



بررسی ترتیب پسوندهای اشتقاقی زبان فارسی در چارچوب رویکرد ترتیب مبتنی بر پیچیدگی

مریم حمصیان‌انفاق^۱

عادل رفیعی^{۲*}

چکیده

ترتیب وندها از جمله مباحث نظری مطرح صرف است. تاکنون رویکرد لایه‌ای (سیگل، ۱۹۷۴)، رویکرد محدودیت‌های گزینشی (فب، ۱۹۸۸) و رویکرد «ترتیب مبتنی بر پیچیدگی» (های و پلاگ، ۲۰۰۰) به بررسی ترتیب وندها پرداخته‌اند. رویکرد ترتیب مبتنی بر پیچیدگی - که چهارچوب نظری مقاله حاضر را تشکیل می‌دهد - بر اساس معیار تقطیع‌پذیری وند از پایه به توجیه ترتیب وندها می‌پردازد. در این نظریه فرض بر آن است هر چه میزان تقطیع‌پذیری بیشتر باشد وند در جایگاه دورتری نسبت به پایه قرار می‌گیرد. هدف از مقاله حاضر محک زدن این فرضیه در ترتیب پسوندهای اشتقاقی زبان فارسی است. برای انجام این بررسی پس از بررسی هم‌آیندی پسوندها و محاسبه نسبت تقطیع‌پذیری پیوستار تقطیع‌پذیری بدست داده شد. با مقایسه ترتیب پسوندها و ترتیب آنها بر روی پیوستار تقطیع‌پذیری مشخص می‌شود که ترتیب پسوندهای اشتقاقی زبان فارسی تا چه میزان با معیار تقطیع‌پذیری قابل توجیه است. بدین منظور ۴۱ پسوند اشتقاقی زبان فارسی که به‌طور مشترک حداقل در دو اثر از سه اثر صادقی (۱۳۷۰)، کشانی (۱۳۷۱) و کلیاسی (۱۳۸۰) مشاهده شد انتخاب و از پیکره پژوهشگاه علوم انسانی و مطالعات فرهنگی ۹۱۷۲ واژه مشتق از آنها استخراج شدند. با استفاده از هم‌آیندی وندها و نمودار پراکندگی، میزان تقطیع‌پذیری وندها محاسبه و پیوستار تقطیع‌پذیری به‌دست آمد. همراستا با نظر های و پلاگ (۲۰۰۴) برای ارائه توجیه کامل‌تری از آرایش پسوندهای اشتقاقی زبان فارسی - علاوه بر معیار میزان تقطیع‌پذیری - می‌بایست محدودیت‌های گزینشی حاکم بر هم‌آیندی وندها را نیز مورد توجه قرار داد.

کلید واژه‌ها: پسوندهای اشتقاقی، رویکرد ترتیب مبتنی بر پیچیدگی، تقطیع‌پذیری، هم‌آیندی

✉ | hemasianmary@yahoo.com

۱- دانشجوی دکتری زبان‌شناسی، دانشگاه اصفهان، اصفهان، ایران

✉ | a.rafieei@fgn.ui.ac.ir

۲- استادیار زبان‌شناسی، دانشگاه اصفهان، اصفهان، ایران (نویسنده مسئول)

۱- مقدمه

ترتیب وندها از جمله مباحثی است که در دهه‌های اخیر مورد توجه زبانشناسان قرار گرفته است. تاکنون بحث‌های زیادی در خصوص وجود اصول یا مکانیسم‌هایی کلی ناظر بر ترتیب وندها صورت گرفته است. در این راستا سه رویکرد مبتنی بر چگونگی قرارگیری وندها مطرح شده است. رویکرد لایه محور^۱ (سیگل^۲، ۱۹۷۴؛ آلن^۳، ۱۹۷۸؛ کیپارسکی^۴، ۱۹۸۲؛ سلکیرک^۵، ۱۹۸۲؛ موهانان^۶، ۱۹۸۶؛ گیگریچ^۷، ۱۹۹۹)، رویکرد محدودیت‌های گزینشی^۸ وند محور (فب^۹، ۱۹۸۸؛ پلاگ^{۱۰}، ۱۹۹۹) و رویکرد ترتیب وند مبتنی بر پیچیدگی^{۱۱} (های^{۱۲}، ۲۰۰۰، ۲۰۰۲) که چهارچوب نظری پژوهش حاضر را تشکیل می‌دهد. این رویکرد با ارائه یک مدل روانشناسی - زبانشناختی مبتنی بر این ایده است که ترکیبات وندی را محدودیت‌های حاکم بر پردازش ساختارهای صرفی کنترل می‌کنند.

پژوهش حاضر به بررسی چهل و یک پسوند اشتقاقی زبان فارسی در چارچوب رویکرد ترتیب مبتنی بر پیچیدگی (باین و پلاگ، ۲۰۰۹) بر اساس داده‌های استخراج شده از پیکره داده‌های زبان فارسی پژوهشگاه علوم انسانی و مطالعات فرهنگی^{۱۳} پرداخته است. برای این کار ابتدا بر اساس هم‌آیندی^{۱۴} پسوندها در پیکره مورد بررسی و ترسیم ماتریس مربعی^{۱۵} و همچنین نمودار گراف جهت‌دار توانسته‌ایم مرتبه هر وند را نیز به دست آوریم. در ادامه برای بررسی مشخصه‌های هر وند بر اساس بسامد هر یک از واژه‌های مشتق و پایه‌های آنها نمودار نقاط پراکنده ویژه هر وند ترسیم شد. سپس با توجه به خط روند هر نمودار و قرار گرفتن داده‌های بالای خط روند نسبت بسامد نوع و نسبت بسامد نمونه تقطیع‌پذیری محاسبه شد. با استفاده از مرتبه هم‌آیندی وندها و مرتبه نسبت بسامد نوع و مرتبه نسبت بسامد نمونه تقطیع‌پذیری پسوندها، پیوستار تقطیع‌پذیری با استفاده از

1. stratum-oriented model
2. Siegel
3. Allen
4. Kiparsky
5. Selkirk
6. Mohanan
7. Giegerich
8. selectional restriction
9. Fabb
10. Plag
11. Complexity-based ordering
12. Hay

۱۳. با تشکر از جناب آقای دکتر عاصی و همکاران در ارائه عضویت برای استفاده از پیکره پایگاه داده‌های زبان فارسی و پژوهشگاه علوم انسانی و مطالعات فرهنگی

14. co-occurrence

۱۵. در علم ریاضیات ماتریس مربعی گونه خاصی از ماتریس‌ها هستند که در آن تعداد سطرها و ستون‌ها با هم برابر بوده و در آن می‌توان رابطه بین اولین سطر و اولین ستون را با قرار دادن عدد یا درایه در محل تلاقی آنها نشان داد. در ماتریس مربعی اگر تمام اعضای زیر خط قطر اصلی ماتریس صفر باشند آن را ماتریس بالا مثلثی و اگر همه اعضای بالای قطر اصلی صفر باشند آن را ماتریس پایین مثلثی می‌نامند. در این مقاله برای نشان دادن هم‌آیندی وندها از ماتریس مربعی و برای رسیدن به مرتبه تقطیع‌پذیری وندها از ماتریس‌های مربعی بالا مثلثی استفاده شده است.

نمودار پراکندگی ترسیم شده است. طبق رویکرد ترتیب مبتنی بر پیچیدگی انتظار می‌رود که تقدم و تأخر پسوندهای اشتقاقی زبان فارسی همراستا با جایگاه وندها در پیوستار تقطیع‌پذیری توصیف شده باشد. یافته‌های پژوهش حاضر تأیید کننده ادعای این رویکرد در توجیه ترتیب پسوندهای اشتقاقی زبان فارسی می‌باشد. مقاله حاضر در شش بخش تدوین شده است. در بخش دوم به مروری بر پیشینه و مبانی تحقیق در خصوص رویکردهای مطرح شده در ترتیب وندها پرداخته‌ایم. در بخش سوم به معرفی رویکرد ترتیب مبتنی بر پیچیدگی های و پلاگ (۲۰۰۲) می‌پردازیم. در بخش چهارم به بررسی پسوندهای اشتقاقی زبان فارسی در چارچوب نظریه ترتیب مبتنی بر پیچیدگی های و پلاگ (۲۰۰۴) پرداخته و با توجه به هم‌آیندی پسوندها و نسبت تقطیع‌پذیری آنها، پیوستار تقطیع‌پذیری پسوندهای اشتقاقی زبان فارسی به دست داده شده است. بخش پنجم به مقایسه پیوستار تقطیع‌پذیری و ترتیب پسوندهای اشتقاقی زبان فارسی پرداخته است. بخش آخر نیز به نتیجه‌گیری و جمع‌بندی اختصاص داده شده است.

۲- پیشینه پژوهش

شیوه ترکیب وندها ابتدا در چارچوب مدل لایه محور مورد توجه قرار گرفت (پلاگ و باین، ۲۰۰۹). از جمله کسانی که در این زمینه به بررسی ترتیب وندها پرداخته‌اند می‌توان به سیگل (۱۹۷۴)، آلن (۱۹۷۸)، کیپارسکی (۱۹۸۲) و موهانان (۱۹۸۶) اشاره کرد. بر اساس این رویکرد وندها اعم از پیشوند و پسوند در دو لایه قرار می‌گیرند. رویکرد لایه‌ای با در نظر گرفتن دو لایه برای وندها دارای نارسایی‌هایی است. از جمله این نارسایی‌ها می‌توان قادر نبودن این رویکرد به توجیه ترکیبات ممکن و غیرممکن درون یک لایه معین، نداشتن دلیل واضح مبنی بر قرار گرفتن وندها در دو لایه مختلف و صرفاً اتکا به معیار منشأ ریشه شناختی و مشخصه‌های واجی جهت قرار دادن وندها در دو لایه اشاره کرد (پلاگ و باین، ۲۰۰۹).

در زبان فارسی نیز با بررسی‌های انجام شده بر طبق این مدل وندها اکثراً در لایه اول قرار گرفته‌اند و تکیه را به سمت خود کشیده‌اند. در واقع اکثر وندهای زبان فارسی تکیه بر هستند (غلامعلی‌زاده، ۱۳۹۰). با توجه به این مطلب مشخص می‌گردد که رویکرد لایه‌ای به آن شکلی که در زبان انگلیسی با معیار تغییرات آوایی به لایه‌بندی وندها می‌پردازد، در زبان فارسی کارایی چندانی ندارد.

دومین رویکرد ناظر بر ترتیب قرارگیری وندها با مطرح ساختن محدودیت‌های گزینشی مانند محدودیت‌های صرفی، واجی و معنایی به بررسی ترکیبات اشتقاقی ممکن و غیرممکن می‌پردازد (فب، ۱۹۸۸). به اعتقاد فب (۱۹۸۸) ترتیب قرارگیری پسوندهای اشتقاقی زبان انگلیسی بر پایه محدودیت‌های گزینشی هر وند شکل می‌گیرد. وی - علاوه بر محدودیت‌های گزینشی هر وند به‌طور مشخص - بر اساس معیار هم‌آیندی، به محدودیت‌هایی نیز اشاره می‌کند که گروه‌هایی از پسوندها بر اساس آن عمل می‌کنند. فب بر اساس این نوع محدودیت‌های گروهی، پسوندهای زبان انگلیسی را در چهار گروه جای می‌دهد: گروه اول پسوندهایی که به پایه دارای پسوند متصل نمی‌شوند. گروه دوم پسوندهایی که بعد از پسوندهای خاص دیگری می‌آیند. گروه

سوم پسوندهایی که به راحتی به هر پسوندی متصل می‌شوند و گروه چهارم پسوندهای به اصطلاح مشکل ساز که در ترکیبات یا تنها به پایه و یا به پسوندهای خاص دیگری متصل می‌شوند.

حمصیان (۱۳۸۹) با اتخاذ رویکرد محدودیت‌های گزینشی با انتخاب ۱۷ پیشوند و ۳۷ پسوند به بررسی ترتیب وندهای اشتقاقی زبان فارسی در این چهارچوب پرداخته است.

اگر چه رویکرد محدودیت‌های گزینشی قادر است ترکیبات ممکن و موجود زیادی را در مقایسه با رویکرد لایه‌ای توجیه کند اما کماکان موارد بسیار زیادی وجود دارد که طبق این رویکرد قابل توجیه نیستند (پلاگ، ۱۹۹۶ و ۱۹۹۹). طبقه‌بندی چهارگانه فب (۱۹۸۸) دارای ضعف‌هایی چون وجود مثال‌های ضد و نقیض زیادی مبنی بر قرار دادن وندها در لایه‌های این طبقه‌بندی است. دوم اینکه طبقه‌بندی وندها اختیاری است و مشخص نیست چرا یک پسوند معین باید متعلق به یک طبقه خاص باشد. سوم اینکه این طبقه‌بندی برای همه محدودیت‌های حاکم بر ترکیبات ممکن و یا ترکیبات غیرممکن به کار نمی‌رود.

دو رویکرد لایه محور و محدودیت‌های گزینشی مطرح شده تا بدین جا تنها به بررسی ساختاری ترتیب وندها پرداخته و صرفاً بسامد نوع واژه‌های مشتق را مدنظر قرار داده‌اند. با توجه به ایرادات وارد بر این دو رویکرد، سومین رویکرد با عنوان ترتیب مبتنی بر پیچیدگی و با داشتن مبنایی روان - زبانشناختی ارائه شده است (های، ۲۰۰۰، ۲۰۰۲؛ های و پلاگ، ۲۰۰۴؛ پلاگ و باین، ۲۰۰۹). طبق این رویکرد ترتیب وندهای اشتقاقی یک زبان تابعی از ترتیب آنها بر روی پیوستار تقطیع‌پذیری است. به بیان ساده‌تر عامل تعیین تقدم و تأخر وندهای اشتقاقی میزان تقطیع‌پذیری آنهاست (های و پلاگ، ۲۰۰۴ و باین، ۲۰۰۲). میزان تقطیع‌پذیری یک وندها را بسامد نمونه واژه‌ها و نسبت بسامد پایه و واژه مشتق مشخص می‌کند. رویکرد ترتیب مبتنی بر پیچیدگی ابتدا در مقاله های (۲۰۰۲) معرفی شده است.

های و پلاگ (۲۰۰۴) با انتخاب ۱۵ پسوند زبان انگلیسی و ترکیب کردن آنها بدون در نظر گرفتن هیچ یک از محدودیت‌های گزینشی قائل به وجود ۲۱۰ ترکیب دو پسوندی شده‌اند. آنها با استفاده از منابعی چون BNC^۱، CELEX^۲ و OED^۴ و اینترنت به این مطلب پرداخته‌اند که آیا وجود ترکیبات ممکن و موجود در هر یک از این منابع را می‌توان براساس رویکرد ترتیب مبتنی بر پیچیدگی توصیف کرد یا آن که رویکرد محدودیت‌های گزینشی تحلیل مناسب‌تری از وجود یا عدم وجود هر یک از ترکیبات دو پسوندی را ارائه خواهد کرد؟ فرضیه های و پلاگ (۲۰۰۴) بر این امر استوار است که محدودیت‌های گزینشی و محدودیت‌های پردازشی هر دو حاکم بر ترکیبات ممکن و موجود پسوندهای اشتقاقی بوده و مکمل یکدیگر می‌باشند.

این پژوهش به بررسی رابطه میزان تقطیع‌پذیری و ترتیب پسوندهای اشتقاقی زبان فارسی پرداخته است. با توجه به اهمیت رویکرد ترتیب مبتنی بر تقطیع‌پذیری در پژوهش حاضر، بخش دوم این مقاله به مبانی نظری این رویکرد اختصاص پیدا کرده است.

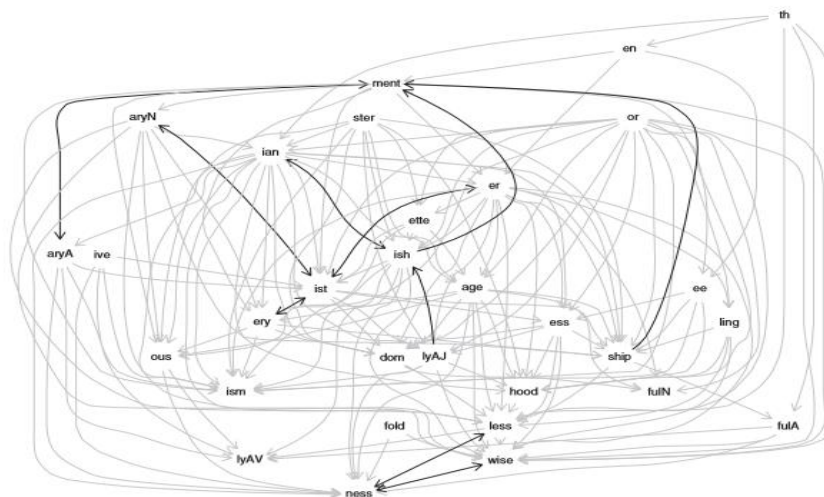
1. Baayen
2. British National Corpus
3. lexical database (Baayen, Piepenbrock, Gullikers 1995)
4. Oxford English Dictionary

۳- رویکرد ترتیب مبتنی بر پیچیدگی

رویکرد ترتیب مبتنی بر پیچیدگی بر این امر استوار است که ترتیب وندها همراه با معیار محدودیت‌های گزینشی حاکم بر وندها تابعی از تقطیع‌پذیری صرفی وند از پایه بوده و این ترتیب به صورت پیوستار قابل توجیه می‌باشد. این رویکرد بر اساس پردازش واژگانی در درک گفتار بنا نهاده شده است به طوری که برای پردازش واژگانی دو راهکار صرفی را در نظر گرفته است: اینکه در هنگام مواجهه با یک واژه مشتق گویشور آن را به صورت یک واحد در نظر بگیرد یا آن که واژه مشتق را به اجزای سازنده تقطیع کرده و با دستیابی به معنای تک تک اجزای واژه مشتق معنای کل واژه را دریابد (های و باین، ۲۰۰۳). در این رویکرد از دو معیار مرتبه تقطیع‌پذیری که بر اساس محدودیت‌های گزینشی وندها و نهایتاً هم‌آیندی آنها تعیین می‌شود و نسبت تقطیع‌پذیری که بر اساس نسبت بسامد نوع و نمونه واژه مشتق به پایه مشخص می‌شود، برای ترسیم پیوستار تقطیع‌پذیری وندها و در نهایت توجیه جایگاه وندهای اشتقاقی زبان استفاده می‌گردد. فرضیه حاکم بر این رویکرد آن است که هر چه یک وند بر روی پیوستار تقطیع‌پذیری میزان تقطیع‌پذیری بیشتری داشته باشد، در مقایسه با وندهایی که تقطیع‌پذیری پایین‌تری دارند، از پایه دورتر است. در ادامه به معرفی و شیوه محاسبه مرتبه تقطیع‌پذیری و نسبت تقطیع‌پذیری در رویکرد ترتیب مبتنی بر تقطیع‌پذیری - به شیوه‌ای که در های و پلاگ (۲۰۰۴) آمده - می‌پردازیم.

برای محاسبه مرتبه تقطیع‌پذیری، های و پلاگ (۲۰۰۴) ابتدا ۳۱ پسوند اشتقاقی زبان انگلیسی را به ترتیب حروف الفبا در یک ماتریس مربعی به این صورت که در اولین ستون، وندهای جایگاه اول نسبت به پایه و در اولین سطر، وندهای جایگاه دوم نسبت به پایه جای گرفته را ترسیم کرده‌اند. آنگاه بر اساس هم‌آیندی وندها در پیکره‌های مورد بررسی، تمامی هم‌آیندی‌های دوگانه را مشخص کرده‌اند. به عنوان مثال برای دو پسوند *-ful* و *-th* در تقاطع دو پسوند در ماتریس مزبور عدد ۱ گذاشته شده که نشان از هم‌آیندی آنها در واژه‌هایی چون *strengthful, lengthful, depthful* می‌باشد. در مواردی که هم‌آیندی وجود نداشته مانند دو پسوند *-ful* و *-th* عدد صفر گذاشته شده است. در ادامه بر اساس اطلاعات ماتریس مربعی با استفاده از نرم‌افزار برنامه نویسی آر^۱، گراف جهت‌دار برای ۳۱ پسوند اشتقاقی زبان انگلیسی که به عنوان راس^۲ و خطوط متصل کننده رئوس که با یال^۳ معرفی می‌شود به دست داده می‌شود. در این گراف پسوندها از کمترین تا بیشترین میزان هم‌آیندی قرار گرفته‌اند. شکل (۱) گراف جهت‌دار ۳۱ پسوند اشتقاقی زبان انگلیسی را نشان می‌دهد.

1. R-software
2. vertice
3. edge



شکل ۱: گراف جهت‌دار ۳۱ پسوند زبان انگلیسی

۳۱ رأس گراف با یال‌های خاکستری رنگ جهت رو به پایین و یال‌های تیره رنگ جهت رو به بالا و حالت چرخشی بودن وندها را نشان می‌دهند.

همان‌طور که در شکل ۱ نشان داده شده بیشترین تعداد یال‌هایی که وندها را به یکدیگر متصل می‌کنند جهت رو به پایین دارند. همان‌طور که مشاهده می‌شود نرم‌افزار به‌گونه‌ای چیدمان پسوندها را قرار داده که کمترین تعداد یال جهت رو به بالا که نشان از موارد استثناء و هم‌آیندی‌های دو طرفه (وندهایی که هم قبل و هم بعد از یکدیگر می‌توانند قرار گیرند) می‌باشد را نشان می‌دهد.

اینک برای مشخص شدن چگونگی قرار گرفتن پسوندها نسبت به یکدیگر با توجه به گراف به‌دست آمده نرم‌افزار آر^۱ می‌تواند ماتریس‌های مربعی بالا مثلی را که نشان از چیدمان گراف باشد ارائه دهد. با توجه به ۳۱ پسوند اشتقاقی زبان انگلیسی و با توجه به ترتیب‌های مختلف که این پسوندها می‌توانند نسبت به یکدیگر به گونه‌ای داشته باشند که کمترین تعداد یال جهت رو به بالا را نشان دهد تعداد ۴۰۳۷ ماتریس مربعی بالا مثلی به‌دست داده می‌شود. جدول (۱) یکی از ماتریس‌های مربعی بالا مثلی ۳۱ پسوند اشتقاقی زبان انگلیسی را نشان می‌دهد.

۱. نرم‌افزار آر یک زبان برنامه‌نویسی و محیط نرم‌افزاری برای محاسبات آماری و علم داده‌ها است. این نرم‌افزار برای اولین بار به‌عنوان یک پروژه تحقیقاتی توسط راس ایهاکا (Ross Ihaka) و رابرت جنتلمن (Robert Gentleman) نوشته شد، و در حال حاضر توسط گروهی از متخصصان علم آمار به نام تیم هسته نرم‌افزار R با صفحه‌ای در آدرس www.r-project.org در حال فعالیت می‌باشد. این نرم‌افزار نه تنها برای محاسبات آماری بلکه در محاسبات ماتریسی به‌کار می‌رود. همچنین ابزار قدرتمندی برای ایجاد اشکال گرافیکی و نمودارها می‌باشد. فونت به‌کار رفته در این نرم‌افزار به زبان انگلیسی می‌باشد.

جدول ۱: ماتریس مربعی بالا مثلثی هم‌آیندی ۳۱ پسوند اشتقاقی زبان انگلیسی (های و پلاگ، ۲۰۰۴)

th	en	ment	aryN	or	ster	ian	er	ette	ish	ive	aryA	ee	ist	age	ess	ery	lyAJ	ling	ship	hood	dom	less	fold	fulA	ous	ism	ness	lyAV	fulN	wise	
0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	1	
0	0	1	0	0	0	1	0	0	0	0	1	1	0	0	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	1	
0	0	0	1	0	0	1	0	0	0	0	1	1	1	0	1	0	0	0	0	0	0	1	1	0	1	1	0	0	1		
0	0	0	0	1	0	1	0	0	0	0	1	1	1	1	0	1	1	1	1	1	0	1	1	0	1	1	0	0	1		
0	0	0	0	0	1	1	0	0	0	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0	1		
0	0	0	0	0	0	1	1	0	0	0	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0	1		
0	0	0	0	0	0	0	1	0	0	0	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0	1		
0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0	1		
0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	1	1	1	1	0	1	1	1	1	1	0	0	1		
0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	1	1	1	0	1	1	1	1	1	0	0	1		
0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	1	1	0	1	1	1	1	1	0	0	1		
0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	1	0	1	1	1	1	1	0	0	1		
0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	0	1	1	1	1	1	0	0	1		
0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	1	1	1	1	1	0	0	1		
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	1	1	1	1	0	0	1		
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	1	1	1	0	0	1		
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	1	1	0	0	1		
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	1	1	0	0	1	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	1	1	0	0	1

این دست از ماتریس‌ها نشان می‌دهند که میزان هم‌آیندی وندها و ایجاد ترکیبات ممکن، بر پایه محدودیت‌های حاکم بر وندها به‌وجود می‌آید. های و پلاگ با استفاده از این ۴۰۳۷ ماتریس بالا مثلثی به‌دست آمده از گراف جهت‌دار و با میانگین گرفتن مرتبه قرارگیری هر وندها در این ماتریس‌ها مرتبه تقطیع‌پذیری هر یک از وندهای مورد مطالعه در این بررسی را مشخص کرده‌اند. با استفاده از مرتبه تقطیع‌پذیری که از هم‌آیندی وندها در ماتریس مربعی به‌دست می‌آید می‌توان ترتیب قرارگیری وندها نسبت به پایه و نسبت به یکدیگر را مورد بررسی قرار داد. بر اساس نظر های و پلاگ (۲۰۰۴) هرچه وندهای تقطیع‌پذیری بالاتری داشته باشد از پایه دورتر است و از میزان تقطیع‌شدگی بالاتری برخوردار است و در مقابل هر چه مرتبه تقطیع‌پذیری یک وندها کمتر باشد به پایه نزدیک‌تر بوده و در پردازش واژگانی کمتر از پایه خود جدا می‌شود. جدول (۲) مرتبه تقطیع‌پذیری ۳۱ پسوند اشتقاقی زبان انگلیسی را نشان می‌دهد.

جدول ۲: مرتبه تقطیع‌پذیری ۳۱ پسوند اشتقاقی زبان انگلیسی

SUFFIX	RANK	SUFFIX	RANK	SUFFIX	RANK	SUFFIX	RANK				
1	th	0.89	9	ette	8.62	17	lyAJ	14.58	25	fulA	21.87
2	en	1.78	10	aryA	9.27	18	ery	16.19	26	fold	21.87
3	ment	2.67	11	ive	9.27	19	ling	16.27	27	wise	25.80
4	or	4.40	12	ist	11.44	20	ship	18.63	28	lyAV	25.83
5	ster	4.42	13	ee	11.54	21	dom	18.71	29	fulN	25.86
6	aryN	4.55	14	ish	12.70	22	hood	18.81	30	ism	25.86
7	ian	6.24	15	ess	13.17	23	less	21.78	31	ness	25.88
8	er	7.13	16	age	14.21	24	ous	21.83			

نکته حائز اهمیت در مرتبه تقطیع‌پذیری که در جدول ۲ آمده اینست که برخی از پسوندها در هم‌آیندی با دیگر پسوندها به‌کار نرفته و امکان مرتبه‌دهی به آنها وجود ندارد. از این رو های (۲۰۰۱) برای تعیین مرتبه این‌گونه وندها از نسبت تقطیع‌پذیری استفاده می‌کند. های و پلاگ (۲۰۰۴) علاوه بر مرتبه تقطیع‌پذیری که از هم‌آیندی وندها حاصل می‌شود و لذا در بعضی موارد قادر به مرتبه‌دهی نیست برای تعیین ترتیب وندها به بررسی معیار نسبت تقطیع‌پذیری هر وند نیز پرداختند تا در نهایت با توجه به هر دو معیار بتوانند پیوستار تقطیع‌پذیری را ترسیم کنند. در ادامه ابتدا نحوه محاسبه نسبت تقطیع‌پذیری بیان شده و سپس نحوه شکل‌گیری پیوستار تقطیع‌پذیری بر اساس دو معیار مرتبه تقطیع‌پذیر و نسبت تقطیع‌پذیر مشخص می‌شود.

های (۲۰۰۱) برای به‌دست آوردن نسبت تقطیع‌پذیری وندهای مورد بررسی از بسامد نوع و بسامد نمونه واژه‌های مشتق از هریک از وندها و پایه‌های آنها استفاده کرده است. نسبت تقطیع‌پذیری به میزان کاربرد و واژه‌سازی هر وند اشاره می‌کند. همان‌طور که های در این رویکرد اشاره می‌کند برخلاف دیگر رویکردها که صرفاً با استفاده از بسامد نوع واژه‌های مشتق به بررسی ترتیب وندها پرداخته‌اند در این رویکرد علاوه بر بسامد نوع از بسامد نمونه واژه‌های مشتق نیز بهره گرفته شده است. بنا به نظر وی هرچه بسامد (اعم از بسامد نوع و بسامد نمونه) واژه مشتق نسبت به پایه آن بیشتر باشد پردازش صرفی سازه محور کمتر و آن واژه مشتق به‌صورت یک واژه کل در نظر گرفته می‌شود. در مقابل هرچه بسامد واژه مشتق نسبت به پایه کمتر باشد پردازش صرفی سازه محور بیشتر و بالطبع تقطیع‌پذیری بیشتر است (های، ۲۰۰۱). این رویکرد برای ایجاد رابطه میان بسامد رخداد واژه مشتق و پایه از نمودار پراکندگی و قرار دادن واژه‌های مشتق بر روی محور افقی و همچنین پایه‌های تشکیل دهنده آنها در محور عمودی و محاسبه میانگین روند آنها (رگرسیون حداقل مربعات معمولی^۱) برای هر وند استفاده کرده است.^۲ های و پلاگ (۲۰۰۴) میانگین روند داده‌ها در نمودار پراکندگی را که در این نمودارها با عنوان رگرسیون حداقل مربعات معمولی شناخته می‌شود به‌عنوان خط تقطیع‌پذیری در نظر گرفته‌اند.^۳ با استفاده از خط تقطیع‌پذیری می‌توان واژه‌های مشتق حاصل از یک وند را به دو دسته تقسیم کرد: واژه‌هایی که در بالای خط قرار گرفته‌اند واژه‌هایی با بسامد رخداد پایه بیشتر هستند به‌طوری که برای درک آنها ابتدا واژه تقطیع شده و سپس معنای آن درک می‌شود و واژه‌هایی که در پایین خط قرار گرفته‌اند دارای بسامد رخداد پایه کمتری بوده و به‌صورت یک واژه درک می‌شوند. در واقع با دو دسته واژه که در این رویکرد آنها را واژه‌های تقطیع‌پذیر^۴ و تقطیع‌ناپذیر^۵ می‌نامیم مواجه خواهیم شد که می‌توان نسبتی برای این واژه‌ها در نظر گرفت و آن را نسبت تقطیع‌پذیری نامید (های و باین، ۲۰۰۲). نسبت تقطیع‌پذیری با دو مقوله

1. least trimmed squares regression

۲. نمودارهای مشابه در بخش تجزیه و تحلیل داده‌های فارسی در نمودار شکل ۳ صفحه ۱۷ آورده شده است.

۳. در نرم‌افزار اکسل با ترسیم نمودار پراکندگی و قرار دادن داده‌ها بر روی آن می‌توان با انتخاب گزینه trend line خط رگرسیون را برای داده‌ها ترسیم کرد. این خط با میانگین گرفتن از مختصات داده‌ها بر روی نمودار ترسیم می‌شود. همچنین این خط روند داده‌ها را نشان می‌دهد.

4. parsable

5. non-parsable

بسامد نوع و بسامد نمونه خبواژه‌های مشتق از هر وند در ارتباط می‌باشد (های و باین، ۲۰۰۲). برای محاسبه نسبت تقطیع‌پذیری لازم است نسبت بسامد نوع و بسامد نمونه واژه‌های مشتق بالای خط روند به بسامد نوع و نمونه کل واژه‌های در بر دارنده وند مورد نظر در پیکره محاسبه گردد. این نسبت مشخص می‌سازد که شنونده برای درک، دریافت و پردازش واژه مشتق از یک وند تا چه میزان به تقطیع‌پذیری و تفکیک اجزای سازنده آن واژه می‌پردازد (های و باین، ۲۰۰۳).

های و پلاگ (۲۰۰۴) با توجه به نسبت تقطیع‌پذیری بسامد نوع و نسبت تقطیع‌پذیری بسامد نمونه حاصل از واژه‌های مشتق از هر وند و مقایسه آن با دیگر وندها توانستند برای هر یک از وندها مرتبه نسبت بسامد نوع و مرتبه نسبت بسامد نمونه را تعیین کنند. در ادامه آنها با توجه به مرتبه تقطیع‌پذیری، مرتبه نسبت بسامد نوع و مرتبه نسبت بسامد نمونه و میانگین گرفتن از آنها و قرار دادن آن در محور عمودی نمودار پراکندگی و مرتبه تقطیع‌پذیری در محور افقی نمودار توانستند به پیوستار تقطیع‌پذیری پسوندهای اشتقاقی زبان انگلیسی برسند. این پژوهش در راستای پاسخ‌گویی به این مسائل است که تا چه میزان در زبان فارسی می‌توان با توجه به معیارهای مطرح شده در رویکرد ترتیب مبتنی بر پیچیدگی اعم از تقطیع‌پذیری، مرتبه تقطیع‌پذیری و نسبت تقطیع‌پذیری ترتیب پسوندهای اشتقاقی زبان فارسی را توجیه کرد و کارایی این رویکرد را محک زد. در ادامه به بررسی داده‌های زبان فارسی در چارچوب رویکرد ترتیب مبتنی بر پیچیدگی می‌پردازیم.

۴- تقطیع‌پذیری پسوندهای اشتقاقی زبان فارسی

این مقاله با هدف بررسی ترتیب پسوندهای اشتقاقی زبان فارسی در چارچوب رویکرد ترتیب مبتنی بر پیچیدگی انجام شده تا مشخص شود این رویکرد تا چه میزان می‌تواند توجیه‌کننده ترتیب‌های ممکن و موجود پسوندهای اشتقاقی زبان فارسی باشد. با توجه به موارد مطرح شده در بخش رویکرد ترتیب مبتنی بر پیچیدگی، برای تعیین مرتبه تقطیع‌پذیری و نسبت تقطیع‌پذیری به پیروی از های و پلاگ (۲۰۰۴) با استفاده از هم‌آیندی پسوندها و میزان تقطیع‌پذیری وند از پایه به بررسی پسوندهای اشتقاقی زبان فارسی پرداخته‌ایم. برای این منظور ابتدا چهل و یک پسوند که حداقل در دو اثر از سه اثر صادقی (۱۳۷۰)، کشانی (۱۳۷۱) و کلباسی (۱۳۸۰) به‌طور مشترک ذکر شده‌اند انتخاب و واژه‌های مشتق از آنها از پیکره پایگاه داده‌های زبان فارسی^۱ استخراج شدند. این پسوندها عبارتند از:

«-ا»، «-اک»، «-اُ»، «-ار»، «-ان»، «-انه»، «-آسا»، «-بان»، «-چه»، «-چی»، «-دان»، «-دانی»، «-زار»، «-سار»، «-ستان»، «-ش»، «-فام»، «-ک»، «-کده»، «-کی»، «-گار»، «-گانه».

۱. پیکره پایگاه داده‌های زبان فارسی دکتر مصطفی عاصی در مجموع دارای شصت میلیون واژه بوده و از طریق اینترنت نیز قابل دسترسی است. این پیکره حاوی هزار و پانصد اثر نظم و نثر ادبیات معاصر ایران و انواع متون زبانی اعم از اثرهای داستانی و غیرداستانی نثر، آثار شعری، نمایشنامه، فیلمنامه، ادبیات کودکان، مجله‌ها و نشریه‌های علمی، ادبی و تخصصی، روزنامه و نشریه‌های خبری، کتاب‌های درسی دانشگاهی و دبیرستانی، نامه‌های اداری، متون حقوقی و دیوار نوشته‌ها می‌باشد.

«-گاه»، «-گر»، «-گری»، «-گون»، «-گین»، «-مان»، «-مند»، «-ن»، «-ناک»، «-نده»، «-و»،
«-وار»، «-واره»، «-ور»، «-یت»، «-ین»، «-ینه»، «-ه»، «-ی»

پس از تعیین پسوندهای مورد بررسی، در ابتدا هم‌آیندی پسوندها را در قالب ماتریس مربعی تنظیم شده بر اساس حروف الفبا ترسیم کردیم. در این ماتریس اولین ستون جدول پسوندهای جایگاه اول نسبت به پایه و اولین سطر پسوندهای جایگاه دوم نسبت به پایه را نشان می‌دهد. همان‌طور که در ماتریس نشان داده شده است در صورت وجود هم‌آیندی پسوندها در محل تلاقی آنها در ماتریس عدد ۱ گذاشته شده است. جدول (۳) ماتریس مربعی هم‌آیندی پسوندهای اشتقاقی زبان فارسی را براساس حروف الفبا نشان می‌دهد.

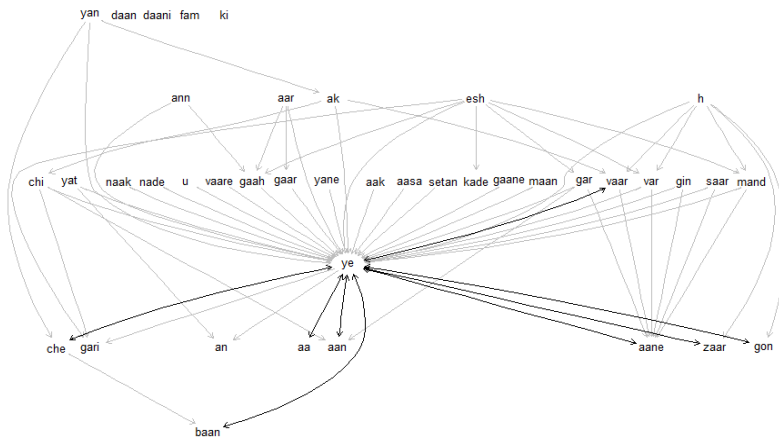
جدول ۳: ماتریس مربعی هم‌آیندی پسوندهای اشتقاقی زبان فارسی

پسوند	ا	آ	از	ار	ان	انه	انسان	انچه	چی	دانی	زار	ساز	استان	ش	فام	ک	کده	کی	کار	کانه	گاه	گر	گری	گون	گین	مان	مند	ن	ناک	نده	وار	واره	ور	ه	ی	یت	ین	ینه		
ا																																								
آ																																								
از																																								
ار																																								
ان																																								
انه																																								
انسان																																								
انچه																																								
چی																																								
دانی																																								
زار																																								
ساز																																								
استان																																								
ش																																								
فام																																								
ک																																								
کده																																								
کی																																								
کار																																								
کانه																																								
گاه																																								
گر																																								
گری																																								
گون																																								
گین																																								
مان																																								
مند																																								
ن																																								
ناک																																								
نده																																								
وار																																								
واره																																								
ور																																								
ه																																								
ی																																								
یت																																								
ین																																								
ینه																																								

همان‌طور که در جدول ۳ مشاهده می‌شود به‌عنوان مثال در واژه مشتق «دانشکده» در محل تلاقی پسوند «ش» در اولین ستون ماتریس و پسوند «کده» در اولین ردیف ماتریس عدد یک گذاشته شده که نشان از هم‌آیندی این دو پسوند با یکدیگر است. هدف استفاده از این ماتریس آن است که نشان دهیم هم‌آیندی وندها به‌صورت اتفاقی نیست و بر اساس محدودیت‌های گزینشی (فب، ۱۹۸۸) حاکم بر آنها صورت می‌گیرد. بر اساس ۴۱ پسوند می‌توان قائل به وجود ۱۶۴۰ ترکیب دو پسوندی بود، اما تعداد ۷۲ هم‌آیندی پسوندها در این ماتریس نشان می‌دهد که محدودیت‌های گزینشی اعم از محدودیت‌های صرفی، واجی، آوایی و نحوی حاکم بر هم‌آیندی وندها در زبان فارسی وجود دارد.

در ادامه با توجه به میزان هم‌آیندی پسوندهای اشتقاقی و با استفاده از نرم‌افزار آر^۱ (گنسنر^۲ و همکاران ۱۹۹۳) به ترسیم گراف جهت‌دار پسوندها پرداخته‌ایم. در این گراف پسوندها به ترتیب از بیشترین هم‌آیندی‌ها در قسمت پایین‌تر گراف تا کمترین هم‌آیندی‌ها در قسمت بالای آن قرار گرفته‌اند. گراف پسوندهای اشتقاقی زبان فارسی از ۴۱ رأس (۴۱ پسوند) و ۷۲ یال (خط متصل کننده دو پسوند) تشکیل شده است. همان‌طور که در شکل ۴ مشاهده می‌شود ۷۲ نوع هم‌آیندی در پسوندهای اشتقاقی دیده می‌شود که بنا بر انتظار باید تمامی یال‌ها به سمت پایین قرار گیرد. در گراف جهت‌دار هم‌آیندی وندها به گونه‌ای است که وندهایی که در بالای گراف قرار می‌گیرند از هم‌آیندی کمتر و هرچه به سمت پایین گراف حرکت کنیم هم‌آیندی بیشتر می‌شود. با این حال در این گراف پسوندهای اشتقاقی به همراه یال‌های دو طرفه را نیز شاهد بوده‌ایم. در این گراف خطوط جهت‌دار یک طرفه با رنگ روشن و خطوط جهت‌دار دو طرفه با مشکی پر رنگ نشان داده شده‌اند. در گراف جهت‌دار پسوندهای اشتقاقی زبان فارسی ۸ مورد از یال‌ها به صورت یال‌های دو طرفه بوده و همگی از پسوند «-ی» خارج شده‌اند. در شکل (۲) گراف جهت‌دار ۴۱ پسوند اشتقاقی زبان فارسی آورده شده است.

بیشترین تعداد هم‌آیندی همان‌طور که در گراف مشاهده می‌شود مربوط به پسوند «-ی» می‌باشد. همان‌طور که در این گراف نشان داده شده پسوندیایی چون «-دان»، «-دانی»، «-فام» و «-کی» در هم‌آیندی با دیگر پسوندها شرکت نداشته و در بالای گراف آورده شده‌اند. از ۴۱ پسوند مورد بررسی ۳۷ پسوند یعنی بیش از ۹۰٪ پسوندها در هم‌آیندی با دیگر پسوندها شرکت داشته‌اند. با توجه به گراف به دست آمده و با استفاده از ماتریس مربعی بالا مثلثی یا همان ماتریسی که تعداد هم‌آیندی زیر خط قطر اصلی آن کمترین تعداد باشد می‌توان مرتبه تقطیع‌پذیری وندها و ترتیب آنها را مشخص کرد.



شکل ۲: گراف جهت‌دار پسوندهای اشتقاقی زبان فارسی با ۴۱ رأس و ۷۲ یال

1. R software
2. Gansner

جدول ۵: مرتبه تقطیع پذیری پسوندهای اشتقاقی زبان فارسی بر اساس گراف جهت‌دار و ماتریس مربعی بالا مثلثی

ردیف	پسوند	مرتبه تقطیع پذیری	ردیف	پسوند	مرتبه تقطیع پذیری	ردیف	پسوند	مرتبه تقطیع پذیری
۱	-ش	۴/۸۶	۱۶	-ستان	۱۴/۷۳	۳۱	-ان	۳۴/۴۰
۲	-ار	۶/۳۵	۱۷	-سار	۱۴/۸۱	۳۲	-زار	۳۴/۴۲
۳	-ین	۶/۵۵	۱۸	-آسا	۱۴/۸۸	۳۳	-	۳۴/۴۲
۴	-ا	۸/۱۲	۱۹	-گاه	۱۴/۹۷	۳۴	-دان	۳۴/۴۶
۵	-ه	۸/۴۹	۲۰	-اک	۱۵/۰۴	۳۵	-فام	۳۴/۴۶
۶	-ک	۱۰/۲۵	۲۱	-یت	۱۵/۳۵	۳۶	-بان	۳۴/۴۸
۷	-ینه	۱۳/۸۴	۲۲	-مند	۱۵/۴۸	۳۷	-کده	۳۴/۶۶
۸	-واره	۱۳/۹۵	۲۳	-ور	۱۵/۵۰	۳۸	-کی	۳۴/۶۷
۹	-و	۱۴/۰۷	۲۴	-گر	۱۵/۶۶	۳۹	-دانی	۳۴/۶۸
۱۰	-گار	۱۴/۱۷	۲۵	-چی	۱۵/۸۹	۴۰	-گری	۳۴/۶۹
۱۱	-نده	۱۴/۱۹	۲۶	-وار	۲۸/۵۱	۴۱	-گون	۳۴/۷۷
۱۲	-ناک	۱۴/۳۲	۲۷	-چه	۲۹/۱۶			
۱۳	-مان	۱۴/۴۵	۲۸	-ی	۲۹/۳۰			
۱۴	-گانه	۱۴/۵۹	۲۹	-انه	۳۴/۳۴			
۱۵	-گین	۱۴/۷۳	۳۰	-ن	۳۴/۳۶			

همان‌طور که در گراف جهت‌دار مشاهده شد تعدادی از پسوندهای اشتقاقی اعم از پسوندهای «-دان»، «-دانی»، «-فام» و «-کی» در هم‌آیندی با دیگر وندها به‌کار نرفته‌اند در نتیجه بر اساس رویکرد ترتیب مبتنی بر پیچیدگی برای ترسیم پیوستار تقطیع‌پذیری نمی‌توان جایگاه دقیقی نسبت به دیگر پسوندها برای آنها بر روی پیوستار قائل شد. لذا در اینجا لازم است از نسبت تقطیع‌پذیری که با توجه به مشخصه‌های هر وندها و بسامد نوع و نمونه واژه‌های ساخته شده با آن به‌دست می‌آید مرتبه نسبت تقطیع‌پذیری را برای هر یک از پسوندهای اشتقاقی زبان فارسی به دست داد.

برای به‌دست آوردن مرتبه نسبت تقطیع‌پذیری واژه‌های مشتق از ۴۱ پسوند مورد بررسی، پایه‌های سازنده همراه با بسامد وقوع واژه‌ها از پیکره استخراج و در نرم‌افزار اکسل^۱ آورده شد. به پیروی از های و باین (۲۰۰۱) برای بسامد رخداد واژه‌های مشتق و بسامد رخداد پایه‌های تشکیل دهنده آنها از مقیاس لگاریتم^۲ مبنای ده^۳

1. Excel
2. logarithm

۳. لگاریتم مبنای ده با استفاده از نرم‌افزار اکسل محاسبه شده است.

استفاده شد. دلیل استفاده از مقیاس لگاریتمی مبنای ده این است که چون میان بسامد وقوع پایه‌ها و بسامد وقوع واژه‌های مشتق حاصل از آنها در بسیاری از موارد فاصله عددی وجود داشته و قابل مقایسه نیست لذا با گرفتن لگاریتم می‌توان این اعداد را به یکدیگر نزدیک‌تر کرده و مقایسه دقیق‌تری بین آنها انجام داد. به‌عنوان مثال واژه‌ای مشتق از پسوند «-آسا» مانند «معجزه آسا» در این پیکره ۵ مرتبه تکرار شده است. این در حالی است که پایه سازنده این واژه «معجزه» که واژه‌ای کاربردی در زبان فارسی است به تعداد ۲۳۰ مرتبه در پیکره حضور داشته است. در نتیجه مقایسه بین دو عدد با اختلاف بسیار زیاد نمی‌تواند به خوبی نتایج آماری را به دست دهد. با گرفتن لگاریتم بسامد رخداد نمونه این دو واژه به اعداد $0/7$ و $2/37$ رسیده‌ایم. در جدول (۶) نمونه‌ای از جدول نرم‌افزار اکسل تهیه شده از پسوند «-آسا» آورده شده است.

جدول ۶: واژه‌های مشتق از پسوند «-آسا»، پایه‌های آن، لگاریتم بسامد وقوع، واژه بالای خط تقطیع پذیر و واژه پایین خط تقطیع پذیر

واژه‌های مشتق از پسوند -آسا	بسامد واژه‌های مشتق	لگاریتم واژه‌های مشتق	پایه واژه‌های مشتق	بسامد پایه	لگاریتم پایه	واژه بالای خط تقطیع پذیری	واژه پایین خط تقطیع پذیری
اسب آسا	۱	۰	اسب	۱۴۳۰	۳,۱۶	اسب آسا	بحر آسا
اسطوره آسا	۲	۰,۳۱	اسطوره	۳۹۹	۲,۶۱	اسطوره آسا	پرنده آسا
بیر آسا	۲	۰,۳۱	بیر	۱۸۲	۲,۲۷	بیر آسا	تندر آسا
بحر آسا	۱	۰	بحر	۶۴	۱,۸۱	برق آسا	خلسه آسا
برق آسا	۵	۰,۷	برق	۱۹۳۹	۳,۲۹	پهلوان آسا	رعد آسا
پرنده آسا	۱	۰	پرنده	۱۰	۱	تن آسا	سپند آسا
پهلوان آسا	۱	۰	پهلوان	۲۷۵	۲,۴۴	جنون آسا	صاعقه آسا
تن آسا	۲	۰,۳۱	تن	۳۴۲۱	۳,۵۴	دیو آسا	طلسم آسا
تندر آسا	۱	۰	تندر	۷	۰,۸۵	سحر آسا	عنکبوت آسا
جنون آسا	۱	۰	جنون	۱۵۶	۲,۲	سیل آسا	غول آسا
خلسه آسا	۱	۰	خلسه	۱۹	۱,۲۸	عطر آسا	معجز آسا
دیو آسا	۱	۰	دیو	۱۳۲	۲,۱۳	مار آسا	هر کول آسا
رعد آسا	۶	۰,۷۸	رعد	۹۳	۱,۹۷		
سپند آسا	۱	۰	سپند	۲۰	۱,۳۱		
سحر آسا	۱	۰	سحر	۲۸۹	۲,۴۷		
سیل آسا	۵	۰,۷	سیل	۲۹۹	۲,۴۸		
صاعقه آسا	۴	۰,۶۱	صاعقه	۵۳	۱,۷۳		



طلسم آسا	۲	۰,۳۱	طلسم	۸۷	۱,۹۴
عطر آسا	۱	۰	عطر	۲۱۲	۲,۳۳
عنکبوت آسا	۱	۰	عنکبوت	۵۲	۱,۷۲
غول آسا	۸	۰,۹۱	غول	۱۸۵	۲,۲۷
مار آسا	۱	۰	مار	۲۷۸	۲,۴۵
معجزه آسا	۵	۰,۷	معجزه	۲۳۰	۲,۳۷
هرکول آسا	۱	۰	هرکول	۲	۰,۳۱

به پیروی از های و باین (۲۰۰۳) برای بررسی میزان بسامد واژه‌های مشتق حاصل از یک وند و ارتباط آن با بسامد پایه‌های همان وند از نمودار پراکندگی (شکل ۳) استفاده کردیم. برای ترسیم این نمودار در نرم‌افزار اکسل لگاریتم بسامد وقوع واژه‌های مشتق از وند مورد بررسی را در محور افقی^۱ و لگاریتم بسامد وقوع پایه‌های سازنده آنها را در محور عمودی^۲ قرار دادیم. هر یک از نقاط ایجاد شده در نمودار پراکندگی با توجه به مختصات آن نسبت به محور افقی (بسامد واژه مشتق) و محور عمودی (بسامد پایه) بیانگر یک واژه مشتق حاصل از پسوند مورد بررسی است. در ادامه با استفاده از نرم‌افزار اکسل خط رگرسیون حداقل مربعات معمولی (خط تقطیع‌پذیری) که رابطه واژه‌های مشتق و پایه‌های آنها را با توجه به میانگین نقاط موجود در فضای بین دو محور نشان می‌دهد ترسیم کردیم. این خط که در نمودارهای شکل ۳ به صورت فلش نشان داده شده واژه‌ها را به دو دسته واژه‌های بالای خط و واژه‌های پایین خط تقسیم کرده است.

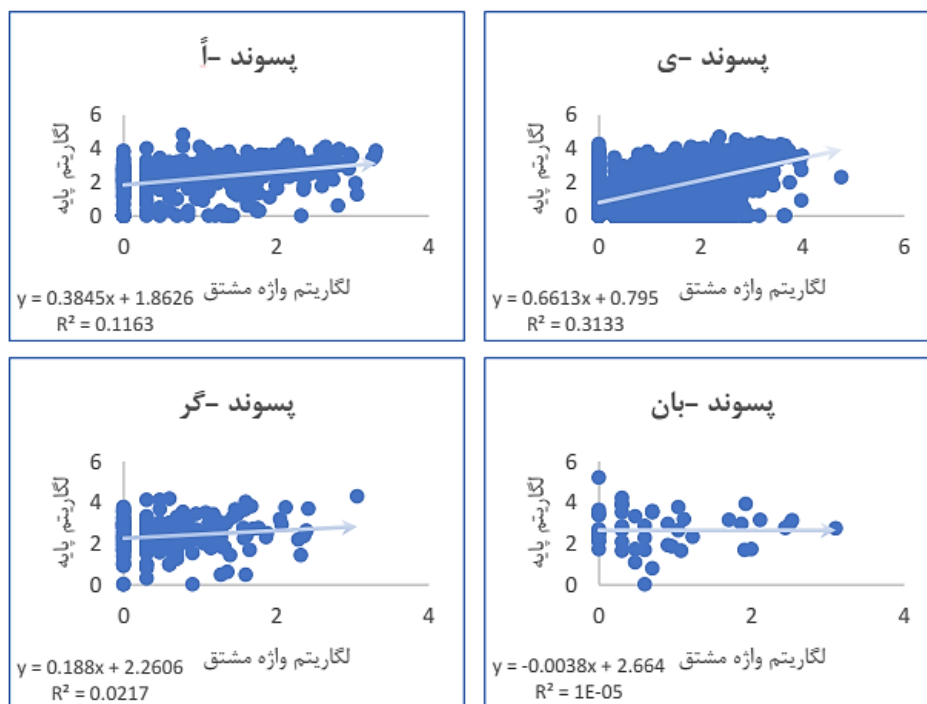
به پیروی از های و باین (۲۰۰۳) داده‌های قرار گرفته در بالای این خط را واژه‌های مشتق تقطیع‌پذیر (واژه‌هایی با بسامد پایه بیشتر از واژه مشتق) و داده‌های قرار گرفته در پایین این خط را واژه‌های مشتق تقطیع‌ناپذیر (واژه‌هایی با بسامد واژه مشتق بیشتر از پایه) نامیده‌ایم. به عنوان مثال در خصوص پسوندهای «-ی»، «-ا»، «-ان»، «-گر» و «-بان» بر طبق نمودار پراکندگی آنها واژه‌های «ابدی»، «ابدأ»، «باغچه بان»، «ایثارگر» که در زیر خط تقطیع‌پذیری قرار گرفته‌اند واژه‌های تقطیع‌ناپذیر و واژه‌های «ابریشمی»، «اخلاقاً»، «جهانبان»، «ارشادگر» که در بالای خط تقطیع‌پذیری قرار گرفته‌اند واژه‌های تقطیع‌پذیر در نظر گرفته می‌شوند. شکل ۳ به عنوان نمونه نمودار پراکندگی چهار پسوند اشتقاقی زبان فارسی را نشان می‌دهد.^۳

1. X- Axis

2. Y- Axis

۳. لازم به توضیح است که این نمودارها شکل کوچک شده نمودارهای اصلی هستند و چون کوچک شده‌اند تراکم داده‌ها زیاد و بعضی نقطه‌ها روی هم قرار گرفته‌اند. صرفاً تمامی واژه‌های حاصل از یک وند در یک پیکره دارای بسامد یکسان و پایه‌هایی با بسامد یکسان نیستند. همچنین با تراکم زیاد داده‌ها روی محور نشان دادن واژگان برای این محورها خوانا نخواهد بود. با استفاده از فرمول خط رگرسیون و محاسبه معادله خط به راحتی می‌توان واژه‌های قرار گرفته در پایین و بالای خط روند را برای هر وند مشخص کرد.

شکل ۳: نمودار پراکندگی پسوندهای «ا»، «ی»، «بان» و «گر». محور افقی بیانگر لگاریتمم واژه مشتق و محور عمودی بیانگر لگاریتمم پایه‌های سازنده واژه مشتق از هریک از وندها می‌باشد. میزان رگرسیون خط نیز با R^2 نشان داده شده است.



همان‌طور که در نمودارهای حاصل از پسوندها در این بررسی مشاهده شد پسوند «ه» و «ی» دارای بیشترین تعداد واژه مشتق بوده و تعداد واژه‌های مشتق از این دو پسوند که در بالای خط روند قرار گرفته‌اند نیز به همان نسبت بیشتر بوده است. با توجه به شکل (۳) شیب خط روند در همه نمودارهای حاصل از پسوندها به جز پسوند «بان» مثبت و به سمت بالا قرار گرفته و در این پسوند مقدار عددی جهت شیب رو به پایین بسیار کم بوده و تقریباً به صورت خط صاف مشاهده می‌شود. این نشان می‌دهد که پسوندهای مورد بررسی پسوندهای فعال در زبان فارسی هستند. همان‌طور که در تصویر نشان داده شده شیب خط روند در پسوند «ی» نسبت به دیگر پسوندها بیشتر است. طبق آنچه در رویکرد ترتیب مبتنی بر پیچیدگی گفته می‌شود، هرچه شیب خط روند بیشتر باشد تقطیع‌پذیری بیشتر و واژه‌سازی با پسوند مورد نظر نیز بیشتر خواهد بود. در مواردی مانند پسوند «فام»، «واره» و «اک» با وجود کم بودن تعداد واژه‌های مشتق ساخته شده از هر یک از این پسوندها تعداد پایه‌های سازنده این واژه‌های مشتق در پیکره بیشتر از واژه‌های مشتق بوده به طوری که باعث شده خط رگرسیون روند صعودی و شیب تندی را نشان دهد. در نمودارهای پراکندگی پسوندهای

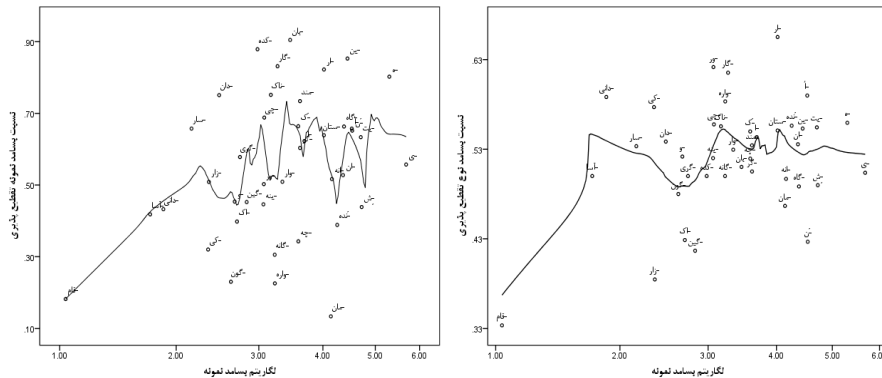
«ا»، «چی»، «دان»، «سار»، «گار»، «کی»، «گون»، «مند»، «ناک»، «ور» خط روند نشان داده شده در رابطه بین واژه‌های مشتق از این پسوندها و پایه‌های آنها به صورتی خطی تقریباً صاف و بدون شیب می‌باشد. این فرایند نشان می‌دهد که بسامد رخداد نمونه واژه‌های ساخته شده با این پسوندها با بسامد رخداد نمونه پایه‌های سازنده آنها در این پیکره به یک اندازه بوده است.

در ادامه با استفاده از فرمول خطی معادله درجه اول خط روند، واژه‌های قرار گرفته در بالای خط تقطیع‌پذیر را مشخص کردیم. آنگاه از تقسیم بسامد نوع و بسامد نمونه واژه‌های مشتق قرار گرفته در بالای خط تقطیع‌پذیر بر بسامد نوع و بسامد نمونه کل واژه‌های مشتق ساخته شده با پسوند مورد بررسی نسبت بسامد رخداد نوع تقطیع‌پذیری و نسبت بسامد رخداد نمونه تقطیع‌پذیری را برای هر پسوند به دست آوردیم. سپس برای هر یک از پسوندها نسبت‌های تقطیع‌پذیری مشخص شده را با یکدیگر مقایسه و مرتبه نسبت بسامد نوع و نمونه هر پسوند را نیز به دست آوردیم. از نظر های و باین (۲۰۰۳) هرچه بسامد نمونه واژه مشتق افزایش یابد، نسبت بسامد نوع و نمونه تقطیع‌پذیر کاهش می‌یابد. استدلال آنها بر آنست که با افزایش بسامد نمونه واژه مشتق، با توجه به اینکه واژه به گوش شنونده آشناتر می‌شود، برای درک و دریافت معنی آن نیازی به تقطیع واژه نیست بلکه آن واژه به صورت یک کل درک می‌شود. برای پی‌بردن به روند کلی تقطیع‌پذیری پسوندهای اشتقاقی زبان فارسی و همچنین ارتباط آنها با بسامد نمونه واژه‌های تشکیل دهنده هر پسوند توانستیم با ترسیم نمودار پراکندگی به روند کلی پسوندهای اشتقاقی زبان فارسی نیز دست یابیم. هدف از این بررسی آن است که دریابیم در واژه‌های مشتق پسوندی زبان فارسی نیز - همان طوری که های و باین (۲۰۰۱) معتقدند - با افزایش بسامد رخداد نمونه واژه‌های مشتق و با توجه به مشخصه روانشناختی و پردازش ذهنی واژگان آیا نسبت بسامد نوع و نسبت بسامد نمونه تقطیع‌پذیر واژه‌های مشتق کاهش می‌یابد.

همان گونه که در نمودارهای شکل (۴) مشاهده می‌شود با جمع‌آوری اطلاعات هر یک از پسوندها در نرم‌افزار اکسل و با ترسیم نمودار با نرم‌افزار اس پی اس اس^۱ روند کلی رابطه بسامد نمونه و نسبت بسامد نوع و نسبت بسامد نمونه تقطیع‌پذیر همراه با خط روند^۲ LOESS را ترسیم کرده‌ایم. شکل (۴) رابطه لگاریتم بسامد نمونه و نسبت بسامد نوع و نسبت بسامد نمونه تقطیع‌پذیری را نشان می‌دهد.

1. SPSS

2. LOcal regrESSion



شکل ۴: الف. نمودار پراکندگی لگاریتم بسامد نمونه و نسبت نمونه تقطیع پذیر،
ب. نمودار پراکندگی لگاریتم بسامد نمونه و نسبت نوع تقطیع پذیر

همان‌گونه که در شکل (۴) مشاهده می‌شود در نمودار الف لگاریتم بسامد نمونه هر یک از پیشوندها در محور افقی و اعداد حاصل از تقسیم بسامد نمونه تقطیع‌پذیر بر بسامد نمونه واژه‌های مشتق از هر وند (نسبت بسامد نمونه تقطیع‌پذیر) در محور عمودی جای گرفته به طوری که محور عمودی را می‌توان نسبتی از محور افقی دانست. همچنین در نمودار ب لگاریتم بسامد نوع هر یک از پسوندها در محور افقی و اعداد حاصل از تقسیم نسبت بسامد نوع تقطیع‌پذیر به بسامد نوع واژه‌های مشتق از هر وند (نسبت بسامد نوع) در محور عمودی جای گرفته است. بر طبق این نمودارها روند کلی به سمت پایین بوده و حالت نزولی به روشنی مشاهده می‌شود. این تصاویر گویای این مطلب هستند که هرچه میزان بسامد نمونه وندهای اشتقاقی زبان فارسی بیشتر می‌شود نسبت تقطیع‌پذیری نمونه واژه‌های مشتق و نسبت تقطیع‌پذیری نوع واژه‌های مشتق نیز کمتر می‌شود. در اصل هر چه میزان بسامد واژه‌های مشتق بالاتر می‌رود آن واژه‌ها به گوش شنونده آشنا تر شده و در دریافت و پردازش معنی واژه‌های ساخته شده با پسوند مورد نظر نیازی به تقطیع کردن واژه‌های مشتق حاصل از آنها نیست. در این صورت این واژه‌های مشتق پربسامد بیشتر به صورت یک واژه کل درک می‌شوند و پردازش آنها کمتر به صورت سازه محور صورت می‌گیرد. این مشاهدات در زبان فارسی نیز همراستا با مشاهدات های و باین (۲۰۰۱) می‌باشد.

برای رسیدن به پیوستار تقطیع‌پذیری از مرتبه نسبت بسامد نوع تقطیع‌پذیری پسوندها و مرتبه نسبت بسامد نمونه تقطیع‌پذیری استفاده کردیم. برای به دست آوردن مرتبه نسبت بسامد نمونه تقطیع‌پذیر ابتدا آنها را بر اساس کمترین میزان تا بیشترین میزان در نرم‌افزار اکسل مرتب کرده و از شماره یک به آنها مرتبه داده شد. جدول (۷) مرتبه تقطیع‌پذیری بر اساس ماتریس مربعی، مرتبه نسبت بسامد نمونه، مرتبه نسبت بسامد نوع تقطیع‌پذیر و در ستون آخر میانگین مرتبه‌بندی مرتبه تقطیع‌پذیر، مرتبه بسامد نمونه و مرتبه بسامد نوع را نشان می‌دهد.

جدول ۷: مرتبه هر پسوند در زبان فارسی بر اساس مرتبه تقطیع‌پذیری، مرتبه نسبت بسامد نمونه تقطیع‌پذیر، مرتبه نسبت بسامد نوع تقطیع‌پذیر و میانگین آنها

مرتبه میانگین (میانگین مرتبه تقطیع‌پذیری مرتبه نمونه و مرتبه نوع)	مرتبه نسبت بسامد نوع	مرتبه نسبت بسامد نمونه	مرتبه تقطیع‌پذیری	پسوند
۷	۸	۱۲	۱	ش-
۲۶,۷	۴۱	۳۷	۲	ار-
۲۴	۳۰	۳۹	۳	ین-
۱۱,۹	۴,۵	۲۷	۴	ن-
۲۵	۳۴	۳۶	۵	ه-
۲۱,۵	۲۹	۲۹,۵	۶	کی-
۱۳,۷	۲۰	۱۴	۷	ینه-
۱۵,۵	۳۵	۳,۵	۸	واره-
۱۴	۱۹	۱۴	۹	و-
۲۹,۴	۴۰	۳۸	۱۰	گار-
۱۶,۷	۳۱	۸	۱۱	نده-
۲۶,۵	۳۳	۳۴,۵	۱۲	ناک-
۶,۷	۶	۱	۱۳	مان-
۹,۹	۱۰,۵	۵	۱۴	گانه-
۱۰,۷	۳	۱۴	۱۵	گین-
۲۳,۴	۲۸	۲۶	۱۶	ستان-
۲۲,۹	۲۲	۲۹,۵	۱۷	سار-
۱۲,۹	۱۰,۵	۱۰	۱۸	آسا-
۱۹,۲	۹	۲۹,۵	۱۹	گاه-
۱۱,۲	۴,۵	۹	۲۰	اک-
۲۴,۴	۲۷	۲۵	۲۱	یت-
۲۶	۲۳	۳۳	۲۲	مند-
۲۶	۳۹	۱۶	۲۳	ور-
۲۱	۱۶	۲۳	۲۴	گر-
۲۹,۷	۳۲	۳۲	۲۵	چی-
۲۱,۵	۲۱	۱۷,۵	۲۶	وار-
۱۷,۴	۱۸	۷	۲۷	چه-
۲۰,۷	۱۳	۲۱	۲۸	ی-

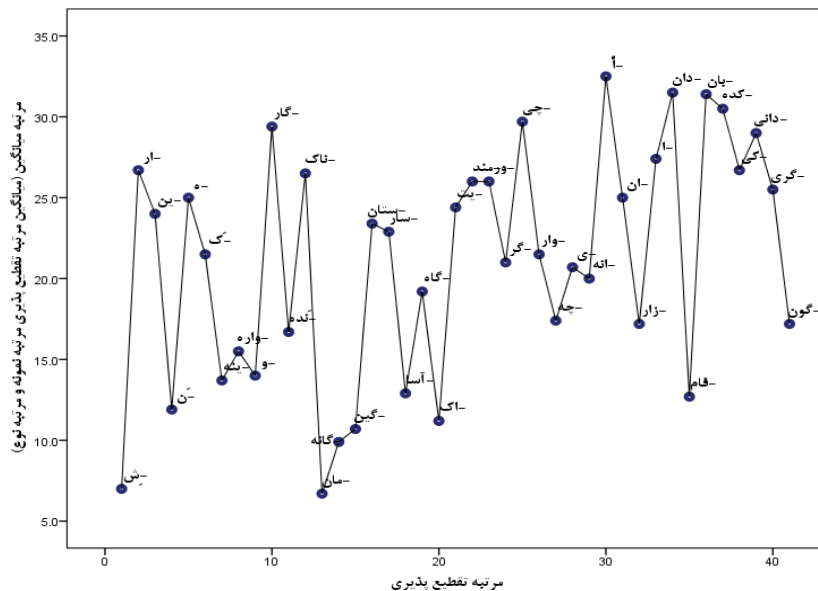
پسوند	مرتبۀ تقطیع‌پذیری	مرتبۀ نسبت بسامد نمونه	مرتبۀ نسبت بسامد نوع	مرتبۀ میانگین (میانگین مرتبۀ تقطیع‌پذیری مرتبۀ نمونه و مرتبۀ نوع)
-انه	۲۹	۱۹	۱۲	۲۰
-ا	۳۰	۲۹,۵	۳۸	۳۲,۵
-ان	۳۱	۲۰	۲۴	۲۵
-زار	۳۲	۱۷,۵	۲	۱۷,۲
-ا	۳۳	۲۴	۲۵	۲۷,۴
-دان	۳۴	۳۴,۵	۲۶	۳۱,۵
-فام	۳۵	۲	۱	۱۲,۷
-بان	۳۶	۴۱	۱۷	۳۱,۴
-کده	۳۷	۴۰	۱۴,۵	۳۰,۵
-کی	۳۸	۶	۳۶	۲۶,۷
-دانی	۳۹	۱۱	۳۷	۲۹
-گری	۴۰	۲۲	۱۴,۵	۲۵,۵
-گون	۴۱	۳,۵	۷	۱۷,۲

بر طبق جدول مرتبۀ نسبت بسامد نوع تقطیع‌پذیری برای دو پسوند «-نده» و «-ناک» عدد ۳۱/۵ است. چون این دو پسوند دارای نسبت بسامد نوع تقطیع‌پذیری یکسان بوده و در جایگاه سی و یکم و سی و دوم قرار گرفته‌اند؛ لذا برای تعیین مرتبۀ آنها از عدد میانگین آنها استفاده کرده‌ایم تا مرتبۀ آنها نسبت به دیگر پسوندها محاسبه گردد. با گرفتن میانگین از سه مرتبۀ به‌دست آمده برای هر پسوند که عبارتند از مرتبۀ تقطیع‌پذیر، مرتبۀ نسبت بسامد نمونه و مرتبۀ نسبت بسامد نوع، مرتبۀ میانگین به‌دست آمده است.

همان‌طور که در جدول (۷) مشاهده می‌شود ترتیب قرارگیری وندها بر اساس هر یک از مرتبۀ متفاوت است. به‌عنوان مثال ترتیب پسوندها بر اساس نسبت بسامد نمونه تقطیع‌پذیری پسوند «-مان» کمترین نسبت و پسوند «-بان» بیشترین نسبت تقطیع‌پذیری را نشان می‌دهد. همچنین بر اساس نسبت بسامد نوع تقطیع‌پذیری کمترین نسبت را پسوند «-فام» و بیشترین نسبت تقطیع‌پذیری را پسوند «-ار» به خود اختصاص داده است. همچنین بر اساس مرتبۀ تقطیع‌پذیری بیشترین مرتبۀ مربوط به پسوند «-گون» و کمترین مرتبۀ را پسوند «-ش» داشته است. بر اساس ادعای‌های و پلاگ (۲۰۰۴) به‌عنوان مثال در زبان فارسی نیز واژه مشتق «دانش» به‌صورت یک واژه کل درک می‌شود و برای رسیدن به معنای آن پسوند «-ش» نیازی به تقطیع ندارد اما در واژه «لاله‌گون» برای رسیدن به معنای آن پسوند «-گون» به راحتی از واژه مشتق جدا شده و معنای آن درک خواهد شد. در ادامه بنا به نظر‌های و پلاگ (۲۰۰۴) بهتر است از این مرتبۀ میانگین گرفته تا در زبان فارسی نیز به یک پیوستار ترتیب پسوندها دست یابیم.

۵- پیوستار ترتیب پسوندهای اشتقاقی زبان فارسی

در این قسمت با توجه به مرتبه‌بندی‌های به‌دست آمده از پسوندهای اشتقاقی زبان فارسی به ترسیم پیوستار ترتیب وندها خواهیم پرداخت. هدف از ترسیم پیوستار آن است که بتوانیم ترتیب قرارگیری پسوندهای اشتقاقی زبان فارسی را در واژه‌های موجود و محتمل نشان دهیم. به تبعیت از های و پلاگ (۲۰۰۴) برای برقراری ارتباط میان مرتبه تقطیع‌پذیری که حاصل از هم‌آیندی پسوندهای اشتقاقی زبان فارسی می‌باشد و مرتبه‌های حاصل از نسبت بسامد نوع و نمونه تقطیع‌پذیر که از ارتباط بسامد واژه‌های مشتق و پایه‌های آنها در نمودار پراکندگی و تقسیم واژه‌های بالای خط تقطیع‌پذیری بر کل واژه‌های مشتق از هر وند به‌دست آمد و با میانگین گرفتن از این مرتبه‌ها و ارتباط دادن آن با مرتبه تقطیع‌پذیری، آنها را در نمودار پراکندگی قرار می‌دهیم تا پیوستار تقطیع‌پذیری به‌دست داده شود. برای ترسیم این پیوستار با استفاده از جدول ۶ مرتبه تقطیع‌پذیری پسوندهای اشتقاقی زبان فارسی را در محور افقی نمودار و مرتبه میانگین را در محور عمودی قرار داده و سپس با استفاده از خط روند پیوستار ایجاد شده را نشان می‌دهیم. با توجه به خط روند ایجاد شده مرتبه تقطیع‌پذیری وندها با میانگین مرتبه‌ها در رابطه مستقیم بوده و با افزایش در محور افقی افزایش در محور عمودی را نیز شاهد بوده‌ایم. در این نمودار پسوندها بر اساس مرتبه تقطیع‌پذیری از کمترین میزان تقطیع‌پذیری تا بیشترین میزان تقطیع‌پذیری بر روی پیوستار قرار گرفته‌اند. شکل ۵. پیوستار ترتیب پسوندهای اشتقاقی زبان فارسی را نشان می‌دهد.



شکل ۵: پیوستار تقطیع‌پذیری با استفاده از نمودارهای پراکندگی مرتبه تقطیع‌پذیری و میانگین مرتبه‌های تقطیع‌پذیری، نسبت بسامد نمونه و نسبت بسامد نوع.

انتظار بر این است وندهایی که در سمت چپ پیوستار قرار گرفته‌اند به پایه نزدیکتر بوده و هر چه به سمت راست پیوستار حرکت کنیم پسوندها در جایگاه دورتری از پایه قرار گیرند. با توجه به نمودار پراکندگی مرتبه تقطیع‌پذیری و میانگین مرتبه‌های حاصل از نسبت بسامد نوع و نسبت بسامد نمونه نشان می‌دهد که وندهایی که در قسمت تقطیع‌پذیر پیوستار قرار گرفته‌اند همان وندهایی هستند که در هم‌آیندی با دیگر وندها نیز شرکت نکرده‌اند و وندهایی که در سمت تقطیع‌ناپذیری پیوستار قرار گرفته‌اند در هم‌آیندی با دیگر وندها شرکت داشته‌اند. خط متصل‌کننده وندها بیانگر ترتیب وندها می‌باشد.

همان‌طور که گفته شد وندهای قرار گرفته در سمت چپ نمودار وندهایی با تقطیع‌پذیری کمتر و وندهای قرار گرفته در سمت راست نمودار وندهایی با تقطیع‌پذیری بیشتر می‌باشند. به‌عنوان مثال در واژه مشتق «دانشمند» پسوند «-ش» قبل از پسوند «-مند» قرار گرفته است و یا در واژه «خشمگینانه» پسوند «-گین» قبل از پسوند «-انه» در این پیوستار قرار گرفته است. بنا به نظر های و پلاگ (۲۰۰۹) هرچه تقطیع‌پذیری بیشتری باشد وند از پایه دورتر است و هرچه تقطیع‌پذیری کمتر باشد وند به پایه نزدیکتر است. همان‌طور که داده‌ها نشان می‌دهد در اغلب موارد بین پیوستار تقطیع‌پذیری و ترتیب وندها در ساختار واژه‌های مستخرج از پیکره همپوشانی وجود دارد.

با توجه به این مطالب می‌توان مدعی شد که معیار تقطیع‌پذیری می‌تواند تعداد قابل ملاحظه‌ای از موارد ترتیب وندها را توجیه کند با این حال در بعضی موارد نمی‌توان ترتیب خاصی را با این معیار توجیه کرد و لاجرم می‌بایست از معیار محدودیت‌های گزینشی فب (۱۹۸۸) بهره برد.

در نهایت نتایج این بخش نشان می‌دهد که هرچه وندها در پردازش زبانی تقطیع‌پذیر باشند احتمال اینکه به‌طور آزادانه به وندهای دیگر متصل شوند بیشتر خواهد بود همچنین وندهایی که به راحتی تقطیع می‌شوند نمی‌توانند درون وندهایی که تقطیع نمی‌شوند جای گیرند. به‌عنوان مثال واژه «لاله گونش» در زبان فارسی مشاهده نشده است.

۶- نتیجه‌گیری

این پژوهش با هدف بررسی ترتیب پسوندهای اشتقاقی زبان فارسی در چارچوب رویکرد ترتیب مبتنی بر پیچیدگی به بررسی ۴۱ پسوند اشتقاقی زبان فارسی و استخراج داده‌ها از پیکره پژوهشگاه علوم انسانی و مطالعات فرهنگی پرداخت. با توجه به مطالعات انجام شده در خصوص پسوندهای اشتقاقی زبان فارسی ابتدا توانستیم بین واژه‌های مشتق از یک وند و پایه‌های سازنده آن با استفاده از نمودار پراکندگی ارتباط برقرار کرده و به نسبت تقطیع‌پذیری هر وند دست پیدا کنیم. با اطلاعات به‌دست آمده از این نمودارها توانستیم سطح فعالیت هر یک از وندهای مورد بررسی را نیز به‌دست آورده و آن را با دیگر وندها مقایسه کنیم. با مقایسه نسبت تقطیع‌پذیری وندها و بسامد نمونه آنها به این نتیجه رسیدیم که هرچه بسامد نمونه بیشتر باشد نسبت تقطیع‌پذیری کمتر بوده که با توجه به ماهیت روانشناختی رویکرد ترتیب مبتنی بر پیچیدگی افزایش بسامد نمونه از آن جهت منجر به تقطیع‌پذیری کمتر می‌شود که شنونده برای دریافت و درک معنی واژه مشتق نیازی

به تقطیع آن ندارد. همچنین با بررسی انجام شده در این پژوهش اینک می‌توان به این پرسش پاسخ داد که پسوندهای اشتقاقی زبان فارسی را می‌توان بر اساس مرتبه تقطیع‌پذیری و مرتبه نسبت بسامد نوع و نسبت بسامد نمونه تقطیع‌پذیری بر روی یک پیوستار قرار داد. این پیوستار می‌تواند بیانگر ترتیب قرارگیری وندها در ترکیبات موجود و ممکن در زبان فارسی باشد. همچنین می‌توان با مقایسه پیوستار ترتیب مبتنی بر پیچیدگی، واژه‌های مشتق در پیکره و ماتریس هم‌آیندی پسوندهای اشتقاقی زبان فارسی ترتیب پسوندهای اشتقاقی را نیز با رویکرد ترتیب مبتنی بر پیچیدگی و معیار تقطیع‌پذیری وند از پایه توجیه کرد. با این حال در مورد پسوند استثنایی «-ی» که در بعضی موارد با پیوستار همپوشانی وجود ندارد می‌توان با استفاده از محدودیت‌های گزینشی حاکم بر این وند آن را توجیه کرد. این ویژگی نیز همراستا با نظر های و پلاگ (۲۰۰۴) مبنی بر مکمل بودن این دو رویکرد نسبت به یکدیگر می‌باشد.

منابع

- حمصیان، مریم (۱۳۸۹). بررسی آرایش وندهای اشتقاقی زبان فارسی. پایان نامه کارشناسی ارشد: دانشگاه اصفهان.
- صادقی، علی اشرف (۱۳۷۰الف). «شیوه‌ها و امکانات واژه‌سازی در زبان فارسی معاصر(۱)»، نشر دانش، خرداد و تیر، تهران: مرکز نشر دانشگاهی.
- صادقی، علی اشرف (۱۳۷۰ب). «شیوه‌ها و امکانات واژه‌سازی در زبان فارسی معاصر(۲)»، نشر دانش، مرداد و شهریور، تهران: مرکز نشر دانشگاهی.
- صادقی، علی اشرف (۱۳۷۰ج). «شیوه‌ها و امکانات واژه‌سازی در زبان فارسی معاصر(۳)»، نشر دانش، آذر و دی، تهران: مرکز نشر دانشگاهی.
- صادقی، علی اشرف (۱۳۷۱الف). «شیوه‌ها و امکانات واژه‌سازی در زبان فارسی معاصر(۴)»، نشر دانش، خرداد و تیر، تهران: مرکز نشر دانشگاهی.
- صادقی، علی اشرف (۱۳۷۱ب). «شیوه‌ها و امکانات واژه‌سازی در زبان فارسی معاصر(۵)»، نشر دانش، مرداد و شهریور، تهران: مرکز نشر دانشگاهی.
- صادقی، علی اشرف (۱۳۷۱ج). «شیوه‌ها و امکانات واژه‌سازی در زبان فارسی معاصر(۶)»، نشر دانش، مهر و آبان، تهران: مرکز نشر دانشگاهی.
- صادقی، علی اشرف (۱۳۷۱د). «شیوه‌ها و امکانات واژه‌سازی در زبان فارسی معاصر(۷)»، نشر دانش، آذر و دی، تهران: مرکز نشر دانشگاهی.
- صادقی، علی اشرف (۱۳۷۱ه). «شیوه‌ها و امکانات واژه‌سازی در زبان فارسی معاصر(۸)»، نشر دانش، بهمن و اسفند، تهران: مرکز نشر دانشگاهی.

- صادقی، علی اشرف (۱۳۷۲ الف). «شیوه‌ها و امکانات واژه‌سازی در زبان فارسی معاصر(۹)»، نشر دانش، فروردین و اردیبهشت، تهران: مرکز نشر دانشگاهی.
- صادقی، علی اشرف (۱۳۷۲ ب). «شیوه‌ها و امکانات واژه‌سازی در زبان فارسی معاصر(۱۰)»، نشر دانش، خرداد و تیر، تهران: مرکز نشر دانشگاهی.
- صادقی، علی اشرف (۱۳۷۲ ج). «شیوه‌ها و امکانات واژه‌سازی در زبان فارسی معاصر(۱۱)»، نشر دانش، مرداد و شهریور، تهران: مرکز نشر دانشگاهی.
- غلامعلی‌زاده، خسرو و یداله فیضی پیرانی (۱۳۹۰). «بررسی لایه‌ها و سطوح واژگانی در وندهای زبان فارسی از نگاه صرف واژگانی»، مجله پژوهش‌های زبان شناسی دانشگاه اصفهان، س ۳. ش ۵. صص. ۶۵-۸۴.
- کشانی، خسرو (۱۳۷۱). *اشتقاق پسوندی در زبان فارسی امروز*، تهران: مرکز نشر دانشگاهی.
- کلیبسی، ایران (۱۳۸۰). *ساخت اشتقاقی واژه در فارسی امروز*، تهران: پژوهشگاه علوم انسانی و مطالعات فرهنگی.
- Allen, M. (1978). *Morphological investigations*. Ph.D. Thesis, University of Connecticut.
- Baayen, R. H. (2002). *Word Frequency Distribution*, Dordrecht, Kluwer Academic Publishers.
- Fabb, N. (1988). “English Suffixation is Constrained only by Selectional Restrictions”, *Natural Language and linguistic Theory* 6, 527-539.
- Gansner, Emden R.; Eleftherios Koutsofios; Stephen C. North; and Kiem-Phon vo. (1993). “A technique for drawing directed graphs”. *IEEE Transactions on Software Engineering* 19.3.214-30.
- Giegerich, H. J. (1999). *Lexical Strata in English. Morphological causes, Phonological Effects*, Cambridge: Cambridge University Press.
- Hay J. (2000). *Causes and Consequences of Word Structure*, Ph.D. dissertation. Northwestern University.
- Hay J. (2001). “Lexical Frequency in Morphology: Is Everything Relative?”, *Linguistics* 39: 1041-1070.
- Hay J. (2002). “From Speech Perception to Morphology: Affix-ordering Revisited”, *Language* 78.3, 527-555.
- Hay, J. & Baayen, H. (2002). “Parsing and Productivity”. *Yearbook of Morphology*, Dordrecht: Kluwer, 203- 255.
- Hay, J. & Baayen, H. (2003). “Phonotactics, Parsing and Productivity”, *Rivista di Linguistica*, 15.1, p. 99-130.
- Hay, J. & Plag, I. (2002). “Suffix Combinations, Grammatical Restrictions, and Parsing”, University of Canterbury.
- Hay, J. & Plag, I. (2004). “What Constraints Possible Suffix Combinations? On the Interaction of Grammatical and Processing Restrictions in Derivational Morphology”, *Natural Language and Linguistic Theory* 22, 565-596.
- Kiparsky, P. (1982a). “From Cyclic Phonology to Lexical Phonology”, in van der Hulst and Smith (1982a)
- Kiparsky, P. (1982b). *Lexical Phonology & Morphology*, Seoul: Linguistics in the Morning Calm.
- Mohanan, K.P. (1986). *The Theory of Lexical Phonology*, Dordrecht: Reidel.

- Plag, I. (1996). "Selectional Restriction in English Suffixation revisited: A Reply to Fabb (1988)", *Linguistics*, 34, 769-798.
- Plag, I. (1999). *Morphological Productivity: Structural Constraints in English Derivation*, Berlin: Mouton de Gruyter.
- Plag, I. (2003). *Word-formation in English*. Cambridge: Cambridge University Press.
- Plag, I. & Baayen, R. H. (2001). *Parsing and Productivity*, University of Canterbury, Interfaculty Research Unit for Language and Speech, University of Nijmegen
- Plag, I. & Baayen, R. H. (2009). "Suffix Ordering and Morphological Processing", *Language*, 85(1), 109-152.
- Selkirk, E. (1982). *The Syntax of Words*, Cambridge: MIT Press.
- Siegel, D. (1974). *Topics in English Morphology*, Cambridge: MIT Press.