) قرار گرفتند.

۱-۲-۵. دیرش واکه

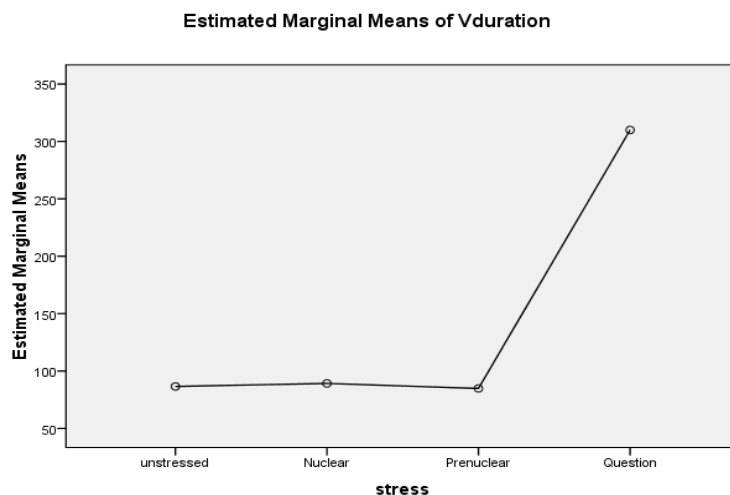
دیرش واکه عبارت است از مدت زمانی که تولید واکه طول کشیده است. دیرش واکه از جمله پارامترهایی است که سبب برجستگی هجا و احساس تکیه می شود. این پارامتر به طور دستی از اولین پالس واکه تا آخرین پالس بر حسب میلی ثانیه بر روی سیگنال محاسبه شد.

جدول ۵-۱ میانگین دیرش واکه هجا بر حسب جایگاه نوایی در کل داده ها را نشان می دهد. (prenuclear همان جایگاه نوایی غیرهسته است زیرا این جایگاه پیش از هسته مورد بررسی قرار گرفت)

جدول ۵-۱: میانگین دیرش واژه هجا بر حسب جایگاه نوایی

stress	Mean(ms)	N	Std. Deviation
unstressed	86.57	138	19.816
Nuclear	89.20	138	18.412
Prenuclear	84.82	138	23.326
Question	309.94	138	42.345
Total	142.63	552	100.565

براساس این جدول میانگین دیرش هجا در جمله سؤالی بسیار بیشتر از سه جایگاه دیگر است. این تفاوت در شکل ۵-۳ نیز به خوبی نشان دیده می شود.



شکل ۵-۳: نمودار دیرش واژه هجا بر حسب جایگاه نوایی

دیرش واژه در هجای سؤالی با دیرش آن در سه هجای دیگر تفاوت معنادار دارد ($p=0$) اما سه جایگاه نوایی دیگر با یکدیگر تفاوت معنایی ندارند ($p>0$). سلسله مراتب کاهش دیرش در هجاها بر حسب جایگاه نوایی به صورت زیر است:

تکیه بر غیر هسته > بی تکیه > تکیه بر هسته > سؤالی

این نتایج نشان می دهد گویشوران مشخصه دیرش واژه را تنها برای انتقال برجستگی در هجاهای CV سؤالی به کار می برند. بنابراین در برنامه تبدیل متن به گفتار تنها با تغییر پارامتر دیرش نمی توان

جملاتی با نوای طبیعی گفتار بدست آورد به نحوی که سایر پارامترهای برجستگی نیز تأثیر بسزایی دارند.

۵-۲-۲. میانگین زیروبمی

زیروبمی یک پارامتر شنیداری است که با محاسبه نرخ ارتعاش تارهای صوتی یا فرکانس پایه (f_0) بدست می‌آید. در واقع گویشور با تغییر ارتعاش تارهای صوتی در گفتار خود سبب احساس زیروبمی‌های متفاوتی در شنونده می‌شود. محاسبه این مشخصه آوایی، با انتخاب سیگنال واکه در نرم‌افزار پرت امکان‌پذیر است. در پژوهش حاضر با انتخاب قطعه واکه به استثنای دو پالس اول (برای کاهش تأثیری که همخوان بافت ماقبل بر آغاز ارتعاش تارهای صوتی و شدت آن دارد) میانگین زیروبمی برحسب هرتز بدست آمد.

نتایج نشان می‌دهد که میانگین زیروبمی یا فرکانس پایه در هجاهای تکیه‌بر غیرهسته همواره به طور معناداری بیشتر از هجاهای بی‌تکیه، تکیه‌بر هسته و سؤالی است (جدول ۵-۲). تحلیل واریانس یک‌طرفه و تست‌های تعقیبی تام‌هن نشان می‌دهد که تفاوت هر چهار جایگاه نوایی با یکدیگر معنادار است ($P=0$).

جدول ۵-۲: میانگین فرکانس پایه برحسب جایگاه نوایی

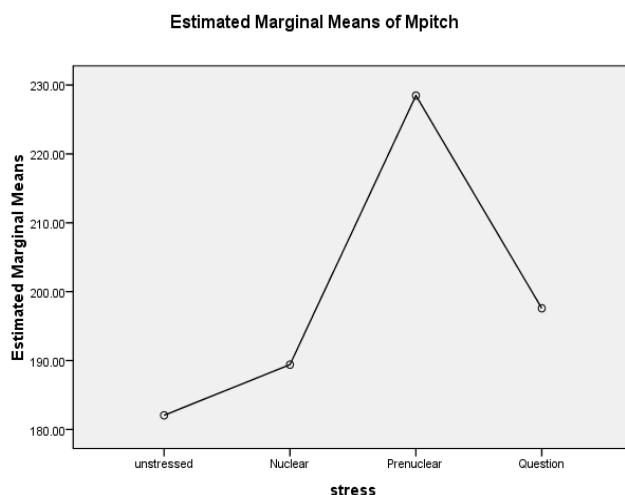
stress	Mean(Hz)	N	Std. Deviation
unstressed	1.8205E2	138	13.05636
Nuclear	1.8941E2	138	16.35340
Prenuclear	2.2847E2	138	22.99031
Question	1.9758E2	138	18.30889
Total	1.9938E2	552	25.22864

سلسله‌مراتب کاهش زیروبمی در هر شش واکه برحسب جایگاه نوایی به صورت زیر است:

بی‌تکیه > تکیه‌بر هسته > سؤالی > تکیه‌بر غیرهسته

همان‌طور که شکل ۵-۴ نیز نشان می‌دهد هجای بی‌تکیه با میانگین فرکانس پایه کمتری نسبت به سه هجای تکیه‌بر هسته، غیرهسته و سؤالی تولید شده است زیرا فاقد تکیه اصلی گروه نوایی است. سه هجای تکیه‌بر نیز با یکدیگر تفاوت دارند: میانگین فرکانس پایه در هجای تولیدشده در جایگاه

غیرهسته بیشتر از دو جایگاه نوایی دیگر است زیرا این هجا با تکیه زیروبمی h بالا (خیزان) تولید شده است و بنابراین با افزایش ارتعاش تارهای صوتی تولید می‌شوند. هجای تولید شده در جایگاه سؤالی پس از آن با فرکانس بیشتری نسبت به جایگاه هسته تولید شده زیرا این هجا نیز با نواخت مرزی H% یا بالا تولید شده است. همان طور که در بخش ۵-۱ دیدیم تأثیر جایگاه نوایی بر هجاهای پرسشی با افزایش پارامتر دیرش واکه روی می‌دهد اما بر اساس این نتایج به نظر می‌رسد افزایش دیرش واکه همراه با افزایش نرخ ارتعاش تارهای صوتی است هرچند میزان این ارتعاش کمتر از جایگاه تکیه‌بر غیرهسته است. هجاهای تولید شده در جایگاه تکیه‌بر هسته با نرخ ارتعاش کمتری نسبت به سایر جایگاه‌های تکیه‌بر تولید شده اند زیرا تکیه زیروبمی این هجاها پایین l (افتان) است و این هجاها با کاهش ارتعاش تارهای صوتی تولید می‌شوند.



شکل ۴-۵: نمودار تغییرات فرکانس پایه برحسب جایگاه نوایی

براساس این نتایج می‌توان نتیجه گرفت که جایگاه نوایی بر مشخصه آوایی فرکانس پایه تأثیر بسزایی دارد. به نظر می‌رسد گویشوران میزان ارتعاش تارهای صوتی را هنگام تولید هجاهای CV برحسب جایگاه نوایی هجا تغییر می‌دهد به طوری که این تغییرات معنادار است و انتظار می‌رود با گنجاندن این نوع هجاهای حساس به بافت نوایی در برنامه تبدیل متن به گفتار سبب افزایش کیفیت گفتار تولید شده و از سویی سبب کاهش فرایندهای لازم برای بدست آوردن نوای طبیعی گفتار شود.

در ادامه میانگین فرکانس پایه (برحسب هرتز) برای هر یک از شش واکه فارسی نیز در هر یک از چهار جایگاه نوایی ارائه شده است. هرچند در هر شش واکه، سلسله مراتب ذکر شده در بالا مشاهده می‌شود اما تفاوت معنایی لزوماً در همه موارد وجود ندارد.

واکه /a/

در واکه /a/ تحلیل واریانس یک طرفه نشان می دهد که تفاوت به طور کلی برحسب جایگاه معنادار است ($p \leq 0.05$) اما بین فرکانس پایه هجاهای تولید شده در جایگاه تکیه بر هسته با بی تکیه تفاوت معناداری مشاهده نمی شود ($p > 0.05$).

vowel [a]

stress	Mean pitch (Hz)	N	Std. Deviation
unstressed	1.7585E2	23	11.43930
Nuclear	1.8294E2	23	11.21151
Prenuclear	2.1494E2	23	17.77339
Question	1.9339E2	23	24.91339
Total	1.9178E2	92	22.55207

واکه /a/

در این واکه نیز تحلیل واریانس یک طرفه نشان می دهد که تفاوت به طور کلی برحسب جایگاه معنادار است ($p \leq 0.05$) اما بین فرکانس پایه هجاهای تولید شده در جایگاه تکیه بر هسته با سؤالی تفاوت معنادار مشاهده نمی شود ($p > 0.05$).

vowel [a]

stress	Mean pitch (Hz)	N	Std. Deviation
unstressed	1.7758E2	23	10.38823
Nuclear	1.8904E2	23	18.49641
Prenuclear	2.2556E2	23	19.74476
Question	1.9748E2	23	18.25216
Total	1.9742E2	92	24.51444

واکه /e/

در این واکه نیز تحلیل واریانس یک طرفه نشان می دهد که تفاوت به طور کلی برحسب جایگاه معنادار است ($p \leq 0.05$) اما بین فرکانس پایه هجاهای تولید شده در جایگاه تکیه بر هسته با جایگاه بی تکیه و نیز بین جایگاه تکیه بر هسته با جایگاه تکیه بر سؤالی تفاوت معناداری مشاهده نمی شود ($p > 0.05$).

vowel [e]

stress	Mean pitch (Hz)	N	Std. Deviation
unstressed	1.7958E2	23	11.85943
Nuclear	1.8489E2	23	18.86691
Prenuclear	2.2482E2	23	19.58349
Question	1.9323E2	23	16.14949
Total	1.9563E2	92	24.22593

واکه /i/

در این واکه نیز تحلیل واریانس یک طرفه نشان می دهد که تفاوت به طور کلی برحسب جایگاه معنادار است ($p \leq 0.05$) اما تنها تفاوت معنادار بین تکیه بر هسته با سؤالی است ($p \leq 0.05$) و سایر جایگاه ها با یکدیگر تفاوت معناداری ندارند ($p > 0.05$).

vowel [i]

stress	Mean pitch (Hz)	N	Std. Deviation
unstressed	1.9628E2	23	10.12016
Nuclear	1.9711E2	23	20.35706
Prenuclear	2.3878E2	23	33.91648
Question	1.9713E2	23	17.38315
Total	2.0733E2	92	28.45387

واکه /o/

در این واکه نیز تحلیل واریانس یک طرفه نشان می دهد که تفاوت به طور کلی برحسب جایگاه معنادار است ($p \leq 0.05$). علاوه بر آن هر چهار جایگاه برحسب این پارامتر با یکدیگر تفاوت معنادار دارند ($p \leq 0.05$).

vowel [o]

stress	Mean pitch (Hz)	N	Std. Deviation
unstressed	1.7909E2	23	10.04198
Nuclear	1.8805E2	23	11.62955
Prenuclear	2.2776E2	23	16.08445
Question	2.0213E2	23	14.79049
Total	1.9926E2	92	22.68195

واکه /u/

در این واکه نیز تحلیل واریانس یک طرفه نشان می دهد که تفاوت به طور کلی برحسب جایگاه معنادار است ($p \leq 0.05$) اما بین فرکانس پایه هجاهای تولید شده در جایگاه تکیه بر هسته با جایگاه تکیه بر سؤالی تفاوت معناداری مشاهده نمی شود ($p > 0.05$).

vowel [u]

stress	Mean pitch (Hz)	N	Std. Deviation
unstressed	1.8394E2	23	13.66816
Nuclear	1.9442E2	23	11.68392
Prenuclear	2.3896E2	23	18.77456
Question	2.0212E2	23	16.44413
Total	2.0486E2	92	25.74220

تجزیه و تحلیل مشخصه میانگین زیروبمی نشان می دهد که جایگاه های نوایی بر هجا تأثیرهای متفاوتی دارند. به عبارتی، تأثیر جایگاه نوایی بر واکه هجا وابسته به بافت واکه ای هجاست و بنابراین از واکه ای به واکه دیگر متفاوت است. از این رو به نظر می رسد استخراج هجاهای نوایی از هر بافت واکه ای و هر جایگاه نوایی مجزا برای تولید گفتار طبیعی تر در برنامه تبدیل متن به گفتار ضروری است.

۵-۲-۳. میانگین شدت انرژی

شدت انرژی نیز همانند زیروبمی یک مشخصه شنیداری است که گوینده با تغییر شدت دامنه انرژی آواها میزان آن را تغییر می دهد. در واقع گوینده با تغییر بلندی صدا، میزان دامنه انرژی آن را به لحاظ آکوستیکی و شدت آن را به لحاظ شنیداری تغییر می دهد. محاسبه این مشخصه نیز همانند زیروبمی با انتخاب قطعه واکه به جز دو پالس اول آن برحسب دسیبل (dB) صورت گرفت.

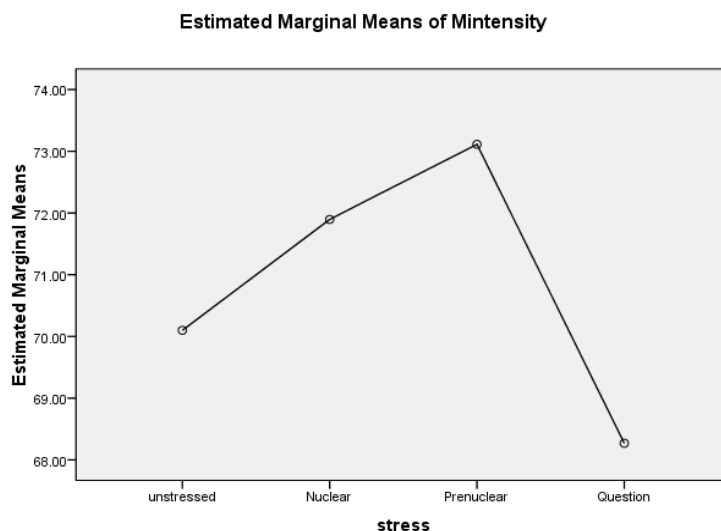
جدول ۵-۳ نشان می دهد که شدت انرژی در هجاهای تولید شده در جایگاه غیرهسته بیشترین است. همان طور که در بخش قبل دیدیم میانگین فرکانس پایه در این هجاها نیز بیشتر از سایرین است. به نظر می رسد که افزایش فرکانس پایه به همراه شدت انرژی سبب تشکیل الگوی نوایی این هجاها به صورت نواخت بالا یا h می شود.

اما در سلسله مراتب شدت انرژی برخلاف سلسله مراتب زیروبمی، شدت انرژی در هجاهای سؤالی کمتر از سه هجای دیگر است. به نظر می رسد با افزایش پارامتر زمانی دیرش واکه، پارامتر دامنه انرژی کاهش پیدا می کند. برطبق جدول ۵-۱ هجای تکیه بر غیرهسته کمترین میزان دیرش را دارد اما با بیشترین میزان شدت انرژی (جدول ۵-۳) تولید شده است.

جدول ۵-۳: میانگین شدت انرژی واکه برحسب جایگاه نوایی

stress	Mean(dB)	N	Std. Deviation
unstressed	70.0987	138	3.43955
Nuclear	71.8958	138	2.97021
Prenuclear	73.1133	138	2.92414
Question	68.2685	138	3.54903
Total	70.8441	552	3.70946

تحلیل واریانس یکطرفه تفاوت معناداری برحسب شدت انرژی بین هر یک از چهار جایگاه نوایی نشان می دهد ($p=0$). تستهای تعقیبی تامهن نشان می دهد هر یک از چهار جایگاه با یکدیگر تفاوت معنادار دارند ($p=0$). میانگین شدت انرژی در هجای بی تکیه کمتر از هجای تکیه بر هسته است (شکل ۵-۵). شدت انرژی بیشتر به همراه نرخ بالاتر ارتعاش تارهای صوتی سبب تمایز هجاهای تکیه بر هسته از هجاهای بی تکیه که شدت انرژی و فرکانس پایه کمتری دارند، می شود. بنابراین برای تولید نوای طبیعی گفتار بین هجاهای بی تکیه و تکیه بر در برنامه های تبدیل متن به گفتار لازم است هجاها براساس هر دو پارامتر فرکانس پایه و شدت انرژی تنظیم شوند حال آن که به نظر می رسد در بیشتر این برنامه تنها تغییراتی در فرکانس پایه صورت می گیرد که در این صورت نوای طبیعی گفتار به درستی و به طور کامل به دست نمی آید.



شکل ۵-۵: نمودار میانگین شدت انرژی برحسب جایگاه نوایی

سلسله مراتب کاهش میانگین شدت انرژی به صورت زیر است:

سؤالی > بی تکیه > تکیه بر هسته > تکیه بر غیرهسته

این سلسله مراتب در همه واژه‌های زیر مشاهده می‌شود (به استثنای /u/ که در آن شدت انرژی در هجای سؤالی بیشتر از هجای بی تکیه است). هرچند در هر شش واژه، سلسله مراتب ذکر شده در بالا مشاهده می‌شود اما تفاوت معنایی لزوماً در همه موارد وجود ندارد.

واژه /a/

در این واژه تحلیل واریانس یک طرفه نشان می‌دهد که تفاوت به طور کلی برحسب جایگاه و براساس پارامتر شدت دامنه انرژی معنادار است ($p \leq 0.05$) اما بین شدت دامنه انرژی هجاهای تولید شده در جایگاه تکیه بر هسته با جایگاه بی تکیه تفاوت معناداری مشاهده نمی‌شود ($p > 0.05$).

vowel [a]

stress	Mean intensity (dB)	N	Std. Deviation
unstressed	71.5526	23	2.87513
Nuclear	73.1265	23	2.07499
Prenuclear	74.8852	23	1.92075
Question	68.0735	23	3.19905
Total	71.9095	92	3.57340

واکه /a/

در این واکه، تحلیل واریانس یک طرفه نشان می دهد که تفاوت به طور کلی برحسب جایگاه و براساس پارامتر شدت دامنه انرژی معنادار است ($p \leq 0.05$) اما بین شدت دامنه انرژی هجاهای تولید شده در جایگاه تکیه بر هسته با جایگاه بی تکیه تفاوت معناداری مشاهده نمی شود ($p > 0.05$).

vowel [a]

stress	Mean intensity (dB)	N	Std. Deviation
unstressed	73.3978	23	2.67195
Nuclear	73.1800	23	3.49347
Prenuclear	75.0009	23	1.89081
Question	69.8383	23	2.83508
Total	72.8542	92	3.32403

واکه /e/

در این واکه، تحلیل واریانس یک طرفه نشان می دهد که تفاوت به طور کلی برحسب جایگاه و براساس پارامتر شدت دامنه انرژی معنادار است ($p \leq 0.05$) اما بین شدت دامنه انرژی هجاهای تولید شده در جایگاه تکیه بر هسته با جایگاه تکیه بر غیرهسته تفاوت معناداری مشاهده نمی شود ($p > 0.05$).

vowel [e]

stress	Mean intensity (dB)	N	Std. Deviation
unstressed	69.6165	23	2.71814
Nuclear	72.2296	23	2.78351
Prenuclear	72.7422	23	2.67555
Question	67.3717	23	3.39576
Total	70.4900	92	3.58817

واکه /i/

در این واکه، تحلیل واریانس یک طرفه نشان می دهد که تفاوت به طور کلی برحسب جایگاه و براساس پارامتر شدت دامنه انرژی معنادار است ($p \leq 0.05$) اما بین شدت دامنه انرژی هجاهای تولید

شده در جایگاه تکیه بر هسته با جایگاه تکیه بر غیرهسته و نیز بین جایگاه سؤالی با جایگاه بی تکیه تفاوت معناداری مشاهده نمی شود ($p > 0.05$).

vowel [i]

stress	Mean intensity (dB)	N	Std. Deviation
unstressed	66.9983	23	2.34660
Nuclear	70.8557	23	2.43730
Prenuclear	71.7735	23	2.23100
Question	66.4117	23	3.13035
Total	69.0098	92	3.44384

واکه /o/

در این واکه، تحلیل واریانس یک طرفه نشان می دهد که تفاوت به طور کلی برحسب جایگاه و براساس پارامتر شدت دامنه انرژی معنادار است ($p \leq 0.05$) اما بین شدت دامنه انرژی هجاهای تولید شده در جایگاه تکیه بر هسته با جایگاه غیرهسته تفاوت معناداری مشاهده نمی شود ($p > 0.05$). بین جایگاه بی تکیه با جایگاه هسته و نیز بین جایگاه بی تکیه با جایگاه غیرهسته تفاوت معناداری مشاهده نمی شود ($p > 0.05$).

vowel [o]

stress	Mean intensity (dB)	N	Std. Deviation
unstressed	71.2865	23	2.31783
Nuclear	71.9391	23	2.42519
Prenuclear	73.6817	23	2.59068
Question	69.0209	23	3.94466
Total	71.4821	92	3.30478

واکه /u/

در این واکه، تحلیل واریانس یک طرفه نشان می دهد که تفاوت به طور کلی برحسب جایگاه و براساس پارامتر شدت دامنه انرژی معنادار است ($p \leq 0.05$) اما تنها تفاوت معنادار بین شدت دامنه انرژی هجاهای تولید شده در جایگاه بی تکیه با جایگاه غیرهسته مشاهده می شود ($p > 0.05$).

vowel [u]

stress	Mean intensity (dB)	N	Std. Deviation
unstressed	67.7404	23	2.97123
Nuclear	70.0439	23	3.27350
Prenuclear	70.5961	23	3.33839
Question	68.8948	23	3.90015
Total	69.3188	92	3.50957

تجزیه و تحلیل مشخصه میانگین شدت دامنه انرژی نشان می دهد که جایگاه های نوایی بر واکه هجا تأثیرهای متفاوتی دارند. به عبارتی، تأثیر جایگاه نوایی بر واکه هجا وابسته به بافت واکه ای هجاست و بنابراین از واکه ای به واکه دیگر متفاوت است. از این رو به نظر می رسد استخراج هجاهای نوایی از هر بافت واکه ای و هر جایگاه نوایی مجزا برای تولید گفتار طبیعی تر در برنامه تبدیل متن به گفتار ضروری است.

۳-۵. تحلیل نتایج برای برنامه تبدیل متن به گفتار

نتایج آماری حاصل از تحلیل مقادیر آکوستیکی نشان می دهد که نوای گفتار فارسی تحت تأثیر جایگاه نوایی هجا است. هجاهای CV که در پایان جملات سؤالی تولید می شوند، با کمترین میزان دیرش نسبت به هجاهایی که در جایگاه بی تکیه کلمات یا جایگاه تکیه بر هسته و غیرهسته جمله قرار دارند، تولید می شوند. پارامتر آکوستیکی دیرش واکه هرچند بین سه جایگاه بی تکیه، تکیه بر هسته و غیرهسته تفاوت ایجاد می کند، اما این تفاوت معنادار نیست.

مشخصه آوایی فرکانس پایه که مهمترین شاخص زیرویمی گفتار قلمداد می شود، تفاوت معناداری بر جایگاه های نوایی و در نتیجه بر نوای گفتار دارد. این مشخصه به طور کلی سبب تمایز معنادار هجاها برحسب جایگاه نوایی و به طور ویژه سبب تمایز هجاهای تکیه بر (هسته، غیرهسته و سؤالی) از هجای بی تکیه می شود. علاوه بر آن از آنجا که دو جایگاه تکیه بر غیرهسته و سؤالی (هر دو با نواخت خیزان h) با زیرویمی بیشتر نسبت به جایگاه تکیه بر هسته با نواخت افتان (l) و جایگاه بی تکیه تولید می شوند، می توان گفت این مشخصه سبب تمایز نواخت بالا از نواخت پایین و حالت بی تکیه می شود.

مشخصه آکوستیکی دیگر میانگین شدت دامنه انرژی واکه هجاست که در جایگاه های نوایی مختلف تأثیر معناداری بر نوای طبیعی گفتار دارد. به نظر می رسد این پارامتر به همراه پارامتر دیرش واکه

همبسته آوایی دو جایگاه نوایی تکیه بر غیرهسته و سؤالی باشد به طوری که هجای سؤالی با بیشترین دیرش واکه و کمترین شدت انرژی و هجای تکیه بر با کمترین دیرش و بیشترین شدت انرژی تولید می شود .

سلسله مراتب کاهش هر یک از سه مقدار آوایی به صورت زیر است:

دیرش واکه > تکیه بر غیرهسته > بی تکیه > تکیه بر هسته > سؤالی
 میانگین فرکانس پایه > بی تکیه > تکیه بر هسته > سؤالی > تکیه بر غیرهسته
 میانگین شدت انرژی دامنه > سؤالی > بی تکیه > تکیه بر هسته > تکیه بر غیرهسته

براساس این سلسله مراتب می توان تفاوت پارامترهای واکه هجا را برحسب جایگاه نوای به صورت کیفی زیر نشان داد (جدولی ۴-۵).

جدول ۴-۵: مقادیر آکوستیکی جایگاه های نوایی به صورت کیفی

جایگاه نوایی	دیرش واکه	میانگین زیروبمی	میانگین شدت انرژی
تکیه بر غیرهسته	کمترین	بیشترین	بیشترین
تکیه بر هسته	بالا	کم	بالا
سؤالی	بیشترین	کم	کمترین
بی تکیه	کم	کمترین	کم

جدول ۴-۵ نشان می دهد که هر پارامتر آکوستیکی به صورت پیوستاری است که مقدار آن بسته به جایگاه های نوایی هجا متفاوت است و بنابراین نوای طبیعی گفتار حاصل کم و زیاد شدن مقادیر این مشخصه های آوایی در هر هجای درون جمله است.

اما همان طور که در فصلهای پیشین نیز ذکر شد، در برنامه هایی که پیکره گفتاری حامل ساختار نوایی زبان نیست و یا با در نظر گرفتن حداقل ساخت ها تهیه می شوند، در گفتار تولید شده اغلب مشکل نوای طبیعی گفتار مشاهده می شود. در برخی از این برنامه ها سعی می شود با گنجاندن قواعدی که دو پارامتر دیرش و فرکانس پایه را تغییر می دهد، تاحدی این مشکل مرتفع شود اما از آنجا که پردازش و تغییر و تنظیم همه این مشخصات ممکن است بر گفتار حاصل تأثیر گذاشته، سبب

غیرطبیعی بودن آن شود، بهتر است هجاهای نوایی با حفظ مشخصه های خود از درون جمله استخراج و در برنامه تبدیل متن به گفتار ذخیره شوند.

نتایج این پژوهش نیز مانند مطالعاتی که در فصل های پیشین به ویژه فصل ۳ به آن ها اشاره شد، نشان می دهد که مشخصات آوایی دیگری نیز به جز دو پارامتر دیرش واکه و فرکانس پایه مانند شدت انرژی بر مشخصه برجستگی و نوای گفتار تأثیر گذار هستند. در این پژوهش تنها این سه پارامتر مورد بررسی آکوستیکی قرار گرفت و همان طور که نتایج نشان می دهد برای بدست نوای طبیعی گفتار باید این سه پارامتر حداقل در نظر گرفته شود هرچند که رفتار این سه پارامتر از واکه به واکه ای دیگر و حتی ممکن است از بافت همخوانی به بافت همخوانی دیگر متفاوت باشد. در این پژوهش نیز تنها تأثیر جایگاه نوایی بر مشخصه های واکه ای به طور کلی و بی آنکه به مشخصات هر واکه جدا از دیگری در نظر گرفته شود (تنها میانگین مقادیر فرکانس پایه و شدت انرژی برای هر واکه ارائه شد) تحلیل آکوستیکی صورت گرفت.

در نهایت با توجه به نتایج می توان چنین نتیجه گرفت که نوای طبیعی گفتار حاصل بر هم کنش مشخصات آوایی مختلفی است که ممکن است گنجاندن همه آنها در برنامه های تبدیل متن به گفتار به شکل قاعده و پردازش های مختلف امکانپذیر نباشد زیرا همان طور که گفته شد علاوه بر مشخصات آکوستیکی مختلف، متغیرهای دیگری نیز مانند بافت واکه ای و همخوانی نیز وجود دارند که هر یک مشخصات نوایی مختص به خود را سبب می شوند. از این رو استخراج این بافت ها از گفتار طبیعی و ذخیره آنها به صورت جایگاه های نوایی مختلف سبب تولید گفتاری طبیعی تر و نزدیک تر به گفتار انسانی می شود.

فصل شش

نتیجه گیری

مهمترین شاخصه یک برنامه تبدیل متن به گفتار، میزان طبیعی بودن گفتار حاصل است. بدین صورت که هر قدر گفتار تولیدشده بوسیله نرم افزار طبیعی تر و به گفتار انسان شبیه تر باشد، آن برنامه از کیفیت بالاتری برخوردار بوده و بنابراین کاربرد آن به مراتب بیشتر خواهد بود.

آزمایش های ادراکی که از سوی افرادی نظیر یون (۲۰۰۷) بر روی زبان کره ای و یا صادقی (۱۳۸۹) بر روی زبان فارسی صورت گرفت نشان می دهد که هرگاه واحدهای بازسازی برنامه تبدیل متن به گفتار با توجه به بافت نوایی جمله در کنار یکدیگر قرار گیرند، گفتار حاصل بسیار طبیعی تر است و شنوندگان و آزمودنی ها واحدهای بازسازی حساس به بافت نوایی را نسبت به واحدهای فاقد این ویژگی ترجیح می دهند و آنها را طبیعی تر تشخیص می دهند.

در برنامه های تبدیل متن به گفتار معمولاً پردازش نوایی با تغییر فرکانس پایه و دیرش واحدهای بازسازی صورت می گیرد، که گفتار حاصل تاحدی ویژگی های نوایی جمله را دارد اما با توجه به آزمون های ادراکی و پژوهش های جدیدی که در زمینه بازسازی واحدهای گفتاری صورت گرفته به نظر می رسد بهتر است واحدهای بازسازی از بافت های نوایی مختلف و حساس به ساختار نوایی انتخاب شده و سپس برای تولید گفتار هم گذاری شوند. در این حالت گفتار حاصل طبیعی تر خواهد بود.

ساختار نوایی زبان فارسی نشان می دهد که هجاها تحت تأثیر نواخت های تکیه زیروبمی L و H و نواخت های مرزنامی جمله H% و L% و گروه تکیه ای l و h قرار دارند. از این رو تهیه واحدهای بازسازی یعنی هجا براین اساس مدنظر قرار گرفت و پیکره نوشتاری هجای CV فارسی براین اساس تهیه شد. علاوه بر آن از آنجا که مشخصه های صوتی هجاهای حاوی همخوان های مشدد متفاوت است، بنابراین این پیکره حاوی این هجاها در بافت های نوایی مذکور است.

از آنجا که پرداختن به همه تفاوت‌های نوایی هجاها و لحاظ کردن همه این تفاوت در برنامه بازسازی گفتار پیچیده و دشوار است، بنابراین هجاهای مورد نظر تنها از چهار بافت نوایی زیر استخراج می‌شوند:

- ۱- هجای بی‌تکیه واژگانی
- ۲- هجای تکیه‌بر واژگانی، تکیه زیروبمی غیرهسته
- ۳- هجای تکیه‌بر واژگانی، تکیه زیروبمی هسته، نواخت مرزی افتان
- ۴- هجای تکیه‌بر واژگانی، تکیه زیروبمی هسته، نواخت مرزی خیزان

در زبان فارسی ۲۳ همخوان و ۶ واکه وجود دارد که از ترکیب همخوان‌ها با واکه‌ها $۱۳۸ (۶ = ۱۳۸ \times ۲۳)$ هجای CV بدست می‌آید که با احتساب چهار بافت نوایی، تعداد نمونه‌ها به $۵۵۲ (۴ = ۵۵۲ \times ۱۳۸)$ خواهد رسید. تولید نمونه‌ها در گفتار پیوسته و در درون جمله صورت می‌گیرد.

بافت تولید هجاهای نوایی نیز کنترل خواهد شد بدین صورت که بافت واکه‌ای قبل یا بعد هر هجای نوایی، واکه‌های افتاده /a/ یا /a/ خواهد بود زیرا مشخصه‌های آکوستیکی همخوان‌ها در بافت واکه‌های افتاده کمتر تحت تأثیر هم‌تولیدی قرار می‌گیرد. در بافت همخوانی قبل یا بعد هجای نوایی، همخوان واگذار لبی /b/ تولید خواهد شد. مشخصه‌های تولیدی همخوان‌های لبی نیز تأثیر کمتری بر واکه هجای نوایی دارد.

هجاهای نوایی در درون کلمه و جمله تولید می‌شوند. اما از آنجا که ممکن است کلمه‌ای در فارسی وجود نداشته باشد که بافت واکه‌ای و همخوانی شرح داده‌شده در بالا را داشته باشد، بنابراین برخی از هجاهای نوایی در درون کلمات بی‌مفهوم و یا جملات بی‌معنا گنجانده شده‌اند.

هجاهای نوایی در کلمات دوهجایی تولید می‌شوند زیرا اگر تعداد هجاها افزایش پیدا کند، به‌ویژه بر هجای بی‌تکیه در آغاز کلمه تأثیر گذاشته و سبب تضعیف بیش‌ازحد آن می‌شود. در کلمات تکیه‌بر نیز با ایجاد تکیه دومین و یا سومین در کلماتی که بیش از دو هجا دارند، سبب کاهش و تضعیف تکیه در هجای تکیه‌بر می‌شود.

پیکره متنی هجاهای مشددی (هجای CV مشدد) تنها هجاهای را شامل می‌شود که احتمال وقوع بیشتری در زبان فارسی دارند. براین اساس تعداد این نمونه‌ها ۹۶ جمله است که در سه موضع تکیه بر

واژگانی (۱) تکیه زیروبمی غیر هسته و (۲) تکیه زیروبمی هسته و نواخت مرزی افتان و (۳) تکیه زیروبمی هسته و نواخت مرزی خیزان تولید می‌شوند.

ضبط و تهیه پیکره گفتاری هجاهای ساده و مرکب براساس سیگنال آکوستیکی و طیف‌نگاشت و با لحاظ معیارهای مشخص صورت گرفت و سپس هجاهای حاصل در پرونده‌های مختلف مبتنی بر جایگاه نوایی ذخیره شد. این داده‌ها می‌توانند برای استفاده در برنامه تبدیل متن به گفتار مورد استفاده قرار گیرد.

تجزیه و تحلیل هجاهای نوایی در این پژوهش نشان می‌دهد که سه پارامتر میانگین زیروبمی واکه (f_0)، میانگین شدت انرژی واکه و دیرش آن عامل تمایز چهار جایگاه نوایی بی‌تکیه، تکیه‌بر هسته، غیرهسته و سؤالی از یکدیگر است. علاوه بر این مشخصات آکوستیکی، متغیرهای دیگری نیز مانند بافت واکه‌ای و همخوانی نیز وجود دارند که هر یک مشخصات نوایی مختص به خود را سبب می‌شوند. از این رو استخراج این بافت‌ها از گفتار طبیعی و ذخیره آنها به صورت جایگاه‌های نوایی مختلف سبب تولید گفتاری طبیعی‌تر و نزدیک‌تر به گفتار انسانی می‌شود.

فهرست منابع

- اسلامی، محرم. ۱۳۸۴. واج‌شناسی: تحلیل نظام آهنگ زبان فارسی. سمت.
- اسلامی، محرم، جواد شیخ‌زادگان، زهرا احمدی‌نیا و علی بهرامی‌راد. ۱۳۸۸. مراحل و نحوه تهیه دادگان‌های صوتی هجایی و دایفونی برای سامانه تبدیل متن به گفتار فارسی. دوفصلنامه پردازش علائم و داده‌ها، سال ششم، شماره ۲، پیاپی ۱۲، ۳-۱۲.
- اسلامی، محرم، و محمود بی‌جن خان. ۱۳۸۱. نظام آهنگ زبان فارسی. مجله زبان‌شناسی، سال ۱۷، شماره ۲، پیاپی ۳۴، ۳۶-۶۱.
- پرمون، یدالله. ۱۳۸۵. یک الگوریتم عروضی برای تکیه پیش روی کلمه در فارسی امروز. مجموعه مقالات دومین کارگاه پژوهشی زبان فارسی و رایانه، دانشگاه تهران، ۲۶۲-۲۸۴.
- سامعی، حسین. ۱۳۷۴. تکیه‌ی فعل در زبان فارسی، یک بررسی مجدد. نامه فرهنگستان، سال ۱، شماره ۴، ۶-۲۱.
- صادقی، وحید. ۱۳۸۹. طراحی و ارزیابی یک سامانه بازسازی گفتار به روش هم‌گذاری واحدهای حساس به بافت نوایی. دوفصلنامه پردازش علائم و داده‌ها، شماره ۲، پیاپی ۱۴، ۶۹-۸۳.
- کرد زعفرانلو کامبوزیا، عالیہ و فرزانه تاج‌آبادی. ۱۳۹۱. مقایسه مشددسازی و تشدیدزایی در زبان فارسی و گویش‌های ایرانی. فصلنامه پژوهش‌های زبان و ادبیات تطبیقی. ۳ ش ۱ (پیاپی ۹)، ۱۲۷-۱۵۱.
- محمودزاده، زهرا. ۱۳۸۸. تجزیه و تحلیل صوت‌شناختی سایشی‌شدگی همخوان‌های انسایشی در فارسی معیار. پایان‌نامه دکتری، دانشگاه تهران.
- وحیدیان کامیار، تقی. ۱۳۷۹. نوای گفتار در فارسی (تکیه، آهنگ، مکث). مشهد، انتشارات دانشگاه فردوسی مشهد.

- Alan W. Black, and Lenzo A. Kevin. 2003. *Building Synthetic Voices*. URL <http://www.festvox.org/bsv/>.
- Alan W. Black, Paul Taylor, and Richard Caley. 1999. *The Festival Speech Synthesis System: System Documentation 1.4 for FestivalVersion1.4.0*. URL http://www.cstr.ed.ac.uk/projects/festival/manual/festival_toc.html.
- Allen, Jonathan, Sharon M. Hunnicutt, and Dennis Klatt. 1987. *From Text to Speech: The MITalk System*. Cambridge University Press, 1987.
- Beckman, Mary E. and Janet B. Pierrehumbert. 1986. *Intonational structure in English and Japanese*. Phonology Yearbook 3: 255-310.
- Bruce, Gösta. 1977. *Swedish word accents in sentence perspective*. Lund: Gleerup.
- Buron, R. H. 1968. *Generation of a 1000-word vocabulary for a pulse-excited vocoder operating as a audio response unit*. IEEE Transactions on Audio and Electroacoustics. AU-16, 21-5.
- Chapman, W.D. 1971. *Techniques for computer voice response*. IEEE International conference Record, 98-90.
- Cho, T. 2002. *The Effects of Prosody on Articulation in English*. New York: Routledge.
- Cho, T. 2005. *Prosodic strengthening and featural enhancement: evidence from acoustic and articulatory realizations of /a, i/ in English*. Journal of acoustic society of America, Vol. 117, P. 3867-3878.
- Diehl, R. L., and W. A. Castleman. 1996. *Integrated perceptual properties: the affricate/ fricative distinction*. in A.P. Simpson and M. Patzold (eds.), sound patterns of connected speech proceedings of the symposium held at Kiel University, P. 191-200.
- Eady, S. J. et al. 1987. *Pitch assignment rules for speech synthesis by word concatenation*. Proceedings of ICASSP 1987. 1473-6.

- Gussenhoven, Carlos. 2004. *The phonology of tone and intonation*. Cambridge: Cambridge University Press.
- Hirschberg, J., and P. Prieto. 1996. *Training intonational phrasing rules automatically from English and Spanish text-to-speech*. Speech Communication, 18:281- 290.
- Kahnemuyipour, Arsalan. 2003. *Syntactic categories and Persian stress*. Natural Language and Linguistic Theory 21(2), 333-379.
- Keating, P. A. 2003. *Phonetic encoding of prosodic structure*. Proceedings of the 6th International Seminar on Speech Production, Sydne: December 7 to 10.
- Keating, P. A., T. Cho, C. Fougeron, and C. S. Hsu. 1998. *Domain-initial articulatory strengthening in four languages*. In Local et al. (eds.), P. 143–161.
- Klatt, D. H. 1987. *Review of text-to-speech conversion for English*. JASA, 82:3, 636-93.
- Kuzla, C., T. Cho, and M. Ernestus. 2007. *Prosodic strengthening of German fricatives in duration and assimilatory devoicing*. Journal of Phonetics, Vol. 35, P.301–320.
- Ladd, D. Robert. 1996. *Intonational phonology*. Cambridge: Cambridge University Press.
- Ladd, D. Robert. 1983. *Phonological features of intonational peaks*. Language 59: 721-759.
- Lavoie, L. M. 2001. *Consonant strength: phonological patterns and phonetic manifestations*, Garland publishing.
- Lazard, Gilbert. 1992. *A grammar of contemporary Persian*. English translation, Costa Mesa, California: Mazda Publishers. (Translated from French by Shirley Lyons; first published in 1957 as *Grammaire du persan contemporain*, Paris: Klincksieck.)
- Lehiste, L. 1970. *Suprasegmentals*, Cambridge: MIT. T. Press.

- Lewis, E., and M. Tatham. 2001. *Automatic Segmentation of recorded speech into syllables for speech synthesis*. Proceedings of Eurospeech, 01, 1703-1707. Aalborg
- Liberman, Mark. 1975. *The intonational system of English*. PhD thesis, MIT, distributed 1978 by Indiana University Linguistics Club.
- Mahjani, Behzad. 2003. *An instrumental study of prosodic features and intonation in Modern Farsi (Persian)*. Ms Thesis.
- Mahootian, Shahrzad. 1997. *Persian. Descriptive grammars*. London: Routledge.
- Nespor, M. and I. Vogel. 1986. *Prosodic phonology*, Dordrecht: Foris.
- Pierrehumbert, Janet B. 1980. *The phonology and phonetics of English intonation*. PhD thesis, MIT, published 1988 by Indiana University Linguistics Club.
- Price, Patti J., Mari Ostendorf, Stefanie Shattuck-Hufnagel, and C. Fong. 1991. *The use of prosody in syntactic disambiguation*. Journal of the Acoustical Society of America 90: 2956-2970.
- Rietveld, A.C.M., J. Kerkhoff and C. Gussenhoven. 2004. *word prosodic structure and vowel duration in Dutch*. Journal of Phonetics, 32, 349-371.
- Roach, Peter. 2000. *English Phonetics and Phonology*, Cambridge University Press.
- Sadat-Tehrani, Nima. 2007. *The Intonational Grammar of Persian*. PhD thesis, University of Manitoba.
- Sato, H. 1992. *Speech synthesis using CVC concatenation units and excitation waveform elements*. Trans. on Acoustical Society of Japan, 83-69.
- Selkirk, E. 1978. *On prosodic structure and its relation to syntactic structures*. in T. Fretheim (ed). Nordic prosody II, TAPIR, Trondheim.
- Sivertsen, E. 1961. *Segment inventories for speech synthesis*. Language and Speech, 4, 27- 90.

- Smith, C. L. 1997. *The devoicing of /z/ in American English: effects of local and prosodic context*. Journal of Phonetics 25, 471–500.
- Sproat, R. 1998. *Multilingual text to speech synthesis: the Bell Labs Approach*. Boston: Kluwer.
- Stevens, K. N. 2000. *Acoustic Phonetics*. MIT press.
- Towhidi, Jalil. 1974. *Studies in the phonetics of Modern Persian, Intonation and related features*. Hamburg: Helmut Buske .
- Wightman, Colin, Stefanie Shattuck-Hufnagel, Mari Ostendorf, and Patti Price. 1992. *Segmental durations in the vicinity of prosodic phrase boundaries*. Journal of the Acoustical Society of America 91: 1707-1717.
- Yoon, K. 2008. *A prosodic phrasing model for a Korean text-to-speech synthesis system*. Journal of Phonetics
- Yoon, K. 2007. *Design and evaluation of prosodically-sensitive concatenative units for a Korean TTS system*. Journal of Computer Speech and Language.
- Zanten, E. van and V.J. van Heuven. 1998. *Word stress in Indonesian: Its communicative relevance*. Journal of the Humanities and Social Sciences of Southeast Asia and Oceania [Bijdragen tot de Taal-, Land- and Volkenkunde] 154, 129–147.
- Zanten, E. van and V.J. van Heuven. 2004. *Word stress in Indonesian: Fixed or free?* NUSA Linguistic Studies of Indonesian and other Languages in Indonesia 53, 1-20.