

زبان‌شناسی گویش‌های ایرانی

نشانی اینترنتی مجله: <http://jill.shirazu.ac.ir>

بررسی پیکره‌بنیاد زایایی پیشوندهای اشتقاقی در گونه نوشتاری زبان فارسی معاصر

حمیرا عابدینی^۱، امیرسعید مولودی^{۲*} و علیرضا خرمايي^۳

۱- کارشناس ارشد زبان‌شناسی، دانشگاه شیراز، شیراز، ایران

۲- استادیار زبان‌شناسی، دانشگاه شیراز، شیراز، ایران

۳- دانشیار زبان‌شناسی، دانشگاه شیراز، شیراز، ایران

چکیده

یکی از مباحث مورد توجه در ساخت‌واژه به طور عام و در فرایندهای واژه‌سازی به طور خاص مطالعه زایایی است. عمده صاحب‌نظران این حوزه، زایایی را یک مفهوم مدرج می‌دانند و آن را بر روی یک پیوستار تعریف می‌کنند. پژوهش حاضر به منظور پاسخ به سه پرسش زیر صورت می‌گیرد: ۱. پیشوندهای مورد مطالعه در این پژوهش از حیث زایایی در چه مرتبه‌ای نسبت به هم قرار دارند؟، ۲. کدام یک از پیشوندها حافظ مقوله و کدام یک تغییردهنده مقوله به شمار می‌روند؟ و ۳. هر یک از پیشوندها به لحاظ سرنمونی به چه مقوله واژگانی متصل می‌شوند؟ برای انجام این مطالعه، ابتدا با استناد به آثاری که در آن‌ها پیشوندهای زبان فارسی معرفی و بررسی شده‌اند، ۱۹ پیشوند اشتقاقی انتخاب شد. در مرحله بعد داده‌های پژوهش توسط نرم‌افزار انت‌کانک از پیکره ۱۵۰ میلیون کلمه‌ای همشهری ۲ استخراج شد و سپس برحسب شاخص زایایی سراسری باین (۲۰۰۹) و نظریه سرنمون رُش (۱۹۷۵، ۱۹۷۷، ۱۹۷۸) مورد تجزیه و تحلیل قرار گرفت. یافته‌های پژوهش نشان دادند که پیشوند «غیر» دارای بیش‌ترین و پیشوند «بل» دارای کم‌ترین میزان زایایی است. برای پاسخ به پرسش‌های دوم و سوم نیز، هر پیشوند توسط آزمون آماری مجذور خی در نرم‌افزار اسپ‌اس‌اس به طور جداگانه بررسی شد و مشخص شد چه پیشوندهایی حافظ مقوله و چه پیشوندهایی تغییردهنده مقوله هستند. همچنین سرنمون مقوله درون‌داد هر یک از پیشوندها نیز مشخص شد.

تاریخچه مقاله:

دریافت: ۲۷ مهرماه ۹۸

پذیرش: ۴ اسفند ۹۸

واژه‌های کلیدی:

پیشوند اشتقاقی

روش پیکره‌بنیاد

زایایی سراسری

حافظ مقوله

تغییردهنده مقوله

* نویسنده مسئول

آدرس ایمیل: amirsaeid.moloodi@shirazu.ac.ir و amirsaeid.moloodi@gmail.com (امیرسعید مولودی)

۱. مقدمه

از جمله مباحث مطرح در ساخت‌واژه به طور عام و در فرایندهای واژه‌سازی به طور خاص مفهوم زایایی است. تعریف‌های کم‌وبیش مشابهی از زایایی توسط محققان مختلف ارائه شده است که می‌توان آن‌ها را این‌گونه جمع‌بندی نمود: زایایی به معنای تولید صورت‌های جدید به وسیله قواعد و فرایندهای موجود در زبان است (آرونوف^۱، ۱۹۷۶؛ لیبر^۲، ۱۹۸۱؛ آل و بویج^۳، ۱۹۸۱؛ بائر^۴، ۲۰۰۴ و بویج^۵، ۲۰۰۷). هرچه فرآیندی زیاتر باشد معنای آن قابل پیش‌بینی‌تر می‌شود و می‌توان با کنار هم گذاشتن معنای تک‌تک اجزا به معنای کل کلمه پی برد، در حالی که هرچه زایایی فرآیندی کم‌تر باشد از میزان ترکیب‌پذیری معنایی آن نیز کاسته می‌شود (بائر، ۲۰۰۴). لازم به ذکر است که عمده محققان این حوزه همچون بائر (۱۹۹۲، ۱۹۹۶)، هی^۶ (۲۰۰۳) و باین^۷ (۱۹۹۲، ۱۹۹۳، ۲۰۰۹) زایایی را مفهومی مدرج می‌دانند و آن را بر روی یک پیوستار مطرح می‌کنند.

فرایند اشتقاق^۸ یکی از مهم‌ترین و پرتکرارترین راه‌های واژه‌سازی به حساب می‌آید که دارای دو نوع پرتکرار پیشوندافزایی و پسوندافزایی است. در نوع نخست، وند به پیش از پایه و در نوع دوم به پس از پایه متصل می‌شود (بویج، ۲۰۰۷). با توجه به مقایسه تعداد پیشوندها و پسوندها در مطالعات محققان ایرانی (فرشیدورد، ۱۳۸۶، رواقی، ۱۳۸۸، کلباسی، ۱۳۹۱ و گیوی و انوری، ۱۳۹۲)، مشاهده شد که تعداد پیشوندها به مراتب بسیار کم‌تر از تعداد پسوندها است. از طرف دیگر، غالب پژوهش‌های صورت‌گرفته در زمینه فرایند اشتقاق به بررسی پسوندهای زبان فارسی پرداخته‌اند و موضوعات مربوط به پیشوندها به ندرت مطالعه شده است. در اثر حاضر، به منظور پرکردن خلأ مذکور، به بررسی زایایی پیشوندهای اشتقاقی زبان فارسی معاصر با روشی پیکره‌بنیاد پرداخته می‌شود. شناسایی درجه زایایی پیشوندها علاوه بر اهمیت نظری آن، از حیث کاربردی به ویژه در زمینه آموزش زبان فارسی به غیرفارسی‌زبانان بسیار کارآمد است؛ به این معنی که با دانستن درجه زایایی پیشوندها می‌توان ترتیب آموزش آن‌ها را طراحی نمود چراکه در حوزه آموزش زبان،

1. Aronoff
2. Lieber
3. Al and Booij
4. Bauer
5. Booij
6. Hay
7. Baayen
8. derivation

ابتدا ساخت‌ها و واژه‌های زایا و سپس ساخت‌ها و واژه‌های کم‌تر زایا آموزش داده می‌شوند (آکیفه^۱ و همکاران، ۲۰۰۷). در اثر حاضر، پس از معرفی پیشینه مطالعاتی، چارچوب نظری و روش پژوهش، داده‌های پژوهش بر حسب شاخص زایایی سراسری^۲ بررسی می‌شوند و از حیث حافظ مقوله بودن یا نبودن مورد مطالعه قرار می‌گیرند. همچنین سرنمون^۳ مقوله درونداد هر پیشوند نیز به صورت مجزا مطالعه می‌شود.

۲. پیشینه پژوهش

در این بخش ذیل دو عنوان کلی پژوهشگران ایرانی و غیرایرانی، پژوهش‌های انجام‌یافته در زمینه زایایی مرور می‌شوند.

۲-۱. پژوهشگران ایرانی

بدخشان (۱۳۸۹) به بررسی روش‌های اندازه‌گیری زایایی صرفی، مشخصات فرآیندهای صرفی و محدودیت‌های موجود بر عملکرد فرآیندهای صرفی در زبان فارسی پرداخته است. وی بر اساس چارچوب نظری پلاگ^۴ (۲۰۰۳) و بررسی پایگاه داده‌های زبان فارسی و نیز تعدادی از فرهنگ‌های زبان فارسی، خاطر نشان کرده است که زایایی امری مطلق نیست بلکه مفهومی پیوستاری است و رابطه مستقیمی بین یک فرآیند صرفی در زمان حال و تعداد صورت‌های زبانی جدید که به وسیله آن فرآیند ساخته می‌شوند وجود دارد. همچنین به عقیده وی، بین زایایی یک فرآیند صرفی و محدودیت‌های ناظر بر آن رابطه معکوس وجود دارد. همچنین وی در اثر خود، محدودیت‌های صرفی، نحوی، آوایی، و معنایی بر عملکرد زایایی فرآیند ترکیب در زبان فارسی را مورد بررسی قرار داده است.

فراگردی (۱۳۸۹) به بررسی پیشوندهای زبان فارسی، تعیین انواع تصریفی و اشتقاقی آن‌ها در زبان فارسی معاصر، مقایسه کارکرد این دو گروه از پیشوندها در زبان فارسی و طبقه‌بندی آن‌ها به لحاظ بسامد وقوع، زایایی، و گستره معنایی پرداخته است. او برای جمع‌آوری داده‌های خود از

1. O'keeffe
2. global productivity
3. prototype
4. Plag

فرهنگ بزرگ سخن استفاده نموده است. چارچوب نظری این پژوهش مفهوم زایایی و نیمه‌زایایی بائر (۲۰۰۴) است و هدف آن مشخص کردن تعداد پیشوندهای کاربردی در فارسی معاصر، نقش پیشوندهای اشتقاقی در واژه‌سازی، تعیین وضعیت پیشوندهای اشتقاقی نسبت به پیشوندهای تصریفی از حیث نزدیکی یا عدم نزدیکی به پایه، و چگونگی توزیع تکواژگونه‌های این تکواژها و گستره معنایی آن‌ها می‌باشد. برحسب یافته‌ها، حداقل و حداکثر بسامد وقوع پیشوندهای سترون از ۲ تا ۱۵ در نوسان است. در این پژوهش، پیشوندهای نیمه‌زایا بسامد بین ۴۷ (پیشوند غیر-) تا ۱۹۷ (پیشوند بر-) داشته‌اند و بسامد پیشوندهای زایا از ۳۲۲ (پیشوند سر-) به بالا بوده‌اند. با توجه به نتایج آماری به‌دست‌آمده از تحلیل داده‌ها، وندهایی با بسامد زیر ۵۰ سترون، کم‌تر از ۲۰۰ نیمه‌زایا و در نهایت بیش از ۲۰۰ زایا قلمداد شده‌اند. البته باید خاطر نشان کرد که مشخص نیست چگونه می‌توان چنین تعبیر آماری را برای محاسبه زایایی پیشوندها معرفی کرد و این اعداد را به عنوان حد و حدود میزان زایایی در نظر گرفت.

امیرارجمندی و عاصی (۱۳۹۲) با استفاده از روش‌های سنجش زایایی باین (۱۹۹۴) و پایگاه داده‌های زبان فارسی، زایایی فرایند ترکیب را در زبان فارسی مطالعه کرده‌اند. آن‌ها معتقدند، برای مطالعه زایایی در پیکره‌های کوچک‌تر، استفاده از فرمول‌هایی که در آن‌ها بسامد موردی نقش مستقیمی نداشته باشد نتیجه مطمئن‌تری را به دست خواهد داد. همچنین، طبق بررسی‌های این پژوهش، در زبان فارسی فرآیند ترکیب «اسم + اسم» بیش‌ترین زایایی را دارد و فرآیند «ضمیر + حرف‌اضافه + ضمیر» از کم‌ترین میزان زایایی برخوردار است. بر حسب یافته‌ها، میان زایایی در زمان‌های حال و گذشته تفاوت وجود دارد. اگر زایایی ساخت‌واژی در زمان گذشته مد نظر باشد بسامد بالای نمونه‌ها نشانه زایایی است، اما اگر زایایی در زمان حال مورد نظر باشد، کثرت نمونه‌ها به معنای زایایی نیست بلکه افزایش واژه‌های با بسامد پایین نمونه، نشانه زایایی است.

۲-۲. پژوهشگران غیرایرانی

باین و رنوف^۱ (۱۹۹۶) به بررسی زایایی پنج‌وند اشتقاقی انگلیسی *ly*، *-ness*، *-ity*، *-un*، و *in-* در در پیکره ۱۸ میلیون کلمه‌ای کوبیلد^۲ پرداخته‌اند. آن‌ها پس از تمرکز بر روی تک‌نمودها و رابطه آن‌ها با نواژه‌ها و همچنین نقش تعداد آن‌ها بر درجه رشد واژگان، به بررسی ویژگی‌های ساختاری

1. Renouf
2. COBUILD

ابداعات واژگانی پرداخته‌اند و سپس از طریق شاخص P باین (۱۹۹۲، ۱۹۹۳، ۱۹۹۴) زایایی این پنج‌وند را مطالعه کرده‌اند. پس از انجام محاسبات مشخص شد که برخلاف ادعای کانن^۱ (۱۹۸۸) که با مطالعه بر روی فرهنگ مریم‌وبستر^۲ (۱۹۶۱-۱۹۸۱) ادعا کرده بود پسوند -ly غیرزایاست، وندهای -ly، -un- و -ness کاملاً زایا هستند. یکی از نتایج مهم این مقاله تأیید فرمول باین (۱۹۹۲، ۱۹۹۳، ۱۹۹۴) در محاسبه زایایی است.

هی (۲۰۰۳) پس از معرفی دو دیدگاه زایایی مدرج و زایایی مطلق، با ارائه شواهدی رویکرد نخست را تأیید می‌کند. او بر این باور است که برخی وندها در زبان انگلیسی (مانند -ness) بسیار زایا هستند و برخی (مانند -th) اصلاً زایا نیستند و برخی دیگر نیز با بعضی پایه‌های خاص می‌توانند زایا باشند. وندهایی که صورت‌های قابل تجزیه‌تری می‌سازند به احتمال زیاد زایاتر از وندهایی هستند که واژه‌هایی با درجه تجزیه‌پذیری کم‌تر می‌سازند. این نکته مؤید این مطلب است که زایایی ساخت‌واژی بازتاب واژگان ذهنی است. او در ادامه، شاخص زایایی P را به نقل از باین (۱۹۹۲، ۱۹۹۳، ۱۹۹۴) ارائه و کارایی آن را در شناسایی زایایی وندها تشریح می‌کند. بائر (۲۰۰۴) در پاسخ به این پرسش که آیا زایایی می‌تواند اندازه‌گیری شود یا خیر، خاطر نشان می‌کند که در این زمینه نمی‌توان زایایی را با فراوانی نوعی یکی دانست زیرا به نقل از آرونوف (۱۹۷۶) محدودیت‌های بسیاری روی پایه‌ها وجود دارند که باعث می‌شوند فقط احتمال وقوع فرآیندهای خاصی وجود داشته باشد. هم‌چنین اتکا به نسبت واژه‌های موجود به واژه‌های ممکن نیز اشکالاتی دارد و در نتیجه باید فراوانی موردی و تک‌نمودها را وارد محاسبات کرد تا بتوان به نتایج قابل‌اعتمادی در تعیین زایایی فرایندهای واژه‌سازی دست یافت. به همین منظور، او دو مقیاس P و P* باین (۱۹۹۲، ۱۹۹۳، ۱۹۹۴) را شاخص‌های مناسبی برای محاسبه زایایی معرفی می‌کند.

باین (۲۰۰۹) در پژوهش خود به پرسش‌های زیر پاسخ می‌دهد: (۱) شرایط تحقق زایایی چیست؟، (۲) آیا زایایی اساساً یک پدیده مدرج است؟، و (۳) درجه زایایی مقوله ساخت‌واژی چگونه باید اندازه‌گیری شود؟ او معتقد است مقوله‌ای که اعضای بیش‌تری دارد نسبت به بقیه زایاتر است. هم‌چنین، برای محاسبه زایایی، فرمول‌هایی را ارائه می‌کند. اولین فرمول، نسبت فراوانی نوعی

1. Cannon
2. Merriam-Webster

(V) به فراوانی موردی (N) واژه‌های پیکره است که به آن زایایی بالفعل^۱ گفته می‌شود. دومین مقیاس زایایی P است که درجه زایایی^۲ یا زایایی بالقوه^۳ نامیده می‌شود. و سومین مقیاس اندازه‌گیری زایایی P* است که با نام زایایی کلی^۴، زایایی سراسری یا زایایی مشروط به تک نمود^۵ شناخته می‌شود. به عقیده باین (۲۰۰۹) زایایی به عوامل متفاوتی بستگی دارد که از آن جمله می‌توان به سیاق^۶، گذر زمان، و واژگان ذهنی اشاره کرد.

لیو و شن^۷ (۲۰۱۲) پژوهشی بر روی ویژگی‌های مختلف پسوند -esque انجام داده‌اند که از جمله این ویژگی‌ها، بررسی زایایی این پسوند به روش پیکره‌بنیاد است. این پژوهش بر روی پیکره COCA^۸ که بیش از ۴۱۰ میلیون واژه (در زمان انجام این پژوهش) را دارد، انجام شده است. چارچوب نظری و محاسباتی این پژوهش نیز اثر پلاگ (۲۰۰۳) است. با استفاده از فرمول محاسبه زایایی برحسب مقیاس P و داده‌های به‌دست‌آمده از پیکره، درجه زایایی پسوند -esque برابر با ۹/۸۰٪ به دست آمده است که بر طبق آرای پلاگ، نسبتاً زیاد است. نتیجه به‌دست‌آمده نشان می‌دهد که پسوند -esque زایا به شمار می‌رود. هم‌چنین این پسوند معمولاً کلمات جدید در بافت‌های رسمی مانند مجلات، روزنامه‌ها و هم‌چنین مطالب دانشگاهی می‌سازد و در بافت‌های گفتاری به ندرت استفاده می‌شود.

همان‌طور که مشاهده شد، تمامی پژوهش‌های انجام‌شده در این زمینه، زایایی را مفهومی مدرج معرفی کرده‌اند و عمدتاً از شاخص‌های باین (۱۹۹۲، ۱۹۹۳، ۱۹۹۴، ۲۰۰۹) برای محاسبه آن استفاده کرده‌اند.

۳. چارچوب نظری پژوهشگاه علوم انسانی و مطالعات فرهنگی

چارچوب نظری پژوهش حاضر دارای دو بخش است. در بخش نخست شاخص‌های محاسبه زایایی به همراه نقاط قوت و ضعف هر یک از آن‌ها معرفی می‌شوند. در بخش دوم نیز نظریه سرنمون

1. realized productivity
2. degree of productivity
3. potential productivity
4. general productivity
5. hapax-conditioned degree of productivity
6. register
7. Liu and Shen
8. Corpus of Contemporary American English

معرفی می‌شود. لازم به ذکر است که بخش نخست چارچوب نظری عمدتاً برای پاسخ به پرسش نخست و بخش دوم برای پاسخ به پرسش‌های دوم و سوم استفاده می‌شود.

۳-۱. روش‌های اندازه‌گیری زایایی

آرونوف (۱۹۷۶) یکی از اولین پژوهشگرانی است که شاخصی برای محاسبه زایایی ارائه می‌کند. او زایایی یک قاعده را به صورت نسبت واژه‌های تولیدشده توسط آن قاعده (واژه‌های موجود) به تعداد واژه‌های ممکن تعریف می‌کند که می‌توانند توسط آن قاعده تولید شوند. فرمولی که باین و لیبر^۱ (۱۹۹۱) برای این مفهوم ارائه داده‌اند $I = \frac{V}{S}$ می‌باشد که در آن، I همان شاخص زایایی، V تعداد واژه‌های موجود و S تعداد واژه‌های ممکن است. از جمله اشکالات اساسی این فرمول این است که با توجه به تعداد محدودیت‌های ممکن بر روی پایه، نمی‌توان تعداد واژه‌های ممکن یک قاعده واژه‌سازی را به طور دقیق مشخص کرد.

از دیگر فرمول‌های محاسبه زایایی، فرمول $P = \frac{n_1}{N}$ است که توسط باین (۱۹۹۲) ارائه شده است. در این فرمول، n_1 تعداد تک‌نمودها و N کل فراوانی موردی کلماتی است که با آن فرآیند خاص در پیکره تولید شده‌اند. از نظر باین (۲۰۰۹) این روش، میزان رشد واژگان در یک مقوله خاص را تخمین می‌زند و میزان احتمال مواجهه با نوواژه‌ها در پیکره را محاسبه می‌کند. به عقیده باین و لیبر (۱۹۹۱)، درجه زایایی که با این فرمول به دست می‌آید وابسته به حجم پیکره است یعنی هرچه پیکره بزرگ‌تر باشد عدد به دست آمده قابل اتکاتر است، اما نقص آن این است که دیگر با استفاده از آن نمی‌توان به مقایسه پیکره‌ها با یکدیگر پرداخت. از این رو، باین (۲۰۰۹) شاخص دیگری را با نام «زایایی سراسری» یا P^* ارائه می‌کند. شاخص مورد نظر با فرمول $P^* = \frac{n_1 \cdot E \cdot t}{h \cdot t}$ محاسبه می‌شود. در این فرمول، E علامت مقوله ساخت‌واژی، t تعداد فراوانی موردی پیکره، و h تعداد تک‌نمودها است. به عقیده باین (۲۰۰۹)، استفاده از پیکره و انجام محاسبات بر روی داده‌های آن به پژوهش‌گر این امکان را می‌دهد تا دریا به دریا در سبب زایایی در سیاق‌های مختلف (از زبان گفتاری به نوشتاری)، و حتی در طی زمان، چگونه می‌تواند متفاوت باشد. در پژوهش حاضر نیز به دلیل برتری پیش‌گفته شاخص P^* نسبت به شاخص P ، از شاخص P^* در محاسبه زایایی پیشوندهای اشتقاقی استفاده می‌شود.

1. Lieber

۲-۳. نظریهٔ سرنمون

نظریهٔ سرنمون در برابر ناکارآمدی نظریهٔ سنتی ارسطویی موسوم به مدل شرایط لازم و کافی^۱ ایجاد شد. بر حسب این مدل، برای این که یک عضو به مقوله‌ای تعلق داشته باشد، باید دارای شرایطی باشد که برای تعریف آن مقوله لازم و در عین حال کافی است. مدل ارسطویی مقوله‌بندی دارای شش ویژگی زیر است:

- ۱) مقوله‌بندی وابسته به شرایط یا ویژگی‌های ثابتی است.
- ۲) هر شرط به طور جداگانه لازم است.
- ۳) شرایط دوگانه (بله یا خیر) هستند.
- ۴) عضویت در مقوله، موضوعی دوگانه (بله یا خیر) است.
- ۵) مقوله‌ها مرز مشخص دارند.

۶) همهٔ اعضای یک مقوله از حیث عضویت در آن مقوله در یک وضعیت برابر قرار دارند (لبنر، ۲۰۰۲؛ اوانز و گرین، ۳، ۲۰۰۶).

ویتگنشتاین^۴ (۱۹۵۸) از اولین پژوهشگرانی بود که با طرح این سؤال که چگونه می‌توان مقوله «بازی» را تعریف کرد و با ارائهٔ پاسخ‌هایی برای این پرسش، نظریهٔ ارسطویی مقوله‌بندی را به چالش کشید. او باور دارد که اعضای مختلف مقولهٔ بازی دارای مجموعه‌ای از ویژگی‌های مشترک نیستند که بتوان به واسطهٔ آن‌ها بازی را از غیربازی متمایز کرد. همچنین، مرز این مقوله مبهم است و مشخصه‌ای قطعی وجود ندارد که به همهٔ اعضا نسبت داده شود. همچنین، اگرچه مسلماً ویژگی‌هایی برای تعریف هر مقوله وجود دارد؛ اما برخی از اعضا دارای تعدادی از این ویژگی‌ها و برخی دیگر در بردارندهٔ تعدادی دیگر هستند و این ویژگی‌ها بین همهٔ اعضا مشترک نیستند. آرای ویتگنشتاین در مورد مقوله‌بندی توسط رُش (۱۹۷۵، ۱۹۷۷، ۱۹۷۸) پی گرفته شد و به صورت منسجم در نظریه‌ای به نام نظریهٔ سرنمون ارائه گردید. سرچشمهٔ این نظریه، دسته‌بندی رنگ‌واژه‌ها بود. این موضوع که در طیف رنگ‌ها و دسته‌بندی آن‌ها برخی رنگ‌ها نادیده گرفته می‌شوند یا نمی‌توان طبق مدل ارسطویی آن‌ها را در یک مقولهٔ مشخص جای داد محققان را به سمت تعریف یک نظریهٔ جدید سوق داد. آن‌ها دریافتند که برخی اعضای یک مقوله نسبت به سایر اعضا، نمونه‌های بهتری

1. Necessary and Sufficient Conditions
2. Löbner
3. Evans and Green
4. Wittgenstein

برای آن مقوله هستند و این نمونه‌های بهتر «سرنمون» نامیده شدند. نتایج پژوهش‌های رُش (۱۹۷۵، ۱۹۷۷، ۱۹۷۸) نشان داد که تعریف مقوله‌ها برحسب شرایط لازم و کافی با واقعیت روان‌شناختی مقوله‌ها در تضاد است چرا که در بازیابی ذهنی اعضای مختلف یک مقوله، سرنمون‌ها سریع‌تر از سایر اعضا بازیابی می‌شوند. وی با انجام آزمایش‌های متعدد تجربی چهار ویژگی مهم نظریه سرنمون را این گونه بیان کرد: ۱. مقوله‌ها دارای ساختار شباهت خانوادگی هستند. ۲. اعضای مقوله به اندازه یکسان نماینده مقوله خود نیستند. ۳. مرز مقوله‌ها در حاشیه‌ها مبهم است و ۴. مقوله‌ها را نمی‌توان بر اساس مجموعه‌های واحدی از مشخصه‌های لازم و کافی تعریف کرد.

۴. روش پژوهش

ابتدا پیشنندهای مورد نظر پژوهش با بررسی دقیق آثاری همچون فرشیدورد (۱۳۸۶)، فراگردی (۱۳۸۹)، فرهنگ بزرگ سخن، فرهنگ عمید، فرهنگ معین و لغت‌نامه دهخدا انتخاب شدند. لازم به ذکر است که در پژوهش حاضر، سعی شده است تا تمام پیشنندهای زنده و در حال استفاده زبان فارسی انتخاب شوند. بنابراین، پیشنندهای منسوخ‌شده (مانند هام- و دش-) یا تصریفی (مانند ن- و می-) در این اثر مورد بررسی قرار نگرفته‌اند. با در نظر گرفتن مؤلفه‌های پیش گفته برای انتخاب پیشنندها، در نهایت ۱۹ پیشنهاد اشتقاقی برای بررسی انتخاب شدند که به قرار زیر می‌باشند:

ابر- با- باز- بل- بلا- بی- پس- پیش- ضد- غیر- فرا- فراز- فرو- فرود- لا- نا- وا- ور- هم-.

پیکره مورد مطالعه در این پژوهش پیکره همشهری ۲ (آل احمد و همکاران، ۲۰۰۹) است. این پیکره ۱۵۰ میلیون کلمه‌ای که حاوی ۳۱۸ هزار سند است، آرشیو روزنامه همشهری از سال‌های ۱۳۷۵ تا ۱۳۸۶ می‌باشد. پیکره مورد نظر یک داده الکترونیکی ماشین‌خوان است. بنابراین می‌توان از یک نرم‌افزار واژه‌نما^۱ به منظور استخراج اطلاعات از آن استفاده نمود. در پژوهش حاضر از نرم‌افزار واژه‌نمای انت‌کانک^۲ (آنتونی^۳، ۲۰۱۴) برای استخراج واژه‌های دارای پیشنندهای مورد نظر استفاده می‌شود.

1. concordancer
2. AntConc
3. Anthony

روش استخراج داده‌ها به این صورت است که ابتدا تمامی واژه‌هایی که ۱۹ پیشوند فوق در ابتدای آن‌ها قرار دارند با استفاده از نرم‌افزار انت‌کانک و با دستور جستجوی پیشوند* و پیشوند # استخراج می‌شوند.^۱ در مرحله بعد، برای حذف داده‌های نامرتب یا اصطلاحاً فیلتر کردن داده‌ها، تمامی داده‌های به دست آمده از جستجوی فوق به همراه بافت متنی‌شان به نرم‌افزار اکسل منتقل و سپس، داده‌های خام به ترتیب حروف الفبا مرتب‌سازی می‌شوند. پیش از حذف موارد نادرست و نامرتب، حجم داده‌های مورد بررسی ۱۰،۳۹۰،۷۷۳ بوده است. پس از بررسی مورد به مورد تمامی این حدوداً ۱۰ میلیون واژه و حذف موارد نادرست، ۱،۵۲۸،۹۱۶ مورد باقی مانده‌اند. فراوانی موردی واژه‌های دارای این پیشوندها به تفکیک هر پیشوند به شرح زیر می‌باشند:

ابر: ۳،۳۵۰ مورد، با: ۴۷،۲۸۸ مورد، باز: ۱۸۹،۰۳۰ مورد، بل: ۸ مورد، بلا: ۱۲،۹۴۳ مورد، بی: ۱۶۱،۱۲۵ مورد، پس: ۱۱،۱۴۳ مورد، پیش: ۱۹۴،۷۸۸ مورد، ضد: ۱۹،۲۹۷ مورد، غیر: ۱۲۶،۵۱۹ مورد، فرا: ۴۰،۱۰۲ مورد، فراز: ۴۸ مورد، فرو: ۱۷،۸۳۳ مورد، فرود: ۱،۶۵۹ مورد، لا: ۳۰،۰۲۲ مورد، نا: ۲۱۳،۴۹۶ مورد، وا: ۸۸،۵۹۸ مورد، ور: ۲،۸۹۵ مورد، هم: ۳۹۶،۶۷۹ مورد.

پس از حذف داده‌های نامرتب، تنوعی از آن‌ها به دست آمد که به قرار زیر است:

– داده‌های بسیط: واژه‌هایی که (از لحاظ هم‌زمانی) بسیط محسوب می‌شدند؛ مانند «ابراهیم، ابرشناسی، بادام، بادمجان، بازرگان، بازی، فروردین، فروشنده، واجد، لابراتوار، لابه»

– اسامی خاص؛ مانند «آلبرت) بازبان، (فروغ) فروهیده، فرووالدتر، لاریجانی»

– موارد دارای غلط املائی و نگارشی؛ مانند «بی‌مسئولیت، ناموزن، ابرترجمه، ابردمرد، بازید (بازدید)، فرونغلطد، فروگاه (فرودگاه)، فروگزار (فروگذار)، واستگی (وارستگی)، واگذاری (واگذاری)، لاقل (لاقل)» یا «کاهش مقاومت مکانیکی خاک و ریزش دیواره‌ها، سرشار از ظرافت و ریزه‌کاری»

۱. از آنجا که در پیکره همشهری، خطاهای مربوط به فاصله‌گذاری نسبتاً زیاد است، بسیاری از پیشوندها به اشتباه به پایه خود نچسبیده‌اند و یک فاصله کامل بین آن‌ها و پایه‌شان وجود دارد. با انتخاب دستور جستجوی فوق، هم پیشوندهایی که به درستی به پایه‌شان چسبیده‌اند و هم آن‌هایی که به اشتباه با یک فاصله کامل نگاشته شده‌اند، بازیابی می‌شوند.

— مواردی که پیشوند نبودند؛ بلکه حرف اضافه بودند؛ مانند «او با آیین نیاکان شما سر جنگ دارد، هنرپیشه اصلی فیلم سیگار خود را با نفرت خاموش کند، ضد آسایش و رفاه عموم است.»

برای پاسخ به پرسش دوم و سوم، باید داده‌ها از حیث درون‌داد و برون‌دادشان برچسب‌زنی می‌شدند. به همین منظور یک ستون در نرم‌افزار اکسل، فقط به مقوله واژگانی درون‌داد و ستونی دیگر به مقوله واژگانی برون‌داد تعلق گرفت. مقوله‌های واژگانی استفاده‌شده برای این هدف عبارتند از: اسم، صفت، قید، فعل، مصدر، ضمیر، شبه‌جمله و ضمیر متقابل.

در نهایت نیز برای تحلیل آمارهای مربوط به پرسش‌های دوم و سوم از آزمون مجذور خی^۱ در نرم‌افزار اسپ‌اس‌اس^۲ استفاده شد.

در این بخش لازم است به برخی از مهم‌ترین نکته‌ها و قراردادهای برچسب‌زنی داده‌ها اشاره کرد:

— کلماتی که یکسان بودند اما دو املای متفاوت داشتند، یک نوع در نظر گرفته شدند. مانند «بی‌محتوی» و «بی‌محتوا»، «غیردمکراتیک» و «غیردموکراتیک».

— کلماتی که تفاوت در سیاق داشتند، مستقل در نظر گرفته شدند. مانند «بی‌معنی» و «بی‌معنا»، «بی‌نشان» و «بی‌نشون».

— صورت‌های تصریف‌شده فعل‌ها در همه صیغه‌ها و همه زمان‌ها، یک مدخل در نظر گرفته شدند.

— در تعیین مقوله درون‌داد و برون‌داد واژه‌ها، و همچنین بسیط و غیربسیط بودن آن‌ها، علاوه بر بافت واژه مورد نظر، از چهار فرهنگ دهخدا، عمید، معین و فرهنگ بزرگ سخن استفاده شد.

— کلماتی که از حیث صورت، یکسان اما از حیث مقوله واژگانی متفاوت بودند، جداگانه فهرست شده و برچسب خوردند. مانند: «باهمیت» که درون‌داد یکسان اسم داشت اما با دو برون‌داد قید و صفت در بافت مشاهده شد. واژه‌های «ضد اخلاق» با درون‌داد اسم و دو برون‌داد اسم و صفت؛ و «بلا تردید» با درون‌داد اسم و دو برون‌داد قید و صفت از جمله سایر موارد بودند.

— برچسب‌زنی برون‌داد بر اساس واژه ساخته‌شده با پیشوند انجام شد. به‌طور مثال، برچسب برون‌داد دو واژه «بی‌دانش» و «بی‌دانشی» هر دو صفت در نظر گرفته شد؛ چراکه ابتدا از اتصال پیشوند

1. chi-square

2. SPSS

به پایه یک واژه ساخته شده، سپس این واژه در ترکیب با وندها یا کلمات دیگر تبدیل به واژه جدید شد.

- صورت کامل کلماتی که به صورت اصطلاح رواج داشتند، به عنوان مدخل در نظر گرفته شد؛ مانند «بی چشم و رو، با جان و دل، بی آب و علف، بی جیره و موجب، بی چون و چرا».
- در برجسب زنی درونداد، علاوه بر در نظر گرفتن سلسله مراتب افزودن وند به پایه، این گونه قرارداد شد که اگر واژه‌ای که وند به آن اضافه شده است، در واژگان موجود باشد و مورد استفاده قرار بگیرد، همان واژه به عنوان درونداد در نظر گرفته می‌شود و برجسب زنی می‌شود. اما اگر پایه‌ای که وند به آن افزوده شده موجود نباشد، در صورتی که صورت کوتاه‌تر آن موجود باشد، آن صورت کوتاه‌تر درونداد محسوب می‌شود. مثل «ناآگاهانه» که هر دو صورت (آگاه) و (آگاهانه) موجود هستند اما درونداد مورد نظر ما (آگاهانه) است. یا در واژه «بازنویسی» چون (نویسی) واژه موجود و مورد استفاده‌ای نیست، صورت کوتاه‌تر یعنی (نویس) درونداد محسوب می‌شود.
- در کلماتی که به صورت هم‌زمان پیشوند و پسوند به پایه افزوده می‌شود، صورت بدون پیشوند و پسوند به عنوان درونداد در نظر گرفته می‌شود؛ مانند «تارضایتی» که درونداد آن «رضایت» است.
- بروندهای دوبرجسبی بر اساس بافت برجسب زنی می‌شوند. مثلاً «غیرآگاهانه» که هم نقش صفت و هم نقش قید را می‌توانست داشته باشد، با توجه به بافت، صفت یا قید برجسب زنی می‌شود.

۵. تحلیل داده‌ها

۵-۱. سنجش درجه زایایی پیشوندهای اشتقاقی در زبان فارسی

همان‌طور که در بخش چارچوب نظری اشاره شد، در پژوهش حاضر، به دلیل محدودیت‌های کم تر و کارایی بیشتر شاخص زایایی سراسری، از این شاخص برای محاسبه درجه زایایی پیشوندهای اشتقاقی زبان فارسی استفاده می‌شود. برای محاسبه این مقیاس به تعداد تک‌نمودهای هر پیشوند به طور جداگانه و هم‌چنین تعداد کل تک‌نمودهای پیکره نیاز است. داده‌های مورد نظر که برای هر پیشوند به صورت مجزا از پیکره استخراج شده‌اند، در جدول (۱) قابل مشاهده است. مجموع تک‌نمودهای کل پیشوندهای مورد بررسی در پیکره برابر با ۳۸۵۴ مورد می‌باشد. با جای‌گذاری آن‌ها در فرمول معرفی شده و محاسبه هرکدام نتایج زیر به دست می‌آید:

جدول (۱) زایایی سراسری پیشوندهای فارسی در مقایسه باهم

P*	V _(1.N)	V _(1.C.N)	پیشوند
۰/۳۵۸۳	۳۸۵۴	۱۳۸۱	غیر
۰/۱۷۲۳	۳۸۵۴	۶۶۴	نا
۰/۱۳۷۰	۳۸۵۴	۵۲۸	بی
۰/۱۳۲۶	۳۸۵۴	۵۱۱	ضد
۰/۰۵۶۰	۳۸۵۴	۲۱۶	فرا
۰/۰۴۱۳	۳۸۵۴	۱۵۹	هم
۰/۰۲۶۷	۳۸۵۴	۱۰۳	ابر
۰/۰۱۹۷	۳۸۵۴	۷۶	باز
۰/۰۱۴۳	۳۸۵۴	۵۵	پیش
۰/۰۱۳۲	۳۸۵۴	۵۱	فرو
۰/۰۰۸۰	۳۸۵۴	۳۱	وا
۰/۰۰۶۷	۳۸۵۴	۲۶	پس
۰/۰۰۵۷	۳۸۵۴	۲۲	بلا
۰/۰۰۲۳	۳۸۵۴	۹	با
۰/۰۰۲۱	۳۸۵۴	۸	فراز
۰/۰۰۱۶	۳۸۵۴	۶	ور
۰/۰۰۰۸	۳۸۵۴	۳	فروود
۰/۰۰۰۸	۳۸۵۴	۳	لا
۰/۰۰۰۵	۳۸۵۴	۲	بل

با توجه به این نتایج می‌توان مشاهده کرد که پیشوند «غیر-» با درجهٔ زایایی ۰/۳۵۸۳ در مقایسه با دیگر پیشوندها زایاترین پیشوند می‌باشد و پیشوند «بل-» با درجهٔ ۰/۰۰۰۵ در آخرین رتبهٔ زایایی قرار دارد.

سنجش میزان زایایی با مقایسهٔ فراوانی نوعی از ابتدایی‌ترین روش‌هایی است که برای اندازه‌گیری زایایی استفاده می‌شود. اگرچه باین (۱۹۹۲، ۱۹۹۳، ۲۰۰۹) در پژوهش‌های خود به این روش اشاره نکرده است، اما به دلیل این که این روش در برخی دیگر از پژوهش‌ها به عنوان ملاک محاسبهٔ زایایی معرفی شده است، در این بخش به آن اشاره می‌شد. بویج (۲۰۰۷) بیان می‌کند که با مقایسهٔ فراوانی نوعی دو یا چند مقوله می‌توان به درجهٔ زایایی آن‌ها پی برد. اما او همچنین تصریح می‌کند که به دلیل وجود برخی محدودیت‌ها در این روش، نمی‌توان آن را یک

روش قابل اعتماد و کامل به حساب آورد. در این پژوهش، به طور مثال با مقایسه فراوانی نوعی پیشوندهای مورد بررسی این نتیجه به دست می‌آید که پیشوند «بل-» دارای کم‌ترین و پیشوند «غیر-» دارای بیش‌ترین میزان زایایی است. درجه‌بندی زایایی پیشوندهای پژوهش حاضر برحسب شاخص فوق را می‌توان در جدول (۲) ملاحظه نمود.

جدول (۲) فراوانی نوعی برای هر پیشوند

پیشوند	فراوانی نوعی (V)	مثال
غیر	۳۲۵۷	غیراتفاقی
بی	۱۸۴۱	بی کیفیت
نا	۱۶۱۶	نا آرام
ضد	۱۲۹۷	ضد زلزله
هم	۶۸۴	همسر
فرا	۵۱۰	فراجنسیتی
باز	۳۳۸	بازنگری
با	۲۷۸	با عرضه
پیش	۲۷۶	پیش ساخته
فرو	۲۰۹	فروخشکیدن
ابر	۲۰۴	ابر تراشه
وا	۱۳۰	واکاوی
پس	۹۴	پس زدن
بلا	۷۹	بلا رقیب
لا	۳۱	لامحال
ور	۲۶	ورقلمبیده
فراز	۱۱	فراز آمدن
فرو	۸	فروداوردن
بل	۵	بلهوس

نکته مهمی که از مقایسه وضعیت دو جدول فوق به دست می‌آید شباهت نسبتاً زیاد درجه زایایی پیشوندهای اشتقاقی است. در هر دو جدول زایاترین و غیرزایاترین پیشوندها به ترتیب «غیر-» و «بل-» هستند. ضمن این که در سایر رده‌بندی‌ها نیز شباهت‌های زیادی بین اعضای دو جدول به چشم می‌خورد.

۵-۲. حافظ مقوله و تغییردهنده مقوله

برای رسیدن به این پاسخ که کدام یک از پیشوندها حافظ مقوله درونداد خود هستند و کدام یک تغییردهنده آن، می‌بایست فراوانی نوعی درونداد و برونداد تمامی واژه‌های دربردارنده پیشوندهای مورد نظر مشخص شود. بنابراین، در فایل اکسل هر پیشوند، در کنار ستون بافت‌های سمت چپ و راست آن پیشوند، دو ستون مجزا برای برچسب درونداد و برونداد و دو ستون دیگر برای نشان‌دادن حافظ‌مقوله‌گی و تغییردهندگی مقوله در نظر گرفته شد و با مشاهده شباهت یا تفاوت درونداد و برونداد واژه‌های هر پیشوند، در سلول‌های پیش‌گفته، اطلاعات مربوط به حافظ مقوله یا تغییردهنده مقوله درج شد. به طور مثال پیشوند «وا-» دارای ۵۳ نوع رخداد تغییردهنده مقوله و ۷۷ نوع رخداد حافظ مقوله است. سپس، با استفاده از نرم‌افزار اسپاس نسخه ۱۶،۰،۰ و آزمون آماری مجذور خی تک‌نمونه‌ای، معنادار بودن یا نبودن رخدادهای مشاهده‌شده بررسی شد. در پیوست (۱) این مقاله، می‌توان نتایج محاسبات را برای همه پیشوندها مشاهده نمود. در این قسمت تنها برای مثال، تحلیل آماری پیشوند «بی» ارائه می‌گردد:

تعداد فراوانی نوعی تغییردهنده مقوله در این پیشوند برابر با ۱۸۲۳ و حافظ مقوله برابر با ۱۸ می‌باشد.

جدول (۳) مجذور خی تک‌نمونه‌ای برای پیشوند «بی»

Test Statistics	
	state
Chi-Square	1769.704 ^a
df	1
Asymp. Sig.	.000

a. 0 cells (.0%) have expected frequencies less than 5. The minimum expected cell frequency is 920.5.

با توجه به عدد به‌دست‌آمده از آزمون مجذور خی، درجه معناداری صفر است که کم‌تر از ۰/۰۵ می‌باشد، به این معنا که در توزیع دو متغیر حافظ مقوله و تغییردهنده مقوله برای پیشوند «بی» تفاوت معناداری وجود دارد.

۱. ارائه جدول‌های آماری تک‌تک پیشوندها، حجم مقاله را به‌شدت افزایش می‌دهد. به همین دلیل در این بخش، تنها جدول آماری مربوط به یکی از پیشوندها معرفی و نتایج تحلیل‌های آماری سایر پیشوندها در قسمت نتیجه‌گیری مقاله ارائه می‌شوند.

جدول (۴) فراوانی موردانتظار برای پیشوند «بی-»

state			
	Observed N	Expected N	Residual
18	18	920.5	-902.5
1823	1823	920.5	902.5
Total	1841		

همان‌طور که در جدول (۴) مشاهده می‌شود، مورد تغییردهنده مقوله رابطه معناداری با پیشوند موردنظر دارد؛ چراکه فراوانی مشاهده‌شده‌اش از فراوانی مورد انتظار بیش‌تر است. پس این پیشوند تغییردهنده مقوله می‌باشد.

برخی پیشوندها نیازی به محاسبات نداشتند زیرا همه رخدادهای آنها حافظ مقوله یا تغییردهنده مقوله بودند. در برخی دیگر از پیشوندها نیز به دلیل این که درجه معناداری بیشتر از ۰/۰۵ به دست آمد نیازی به مرحله بعدی نبود و این پیشوندها در دسته جداگانه‌ای قرار گرفتند که نمی‌توان در مورد حافظ مقوله بودن یا نبودنشان اظهار نظر کرد.

مقدمات پاسخ به پرسش سوم، تعیین مقوله درونداد بود که در بخش روش پژوهش توضیح داده شد. مرحله بعد برای تعیین این مطلب که هر پیشوند به صورت سرنمونی و حاشیه‌ای به چه پایه‌ای متصل می‌شود، تحلیل آماری فراوانی هر «پیشوند + مقوله پایه» است. به طور مثال پیشوند «فرو-» در ۱۰۷ نوع به فعل، در ۶۱ نوع به صفت، و در ۳۱ نوع به اسم پیوسته است و در هیچ رخدادی به قید، ضمیر، شبه‌جمله و ضمیر پرسشی متصل نشده است. برای به‌دست آوردن این که آیا بین این چند مقوله درونداد تفاوت معناداری وجود دارد یا خیر، و این که بین چه مقوله‌ای با پیشوند مورد نظر رابطه قوی‌تری وجود دارد از آزمون مجذور خی تک‌نمونه‌ای استفاده شد.

جدول (۵) مجذور خی تک‌نمونه‌ای برای بررسی سرنمون مقوله درونداد پیشوند «فرو-»

Test Statistics

	POS
Chi-Square	32.880 ^a
Df	2
Asymp. Sig.	.000

a. 0 cells (.0%) have expected frequencies less than 5. The minimum expected cell frequency is 69.7.

با توجه به عدد به دست آمده از آزمون مجذور خی، درجه معناداری صفر است که کم تر از ۰/۰۵ می باشد، به این معنا که در توزیع مقوله های درون داد برای پیشوند «فرو» تفاوت معناداری وجود دارد. برای این که مشخص شود کدام مقوله سرنمون و کدام مقوله حاشیه درون دادهای این پیشوند است، از تفاضل فراوانی مشاهده شده از فراوانی مورد انتظار بهره برده شد. آن مقوله که عدد به دست آمده از این تفاضل برای آن مثبت و بیش تر از سایر مقوله ها باشد قطعاً سرنمون است و آن مقوله که عدد به دست آمده از این تفاضل برای آن منفی و کم تر از سایر مقوله ها باشد قطعاً حاشیه است. لازم به ذکر است که در مورد سایر مقوله ها نمی توان با قطعیت و بر اساس آزمون آماری اظهار نظر کرد.

جدول (۶) فراوانی مشاهده شده و مورد انتظار برای بررسی سرنمون مقوله درون داد پیشوند «فرو»

POS			
	Observed N	Expected N	Residual
41	41	69.7	-28.7
61	61	69.7	-8.7
107	107	69.7	37.3
Total	209		

همان طور که در جدول مشاهده می شود، با توجه به اعداد به دست آمده از تفاضل فراوانی مورد انتظار و فراوانی مشاهده شده، درون داد فعل، سرنمون درون دادهای این پیشوند می باشد.

۶. نتیجه‌گیری

در اثر حاضر، پس از بررسی هر یک از شاخص‌های محاسبه زایایی، شاخص زایایی سراسری به عنوان شاخص موردنظر این پژوهش انتخاب شد و طبق داده‌های بیکره همشهری ۲، به سه پرسش پژوهش پاسخ داده شد. پرسش اول معطوف به مشخص کردن درجه زایایی پیشوندهای اشتقاقی بود. پس از بررسی داده‌ها با استفاده از فرمول زایایی سراسری، درجه زایایی پیشوندهای اشتقاقی زبان فارسی از بیش‌ترین تا کم‌ترین به صورت زیر به دست آمد:

غیر-، نا-، بی-، ضد-، فرا-، هم-، ابر-، باز-، پیش-، فرو-، وا-، پس-، بلا-، با-، فراز-، ور-، فرود-، لا-، بل-.

هدف دوم این پژوهش بررسی حافظ مقوله بودن یا نبودن پیشوندهای اشتقاقی بود. پس از انجام محاسبات آماری مجذور خی در این زمینه، نتایج زیر به دست آمد:

پیشوندهای حافظ مقوله: ابر-، باز-، پیش-، ضد-، غیر-، فرا-، فرو-، نا-، وا-، ور-، پیشوندهای تغییردهنده مقوله: با-، بلا-، بی-، لا-، هم-.

از میان ۱۹ پیشوند مورد بررسی، پیشوندهای «بل-، پس-، فراز-، فرود-» مواردی هستند که آزمون آماری مجذور خی تفاوت معناداری را بین متغیرهای آنها نشان نمی‌دهد. در نتیجه نمی‌توان در مورد حافظ مقوله بودن یا نبودنشان اظهارنظر کرد.

هدف سوم پژوهش حاضر نیز، مشخص کردن سرنمون مقوله درونداد برای پیشوندهای مورد مطالعه بود. پس از بررسی‌های آماری صورت‌گرفته نتایج زیر به دست آمد:

پیشوندهای با سرنمون صفت: غیر-، فرا-، نا-، پیشوندهای با سرنمون اسم: ابر-، با-، بلا-، بی-، پیش-، ضد-، فرود-، لا-، هم-.

پیشوندهای با سرنمون فعل: باز-، پس-، فرو-، وا-، ور-، پیشوندهایی که نمی‌توان در مورد سرنمون مقوله درونداد آنها بر حسب آزمون آماری اظهار نظر کرد: بل-، فراز-.

فهرست منابع

- احمدی گیوی، حسن. و انوری، حسن. (۱۳۹۲). دستور زبان فارسی ۲. تهران: انتشارات فاطمی.
امیراجمندی، سیده‌نازنین. و عاصی، مصطفی. (۱۳۹۲). زایایی فرآیند ترکیب در زبان فارسی. *زبان‌شناخت*، ۴ (۱)، ۱-۱۴.

- انوری، حسن. (۱۳۸۲). فرهنگ بزرگ سخن، تهران: سخن.
- بدخشان، ابراهیم. (۱۳۸۹). بررسی پیکره‌بنیاد میزان زایایی فرآیند ترکیب در زبان فارسی امروز. پایان‌نامه کارشناسی ارشد. تهران: دانشگاه علامه طباطبایی.
- دهخدا، علی‌اکبر. (۱۳۷۷). لغت‌نامه. تهران: دانشگاه تهران.
- راسخ مهند، محمد. (۱۳۹۳). درآمدی بر زبان‌شناسی شناختی: نظریه‌ها و مفاهیم. تهران: انتشارات سمت.
- رواقی، علی. (۱۳۸۸). بررسی پسوندها در زبان فارسی. تهران: فرهنگستان زبان و ادب فارسی.
- شقاقی، ویدا. (۱۳۸۹). مبانی صرف. تهران: انتشارات سمت.
- عمید، حسن. (۱۳۷۱). فرهنگ عمید. چاپ هشتم، تهران: انتشارات امیرکبیر.
- فراگردی، راحله. (۱۳۸۹). پیشوندهای زبان فارسی. پایان‌نامه کارشناسی ارشد. تهران: دانشگاه علامه طباطبایی.
- فرشیدورد، خسرو. (۱۳۸۶). فرهنگ پیشوندها و پسوندهای زبان فارسی همراه گفتارهایی درباره دستور زبان فارسی. تهران: زوآر.
- کلباسی، ایران. (۱۳۹۱). ساخت اشتقاقی واژه در فارسی امروز. تهران: پژوهشگاه علوم انسانی و مطالعات فرهنگی.
- معین، محمد. (۱۳۷۸). فرهنگ فارسی. تهران: امیرکبیر.
- Al, B.P.F. and Booij, G. E. (1981). De productiviteit van woordvormingsregels. *Forum der Letteren*, 22, 26–38.
- AleAhmad, A., Amiri, H., Darrudi, E., Rahgozar, M., and Oroumchian, F. (2009). Hamshahri: A standard Persian text collection. *Knowledge-Based Systems*, 22 (5), 382-387.
- Anthony, L. (2014). AntConc (Version 3.4.3.0) [Computer Software]. Tokyo, Japan: Waseda University. Retrieved from <http://www.laurenceanthony.net/software/antconc/>
- Aronoff, M. (1976). *Word Formation in Generative Grammar*. Cambridge: MIT Press.
- Baayen, R. H. (1992). Quantitative aspects of morphological productivity. In: Geert E. Booij and J. van Marle (Eds.), *Yearbook of Morphology*. Dordrecht: Kluwer Academic Publishers. pp. 109-149.

- Baayen, R. H. (1993). On frequency, transparency, and productivity. In: Geert E. Booij and J. van Marle (Eds.), *Yearbook of Morphology*. Dordrecht: Kluwer Academic Publishers. pp. 181-208.
- Baayen, R. H. (1994). Derivational productivity and text typology. *Journal of Quantitative Linguistics*, 1, pp. 16-34.
- Baayen, R. H. (2009). Corpus linguistics in morphology: Morphological productivity. In: Anke Luedeling and Merjä Kyto (Eds.), *Corpus Linguistics. An International Handbook*. Berlin: Mouton De Gruyter. pp. 900-919.
- Baayen, R. H. and Renouf, A. (1996). Chronicling the times: Productive lexical innovations in an English newspaper. *Language*, 72, pp. 69-96.
- Baayen, R. H. & Lieber, R. (1991). Productivity and English derivation: A Corpus based study. *Linguistics*, 29, pp. 801-843.
- Bauer, L. (1992). Scalar productivity and -lily adverbs. In: Geert E. Booij and J. van Marle (Eds.), *Yearbook of Morphology*. Dordrecht: Kluwer Academic Publishers. pp. 185-191.
- Bauer, L. (1996). Is morphological productivity non-linguistic?. *Acta Linguistica Hungarica*, 43, pp. 19-31.
- Bauer, L. (2004). *Morphological productivity*. Cambridge: Cambridge University Press.
- Booij, G. E. (2007). *The Grammar of Words: An Introduction to Linguistic morphology*. Oxford: Oxford University Press.
- Cannon, G. (1988). *Historical Change and English Word-formation*. New York: Peter Lang.
- Evans, V. and Green, M. (2006). *Cognitive Linguistics: An Introduction*. Edinburgh: Edinburgh University Press.
- Hay, J. B. (2003). *Causes and Consequences of Word Structure*. New York: Routledge.

Lieber, R. (1981). *On the Organization of the Lexicon*. Doctoral dissertation, Massachusetts Institute of Technology.

Liu, W. and Shen, H. (2012). A corpus-based analysis of English suffix –esque. *Theory and Practice in Language Studies*, 2(4), pp. 767-772.

Löbner, S. (2002). *Understanding Semantics*. New York: Routledge.

Merriam, G. and Merriam, C. (1961). *Webster's Third International Dictionary of the English Language*. Massachusetts: Merriam Webster Inc.

O'keeffe, A., McCarthy, M., and Carter, R. (2007). *From corpus to classroom: Language use and language teaching*. Cambridge: Cambridge University Press.

Plag, I. (2003). *Word-formation in English*. Cambridge: Cambridge University Press.

Rosch, E. (1975). Cognitive representation of semantic categories. *Journal of Experimental Psychology*, 104 (3), pp. 192-233.

Rosch, E. (1977). Human Categorization. In: Nail Warren (Eds.), *Studies in Cross-Linguistic Psychology*. London: Academic Press. pp. 1-49.

Rosch, E. (1978). Principles of categorization. In: B. Lloyd & E. Rosch (Eds.), *Cognition and Categorization*. Hillsdale, NJ: Erlbaum. Pp. 27-48.

Wittgenstein, L. (1958). *Philosophical Investigations: The German Text, with a Revised English Translation*. Translated by GEM Anscombe. Oxford: Blackwell.

بیوست (۱) داده‌های نهایی به‌دست آمده از پژوهش برای محاسبات آماری

Prefix	token (N)	hapax	P*=ha(pre)/Σha	changing	preserving	Adj	Adv	N	V	Pronoun	Pseudo-sentence
۱ ابر	204	103	0.0267	7	197	50	1	153	0	0	0
۲ با	278	9	0.0023	277	1	0	0	278	0	0	0
۳ باز	338	76	0.0197	129	209	42	0	118	178	0	0
۴ بل	5	2	0.0005	4	1	1	0	3	0	0	1
۵ بلا	79	22	0.0057	79	0	0	0	79	0	0	0
۶ بی	1841	528	0.1370	1823	18	1	0	1840	0	0	0
۷ پس	94	26	0.0067	40	54	14	0	24	56	0	0
۸ پیش	276	55	0.0143	114	162	42	2	145	87	0	0
۹ ضد	1297	511	0.1326	522	775	378	1	915	0	3	0
۱۰ غیر	3257	1381	0.3582	27	3230	2775	9	473	0	0	0
۱۱ فرا	510	216	0.0560	93	417	242	0	180	87	1	0
۱۲ فراز	11	8	0.0021	8	3	1	0	7	3	0	0
۱۳ فرو	209	51	0.0132	49	160	41	0	61	107	0	0
۱۴ فرود	8	3	0.0008	4	4	1	0	7	0	0	0
۱۵ لا	31	3	0.0008	22	9	4	0	24	3	0	0
۱۶ نا	1616	664	0.1722	141	1475	1185	5	382	39	0	0
۱۷ وا	130	31	0.0080	53	77	18	0	32	80	0	0
۱۸ ور	26	6	0.0016	7	19	4	0	8	14	0	0
۱۹ هم	684	159	0.0412	612	72	26	0	573	85	0	0

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

A Corpus-based Study of Productivity of Derivational Prefixes in the Written Variety of Contemporary Persian

H. Abedini, A. Moloodi and A. Khormae

Shiraz University, Shiraz, Iran

Abstract

One of the significant issues in morphology, generally and in word-formation processes, especially is productivity. Most researchers consider productivity as a gradient concept and define it in a continuum. The present research is about to answer three questions below: 1. what are the rankings of the studied prefixes in terms of productivity? 2. Which of the prefixes is category-preserving and which one is category-changing? 3. What is the prototypical input of each prefix? In order to answer the questions, first, 19 derivational prefixes of Persian were selected based on the works in which Persian derivational prefixes had been presented. Then, the data of the research were extracted from the 150 million words corpus of Hamshahri 2 by the AntConc software and after that they were transferred to Excel program to be tagged and analyzed based on Baayen's (2009) global productivity measure and Rosch's (1975, 1977, 1978) prototype theory. The results showed that /Gejr/ is the most and /bol/ is the least productive prefixes of Persian language. For answering the second and third questions, each prefix was studied independently by the chi-square statistical test in SPSS software and the category-changing or category-preserving features of each prefix were specified. Also, the prototypical lexical category of each prefix input were identified.

Keywords: Derivational Prefix, Corpus-based Method, Global Productivity, Category-preserving, Category-changing
