



واژه‌گزینی و دانش زبانی

محرم اسلامی (عضو هیئت علمی دانشگاه زنجان)

سمانه ابهری (دانشجوی دکتری زبان و ادبیات فارسی، دانشگاه شهرکرد)

صفورا فدایی‌زاده (دانش‌آموخته کارشناسی ارشد زبان و ادبیات فارسی، دانشگاه زنجان)

مقدمه

بیان مفاهیم جدید مستلزم وضع واژه‌های نو است. زبان چنان سامان یافته است که از عهده بیان این مفاهیم برآید. هیچ زبانی نیست که فاقد این توانایی باشد. در واژه‌سازی یا واژه‌گزینی برای مفاهیم جدید که در زبان‌های بیگانه لفظ دارند، یکی از شقوق سه‌گانه زیر را می‌توان اختیار کرد: وام‌گیری که جز در شرایط خاص مطلوب به نظر نمی‌رسد؛ گزینش واژه‌ای برابر لفظ بیگانه از میان واژه‌های زبان که می‌تواند با گسترش معنایی آن قرین باشد؛ واژه‌سازی در چارچوب اصول و امکانات زبان. واژه‌سازی آگاهانه یا ناآگاهانه صورت می‌گیرد. در واژه‌سازی ناآگاهانه، اهل زبان با تکیه بر دانش زبانی خود واژه می‌سازند و این واژه‌ها غالباً از آشناترین صورت‌های زبانی به لحاظ آوایی، صرفی، نحوی و معنایی و کاربردی‌اند و، به همین دلیل، گویشوران بی‌هیچ مقاومتی آنها را می‌پذیرند. واژه‌هایی مانند گِلگیر، برف‌پاک‌کن از این دست‌اند. کودکان، شعرا و نویسندگان، صاحبان حرف و مشاغل عاملان اصلی واژه‌سازی ناآگاهانه‌اند. واژه‌سازی آگاهانه را نهاد یا شخص معینی انجام می‌دهد (طباطبائی، ص ۱۵). گسترش روابط فرهنگ‌ها و جوامع و رشد روزافزون علوم و فناوری به‌ویژه بومی نبودن آنها

واژه‌گزینی و واژه‌سازی آگاهانه را ایجاب می‌کند. در فرهنگستان زبان و ادب فارسی، واژه‌گزینی چنین تعریف شده است:

واژه‌گزینی فرایندی است که طی آن برای یک مفهوم مشخص علمی یا فنی یا حرفه‌ای یا هنری معمولاً یک و در مواردی بیش از یک لفظ برگزیده یا ساخته می‌شود. (اصول و ضوابط واژه‌گزینی، ص ۹)

پیداست که، در این تعریف، شقوق دوم و سوم -گزینش واژه از میان واژه‌های موجود و ساختن واژه نو- لحاظ شده است. واژه‌گزینی و، از آن حساس‌تر، واژه‌سازی، هرگاه پشتوانه علمی نداشته باشد، تلاشی ناموفق خواهد بود. در فراهم آوردن این پشتوانه، دستیابی به رمز و راز موفقیت واژه‌سازی ناآگاهانه کمک مؤثری است و می‌توان واژه‌گزینی را حتی الامکان بر پایه نتایج آن استوار ساخت. به این منظور، توصیف دقیقی جنبه‌های گوناگون توانش زبانی اهل زبان ضرورت می‌یابد.

در این تحقیق، پاره‌ای از خصوصیات نظام آوایی زبان فارسی به حیث جزئی از دانش زبانی اهل زبان توصیف و، بر اساس آن، واژه‌های مصوب فرهنگستان زبان و ادب فارسی در دفترهای اول تا پنجم فرهنگ واژه‌های مصوب ارزش‌سنجی می‌شود. این ارزش‌سنجی از راه مقایسه داده‌های طبیعی زبان با واژه‌های مصوب فرهنگستان زبان و ادب فارسی و میزان همسوئی واژه‌گزینی در فرهنگستان با مبانی آوایی حاکم بر زبان فارسی صورت می‌گیرد و پیش‌بینی میزان پذیرش واژه‌های مصوب را تنها از این نظر میسر می‌سازد؛ چون، در واژه‌گزینی، ملاک‌های آوایی به‌تنهایی تعیین‌کننده نیستند.

تحقیق حاضر مبتنی بر پژوهشی است که محرم اسلامی و محمود بی‌جن خان در سال ۱۳۸۱ انجام داده‌اند و نتایج آن در دومین هم‌اندیشی واژه‌گزینی و اصطلاح‌شناسی عرضه شده است (← اسلامی و بی‌جن خان). این پژوهش، با استفاده از پیکره بزرگ گفتاری زبان فارسی (← منابع)، برای بازنمایی محدودیت‌های نظام واج‌آرایی زبان فارسی، گرایش‌های بی‌نشان را در این واج‌آرایی نشان می‌دهد. نتایج پژوهش می‌تواند یاری‌رسان واژه‌گزینان باشد که می‌توانند، با مراجعه به آن، از میزان همسوئی حاصل کار خود با اصول حاکم بر واج‌آرایی زبان واقف شوند و جاافتادن واژه‌های مصوب در زبان را پیش‌بینی کنند. در این پژوهش، سه کارشناس ارشد زبان‌شناسی طی دو سال

۵۰۰'۰۰۰ واژه را در قالب پرونده‌های متنی رایانه‌ای واج‌نویسی کردند. داده‌های واج‌نویسی شده از مقوله‌های متنوع اختیار شده بود. بررسی این پیکره نشان داد که آن از ۳۰۰'۰۰۰ واژه با بسامدهای گوناگون تشکیل شده است. سپس صورت‌های تصریفی و واژه‌ها و ترکیب‌های خارجی از فهرست ۳۰۰'۰۰۰ واژه‌ای حذف شدند و، در نتیجه، ۱۶'۸۸۳ واحد واژگانی با بسامدهای متفاوت باقی ماند که بسامد وزنی^۱ اجزای آنها مجموعاً ۴۱۹'۰۸۲ است. واحدهای واژگانی مشخصه‌های آوایی، صرفی، نحوی، و معنایی معینی دارند که اطلاع از آنها بخشی از دانش زبانی ناخودآگاه اهل زبان شمرده می‌شود و همین دانش ناخودآگاه و علم به الگوهای بی‌نشان آوایی و واژه‌سازی در قالب این الگوها موجب می‌شود که واژه‌های مصوب از احتمال پذیرفتاری قوی‌تری نزد اهل زبان برخوردار باشند. از تحلیل واج‌شناختی و وزنی ۱۶'۸۸۳ واحد واژگانی یادشده نتایجی به دست آمد که توجه به آنها در واژه‌گزینی بسیار حایز اهمیت است.

واج آرائی واحدهای واژگانی پیکره

از توزیع ۱۶'۸۸۳ واحد واژگانی در هفت گروه از یک هجایی تا بیش از شش هجایی نتایجی به دست آمد که جدول ۱ و نمودار ۱ آنها را نشان می‌دهد.

جدول ۱- پراکندگی واحدهای واژگانی پیکره برحسب شمار هجاهای آنها

۱) برای هر گروه هجائی در پیکره ما، چند واژه وجود دارد. مثلاً، در جدول ۱، ۱۲۹۵ واژه تک‌هجایی مانند «به، در، باد، بار، بیم» شمرده شده یعنی ۷/۶۷ درصد از کل واحدهای واژگانی پیکره تک‌هجایی است. شمار این ۱۲۹۵ واحد واژگانی تک‌هجایی، با محسوب داشتن تکرار در آنها در پیکره، ۱۸۱۳۵۱ است که همان بسامد وزنی است و ۴۳/۲۷ درصد کل واژه‌ها را شامل است. مثلاً «به، در، از، که، و» (حروف اضافه و ربط و عطف)، درست است که پنج واژه‌اند، ولی شمار وقوع آنها در متن بسیار زیاد است و همین باعث می‌شود بسامد وزنی واژه‌های تک‌هجایی بسیار بالا باشد.

ادامۀ جدول ۱

نمودار ۱- پراکندگی واحدهای واژگانی پیکره برحسب شمار هجاهای آنها



جدول ۱ و نمودار ۱ نشان می‌دهد که ترکیب واحدهای واژگانی سه و دوهجایی نزدیکترین بسامد به یکدیگر را دارند. اگر بسامد را به عنوان شاخص بی‌نشانی در زبان بپذیریم، می‌توان نتیجه گرفت که واژه‌های سه و دوهجایی، از نظر شمار هجا، بی‌نشانی‌ترین ساخت‌های هجائی‌اند. در واژه‌گزینی و واژه‌سازی آگاهانه توجه به این امر مهم است و، براساس آن، انتظار می‌رود واحدهای واژگانی مصوب سه و دوهجایی، در مقایسه با سایر گروه‌ها، از مقبولیت بیشتری برخوردار شوند. واژه‌های تک‌هجایی

عموماً متعلق به واژه‌های دستوری مانند حروف اضافه و حروف ربط و واژه‌های بنیادی زبان‌اند. به همین دلیل، این واژه‌ها نزدیک به نیمی از بسامد وزنی کل (۴۱'۹۰۸۲) را شامل‌اند. بسامد وزنی واژه‌های دو هجایی نیز حدود ۳۰ درصد است که دلیل آن حضور واژه‌های دستوری دو هجایی و نیز عموم واژه‌های بنیادی زبان در این گروه است. جدول ۱ همچنین نشان می‌دهد که بیش از ۹۳ درصد از بسامد وزنی پیکره به واژه‌های یک و دو و سه هجایی تعلق دارد که در داده‌های گفتاری چه بسا این نسبت بیشتر هم بشود. نمودار ۱ نشان می‌دهد که بسامد واحدهای واژگانی در گروه واژه‌های سه هجایی به اوج می‌رسد و، پس از آن، سیر نزولی طی می‌کند و در گروه واژه‌های بیش از شش هجایی به کمترین حد می‌رسد. از سوی دیگر، بسامد وزنی در گروه تک هجایی بالاترین بسامد را دارد و از آن پس مرتباً سیر نزولی دارد تا، در گروه بیش از شش هجایی، به کمترین حد خود می‌رسد. ناگفته نماند که واژه‌های تک هجایی بسیط‌اند و فهرست آنها، صرف نظر از وام‌گیری‌های جدید، بسته است. لذا، در واژه‌سازی آگاهانه و، به دنبال آن، در واژه‌گزینی عام، به راحتی نمی‌توان، از میان واژه‌های موجود در زبان، برابری تک هجایی برای بیان مفهوم جدید برگزید. برخی از واژه‌های بنیادی زبان نیز که بسامد بالایی در کاربردهای روزانه دارند، از ابتدا، تک هجایی بوده‌اند و بسیاری از آنها نیز به تدریج به صورت تک هجایی درآمده‌اند. از آنجایی که واژه‌سازی برای مفاهیم جدید با استفاده از امکانات زبان و عموماً با ترکیب تکواژهای آن صورت می‌گیرد، طبعاً حاصل ترکیب به واژه‌های بیش از یک هجایی منجر می‌گردد.

بسامد وزنی بسیار بالای واژه‌های تک هجایی در کاربردهای روزانه نمودار تحقق اصل کم‌کوشی در زبان است. هرگاه واژه‌های پربسامد زبان، عموماً واژه‌های دستوری یا بنیادی، فرضاً پنج هجایی یا بیشتر می‌بودند، برقراری ارتباط بس دشوار می‌شد.

مقایسه ساخت آوایی داده‌های زبانی پیکره با واژه‌های مصوب

از آنجایی که در باب واژه‌های مصوب بسامد وزنی مطرح نیست، مقایسه را تنها از حیث بسامد واحدهای واژگانی انجام می‌دهیم. در این تحقیق، واژه‌ها از فرهنگ واژه‌های مصوب فرهنگستان (دفترهای اول تا پنجم) با توجه به شرایط زیر انتخاب شده‌اند:

– واژه‌های بسیط یا با ترکیب صرفی و نه دارای ساختار نحوی. بدین سان، واحدهایی مانند برات بانک‌پذیر، (دفتر اول ص ۱۶)، موج P (دفتر دوم، ص ۱۸۵)، راه درون‌شهری (دفتر دوم، ص ۹۷)، تفریحگاه ساحلی (دفتر چهارم، ص ۶۲)، سینمای جنگی (دفتر، ص ۲۰۹) که دارای ترکیب نحوی اند بیرون از پیکره تحقیق مانده‌اند.

– الفاظ تکراری حتی اگر، در علوم گوناگون، معادل اصطلاحات متفاوت اختیار شده باشند. مانند آمیزه به ترتیب معادل compound در شیمی، cocktail در گردشگری، و blend در زبان‌شناسی. فقط یک بار به حساب آمده‌اند چون هدف سنجش احتمال جافتادن آنها به لحاظ آوایی بوده است.

– سرواژه‌ها، هر چند جانشین چند واژه اختیار شده‌اند، یک واژه به حساب آمده‌اند؛ مانند زُخک (رساننده خدمات کاربردی) (دفتر دوم، ص ۹۷)، ساناد (سامانه نام دامنه) (دفتر چهارم، ص ۱۳۴)، پهباد (پرنده هدایت‌پذیر از دور) (دفتر پنجم، ص ۶۵).

– مصدرها در تحلیل داده‌ها لحاظ نشده‌اند؛ مانند بازبَلوری شدن (دفتر اول، ص ۱۴)، میانداری کردن (دفتر چهارم، ص ۲۱۱)، زمستانه کردن (دفتر پنجم، ص ۱۸۴). اما اسم مصدرها و صفت‌های مفعولی، که از راه ترکیب و اشتقاق ساخته می‌شوند، در تحلیل داده‌ها لحاظ شده‌اند؛ مانند پوش‌زدایی (دفتر دوم، ص ۳۸)، توشه‌گیری (دفتر پنجم، ص ۹۷)، بازبَلوری شده (دفتر اول، ص ۱۴).

– ترکیب‌هایی که اجزای آنها با نیم‌خط از هم جدا شده‌اند یک واحد واژگانی به حساب آمده‌اند؛ مانند ماهی-گیاهخوار (دفتر سوم، ص ۱۰۳)، چرخ-محور (دفتر چهارم، ص ۸۰)، زهدان-لوله‌نگاری (دفتر پنجم، ص ۱۸۷).

– اختصارات مانند ام. اس (دفتر اول، ص ۱۰) لحاظ نشده‌اند.
– داده‌ها با توجه به صورت آوایی آنها واج‌نویسی شده‌اند؛ مثلاً چهارتمبر به صورت تلفظی چهارتمر واج‌نویسی شده است.

– واکه [ow] به صورت /o/ واج‌نویسی شده است؛ مانند قوس‌نما/qosnamâ/ (دفتر اول، ص ۸۷)، اوج/oj/ (دفتر چهارم، ص ۱۸).

– وام‌واژه‌های مصوب لحاظ شده‌اند؛ مانند کالیستو (دفتر دوم، ص ۱۵۴)، سنکروترون (دفتر سوم، ص ۷۲).

— واحدهای دارای دو صورت تلفظی مانند سازمان، آسمان به صورت /sâzemân/ و /âsemân/ (با اختیار دو هجای CV /CV به جای یک هجای CVC آغازی) واج‌نویسی شده‌اند؛ اما واحدهای غیرمشتمل را به صورت یک هجای آغازی (CVC) واج‌نویسی کرده‌ایم؛ مانند /jaryân/ و نه /jarayân/.

در جدول ۲ پراکندگی واحدهای واژگانی و واژه‌های مصوب برحسب شمار هجا مقایسه می‌شود. در این جدول می‌بینیم که درصد بسامد واحدهای واژگانی در واژه‌های مصوب یک و دو و سه هجایی کمتر از متناظرهای خود در داده‌های زبانی است، یعنی از وضعیت طبیعی زبان در این هر سه گروه فاصله دارد. در گروه واژه‌های تک هجایی این فاصله موجه است؛ چون، در این گروه، واژه‌های مصوب تنها از میان واژه‌های موجود در زبان، مانند نقش، نت، مین، لک، طیف، صرع، رمز برگرفته شده‌اند و، در این حالت، واژه‌سازی از طریق ترکیب یا اشتقاق و یا هر دو طبعاً منتفی است و شمار برگرفته‌ها البته از شمار تک هجایی‌های پیکره به مراتب نازل‌تر است. بسامد واژه‌های مصوب دوهجائی ۱۱/۴۴ درصد از درصد بسامد متناظر آنها در پیکره ما کمتر است. دلیل این فاصله آن است که معمولاً مفاهیم جدید در قالب واژه‌هایی بیان شوند که از راه اشتقاق و ترکیب ساخته می‌شوند و حاصل این عمل معمولاً واژه‌های بیش از دو هجایی است. در واژه‌های سه هجایی، درصدهای بسامد واحدهای واژگانی پیکره ما و بسامد واژه‌های مصوب تقریباً نزدیک به هم است و نشان می‌دهد که، از حیث شمار، واژه‌های مصوب سه هجایی با داده‌های طبیعی همسوئی نزدیکی دارد و این بدان معناست که احتمال جا افتادن واژه‌های مصوب سه هجایی از این لحاظ قوی‌تر است.

جدول ۲- پراکندگی واحدهای واژگانی پیکره و واژه‌های مصوب برحسب شمار هجاهای آنها

پرتال جامع علوم انسانی

ادامۀ جدول ۲

در نمودار ۲، با مقایسه واحدهای واژگانی پیکره و واژه‌های مصوّب، نسبت درصدهای آنها را در هر یک از گروه‌های هجائی هفت‌گانه می‌توان به راحتی مشاهده کرد. چنان‌که ملاحظه می‌شود فاصله این نسبت درصدها، با افزایش هجاها به تدریج کاهش می‌یابد به طوری که در گروه‌های از چهار هجایی به بالا نسبت درصد واژه‌های مصوّب از نسبت درصد در واحدهای واژگانی پیکره بالاتر می‌رود. اما در همه گروه‌های هجایی، انحراف به درجات متفاوت وجود دارد.

درخور توجه است که این تحقیق همه واژه‌های مصوّب، اعم از واژه‌های برگرفته از واژگان موجود زبان فارسی و نوساخته، را در بر می‌گیرد و هرگاه صرفاً به واژه‌های نوساخته محدود گردد چه بسا نتایج متفاوتی به دست آید و محتمل آن است که، در واژه‌های نوساخته، انحراف از معیار طبیعی زبان بیشتر باشد.

نمودار ۲- پراکندگی واحدهای واژگانی پیکره و واژه‌های مصوّب بر حسب شمار هجاهاى آنها

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

پیش‌تر گفته شد که واژه‌های سه‌هجایی بی‌نشان‌ترین الگوی هجائی واژه در زبان فارسی هستند و این را بسامد واژگانی ۳۴/۲۹ درصد آنها در جدول ۲ نشان می‌دهد. درصد بسامد واژه‌های چهارهجایی در پیکره ۱۸/۶۰ و در واژه‌های مصوّب ۲۴/۶۵ درصد است. یعنی واژه‌های مصوّب چهارهجایی، در مقایسه با داده‌های طبیعی، حدود ۶ درصد بیشتر ساخت نشان‌دار دارد. همین معنی را با تفاوت‌های بالاتر در واژه‌های پنج و شش و بیش از شش‌هجایی می‌بینیم که عمدتاً از نوواژه‌سازی آگاهانه ناشی می‌گردد. در واژه‌سازی آگاهانه، خواه ناخواه این گرایش وجود دارد که تمام یا حداقل بیشتر مؤلفه‌های معنائی مفهوم در واژه بازتاب یابد و این رویکرد چه بسا باعث افزایش شمار هجاها در واژه‌های نوساخته گردد. این قبیل واژه‌ها عمدتاً خصلت تخصصی ظریف و کم‌بسامد دارند و اهل فن، به دلایل گوناگون، آنها را کمتر به کار می‌برند و ترجیح می‌دهند که همان اصطلاح بیگانه را اتخاذ کنند. درصد واژه‌های مصوّب بیش از شش‌هجایی حدود هفت برابر داده‌های طبیعی است که انحراف فاحشی از معیار زبانی شمرده می‌شود و نشان می‌دهد که واژه‌سازی در این قالب اصلاً مطلوب اهل زبان نیست و احتمال جا افتادن این قبیل نوواژه‌ها ضعیف است.

مقایسه ساخت هجائی واحدهای واژگانی پیکره و واژه‌های مصوّب

در این پاره از گزارش، به مقایسه واحدهای واژگانی پیکره و واژه‌های مصوّب از نظر نوع و آرایش هجاها در گروه‌های هفت‌گانه می‌پردازیم. در نظام آوایی زبان فارسی هجا دارای سه ساخت CV، CVC، CVCC است (ثمره، ص ۱۰۹) که در آن C نماینده همخوان (صامت) و V نماینده واکه (مصوّت) است.

ساخت هجائی واحدهای واژگانی تک‌هجایی

جدول ۳ پراکنندگی ساخت هجائی واژه‌های تک‌هجایی را در واحدهای واژگانی پیکره و واژه‌های مصوّب نشان می‌دهد. در هر دو دسته، واژه‌های تک‌هجائی با ساخت CVCC بیشترین و واژه‌های تک‌هجائی با ساخت CV کمترین بسامد را دارند. تفاوت درصدهای بسامد واحدهای واژگانی پیکره با واژه‌های مصوّب در الگوی هجائی CV را می‌توان

موجه دانست؛ چون گزینش واژه‌ای برای مفهوم جدید با ساخت CV کمتر میسر است. نسبت درصد واژه‌های تک‌هجایی با ساخت‌های CVC و CVCC در داده‌های طبیعی و واژه‌های مصوب بسیار نزدیک به هم است و دلیل آن چه بسا آن باشد که واژه مصوب، در این حالات، از میان واژه‌های موجود در زبان برگرفته می‌شود.

جدول ۳- پراکندگی ساخت‌های تک‌هجایی

ساخت‌های تک‌هجایی

جدول ۴ پراکندگی ساخت‌های تک‌هجایی در واحدهای واژگانی دو‌هجایی را در واحدهای واژگانی پیکره و واژه‌های مصوب نشان می‌دهد. در اینجا، تفاوت‌های فاحشی در درصد‌های بسامد آن دو دیده می‌شود. به عنوان مثال، ساخت CV.CVC بالاترین بسامد (۳۲/۷۱ درصد) را در واحد واژگانی پیکره در مقابل ۲۲/۸۹ درصد در واژه‌های مصوب دارد. واژه‌هایی مانند مادر، خواهر، پدر و عموماً واژه‌های بنیادی دارای ساخت CV.CVC هستند و این ساخت بیشترین بسامد در واژگان زبان را داراست.

از ترکیب ساخت‌های سه‌گانه‌ی (CV، CVC، CVCC) با دو جایگاه، بالقوه‌ی نه ساخت‌های تک‌هجایی پدید می‌آید که در جدول ۴ نشان داده شده‌اند. در این ساخت‌های تک‌هجایی نه‌گانه، درصد بسامد ساخت CV.CVC نماینده‌ی بی‌نشانی است. هرچه ساخت‌های تک‌هجایی بی‌نشانی‌تر باشد، ساخت آوائی‌ی واژه در نزد اهل زبان مطلوب‌تر و احتمال جافتادن فرآورده‌ی واژه‌سازی در قالب این ساخت‌ها قوی‌تر است. جدول ۴ نشان می‌دهد که ساخت CV.CVC با ۳۲/۷۱ درصد بسامد در واحدهای واژگانی پیکره بی‌نشانی‌ترین و ساخت CVCC.CV با ۰/۲۴ درصد بسامد نشان‌دارترین است. ساخت CVCC.CVCC با ۰/۴ درصد در مرتبه‌ی نشان‌داری پایین‌تر از آن قرار دارد؛ اما باز،

واژه‌سازی در این الگوی هجایی، دست‌کم به لحاظ آوایی مطلوب نیست. هرچند ساخت CVCC.CVCC کمترین بسامد (۴/۰ درصد) را در واژگان زبان فارسی دارد، اهل زبان، با تغییرات تلفظی، الگوی هجائی آن را عوض می‌کنند و به صورت مطلوب‌تری درمی‌آورند؛ مثلاً (CVCC.CVCC) /dastband/ را (CVC.CVCC) [dasband] تلفظ می‌کنند که بسامد نسبتاً بالاتری (۳۳/۴ درصد) دارد. از واژه‌های مصوب مثلاً برق‌کافت (CVCC.CVCC) را نمی‌توان برکافت تلفظ کرد و همین امر پذیرش آن و نظایر آن را دشوار می‌سازد.

جدول ۴- پراکندگی ساخت هجائی واحدهای واژگانی دو هجایی



حال اگر نهادی مانند فرهنگستان میخیزد به انتخاب میان بالگرد و چرخ‌بال باشد، با مراجعه به جدول ۴، می‌توان پیش‌بینی کرد که بالگرد با ساخت CVC.CVCC را بر چرخ‌بال با ساخت CVCC.CVC ترجیح خواهد داد، چون ساخت آوایی آن بی‌نشان‌تر و احتمال پذیرش آن قوی‌تر است. اعداد و ارقام جدول ۴ میزان همسویی یا انحراف واژه‌های مصوب از وضعیت طبیعی در زبان را نشان می‌دهد. عملکرد توأم دو محدودیت متضاد بسیار مهم - نشان‌داری^۲ و پایداری^۳ (KAGER 1999, pp. 1-46) - در نظام زبان در همه سطوح آن توازن پدید می‌آورد. محدودیت نشان‌داری به این معناست که مفاهیم زبانی گرایش

2) markedness

3) faithfulness

دارند بیشتر در قالب‌های بی‌نشان تولید شوند، مثلاً هجای باز نسبت به هجای بسته ساخت بی‌نشان‌تر محسوب می‌شود. بر اساس محدودیت نشان‌داری، بسامد واحدهای واژگانی تک‌هجائی دارای ساخت CV باید بیشتر از سایر ساخت‌ها باشد. محدودیت پایداری ناظر بر آن است که، در تبدیل صورت نشان‌دار به صورت بی‌نشان، لازم است تغییر به اندازه‌ای باشد که امر ارتباط را مخدوش یا مختل نسازد. اگر واژه‌ای با ساخت CVCC مثلاً /dast/ در گفتار به صورتی با ساخت CV مثلاً [da] تقلیل یابد، امر ارتباط مختل می‌گردد. در چنین مواردی، محدودیت پایداری برای ایجاد توازن در نظام زبان وارد عمل می‌شود و در مقابل چنین تغییری ایستادگی می‌کند. به همین دلیل، هجای CV، هر چند ساخت بی‌نشان شمرده می‌شود، بسامد واحدهای واژگانی در آن کمتر از گروه‌های نشان‌دارتر یک‌هجایی است. (← جدول ۳)

پراکندگی ساخت هجائی واحدهای واژگانی سه‌هجایی

جدول ۵ پراکندگی ساخت هجائی واحدهای واژگانی پیکره و واژه‌های مصوب سه‌هجایی را نشان می‌دهد. برای واژه‌های سه‌هجایی، بالقوه ۲۷ ترکیب هجایی وجود دارد که، در داده‌های زبانی پیکره، تنها ۲۴ ترکیب آن دارای نمونه‌هایی با نسبت درصدهای متفاوت است. یکی از سه ترکیب بالقوه فاقد نمونه در واژگان زبان (← جدول ۵، سه ترکیب آخر) با واژه‌های مصوب فعلیت یافته است. در این ترکیب (CVCC.CVCC.CVC)، ترکیب ماقبل آخر در جدول ۵، دو واژه مصوب شده است. فقدان این ترکیب در واحدهای واژگانی پیکره، به عبارت دیگر، در ذهن اهل زبان چه بسا پذیرش این دو واژه را دشوار سازد.

در جدول ۵ ملاحظه می‌شود که، در واحدهای واژگانی سه‌هجایی، بالاترین بسامد از آن ساخت CV.CV.CV و در واژه‌های سه‌هجائی مصوب بالاترین بسامد از آن ساخت CVC.CV.CV است. بسامد واژگانی، از بالا به پایین، پیوستار بی‌نشانی را در واژه‌های سه‌هجائی نشان می‌دهد و مقایسه آنها با واژه‌های مصوب موارد عدول از این پیوستار را در واژه‌های مصوب مشخص می‌سازد. چنانچه پیش‌تر گفته شد، هجای CV بی‌نشان‌ترین و هجای CVCC نشان‌دارترین ساخت هجایی در زبان فارسی است که

معمولاً، در گفتار همچنین در تحوّل تاریخی، هجای CVCC به هجای CV تبدیل یا به آن نزدیک می‌شود و تبدیل /dastband/ به [dasband] نمونه‌ای از آن است که، در آن، CVCC اوّل به CV بدل نشده امّا به صورت CVC درآمده که به CV نزدیک است. هجای CVC نیز در گفتار و در تحوّل تاریخی به CV بدل می‌شود - مثلاً /kâr/ (CVC) در /kârgar/ به [kâ .re] (CV.CV) - که از بی‌نشانی ساخت هجائی CV در زبان فارسی حکایت دارد. با این مقدمه، می‌بینیم که پربسامدترین و البته بی‌نشان‌ترین ترکیب واژه‌های سه‌هجایی CV.CV.CV است، در حالی که، در واژه‌های مصوّب، پربسامدترین ساخت CVC.CV.CV است که عدول از معیارهای زبانی شمرده می‌شود. البته، بر اساس جدول ۵، مشاهده می‌شود که واژه‌های مصوّب، در بیشتر موارد و در مجموع، با واقعیات زبانی همسویی دارند؛ ولی توجّه به موارد عدول از این همسویی می‌تواند امر واژه‌گزینی را بیشتر قرین موقّیّت سازد.

جدول ۵- پراکنندگی ساخت هجائی واحدهای واژگانی سه‌هجایی

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

ادامۀ جدول ۵



ساخت هجائی واحدهای واژگانی چهارهجایی

شمار بالقوۀ ترکیبات هجایی در واژه‌های چهارهجایی ۷۱ است که، از این مقدار، در واژگان زبان فارسی فقط ۳۹ ساخت فعلیت یافته‌اند و، در بقیۀ موارد، خلأ وجود دارد. این در حالی است که، در واژه‌های مصوب، شمار بالفعل ترکیبات هجا در واژه‌های چهارهجایی به ۵۲ می‌رسد. بدین قرار، از این ۵۲ ترکیب، ۱۳ فقره، هرچند با بسامد کم، فاقد نمونه در داده‌های زبانی است. می‌توان احتمال داد که واژه‌های مصوب، در این ۱۳ ترکیب، نوساخته باشند.

جدول ۶- پراکندگی ساخت همجائی واحدهای واژگانی چهارهجایی



ادامۀ جدول ۶

در جدول ۶ نیز، ترکیبات ساخت‌های هجایی در واحدهای واژگانی چهارهجایی، به ترتیب بسامد و نسبت درصد بسامد و میزان بی‌نشانی، از بالا به پایین مرتب شده‌اند و مقایسه آنها با بسامد و درصد بسامد واژه‌های مصوب در برخی موارد انحراف‌هایی را نشان می‌دهد. در اینجا، بی‌نشان‌ترین ترکیب ساخت‌های هجایی در واحدهای واژگانی پیکره CV.CV.CV.CV و در واژه‌های مصوب CVC.CV.CV.CV است که از وجود انحراف از معیار زبان خبر می‌دهد.

ساخت هجائی واحدهای واژگانی پنج‌هجایی

واحدهای واژگانی پنج‌هجایی نیز شرایط و رفتارهای واحدهای واژگانی چهارهجایی را دارند و، در جدول ۷، برحسب بسامد و درصد بسامد مرتب شده‌اند. برخلاف واحدهای

واژگانی چهارهجایی، در اینجنا، واحدهای واژگانی پیکره و واژه‌های مصوّب در پیرسامدترین ترکیب ساخت‌های هجایی مشترک‌اند. شمار بالقوه ترکیبات ساخت‌های هجایی در واحدهای واژگانی پنج‌هجایی ۲۱۳ است که در داده‌های زبانی تنها برای ۵۰ ترکیب نمونه‌هایی با بسامدها و درصد بسامدهای متفاوت وجود دارد و این در حالی است که، در واژه‌های مصوّب، ۷۱ ترکیب ساخت‌های هجایی را شاهدیم. بدین قرار، واژه‌هایی پنج‌هجایی در ۲۱ ترکیب ساخت هجایی مصوّب شده‌اند که در داده‌های طبیعی زبان فارسی فاقد نمونه‌اند. این واژه‌های نوساخته، به احتمال قوی، با مقاومت اهل زبان مواجه خواهند شد.

جدول ۷- پراکندگی ساخت هجایی واحدهای واژگانی پنج‌هجایی



ادامۀ جدول ۷



ساخت هجائی واحدهای واژگانی شش هجایی
جدول ۸ نشان می‌دهد که واژه‌های شش هجایی کم و بیش نوعاً همان اطلاعات
جدول‌های پیشین را به دست می‌دهند. در اینجا نیز، ترکیبات ساخت‌های هجایی
به ترتیب میزان نشان‌داری آنها مرتب شده‌اند که، در مواردی، واژه‌های مصوب
از این حیث با متناظر آنها در واحدهای واژگانی پیکره تفاوت دارند.
شمار ترکیبات بالقوه ساخت‌های هجائی واحدهای واژگانی شش هجایی ۶۳۹ است
که تنها ۴۸ فقره از آنها در زبان فعلیت یافته‌اند؛ در حالی که، در واژه‌های مصوب، این رقم

به ۷۴ بالغ می‌گردد. بدین قرار، در واژه‌های مصوّب شش هجایی ۲۶ ترکیب ساخت‌های هجائی جدید دیده می‌شود که نمودار انحراف از معیار زبان است.

جدول ۸- پراکنندگی ساخت هجائی واحدهای واژگانی شش هجایی



ادامۀ جدول ۸



نتیجه

واژه‌گزینی در نهادهایی مانند فرهنگستان زبان و ادب فارسی به منزله ورود آگاهانه و انگیخته به ساحت زبان است و آن را می‌توان نوعی برنامه‌ریزی زبان شمرد. برنامه‌ریزی زبان مستلزم تلاش‌های سنجیده، نظام‌مند، و نظریه‌بنیاد است که با مطالعه زبان‌ها و گویش‌های جامعه و اتخاذ سیاست زبانی خاص در حل مسائل و مشکلات ارتباطی آن جامعه صورت می‌گیرد (Crystal 2008, p. 268). تحقیق حاضر برای ارزش‌سنجی واژه‌گزینی در فرهنگستان انجام گرفته و در آن سعی شده است، با تکیه بر نظام آرایش‌های واجی در داده‌های طبیعی، که بخشی از توانش زبانی اهل زبان شمرده می‌شود، گرایش‌های بی‌نشان در این نظام استخراج و نتایج آن با واژه‌های مصوب فرهنگستان زبان و ادب فارسی مقایسه شود. نتایج این تحقیق صرفاً از تحلیل اطلاعات آماری به دست آمده و کاملاً عینی است و دخل و تصرفی در آن

صورت نگرفته است. این نتایج در جدول‌ها و نمودارهایی نشان داده شده‌اند. از مقایسه واج‌آرایی در داده‌های طبیعی و واژه‌های مصوّب، میزان همسوئی فعالیت‌های واژه‌گزینی در فرهنگستان با روح حاکم بر زبان معلوم می‌گردد. واژه‌های مصوّب، هرچه بیشتر در قالب ساخت‌های هجائی بی‌نشان باشند، احتمال پذیرش آنها از سوی اهل زبان قوی‌تر خواهد بود. ارائه پیوستاریبی‌نشانی در ساخت هجائی گروه‌های متفاوت واژه‌های زبان فارسی از نتایج مهمّ این تحقیق است که، با مراجعه به آن، دست‌کم از نظر آوایی، می‌توان از میزان احتمال پذیرفته شدن واژه‌های پیشنهادی آگاهی حاصل کرد. از سوی دیگر، با مقایسه واژه‌های مصوّب پیشین با آن پیوستار می‌توان عملکرد فرهنگستان را در واژه‌گزینی ارزش‌سنجی کرد.

همچنان‌که جدول‌های متعدّد نشان می‌دهند، واژه‌گزینی در فرهنگستان مجموعاً با نظام آوایی زبان فارسی همسویی دارد هرچند، در مواردی - از راه اختیار واژه‌ای با ساخت آوایی نشان‌دار یا در نوواژه‌سازی - انحراف‌های نسبی از معیار زبان روی داده است. نتیجه آنکه، در واژه‌سازی، شایسته است ساخت آوایی نوواژه‌ها بی‌نشان و حتّی‌المقدور در قالب ساخت‌های آوایی بالفعل زبان باشد. با نگاهی به جدول‌ها، در واژه‌های مصوّب، نمونه‌هایی را می‌توان نشان کرد که با الگویی خارج از قالب‌های بالفعل زبان ساخته شده‌اند و احتمال بُر خوردن آنها در زبان ضعیف است. ناگفته نماند که از اطلاعات آماری مندرج در جدول‌ها می‌توان در مطالعات زبانی دیگر نیز بهره جست.

منابع

- اسلامی، محرم و محمود بی‌جن‌خان، «واژه‌گزینی و واج‌آرایی»، مجموعه مقالات دومین هم‌اندیشی واژه‌گزینی و اصطلاح‌شناسی، فرهنگستان زبان و ادب فارسی، تهران ۱۳۸۴، ص ۴۳۰-۴۴۸.
- اصول و ضوابط واژه‌گزینی، فرهنگستان زبان و ادب فارسی، تهران ۱۳۸۶.
- پیکره بزرگ‌گفتاری زبان فارسی، پژوهشکده پردازش هوشمند علائم، تهران ۱۳۸۱.
- ثمره، یدالله، آواشناسی زبان فارسی، ویرایش دوم، مرکز نشر دانشگاهی، تهران ۱۳۷۸.
- طباطبائی، سیدعلاءالدین، اسم و صفت مرکّب در زبان فارسی، مرکز نشر دانشگاهی، تهران ۱۳۸۲.
- فرهنگ واژه‌های مصوّب فرهنگستان (دفتر اوّل)، فرهنگستان زبان و ادب فارسی، تهران ۱۳۸۴.

- (دفتر دوم)، فرهنگستان زبان و ادب فارسی، تهران ۱۳۸۴.
- (دفتر سوم)، فرهنگستان زبان و ادب فارسی، تهران ۱۳۸۵.
- (دفتر چهارم)، فرهنگستان زبان و ادب فارسی، تهران ۱۳۸۶.
- (دفتر پنجم)، فرهنگستان زبان و ادب فارسی، تهران ۱۳۸۷.

CRYSTAL, David (2008), *A Dictionary of Linguistics and Phonetics*, 6th ed., Blackwell Publishing, Oxford.

KAGER, René (1999), *Optimality Theory*, Cambridge University Press, Cambridge.

□

