

## واژگانی برای زبان فارسی\*

امید طیب‌زاده

چندی پیش کتابی با عنوان *ساختمان واژه و مقوله دستوری: تشخیص مقوله دستوری واژه‌ها بر اساس ملاک‌های صرفی*، اثر جناب آقای دکتر علاء‌الدین طباطبایی به دستم رسید. هم‌عنوان کتاب و هم‌نام نویسنده آن، مرا بر آن داشت تا در اولین فرصت آن را مطالعه کنم. نویسنده در همان خط اول کتاب در پیشگفتار تصریح کرده است که «هدف از پژوهش حاضر، به‌دست‌دادن ملاک‌های صرفی برای تشخیص مقوله‌های دستوری واژه‌هاست» (ص ۹). این جمله به قدری ساده و روشن است که ظاهراً نیازی به هیچ شرح و تفسیری ندارد، اما واقعیت این است که تا زمانی که ۷۰-۸۰ صفحه از کتاب را نخواندم در نیافتم که با چه اثر پژوهشی شگفت‌انگیز و مهمی روبه‌رو هستم. زبان بسیار روشن و موجز نویسنده و اجتناب او از پیچیده‌نویسی و نیز احتراز فروتنانه وی از پرداختن به مباحث مقدماتی نظری — که در حال از ملزومات چنین اثری است — همه و همه باعث شد تا اندکی دیر به اهمیت و نقش واقعی کتاب پی ببرم، و عجباً که فصاحت، ایجاز، و خفّض جناح نویسنده هم می‌تواند گاه هم‌چون عوامل مخل بلاغت، مانع از ایجاد ارتباط سریع خواننده با اثر بشود! در بیان نپرداختن عامدانه نویسنده کتاب به مآخذ و مفاهیم و اصطلاحاتی که البته اشراف کاملی به تمامی آن‌ها داشته است (مثلاً ر.ک. طباطبایی ۱۳۸۲). تذکر همین یک نکته بس که او تقریباً هیچ‌ذکری از هیچ‌یک از آثاری که در این نقد و معرفی به آن‌ها اشاره کرده‌ام به میان نیاورده است، و نیز از هیچ‌کدام از اصطلاحاتی که معادل‌های فرنگی آن‌ها را در این‌جا به دست داده‌ام استفاده نکرده و یا دست‌کم به گونه‌ای استفاده کرده است که گویی از مفاهیم و اصطلاحات کاملاً جاافتاده و

---

\* «دستور» این مقاله، نقدی است بر کتاب *ساختمان واژه و مقوله دستوری*، نوشته علاء‌الدین طباطبایی، تهران: پژوهشگاه فرهنگ، هنر و ارتباطات، شماره ۶، اسفند ۱۳۸۹.

بسیار رایج دستور زبان است. در هر حال، در این جا کوشیده‌ام ابتدا جای خالی مربوط به مقدمات نظری کتاب را به اختصار هرچه تمام‌تر پر کنم تا معلوم شود این اثر متعلق به کدام حوزه از زبان‌شناسی است، و سپس خود اثر را به اجمال معرفی کنم، با این امید که توجه علاقه‌مندان به صرف<sup>۱</sup> و نیز علاقه‌مندان به پردازش زبان‌های طبیعی<sup>۲</sup> و مخصوصاً طرفداران رویکرد قاعده‌بنیاد<sup>۳</sup> در این زمینه را به این اثر ارزشمند و راهگشا جلب کنم. در پایان نیز با ارایه<sup>۴</sup> یکی دو مثال نشان داده‌ام که چگونه می‌توان در صورت لزوم، با ایجاد تغییراتی جزئی در واژگان، کاربری‌های آن را بسط داد.

واژگان<sup>۴</sup> و نظریه‌های گوناگون مربوط به چگونگی تنظیم اقلام واژگانی<sup>۵</sup> در آن، از جمله بحث‌برانگیزترین مباحث مربوط به صرف است. بحث درباره<sup>۶</sup> واژگان و جایگاه آن در دستور زبان، با کتاب جنبه‌هایی از نظریه نحو، اثر چامسکی (Chomsky 1965) و اظهار نظرهای بسیار کلی و تقریباً غیرصریح او در این باره آغاز شد، سپس نظریات افراطی معناشناسان زایشی<sup>۶</sup> طرح شد که می‌کوشیدند عملکرد واژگان زبان را با گشتارهایی همانند گشتارهای نحوی توجیه کنند. معناشناسان زایشی قواعد واژه‌سازی را بخشی از قواعد نحو قلمداد می‌کردند و با زیر سؤال بردن رویکرد حوزه‌ای<sup>۷</sup> چامسکی، می‌کوشیدند نظام زبان را صرفاً با دو مؤلفه نحوی و معناشناختی توصیف کنند. چامسکی در اثر معروف خود با عنوان «ملاحظات درباره<sup>۸</sup> اسم‌سازی» (Ibid 1970) موضع روشن‌تری را در مورد واژگان اتخاذ کرد و ضمن رد کردن احتجاجات معناشناسان زایشی، صریحاً از لزوم قائل شدن به حوزه مستقلی با عنوان واژگان سخن راند (برای اطلاعات بیشتر در این زمینه ر.ک. دبیرمقدم ۱۳۸۳: ۱۵۹-۳۱۲). پس از این اثر، هله (Halle 1973) در مقاله «شروحنی بر یک نظریه ساخت واژه»، در پشتیبانی از آرای چامسکی، به تعدیل آرای او پرداخت و بر اهمیت حوزه<sup>۹</sup> واژگان تأکید کرد و مدلی را برای واژگان، به عنوان یکی از حوزه‌های زبانی، پیشنهاد کرد. از آن پس مدل‌های گوناگون دیگری برای توضیح چگونگی ساخت و کار حوزه<sup>۱۰</sup> واژگان در چهارچوب نظریه زایشی و بعد از آن در چهارچوب نظریه‌های دیگر پیشنهاد شد (برای آشنایی با بعضی از این مدل‌ها ر.ک. Carstairs-McCarthy 1992: 11-52، Haspelmath 2002: 39-58)، اما علاوه بر زبان‌شناسان نظری، برنامه‌نویسان و مهندسان رایانه نیز که عملاً درگیر کارهای مربوط

به پردازش زبان شده‌بودند، به تدریج و بالاجبار به مباحث مربوط به ویژگی‌های واژگان روی آوردند و مدل‌های گوناگونی را پیشنهاد کردند و مورد استفاده قرار دادند. با افزایش یافتن حافظه و سرعت پردازش رایانه، برای مدتی این تصور قوت گرفت که شاید دیگر نیازی به در نظر گرفتن بخش‌ها و حوزه‌های گوناگون برای واژگان نباشد، اما دیری نپایید که پیشرفت مطالعات نشان داد که حتی در زمینه‌های کاملاً عمل‌گرا و فن‌مداری هم چون پردازش زبان و زبان‌شناسی رایانه‌ای نیز نمی‌توان واژگان را صرفاً انسانی از کلمات<sup>۱۰</sup> در نظر گرفت و تمام واژه‌ها<sup>۱۱</sup> و صورت‌واژه‌ها<sup>۱۲</sup> را در یک فهرست عام و مشترک گنجانند. در واقع، نظریه پردازان مدل‌های واژگانی در حوزه پردازش زبان نیز تدریجاً و خودبه‌خود به این نتیجه رسیدند که بهتر است به جنبه‌های شناختی واژگان نیز توجه داشته باشند و در ارایه مدلی برای واژگان، صرفاً به امکانات رایانه‌های پرسرعت و خوش حافظه دل خوش نکنند، بلکه تلقی خود از واژگان ذهنی<sup>۱۱</sup> یا چیزی شبیه به آن را سرمشق قرار بدهند (به عنوان مثال ر.ک. Brill 1992).

امروزه در بسیاری از مدل‌های واژگانی<sup>۱۲</sup> اقلام واژگانی را غالباً به دو بخش بسیار کلی تقسیم می‌کنند: یکی بخش فهرست واژه‌ها و دیگری بخش قواعد. به دنبال تفکیک فوق، این بحث نیز خاصه در مباحث مربوط به پردازش زبان طرح می‌شود که مقوله دستوری یا جزو کلام<sup>۱۳</sup> کلمات گروه نخست غیر قابل پیش‌بینی است و در نتیجه اطلاعات مربوط به آن‌ها نیز در این زمینه، باید در مقابل هر واژه در فهرست واژه‌ها ثبت شود، اما مقوله دستوری کلمات گروه دوم، بر اساس قواعد مندرج در بخش قواعد، قابل پیش‌بینی است. در زبان‌هایی چون انگلیسی و آلمانی مدل‌های گوناگونی برای نمایش دو بخش کلی واژگان و زیربخش‌های آن‌ها عرضه شده‌است، اما در مورد زبان فارسی، جز اشارات و توضیحات کلی نظری، به کاری نمی‌توان اشاره کرد. کتاب حاضر را می‌توان اولین مدل کاملی دانست که برای نمایش واژگان ذهنی فارسی‌زبانان ارایه شده است.

مؤلف کتاب حاضر از حدود ۳۲۰۰۰ واژه مندرج در فرهنگ معاصر فارسی (صدری افشار ۱۳۸۲) و فرهنگ واژه‌های مصوب فرهنگستان (۱۳۷۶-۱۳۸۲)، به‌عنوان پیکره خود استفاده کرده و آن‌ها را برحسب ساختارشان به دو گروه بسیط و پیچیده تقسیم کرده است. واژه‌های بسیط واژه‌هایی هستند که از حیث هم‌زمانی قابل تجزیه به عناصر نقش‌مند نیستند (مانند:

کتاب، زن، مرد، چوپان، داور)، و واژه‌های پیچیده واژه‌هایی هستند که عناصر تشکیل‌دهنده آن‌ها یا به صورت مستقل و یا در ساختمان واژه‌های دیگر به کار می‌رود (مانند: خطرناک، آب‌انبار، هنرمند). وی سپس مدعی شده است که مقولۀ دستوری واژه‌های بسیط برحسب ساختارشان قابل تشخیص نیست و از این رو باید آن‌ها را همراه با برجسیبی که مقولۀ‌شان را مشخص می‌کند، در بخش فهرست واژه‌های واژگان درج کرد، اما مقولۀ دستوری غالب واژه‌های پیچیده را می‌توان برحسب ساختارشان تشخیص داد و از این رو لازم نیست این دسته از واژه‌ها را در واژگان آورد، بلکه تنها باید قواعد ناظر بر شکل‌گیری آن‌ها را در بخش مربوط به قواعد درج کرد. این ۳۲۰۰۰ واژه شامل تقریباً تمامی واژه‌های فارسی معیار و رسمی امروز است و از این رو می‌توان گفت پیکره‌ای که مؤلف آن را برگزیده، در نوع خود کامل است. مؤلف هم‌چنین مرز میان واژه‌ها را با تمام‌فاصله، و مرز میان واژه‌های تشکیل‌دهنده واژه‌های مرکب را با نیم‌فاصله مشخص ساخته، در این کار از روش توصیه‌شده در فرهنگ *املایی خط فارسی* (صادقی و زندی‌مقدم، ۱۳۸۵) پیروی کرده است. البته وی در هیچ‌جای کتاب متذکر استفاده‌اش از این فرهنگ نشده، اما بررسی املای کلمات مرکب در واژگان او، مبین استفاده وی از رسم‌الخط فرهنگ مذکور است.<sup>۱۴</sup>

روش عمل‌گرا، توصیفی و کاملاً داده‌بنیاد مؤلف، باعث شده است که تکلیف بسیاری از مباحثی که در صرف فارسی همواره با ابهام یا سکوت روبه‌رو بوده است، در این جا به‌سادگی مشخص شود. مثلاً او بسیاری از وام‌واژه‌های غیربسیطی را که در مقیاس وسیع وارد زبان فارسی شده است، به‌عنوان کلمات پیچیده در نظر گرفته و در نتیجه وندهای جدیدی را به فهرست وندهای فارسی افزوده است که دستورنویسان سنتی و حتی زبان‌شناسان، معمولاً در مورد آن‌ها سکوت کرده‌اند. مثلاً او اسم‌هایی چون مداومت، مزاحمت و مناسبت را کلماتی پیچیده در نظر گرفته که از ساختار «صفت + پسوند - ت» حاصل شده‌اند. به بیان دیگر، در واژگان حاضر، صفت‌هایی چون مداوم، مزاحم و مناسب در فهرست واژه‌ها درج شده است، اما اسم‌هایی چون مداومت، مزاحمت و مناسبت، در فهرست واژه‌ها درج نشده و فقط در بخش قواعد واژگان تصریح شده است که از ترکیب «صفت + ت» همواره اسم به وجود می‌آید. همین بحث در مورد اسم‌های گرایش و مکتبی چون ارگانسیم، فتودالیسم، لیبرالیسم و

ماشینیسم نیز صادق است؛ به این معنا که این اسم‌ها در فهرست واژه‌ها درج نمی‌شود، بلکه اسم‌های ارگان، فتودال، لیبرال و ماشین، در فهرست واژه‌ها درج و در بخش قواعد نیز تصریح می‌شود که از ساختار «اسم + پسوند - یسم» اسم‌گرایش و مکتب به وجود می‌آید.

در کلمات پیچیده، گاهی یک ساختار واحد دو مقوله متفاوت تولید می‌کند؛ مثلاً ساختار «اسم + اسم»، هم اسم می‌سازد (مانند: کتابخانه، دامپزشک، دندان‌درد) و هم صفت (مانند: شیردل، وطن‌دوست، سیاست‌مدار). مؤلف در مورد عرضه این واژه‌ها در واژگان راه‌حل ساده و روشنی را برگزیده است. او واژه‌هایی را که بر اساس الگوی کم‌بسامدتر ساخته شده‌اند و واژه‌های نامتعارف نامیده و در درون فهرست واژه‌ها قرار داده است، و واژه‌های پربسامدتر را در واژگان درج نکرده، بلکه الگوی ساخت آن‌ها را در بخش قواعد عرضه کرده است. مثلاً او در مورد همین ساختار «اسم + اسم»، صفات مرکبی چون شیردل، وطن‌دوست، و سیاست‌مدار را که تعدادشان در پیکره مورد استفاده وی بسیار کمتر از تعداد اسم‌های مرکبی با همین ساختار است، به‌عنوان واژه‌های نامتعارف در درون واژگان آورده، و اسم‌های مرکبی چون کتابخانه، دامپزشک، و دندان‌درد را در واژگان درج نکرده، اما در بخش قواعد تصریح کرده است که ساختار «اسم + اسم» دارای مقوله اسم مرکب است. از آن‌جا که مؤلف از دادگان مشخص و ثابتی استفاده کرده و راه‌حل وی مبتنی بر آمار دقیقی از واژه‌های مورد بررسی اوست، می‌توان روش وی را منطبق با شمّ زبانی و اصل اقتصاد زبانی دانست.

مؤلف، پیش از پرداختن به بحث اصلی خود، که عبارت است از تشخیص مقوله دستوری واژه‌های فارسی در واژگان، اجزای کلام ذیل را به‌عنوان مقوله‌های دستوری زبان فارسی معرفی کرده است: اسم، ضمیر، صفت، اسم - صفت عامل، قید، فعل، عدد، حروف اضافه، و پیوندها. در این میان مقوله اسم - صفت عامل که مؤلف برای اولین بار آن را به‌عنوان یکی از اجزای کلام فارسی معرفی کرده است، نیاز به توضیح مختصری دارد:

شماری از اسم‌ها و صفت‌های فارسی دارای ویژگی‌های صرفی‌ای است که ایجاب می‌کند آن‌ها را در یک مقوله واحد بگنجانیم (ص ۲۷).

او تمام اسم‌ها و صفاتی را که با پسوند - ی، اسم عمل می‌سازند، از مقوله اسم و صفت بیرون آورده و درون مقوله دستوری مشترکی به نام اسم - صفت عامل گنجانده است. در این

معنا، اسم‌هایی چون شاعر، آهنگر، و قالی‌باف اسم عامل هستند، و صفاتی فاعلی هم چون فریبنده، بخشنده، و زنده، و صفات فاعلی مرخمی چون دل‌فریب، جوان‌گرا، و صلح‌طلب صفت عامل هستند، و کل آن‌ها درون مقوله اسم - صفت عامل می‌گنجند. وجه مشترک تمام اسم‌ها و صفت‌های عامل این است که با پسوند - ی اسم عمل می‌سازند: شاعری، آهنگری، قالی‌بافی - فریبندگی، بخشندگی، زندگی - دل‌فریبی، جوان‌گرایی، و صلح‌طلبی. اما استدلالی که مؤلف در لزوم قائل شدن به مقوله اسم - صفت عامل می‌آورد، دقیقاً به مسئله اقتصاد زبانی و مقتضیات واژگان ذهنی مربوط می‌شود. او ابتدا متذکر وجود چندین فرایند واژه‌سازی گوناگون و بسیار فعال می‌شود که حاصل هر کدام از آن‌ها هم صفت عامل می‌تواند باشد و هم اسم عامل. مثلاً «اسم + گر» (اسم عامل: آهنگر، بیمه‌گر؛ صفت عامل: ستمگر، آشوبگر)، «اسم + کار» (اسم عامل: خدمتکار، پیمانکار؛ صفت عامل: تبهکار، گناهکار)، «ستاک حال + - نده» (اسم عامل: راننده، نویسنده؛ صفت عامل: فریبنده، کوبنده)، «اسم + ستاک حال» (اسم عامل: بادسنج، جغرافی‌دان؛ صفت عامل: سنت‌گرا، سلطه‌پذیر). وی سپس می‌افزاید:

[چون] چندین فرایند واژه‌سازی وجود دارد که هم صفت [عامل] می‌سازد و هم اسم [عامل]... اگر به مقوله اسم - صفت عامل قائل نباشیم باید در هر فرایند یکی از مقوله‌های اسم یا صفت را در واژگان ذکر کنیم. چنین روشی با اصل اقتصاد در تعارض است. بنابراین، با توجه به این‌که در فارسی، بسیاری از صفت‌ها از طریق تغییر مقوله به اسم تبدیل می‌شود، مقوله اسم - صفت عامل به لحاظ نظری موجه به نظر می‌رسد (صص ۲۹-۳۰).

مؤلف هم‌چنین یادآور می‌شود که تعداد کمی اسم - صفت عامل بسیط در فارسی وجود دارد، اما مثال‌هایی که به دست می‌دهد همه اسم عامل هستند و هیچ صفت عاملی در میان آن‌ها مشاهده نمی‌شود، مانند بزاز، شاعر، و نجار (ص ۳۰). وی هم‌چنین تصریح می‌کند که این اسم‌ها را، هم در زمره اسم‌های بسیط واژگان خود آورده است، و هم در فهرست جداگانه‌ای با عنوان اسم‌های عامل بسیط، اما هیچ توضیحی نمی‌دهد که چرا این موارد را در دو فهرست جداگانه درج کرده است. در هر حال، همه اسم‌های عامل بسیط با گرفتن پسوند - ی به اسم عمل تبدیل می‌شوند (مانند: بزازی، شاعری و نجاری)، اما هیچ‌کدام از این اسم‌های عمل در واژگان درج نمی‌شوند، بلکه فقط در بخش قواعد واژگان تصریح می‌شود که از ترکیب «اسم

عامل + - ی» همواره اسم عمل به وجود می‌آید. در واقع، مؤلف دو مقوله متفاوت اسم عامل و صفت عامل را در درون مقوله واحدی با عنوان «اسم - صفت عامل» گنجانده است تا مسئله پردازش واژه‌ها در واژگان با هزینه کمتری صورت بگیرد. با استدلال‌های صوری دیگر هم می‌توان نشان داد که گنجاندن این دو مقوله متفاوت در درون یک مقوله واحد، منطبق با شمّ زبانی فارسی‌زبانان است. مثلاً اکثر اسم‌های فارسی با گرفتن پسوند - ی، مبدل به صفت می‌شوند (آب: آبی؛ خوب: خوبی؛ قلم: قلمی؛ کاغذ: کاغذی)، و بسیاری از صفات نیز با گرفتن پسوند - ی مبدل به اسم می‌گردند (بد: بدی؛ خوب: خوبی؛ زرد: زردی). اما چنان‌که دیدیم، تمام اسم‌ها و صفاتی که مؤلف آن‌ها را درون مقوله اسم - صفت عامل قرار داده‌است، با گرفتن پسوند - ی به اسم عمل تبدیل می‌شوند؛ یعنی زبان فارسی با این دسته از اسم‌ها و صفات رفتار مشابهی دارد و به همین دلیل، می‌توان آن‌ها را یک طبقه طبیعی در ذهن فارسی‌زبانان دانست و در درون مقوله واحدی قرار داد.

پس از توضیحات مقدماتی نویسنده، به فصل سوم کتاب با عنوان ساختمان واژه‌های مرکب (ص ۶۱) می‌رسیم که می‌توان آن را نخستین بخش از واژگان حاضر دانست. مؤلف در این بخش حدود ۹۰ ساختار گوناگون را برای کلمات مرکب فارسی به دست داده و در عین حال متذکر شده است که کدامیک از کلماتی که از این ساختارها حاصل می‌آیند لازم نیست در فهرست واژه‌ها ذکر شوند و کدامیک باید به‌عنوان واژه‌های نامتعارف در فهرست واژه‌ها قرار گیرند. مثلاً دومین ساختاری که او معرفی کرده «اسم + صفت فاعلی» است که حاصل آن اسم - صفت عامل (مانند: آغازکننده، شرکت‌کننده، گردآورنده، وام‌دهنده، هستی‌بخشنده، یاری‌دهنده) است. نویسنده تصریح می‌کند که تمام واژه‌های حاصل از این ساخت، اسم - صفت عامل است، بنابراین هیچ‌یک از این واژه‌ها نباید در فهرست واژه‌های واژگان درج شود، یا پنجمین ساختی که در این قسمت معرفی شده ساخت «اسم + صفت حالی» است که حاصل آن دو دسته از کلمات است؛ یکی صفت یا قید (مانند: رقص‌کنان، دل‌هراسان، چشم‌گریبان)، و دیگری اسم (مانند: عقدکنان، برگ‌ریزان، حنابندان). مؤلف متذکر می‌شود که چون این ساخت، به‌عنوان یک قاعده کلی، صفت یا قید تولید می‌کند، لازم نیست آن‌ها را در واژگان درج کنیم، اما اسم‌هایی که از این ساخت حاصل می‌شوند، همه واژه‌های نامتعارف هستند و باید در

واژگان درج شوند.

فصل چهارم کتاب با عنوان ساختمان واژه‌های مشتق (ص ۹۷) به بررسی همین نوع قواعد و مسائل در حوزه فرایند واژه‌سازی اشتقاقی<sup>۱۵</sup> اختصاص دارد. مؤلف واژه‌های مشتق رابه دو گروه مشتق پسوندی و مشتق پیشوندی تقسیم کرده و همه را یک به یک بررسی کرده و مشخص ساخته است که کدامیک از این مشتقات باید در واژگان درج شود و کدامیک توسط بخش قواعد واژگان تولید گردد. در این بخش با بیش از ۱۲۰ ساختار اشتقاقی گوناگون آشنا می‌شویم.

فصل پنجم با عنوان تصریف واژه‌ها (ص ۱۴۱)، به بررسی وندهای تصریفی و واژه‌بست‌ها اختصاص دارد. بدیهی است که هیچ‌یک از صورت‌واژه‌های حاصل از واژه‌سازی تصریفی<sup>۱۶</sup> در فهرست واژه‌های واژگان درج نمی‌شود. در هر حال، رایۀ ساختارهای مولد واژه‌های مرکب و مشتق و نیز صورت‌های تصریفی در این اثر، به همراه مثال‌های متعدد برای هر ساختار، کتاب حاضر را به متن درسی کاملی برای آموزش صرف فارسی نیز تبدیل کرده است. شایسته است که استادان درس صرف فارسی به این جنبه از کتاب حاضر نیز توجه کنند.

از صفحه ۱۶۹ فهرست واژه‌های کتاب آغاز می‌شود که به ترتیب عبارت‌اند از فهرست اسم‌های بسیط، صفت‌های بسیط، فهرست افعال بسیط: ستاک‌های حال و ستاک‌های گذشته، فهرست ضمائر، اصوات، نام‌آواها و شبه‌جمله‌های دعایی، و فهرست واژه‌های پیچیده نامتعارف: اسم‌های پیچیده، صفت‌های پیچیده، قیده‌های پیچیده. از این فهرست‌ها می‌توان به‌عنوان داده‌های جامع و مطمئنی برای انجام تحقیقات گوناگون در زمینه مقوله‌های دستوری فارسی نیز استفاده کرد.

اما از آنچه گذشت پیداست که این فهرست‌ها فقط شامل تکواژهای آزاد<sup>۱۷</sup> و واژه‌های پیچیده‌ای است که بخش قواعد نمی‌تواند مقوله دستوری آن‌ها را تعیین کند. به عبارت دیگر، بر اساس این واژگان، اهل زبان باید مقوله دستوری موارد فوق را به خاطر بسپارند، اما می‌توانند مقوله بقیۀ واژه‌ها را که در فهرست‌ها ذکر نشده است، بر اساس قراین صوری مندرج در بخش قواعد، پیش‌بینی کنند. این واژگان عملاً فاقد فهرستی از تکواژهای وابسته،<sup>۱۸</sup> مثلاً پسوندها و پیشوندهای اشتقاقی و تصریفی است. چنان‌که دیدیم، وضعیت این‌گونه تکواژها در

بخش‌های مربوط به قواعد روشن می‌شود. واژگان حاضر به شکلی مطلوب و منطقی آماده شده است، به گونه‌ای که می‌توان از آن به عنوان مدل یا مرجع مناسب و قابل اعتمادی برای تهیه واژگان پیشرفته‌تر در برنامه‌های گوناگون پردازش زبان و مخصوصاً مباحث مربوط به برچسب‌زنی مقوله‌های دستوری<sup>۱۹</sup> استفاده کرد. بدیهی است که در چنین مواردی، بسته به نوع کاربری، باید تغییراتی در آن به وجود آورد. ساختار واژگان حاضر به گونه‌ای طراحی شده است که به راحتی و بدون نیاز به دگرگون کردن اساس آن، می‌توان چنین تغییراتی را در آن پدید آورد. در این قسمت به اختصار به یکی دو مورد از چنین تغییراتی اشاره می‌کنیم.

از جمله سؤال‌هایی که در مورد مدل‌های گوناگون واژگان ذهنی طرح می‌شود این است که ذهن انسان کدام دسته از کلمات را تک‌تک به خاطر می‌سپارد و معنای کدام دسته را بنا به قواعد صرف زبان پیش‌بینی می‌کند. پاسخ هرچه باشد، نتیجه این خواهد بود که کلمات دسته نخست در فهرست واژه‌ها گنجانده می‌شوند، اما کلمات دسته دوم در آن فهرست درج نمی‌شوند، بلکه تنها قواعد ناظر بر شکل‌گیری آن‌ها در بخش مربوط به قواعد درج می‌شود. چنان‌که دیدیم، واژگان حاضر لزوماً چنین هدفی ندارد، بلکه بیش از هر چیز می‌کوشد تا مقوله دستوری کلمات پیچیده را با استفاده از قراین صوری پیش‌بینی کند. البته این دو هدف تا حد زیادی همپوشی دارند، اما در مواردی نیز این همپوشی چندان به چشم نمی‌خورد یا بسیار کلی‌تر از آن می‌شود که بتوان به قطعیت بدان استناد کرد. به عنوان مثالی از همپوشی کامل این دو بحث می‌توان به ساخت «اسم + مند» اشاره کرد که در فصل چهارم کتاب با عنوان ساختمان واژه‌های مشتق از آن یاد شده است. از این ساخت همواره صفت حاصل می‌شود، و پسوند - مند نیز در آن همواره به معنای دارنده است. مثلاً ثروتمند یعنی دارنده ثروت، خردمند یعنی دارای خرد، قاعده‌مند یعنی دارای قاعده، و غیره (ص ۱۰۸). به عبارت دیگر، بر اساس این ساخت، هم مقوله دستوری و هم معنای واژه‌های حاصل از آن قابل پیش‌بینی است، به همین دلیل لازم نیست که واژه‌های حاصل از این ساخت در فهرست واژه‌های واژگان درج شوند. بدیهی است که واژه‌هایی چون کمند و سمند که ربطی به ساخت فوق ندارند، باید در فهرست واژه‌ها درج شوند. اما چنان‌که گفتیم این همپوشی همواره دقیق و قطعی نیست. مثلاً باز هم در همان فصل چهارم می‌خوانیم که با ساختار «اسم + ک»، اسم‌های گوناگونی

هم چون اسم مکان (لشکرک)، اسم گیاه (میخک)، نام بیماری (خروسک)، نام ابزار (بادامک)، اسم خوردنی (پشمک)، و برخی از اسم‌هایی را که مطلق وابستگی را با اسم پایه می‌رسانند (سنگک)، ساخته می‌شود. مؤلف تصریح می‌کند که این ساخت، به‌عنوان یک قاعده کلی، همواره اسم تولید می‌کند؛ بنابراین، هیچ‌یک از کلمات حاصل از این ساخت در واژگان ذکر نمی‌شود (ص ۱۰۲). بدیهی است که واژه‌های بسیطی چون تَرک، خشتک، ذرک، غلغلک و کپنک، که اسم هستند اما ربطی به ساخت فوق ندارند، باید در فهرست واژه‌ها درج شوند. در هر حال، چنان‌که می‌بینیم این ساخت به‌دقت مقوله دستوری واژه‌های حاصل از آن را پیش‌بینی می‌کند، اما در مورد پیش‌بینی معنای آن چندان دقیق نیست. مثلاً بر اساس این ساخت به قطعیت می‌توان گفت که واژه‌هایی چون خارک (نوعی خرما) و خرسک (نوعی قالی درشت‌بافت نامرغوب)، اسم هستند، اما صرفاً بر اساس این ساخت، معنای دقیق یا حتی معنای کلی آن‌ها را نمی‌توان تعیین کرد. وقتی فقط مقوله دستوری واژه‌ها از روی ساختشان قابل پیش‌بینی است اما معنای آن‌ها قابل پیش‌بینی نیست، می‌توان در صورت لزوم تمهیداتی اندیشید تا ناهمپوشی‌های این دو حوزه، تا حد امکان، جبران شود. مثلاً می‌توان حوزه‌های معنایی را هرچه محدودتر گرفت تا تعداد واژه‌های بیشتری را در خود بپوشانند، و یا می‌توان تعدادی از واژه‌هایی را که معنایشان در هیچ‌کدام از حوزه‌ها نمی‌گنجد یا به هیچ وجه قابل پیش‌بینی نیست، به‌عنوان واژه‌هایی که از حیث معنایی نامتعارف هستند، درون فهرست واژه‌ها ذکر کرد. البته چنین بحث‌هایی در این مختصر نمی‌گنجد، اما قدر مسلم این است که واژگان حاضر برای اولین بار این امکان را به ما می‌دهد تا با استناد به پیکره‌ای انبوه و واقعی از واژه‌های فارسی و نیز فهرست جامعی از قواعد صوری واژه‌سازی در این زبان، بتوانیم درباره شکل واژگان فارسی بحث کرده، نسبت به پیش‌بردن آن اقدام کنیم.

در میان فهرست واژه‌های کتاب حاضر جای فهرست افعال مرکب خالی است. احتمالاً مؤلف کتاب به دو علت خود را درگیر ضبط افعال مرکب فارسی نکرده است: یکی این که افعال مرکب ضبط‌شده در فرهنگ معاصر فارسی (صدری افشاری ۱۳۸۲) و فرهنگ واژه‌های مصوب فرهنگستان (۱۳۷۶-۱۳۸۲)، چنان‌که باید مبین تعداد واقعی این افعال در زبان فارسی امروز نیست، و دیگر این که ضبط دقیق این افعال خود به کتاب دیگری در حجم و اندازه همین

کتاب نیاز دارد. بدیهی است که تهیه فهرستی از افعال مرکب فارسی کار ساده‌ای نیست و انجام چنین مهمی مستلزم روشن شدن بحث‌های نظری بسیاری در زمینه افعال مرکب فارسی و ویژگی‌های آن است و نگارنده این سطور یقین دارد که اگر فقط یک نفر از عهده این کار برآید، آن یک نفر کسی جز مؤلف همین کتاب نخواهد بود. امیدوارم دکتر طباطبایی در ویرایش‌های بعدی کتاب حاضر، فهرست افعال مرکب را نیز به آن بیفزاید و به این ترتیب هم بر فواید واژگان‌نگاری کتاب بیفزاید و هم باب تازه‌ای را برای رسیدن به ماهیت واقعی این افعال در زبان فارسی فتح کند.

از این پس، تمام کسانی که در حوزه واژگان فارسی کار می‌کنند، چه آنان که به مطالعات نظری و شناختی می‌پردازند و چه آنان که درگیر پژوهش‌های مربوط به پردازش زبان و برچسب‌دهی خودکار مقوله‌های دستوری واژه‌های فارسی هستند، از این اثر ارزشمند بی‌نیاز نخواهند بود. به اعتقاد نگارنده، حتی بهتر است از این پس تمام کسانی که به بررسی‌های داده‌بنیاد و هم‌زمانی در زمینه صرف فارسی می‌پردازند، برای یک‌دست کردن تحقیقات و نتایج کار خود، از داده‌ها و رسم‌الخط کتاب حاضر استفاده کنند. چنان‌که اشاره شد، داده‌های این کتاب برگرفته از دو مأخذ معتبر فرهنگ معاصر فارسی (صدری افشاری ۱۳۸۲) و فرهنگ واژه‌های مصوب فرهنگستان (۱۳۷۶-۱۳۸۲) است، و رسم‌الخط آن نیز منطبق با دستورالعمل‌های فرهنگ املائی خط فارسی (صادقی و زندی مقدم ۱۳۸۵) است.

این کتاب با شمارگان بسیار کم ۵۰۰ نسخه چاپ شده است و متأسفانه افتادگی‌ها و اغلاط مطبعی آن نیز کم نیست. از طرفی، ناشر کتاب نیز اهتمام درخوری برای توزیع چنین کتابی به خرج نداده است و تهیه آن، جز از طریق تماس مستقیم با خود ناشر، میسر نیست! امیدواریم این کتاب ارزشمند به صورتی شسته‌رفته‌تر و با شمارگانی بیشتر و توزیعی بهتر، تجدید ویرایش و چاپ شود و در دسترس علاقه‌مندان قرار گیرد.

پی‌نوشت‌ها

1. NLP=Processing Language ralNatu
2. rule based approach
3. lexicon
4. items lexical

5. generative semanticists
6. modular
7. words
8. lexemes
9. forms word
10. mental lexicon
11. lexical models
12. part of speech

۱۳. انتشار فرهنگ *املائی خط فارسی* را به همت دکتر علی اشرف صادقی، حقاً باید ادامه راهی دانست که از اواسط قرن سوم هجری قمری، یعنی از بیش از هزار سال پیش، برای معیارسازی زبان فارسی در ایران آغاز شد و هنوز نیز خوشبختانه با شدت و قوت بسیار ادامه دارد. جای تأسف است که بعضی افراد غیرمتخصص و سیاست‌پیشه در یادداشت‌های شتاب‌زده و جانب‌دارانه خود، به دیده تحقیر به این اثر مهم نگریسته‌اند. اگر قرار باشد که زبان فارسی را در قالب خط فعلی آن وارد عرصه پژوهش‌های پیشرفته و یک‌دست رایانه‌ای بکنیم، چاره‌ای جز پیروی از یک فرهنگ *املائی* معتبر و جامع مانند فرهنگ *املائی خط فارسی* نخواهیم داشت. از آن‌جا که گرایش غالب در این فرهنگ در املائی کلمات مرکب معطوف به جدانویسی آن‌هاست، می‌توان در صورت لزوم و با استفاده از چند دستورالعمل ساده، تکواژهای کلمات را نیز در اقلام پیچیده این فرهنگ نمایش داد.

14. derivational word formation
15. inflectional word formation
16. free morphemes
17. bound morphemes
18. part of speech tagging

### منابع و مأخذ

۱. دبیرمقدم، محمد (۱۳۸۳)، *زبان‌شناسی نظری؛ پیدایش و تکوین دستور زایشی* (ویراست دوم)، تهران: سمت.
۲. صادقی، علی اشرف و زهرا زندی‌مقدم (۱۳۸۵)، *فرهنگ املائی خط فارسی*، تهران: فرهنگستان زبان و ادب فارسی.
۳. صدری افشار، غلامحسین (۱۳۸۲)، *فرهنگ معاصر فارسی*، تهران: فرهنگ معاصر.
۴. طباطبایی، علاءالدین (۱۳۸۲)، *اسم و صفت مرکب در زبان فارسی*، تهران: مرکز نشر دانشگاهی.
۵. فرهنگستان زبان و ادب فارسی (۱۳۸۲)، *فرهنگ واژه‌های مصوب فرهنگستان* (۱۳۷۶-۱۳۸۲)، تهران: فرهنگستان زبان و ادب فارسی.

- Brill, E. (1992), "A simple rule-based part of speech tagger", *Proceedings of the Third Conference on Applied Computational Linguistics (ACL)*, Trento: Italy;
- Carstairs-McCarthy, A. (1992), *Current Morphology*, London: Routledge;
- Chomsky, N. (1965), *Aspects of the Theory of Syntax*, Cambridge: MA: MIT Press;
- (1970), "Remarks on nominalization", in R. A. Jacobs and P. S. Rosenbaum (eds), *Reading in English Transformational Grammar*, Waltham: 184-221;
- Halle, M. (1973), "Prolegomena to a theory of word-formation", *Linguistic Inquiry*, 4: 3-16;
- Haspelmath, M. (2002), *Understanding Morphology*, London: Arnold