

## توزیع واج‌های خیشومی در خوشه‌ دو همخوانی و ویژگی‌های توزیعی - واج‌شناختی وام‌واژه‌ها در زبان فارسی

۱- عباس صفر دوست\*؛ ۲- عالیه کرد زعفرانلو کامبوزیا\*\*

۱- دانشجوی دکترای زبان‌شناسی دانشگاه تربیت مدرس، تهران، ایران

۲- دانشیار زبان‌شناسی همگانی دانشگاه تربیت مدرس، تهران، ایران

(تاریخ دریافت: ۱۳۹۷/۰۷/۲۸؛ تاریخ پذیرش: ۱۳۹۷/۱۰/۲۵)

### چکیده

بخش نخست این مقاله به بررسی توزیع خیشومی‌ها در خوشه‌های دوهمخوانی پرداخته است. در این بخش قصد داریم الگوهای توزیعی واج‌های خیشومی را در خوشه‌های همخوانی بشناسیم. فرض ما این است که خیشومی‌ها در واژه‌هایی که به لحاظ ریشه شناختی فارسی هستند، در دو موضع C1 و C2 گرایش قوی به توزیع تکمیلی دارند؛ به این معنی که حضور واج m و n به ترتیب، در دو جایگاه C1 و C2 از خوشه همخوانی با محدودیت همراه است. از آنجا که چنین هدفی مستلزم تمایز میان واژه‌های فارسی (به لحاظ ریشه‌شناختی) و وام‌واژه‌هاست، بخش دوم این مقاله به بررسی توزیع مختصه‌های واج‌شناختی در واژه‌های (به لحاظ ریشه‌شناختی فارسی) و وام‌واژه‌ها پرداخته است. فرض ما در این بخش این بود که توزیع مختصه‌های واجی در وام‌واژه‌ها، نسبت به واژه‌های فارسی، از پراکندگی بیشتر و توزیع متوازن‌تری برخوردار است. با این حال، اگر وام‌واژه‌ها به طبقه ساختارهای مشخصی از زبان مبداء تعلق داشته باشند، این موضوع موجب تجمع مختصه‌های واجی در وام‌واژه‌ها و در نتیجه، پراکندگی کمتر و عدم توازن در پراکندگی‌شان خواهد شد. در ادامه مبادی مختلف ورود وام‌واژه‌ها و تصادفی بودن فرایند وام‌گیری را دلیل چنین ویژگی توزیعی‌ای عنوان کردیم. در انتها دلایلی را برای پربسامد بودن خوشه nC در زبان فارسی مطرح کردیم که تبیین ارائه شده بر اهمیت مختصه انسدادی در توصیف واج‌های خیشومی تأکید دارد. داده‌های این پژوهش از فهرست واژه‌های زایا که حاوی حدود ۵۵۰۰۰ واژه است، به دست آمده است.

**کلیدواژه‌ها:** واج خیشومی، خوشه همخوانی، واج‌آرایی، واج‌آماری، وام‌واژه.

\*E-mail: a.safardoost@hotmail.com (نویسنده مسئول)

\*\*E-mail: akord@modares.ac.ir

## ۱. مقدمه

توزیع واج‌های زبان در جایگاه‌های مختلف تابع محدودیت‌هایی است که این محدودیت‌ها، آرایش‌های متفاوتی از توالی واج‌ها را در واژه‌های زبان به همراه دارد؛ آرایشی که از آن تحت عنوان واج‌آرایی<sup>۱</sup> نام برده می‌شود. واج‌آرایی، در معنایی دیگر، به حوزه‌ای مطالعاتی در واج‌شناسی اطلاق می‌شود که به توصیف و تبیین محدودیت‌های آرایش واجگان زبان، در توالی‌های واجی می‌پردازد. روش عمده در مطالعه واج‌آرایی زبان، یافتن الگوهای آماری از توزیع واج‌ها و مختصه‌های واجی در مواضع مختلف است که از آن با عنوان واج‌آماري<sup>۲</sup> یاد می‌شود. در این پژوهش قصد داریم تا با استفاده از این روش، به مطالعه توزیع واج‌های خیشومی در خوشه‌های دو همخوانی زبان فارسی بپردازیم.

انگیزه اصلی انجام این پژوهش نتایجی است که از پژوهش علم‌الهدی (Alamolhoda, 2000: 233-238) بر روی واج‌آرایی زبان فارسی به دست آمده است. در این پژوهش، علم‌الهدی با تمرکز بر خوشه‌های همخوانی، محدودیت‌های موجود در این خوشه‌ها را با تفکیک واژه‌هایی که به لحاظ ریشه‌شناختی فارسی<sup>۳</sup> هستند و وام‌واژه‌ها، مورد مطالعه قرار داده است. در نتایج به دست آمده از این پژوهش، شاهد توزیع نامتوازن برخی واج‌ها در خوشه‌های همخوانی هستیم. با این حال، یکی از این توزیع‌های نامتوازی که نظر ما را جلب کرد، توزیع نامتوازن واج‌های خیشومی در خوشه همخوانی بود. از آنجا که واج‌های خیشومی به طبقه‌ای طبیعی تعلق دارند و تمایز مشاهده شده در این طبقه در دیگر طبقات طبیعی واجی مشاهده نشده است،

1. phonotactics

2. phonostatistics

۳. در ادامه هر جا واژه‌ای با صفت فارسی توصیف شده است، منظور ما از نظر ریشه‌شناختی فارسی است که قید «از نظر ریشه‌شناختی» به جهت اختصار حذف شده است.

موضوع این پژوهش را به بررسی توزیع دو واج خیشومی  $m$  و  $n$  در خوشه‌های دو همخوانی و علل چنین توزیعی اختصاص داده‌ایم. بر این اساس پرسش نخست ما در این پژوهش این است که چه الگویی در توزیع واج‌های خیشومی در خوشه‌های دوهمخوانی وجود دارد؟ فرضیه ما این است که واج‌های خیشومی، در خوشه‌های همخوانی واژه‌های فارسی، گرایشی قوی به توزیع تکمیلی از خود نشان می‌دهند؛ به این معنی که در واژه‌های فارسی حضور واج  $m$  در جایگاه  $C_1$  و حضور واج  $n$  در جایگاه  $C_2$  دارای محدودیت است.<sup>۱</sup>

همانطور که از پرسش و فرضیه نخست این پژوهش مشهود است، در سنجش این فرضیه میان واژه‌های فارسی و وام‌واژه‌ها تمایز قائل خواهیم شد. از آنجا که در فرضیه نخست، واژه‌هایی مورد مطالعه قرار خواهند گرفت که حاوی واج خیشومی در خوشه همخوانی‌شان هستند و از طرفی خیشومی‌ها طبقه‌ای طبیعی از واج‌ها را تشکیل می‌دهند، تمایز میان واژه‌های فارسی و وام‌واژه‌ها فرصتی را برای ما مهیا می‌کند تا ویژگی‌های توزیعی - واجشناختی وام‌واژه‌ها را در مقیاسی محدود و با در اختیار داشتن نمونه آماری مناسب، مورد مطالعه قرار دهیم. بنابراین، بخش دوم این پژوهش به مطالعه ویژگی‌های توزیعی - واجشناختی واژه‌های فارسی و وام‌واژه‌ها اختصاص دارد. پرسش اصلی ما در این بخش این خواهد بود که مختصه‌های واجی در وام‌واژه‌ها نسبت به واژه‌های فارسی، در سه سطح شیوه تولید، محل تولید، و وضعیت واکداری، از منظر توزیعی، دارای چه ویژگی‌هایی هستند؟ فرضیه ما این است که در وام‌واژه‌ها مختصه‌های واجی، از منظر توزیعی، دارای توزیع پراکنده‌تر و

۱. نشانه  $C$  در متن برای نشان همخوان به کار می‌رود و دو علامت  $C_1$ ,  $C_2$  به ترتیب برای نشان دادن همخوان اول و دوم در یک خوشه همخوانی ( $CC$ ) به کار می‌رود.

متوازن تری نسبت به مختصه‌های متناظر در واژه‌های فارسی هستند.<sup>۱</sup> با این حال، اگر در فرایند وام‌گیری، طبقه ساختواژی خاصی از زبان مبدا به‌طور گسترده‌ای وارد زبان مقصد شود، این موضوع موجب می‌شود که شاهد تجمع مختصه‌های واجی در وام‌واژه‌ها باشیم و تمایز واجشناختی - توزیعی میان وام‌واژه‌ها و واژه‌های فارسی ضعیف شود. در انتهای این بخش، به تبیین علل توزیع پراکنده‌تر و متوازن‌تر مختصه‌های واجی در وام‌واژه‌ها خواهیم پرداخت.

ترکیب واج‌های خیشومی با واجگان زبان فارسی در خوشه همخوانی چهار توالی مختلف را به همراه خواهد داشت: *mC*، *Cm*، *nC* و *Cn*. نتایج علم‌الهدی (Alamolhoda, 2000) نشان می‌دهد که توالی *nC* یکی از پربسامدترین توالی‌های زبان فارسی است. ما بخشی از این پژوهش را نیز به تبیین علل پربسامد بودن این توالی در خوشه‌های همخوانی اختصاص داده‌ایم.

## ۲. پیشینه پژوهش

مطالعه واج‌آرایی هر زبانی مستلزم این است که واحد معینی را به عنوان واحد مطالعه مشخص کنیم. پایک (Pike, 1947) پنج واحد زبانی را برای مطالعه واج‌آرایی زبان پیشنهاد می‌کند: پاره‌گفتار<sup>۲</sup>، واژه، تکواژ، هجا و مختصه‌های غیرزنجیری<sup>۳</sup>. از میان پنج واحد ذکرشده، غالباً هجا را می‌توان مناسب‌ترین گزینه در مطالعه واج‌آرایی زبان دانست، زیرا واحدهای دیگر زبانی دارای تنوع و تکرار بوده و از طرفی، از آنجا که هر زبانی از ساختار هجایی مشخصی پیروی

۱. ذکر این نکته ضروری است که دو مفهوم پراکنده‌تربودن و متوازن‌تربودن در این فرضیه مفاهیمی نسبی هستند که از مقایسه توزیع مختصه‌های واجی واژه‌های فارسی و وام‌واژه‌ها استنتاج می‌شوند.

2. utterance

3. nonsegmental characteristics

می‌کند، مطالعه واج‌آرایی در هجا می‌تواند زمینه مقابله میان زبانی را نیز فراهم کند. علاوه بر این، هجا به عنوان نخستین واحد زبانی که در فرایند فراگیری زبان، پیش از واج ظاهر می‌شود (Jusczyk, et al., 1993, Liberman, et al., 1974)، و نقش آن در وزن شعر در برخی زبان‌ها، بر اهمیت مطالعه واج‌آرایی این واحد زبانی می‌افزاید.

از بین این پنج واحد زبانی، پژوهش‌هایی که در زبان فارسی انجام گرفته‌اند، واج‌آرایی زبان فارسی را در سه سطح واژه (Kramsky, 1939, 1948؛ کردزعفرانلو کامبوزیا و همکاران، ۱۳۹۵)، تکواژ (کردزعفرانلو کامبوزیا و خیرآبادی، ۱۳۹۱)، و هجا (ثمره، ۱۳۷۸؛ اسلامی و بی‌جن‌خان، ۱۳۸۴؛ اسلامی، ۱۳۸۷؛ Alamolhoda, 2000؛ Zolfaghari, 2004؛ و اسلامی و همکاران، ۱۳۹۲) مورد بررسی قرار داده‌اند. در این میان، برخی از این آثار هجابندی واژه‌ها را نیز مورد مطالعه قرار داده‌اند (اسلامی و بی‌جن‌خان، ۱۳۸۴؛ اسلامی، ۱۳۸۷). در برخی از این آثار هم به‌طور خاص خوشه‌های همخوانی مورد توجه قرار گرفته است (Alamolhoda, 2000؛ Zolfaghari, 2004؛ اسلامی و همکاران، ۱۳۹۲).

### ۳. روش تحقیق و ملاحظات نظری

داده‌های پژوهش حاضر از فهرست واژه‌های زبانی فارسی (اسلامی و همکاران، ۱۳۸۳) استخراج شده است. این فهرست حاوی حدود ۵۵ هزار واژه لماتایز شده<sup>۱</sup> است که گردآورندگان آن ۴۴ هزار واژه آن را از پیکره‌ای ۱۰ میلیون لغتی استخراج کرده‌اند و سپس، به دلیل عدم حضور برخی واژه‌های عامیانه و تخصصی، ۱۱ هزار

---

1. lemmatized

واژه را نیز از لغتنامه فرهنگ فارسی امروز که حاوی واژه‌های فارسی معاصر است، به آن افزوده‌اند. هرچند افزودن ۱۱ هزار واژه از لغتنامه ماهیت این فهرست را به‌عنوان فهرستی پیکره‌بنیاد تغییر خواهد داد، اما انتخاب این فهرست به این دلیل بوده‌است که تمامی واژه‌های این فهرست دارای واج‌نویسی هستند. از این فهرست واژه‌هایی که به مقوله اسم خاص (شخص و مکان) و همین‌طور به طبقه عبارات<sup>۱</sup> تعلق داشتند، حذف شدند. معیار این حذف، برچسبی بود که در این فهرست به واژه‌ها اختصاص داده شده‌است.<sup>۲</sup> همچنین واژه‌های ابداعی فرهنگستان زبان نیز که کاربرد ندارند، به دلیل عدم طبیعی‌بودنشان، از فهرست حذف شد. متأسفانه اطلاعاتی از ترکیب گونه‌های زبانی پیکره‌ای که این فهرست از آن استخراج شده ارائه نشده است و به نظر می‌رسد عدم ارائه چنین اطلاعاتی در مقالاتی که به معرفی پیکره‌ها می‌پردازند، به یک سنت تبدیل شده است. معرفی سهم گونه‌های مختلف زبانی در پیکره از این نظر حائز اهمیت است که براساس آن بتوان در مورد اینکه پیکره تا چه اندازه زبان را نمایندگی می‌کند، قضاوت کرد. از طرفی، در پژوهش‌هایی که با استفاده از داده‌های پیکره صورت می‌گیرد، اطلاع از سهم گونه‌های زبانی پیکره کمک خواهد کرد تا اگر نتایجی غیرعادی در پژوهش مشاهده شد، درمورد اینکه آیا این نتایج از گونه زبانی خاصی ناشی شده است یا نه، قضاوت کرد.

غالباً آثاری که در آنها میان واژه‌های فارسی و وام‌واژه‌ها تمایز قائل می‌شوند، ماهیتی تجویزگرایانه دارند. چنین موضوعی به‌این دلیل است که در رویکردی تجویزگرایانه، فارسیت ویژگی‌ای ذاتی است که عناصر زبانی از آن برخوردارند، ویژگی‌ای که در زبان‌شناسی مدرن نسبی دانسته می‌شود. از آنجا که در این پژوهش

1. expression

۲. فهرست اصلاح‌شده مورد نظر در آدرس زیر در دسترس است:

[https://github.com/asdoost/Enhanced\\_Flexicon](https://github.com/asdoost/Enhanced_Flexicon)

رویکردی توصیف‌گرایانه در مطالعه واج‌آرایی واژه‌ها خواهیم داشت، لازم است برای رفع هرگونه سوءبرداشتی موضع خود را در مورد مفهوم فارسیت مشخص کنیم. از نظر همزمانی، هر واژه‌ای که برای دلالت بر چیزی بین سخنگویان زبان فارسی به کار برود، فارسی خواهد بود. از این منظر، وام‌واژه‌هایی هم که در گفتار سخنگویان زبان فارسی امروز به کار می‌روند، فارسی محسوب می‌شوند و نمی‌توان تمایزی میان واژه‌های فارسی و وام‌واژه‌ها قائل شد. در مقابل و از منظر در زمانی، فارسیت یک واژه وابسته به کاربرد آن نزد سخنگویان پیشین است. در این پژوهش، منظور ما از فارسی‌بودن یک واژه، فارسیت آن از نظر ریشه‌شناختی است. بنابراین، در ادامه هرجا واژه یا هجایی با صفت «فارسی» توصیف شده است، منظور ما «به‌لحاظ ریشه‌شناختی فارسی» است که قید «به‌لحاظ ریشه‌شناختی» به‌جهت اختصار حذف شده است.

داده‌های این پژوهش، چارچوبی همزمانی را برای این پژوهش فراهم می‌کنند، این درحالی است که در این پژوهش، میان واژه‌های به‌لحاظ ریشه‌شناختی فارسی و وام‌واژه‌ها تمایز قائل شده‌ایم و فرضیه نخست این پژوهش به توصیف و تبیین واژه‌های به‌لحاظ ریشه‌شناختی فارسی می‌پردازد. بنابراین، نتایج این پژوهش تنها توصیف‌گر آن دسته از واژه‌های به‌لحاظ ریشه‌شناختی فارسی‌ای است که در زمان حال کاربرد دارند و فرضیات این پژوهش نیز تنها در مورد این دسته از واژه‌ها صادق بوده و این امکان وجود دارد که در مطالعه‌ای با چارچوب در زمانی، داده‌هایی را علیه فرضیات این پژوهش یافت. هرچند تفاوت‌های چشمگیری که در داده‌ها مشاهده خواهد شد، احتمال وجود چنین داده‌هایی را ضعیف می‌کند. ذکر این نکات به این جهت بود که گستره ادعاهای مطرح‌شده در این پژوهش برای خواننده روشن شود.

در این پژوهش توزیع واج‌های خوشه همخوانی را در چهار سطح بررسی خواهیم کرد: هجا، هجای منحصربه‌فرد، قافیه<sup>۱</sup> منحصربه‌فرد، و خوشه منحصربه‌فرد. منظور از هجای منحصربه‌فرد، توالی‌ای منحصربه‌فرد از چهار واج است که یک توالی CVCC را تشکیل می‌دهد؛ منظور از قافیه منحصربه‌فرد، توالی منحصربه‌فردی از سه واج است که توالی VCC را در هجای CVCC تشکیل می‌دهد و منظور از خوشه منحصربه‌فرد توالی منحصربه‌فردی از دو همخوان است که توالی CC را در هجای CVCC تشکیل می‌دهد. در ادامه، برای اختصار به جای «منحصربه‌فرد» از واژه «منحصر» استفاده خواهد شد.

در این پژوهش از روش‌های آمار توصیفی برای سنجش فرضیات بهره خواهیم برد. در فرضیات بخش دوم که به توزیع پراکنده و متوازن مختصه‌های واجی وام‌واژه‌ها در مقایسه با واژه‌های فارسی اختصاص دارند، توزیع مختصه‌های واجی در سه سطح شیوه تولید، جایگاه تولید و وضعیت واک مورد بررسی قرار خواهند گرفت. برای محاسبه پراکندگی و توازن از مفهوم دامنه<sup>۲</sup> می‌توان بهره برد. دامنه از تفاضل پربسامدترین و کم‌بسامدترین عنصر به دست می‌آید. بر این اساس، سنجش توازن مختصه‌های واجی از مقایسه تفاضل پربسامدترین و کم‌بسامدترین مختصه واجی اشغال شده در وام‌واژه‌ها و واژه‌های فارسی به دست می‌آید. قید «اشغال شده» به این معنی است که تنها بسامد مختصه‌هایی مورد محاسبه قرار خواهد گرفت که در واژه‌ها حضور داشته باشد. بنابراین، اگر مختصه واجی‌ای در داده‌ها حضور نداشته باشد، یعنی صفر باشد، در محاسبات به عنوان کم‌بسامدترین لحاظ نخواهد شد، مگر اینکه تنها یک ویژگی واجشناختی وجود داشته باشد.

---

1. rhyme/rime

2. range



برای محاسبه پراکندگی از دو مفهوم گستره پراکندگی و میزان پراکندگی بهره خواهیم برد. به این ترتیب، اگر مختصه‌های شیوه تولید و محل تولید را به ترتیب، براساس میزان بست و محل بست به شکل زیر ارزش عددی یک تا ۷ (برای شیوه تولید) و یک تا ۸ (برای محل تولید) بدهیم، گستره پراکندگی حاصل تفاضل دو منتهی‌الیه در پیوستار شیوه تولید و محل تولید خواهد بود:

انسدادی > انسایشی > سایشی > خیشومی > لرزشی > کناری > غلت  
دولبی > لبی‌دندانی > دندانی > لثوی > لثوی کامی > کامی > ملازی > چاکنایی

مفهوم میزان پراکندگی به تعداد مختصه‌های واجی اشاره دارد که یک جایگاه از آن برخوردار است. بنابراین، عدد بزرگتر در گستره پراکندگی و میزان پراکندگی به معنی پراکندگی بیشتر در آن جایگاه خواهد بود. از آنجا که مختصه واکداری دو ارزشی است، هیچ‌یک از دو مفهوم گستره پراکندگی و میزان پراکندگی را نمی‌توان برای سنجش پراکندگی در این مختصه به کار برد.

از آنجا که طبقات مختلف مختصه‌های واجی دارای تعداد اعضای متفاوتی هستند (انسدادی‌ها که ۶ عضو دارند در برابر لرزشی‌ها که تنها ۱ عضو دارند)، بیشتر بودن بسامد یک مختصه واجی در یک موضع می‌تواند به دلیل بیشتر بودن تعداد اعضای آن طبقه باشد. برای غلبه بر چنین مشکلی بسامد هر طبقه از ویژگی‌های واجی به صورت نسبی محاسبه شده که حاصل تقسیم بسامد بر تعداد اعضای طبقه و سپس، محاسبه درصد آن است. ما از عدد حاصل از چنین فرایندی با عنوان درصد نسبی نام خواهیم برد.

#### ۴. ارائه و تحلیل داده‌ها

مباحث این قسمت از چهار بخش تشکیل شده است. ابتدا در بخش ۴-۱، توزیع عضو خیشومی خوشه همخوانی را مورد بررسی قرار می‌دهیم و سپس، در بخش ۴-۲، به بررسی عضو غیرخیشومی در خوشه همخوانی می‌پردازیم. ذکر این نکته ضروری است که اصطلاح عضو خیشومی و غیرخیشومی ناظر بر نقطه تمرکز مبحث است. خوشه‌های دارای واج خیشومی چهار ترکیب  $mC$ ،  $Cm$ ،  $nC$  و  $Cn$  را تشکیل می‌دهند. زمانی که در مورد عضو خیشومی صحبت می‌کنیم، منظور ما عنصر ثابت این چهار ترکیب، یعنی  $m$  و  $n$  است و زمانی که در مورد عضو غیرخیشومی صحبت می‌کنیم، منظور ما عنصر متغیر این چهار ترکیب، یعنی  $C$  (همخوان) است، هرچند در این جایگاه نیز واج خیشومی ممکن است ظاهر شود. در بخش ۴-۲-۴، به علل پراکندگی و توزیع متوازن مختصه‌های واجی در خوشه همخوانی وام‌واژه‌ها می‌پردازیم. و در نهایت، در بخش ۴-۳-۳ فهرست بسامد خوشه‌ها را مطالعه می‌کنیم و در بخش ۴-۳-۱ به تبیین علل پربسامد بودن خوشه  $nC$  در زبان فارسی خواهیم پرداخت.

#### ۴-۱. توزیع عضو خیشومی خوشه دوهمخوانی

بررسی فهرست واژه‌های زایا نشان می‌دهد که در زبان فارسی ۲۳۲۹ واژه وجود دارد که دارای هجای CVCC بوده و یکی از دو واج خیشومی  $m$  و  $n$  در خوشه همخوانی آنها حضور دارد. از این ۲۳۲۹ واژه، ۲۳۴۹ هجای CVCC حاوی واج خیشومی در خوشه همخوانی، به دست می‌آید. علت تمایز تعداد واژه و تعداد هجا به دلیل وجود ۲۰ واژه است که در آنها دو هجای CVCC حاوی واج خیشومی حضور دارد

(چشم‌تنگ، کندذهن، عمق‌سنج). ۲۳۴۹ هجای به‌دست‌آمده از ۳۹۰ هجای منحصراً، ۱۲۵ قافیه منحصراً، و ۵۴ خوشه منحصراً تشکیل شده‌اند.

از بین ۲۳۲۹ واژه مستخرج‌شده، ۱۴۶۳ واژه فارسی (۶۲٫۸۲ درصد)؛ ۸۶۲ واژه (۳۷٫۰۱ درصد) وام‌واژه و ۴ واژه (۰٫۱۷ درصد) ترکیبی از یک واژه فارسی و یک وام‌واژه است.<sup>۱</sup> از هریک از واژه‌های فارسی و وام‌واژه‌ها، به ترتیب، ۱۴۷۶ و ۸۷۳ هجای CVCC به دست آمد. تعداد هجای منحصراً، قافیه منحصراً، و خوشه منحصراً برای واژه‌های فارسی، ۱۲۲، ۳۰، و ۱۶، و برای وام‌واژه‌ها، به ترتیب، ۲۸۸، ۱۱۴ و ۵۲ است. مقایسه سرانگشتی این اعداد نشان می‌دهد که آرایش واجی هجای CVCC وام‌واژه‌ها، در سطوح مختلف نسبت به واژه‌های فارسی از تنوع بیشتری برخوردار است. خلاصه‌ای از آمار واج‌های خیشومی بر اساس نوع واج و جایگاهشان در خوشه همخوانی در واژه‌های فارسی و وام‌واژه‌ها را می‌توانید در جدول و نمودار ۱ مشاهده کنید.

مطالعه جدول و نمودار ۱ نشان می‌دهد تمایز عمده‌ای در توزیع دو واج خیشومی و جایگاه آنها در خوشه همخوانی وجود دارد: درحالی که دامنه تغییرات (تفاضل) میان دو واج خیشومی و میان دو جایگاه خوشه همخوانی در کل هجاهای واژه‌های فارسی، به ترتیب ۷۵٫۸۸ و ۷۳٫۷۲ درصد است، این رقم در وام‌واژه‌ها، به ترتیب،

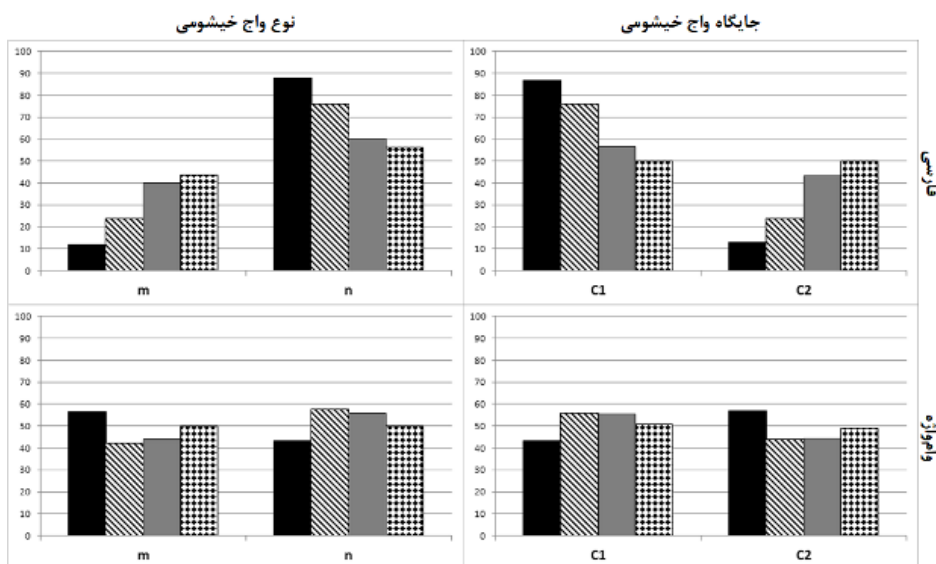
۱. ما در تفکیک واژه‌های فارسی و وام‌واژه‌ها اگر یک واژه مرکب تنها حاوی یک هجای CVCC بود و هجای CVCC در عضو فارسی واژه مرکب حضور داشت، آن را به‌عنوان فارسی و اگر در عضو وام‌واژه واژه مرکب حضور داشت، آن را به‌عنوان وام‌واژه در نظر گرفته‌ایم، برای مثال واژه مرکب زودفهم هرچند از ترکیب واژه فارسی زود و وام‌واژه فهم تشکیل شده است، به‌اعتبار حضور هجای CVCC در وام‌واژه (فهم)، کل ترکیب در دسته وام‌واژه‌ها جای گرفته است. ترکیبی بودن در اینجا به این معنی است که در یک واژه مرکب که از یک واژه فارسی و یک وام‌واژه تشکیل شده است، هریک از این دو واژه دارای یک هجای CVCC حاوی واج خیشومی در خوشه همخوانی هستند. برای مثال از واژه عمق‌سنج دو هجای CVCC به دست می‌آید که هجای *omG* به وام‌واژه عمق و هجای *sændʒ* به واژه فارسی سنج تعلق دارد.

۱۳،۳۸ و ۱۳،۶ درصد است. چنین تمایز چشمگیری نشان می‌دهد که در وام‌واژه‌ها، علی‌رغم تنوع هجایی بیشتری که دارند، این تنوع از نظر نوع واج خیشومی و جایگاه آن، نسبت به واژه‌های فارسی، از توزیع متوازن‌تری برخوردار است که این توزیع متوازن در نوع واج خیشومی و جایگاه واج خیشومی (همانطور که در بخش‌های بعدی نشان خواهیم داد) می‌تواند به توزیع متوازن‌تر مختصه‌های واجی نسبت به واژه‌های فارسی، منجر شود. علاوه بر این، در این نمودار با حرکت به سمت انتهای هجا، چه در واژه‌های فارسی و چه در وام‌واژه‌ها، شاهد متعادل شدن مشارکت واج‌های خیشومی (چه از نظر نوع واج و چه از نظر جایگاه آنها در خوشه همخوانی) هستیم که چنین تعادلی بر طبیعی بودن طبقه واجی خیشومی‌ها در زبان فارسی دلالت دارد.

برای یافتن منشاء توزیع نامتوازن در نوع واج خیشومی و جایگاه آن در واژه‌های فارسی، میزان مشارکت هریک از واج‌های خیشومی را در ترکیبات چهارگانه خوشه همخوانی، به تفکیک نوع واژه (فارسی/ وام‌واژه)، در جدول و نمودار ۲ به تصویر کشیده‌ایم. جدول و نمودار ۲ همچنین به ما کمک می‌کند تا ببینیم آیا توزیع واج‌های خیشومی در ترکیبات چهارگانه خوشه همخوانی وام‌واژه‌ها نیز دارای توزیع متوازن‌تری نسبت به واژه‌های فارسی هستند.

جدول (۱): بسامد واج‌های خیشومی به تفکیک نوع واج و جایگاه آن در خوشه همخوانی در واژه‌های فارسی و وام‌واژه‌ها

خوشه منحصر		قافیه منحصر		هجای منحصر		کل هجا		جایگاه /	نوع واژه
درصد	تعداد	درصد	تعداد	درصد	تعداد	درصد	تعداد	نوع واج	
۵۰	۸	۵۶٫۶۷	۱۷	۷۶٫۲۳	۹۳	۸۶٫۸۶	۱۲۸۲	C <sub>1</sub>	فارسی
۵۰	۸	۴۳٫۳۳	۱۳	۳۳٫۷۷	۲۹	۱۳٫۱۴	۱۹۴	C <sub>2</sub>	
۴۳٫۷۵	۷	۴۰	۱۲	۳۳٫۷۷	۲۹	۱۲٫۰۶	۱۷۸	m	فارسی
۵۶٫۲۵	۹	۶۰	۱۸	۷۶٫۲۳	۹۳	۸۷٫۹۴	۱۲۹۸	n	
۵۰٫۹۴	۳۷	۵۵٫۵۶	۶۵	۵۶٫۰۱	۱۶۳	۴۳٫۲۰	۳۸۱	C <sub>1</sub>	وام‌واژه
۴۹٫۰۶	۲۶	۴۴٫۴۴	۵۲	۴۳٫۹۹	۱۲۸	۵۶٫۸۰	۵۰۱	C <sub>2</sub>	
۵۰	۲۷	۴۴٫۰۷	۵۲	۴۲٫۱۲	۱۲۳	۵۶٫۶۹	۵۰۰	m	وام‌واژه
۵۰	۲۷	۵۵٫۹۳	۶۶	۵۷٫۸۸	۱۶۹	۴۳٫۳۱	۳۸۲	n	



ستون‌ها از چپ به راست، به ترتیب: ستون اول کل هجا، ستون دوم هجای منحصراً، ستون سوم قافیه منحصراً، و ستون چهارم خوشه منحصراً

نمودار (۱): بسامد واچ‌های خیشومی به تفکیک نوع واچ خیشومی و جایگاه آن در واژه‌های فارسی و وام‌واژه‌ها

داده‌هایی که در جدول ۲ آمده منجر به کشف الگویی در توزیع واچ‌های خیشومی در خوشه همخوانی واژه‌های فارسی می‌شود. این الگو که در قالب نمودار ۲ می‌توان آن را به راحتی تشخیص داد، حاکی از حضور چشمگیر واچ m در جایگاه C<sub>2</sub> و حضور بسیار ضعیف این واچ در جایگاه C<sub>1</sub> است. در مقابل، در توزیع واچ n، عکس چنین رابطه‌ای مشاهده می‌شود، یعنی واچ n در جایگاه C<sub>1</sub> حضور چشمگیر داشته و در جایگاه C<sub>2</sub> شاهد حضور بسیار ضعیف این واچ هستیم. این درحالیست که در وام‌واژه‌ها، هرچند چنین الگویی مشاهده می‌شود، با این حال توزیع واچ‌های خیشومی، نسبت به واژه‌های فارسی، دارای توزیع متوازن تری است. توزیع دو واچ خیشومی m و n در دو جایگاه C<sub>1</sub> و C<sub>2</sub> واژه‌های فارسی به قدری متمایز است که

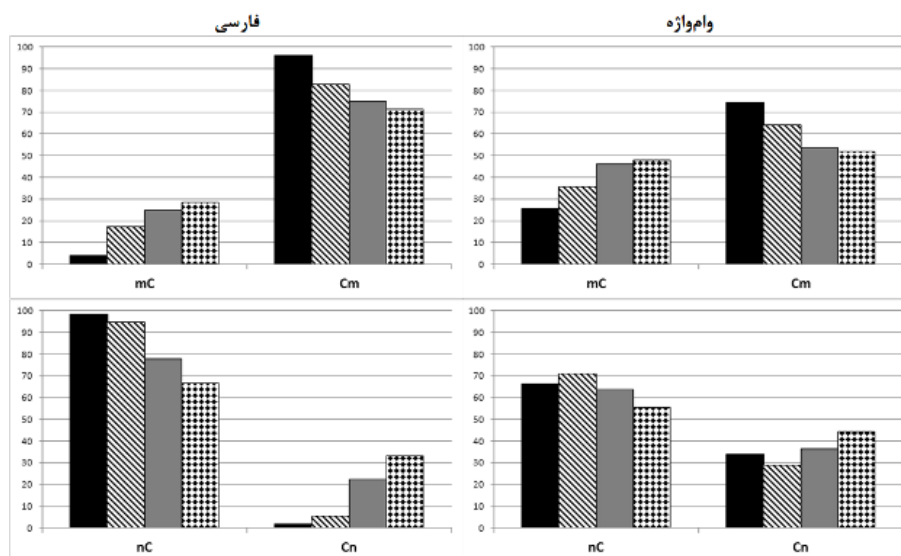
می‌توان از مفهوم توزیع تکمیلی در توصیف این رابطه استفاده کرد. رابطه توزیع تکمیلی در چنین موقعیتی به این معنی است که در خوشه همخوانی واژه‌های فارسی، هر یک از دو واج خیشومی در موقعیتی ظاهر می‌شود که حضور واج دیگر در آن موقعیت دارای محدودیت است. براین اساس واج  $m$  نمی‌تواند در جایگاهی که  $n$  حضور دارد ( $C_1$ ) ظاهر شود و واج  $n$  نمی‌تواند در جایگاهی که واج  $m$  حضور دارد ( $C_2$ ) حاضر شود. باین حال، رابطه توزیع تکمیلی‌ای که در توزیع واج‌های خیشومی در خوشه همخوانی واژه‌های فارسی مشاهده می‌شود، رابطه‌ای مطلق نبوده و گرایشی قوی در واژه‌های اصالتاً فارسی است که تعداد معدودی از داده‌ها شکل مطلق چنین رابطه‌ای را نقض می‌کنند.

همانطور که در جدول ۲ مشاهده می‌کنید، واج  $m$  در جایگاه  $C_1$  خوشه همخوانی ۷ هجا ظاهر شده است که این ۷ هجا از ۷ واژه مستخرج شده‌اند. ۷ واژه فارسی‌ای که در آنها واج  $m$  در جایگاه  $C_1$  ظاهر شده است، عبارتند از: رُمب، گرمب، گرمب، دمب، تلمب، دالامب و دولومب، و شمش. از بین این ۷ واژه دو واژه گرمب و گرمب در اصل دو گونه نوشتاری متفاوت از یک واژه هستند، بنابراین در اینجا، ما در اصل با ۶ واژه مواجه هستیم. از ۷ هجای دارای واج  $m$  در جایگاه  $C_1$ ، ۵ هجای منحصر، ۳ قافیه منحصر، و ۲ خوشه منحصر به دست می‌آید. درمقابل واج  $n$  در جایگاه  $C_2$  خوشه همخوانی ۲۳ واژه ظاهر شده است که ۲۳ واژه مذکور صورت‌های ساده، مرکب، و مشتق ۵ واژه جشن، پهن، گوزن، رشن و گشن هستند. از این ۲۳ واژه، ۲۳ هجا مستخرج شده است که حاصل تکرار ۵ هجای منحصر، ۴ قافیه منحصر، و ۳ خوشه منحصر است. به این ترتیب، تنها ۱۱ واژه وجود دارد (۶ واژه برای واج  $m$  و ۵ واژه برای واج  $n$ ) که شکل مطلق رابطه توزیع تکمیلی در توزیع واج‌های خیشومی را در خوشه واژه‌های فارسی نقض می‌کند.

جدول (۲): بسامد واج‌های خیشومی در خوشه همخوانی به‌تفکیک واژه‌های فارسی و وام‌واژه‌ها

نوع واژه	جایگاه واج	کل هجا		هجای متعصمر		قائیه متعصمر		خوشه متعصمر	
		تعداد	درصد	تعداد	درصد	تعداد	درصد	تعداد	درصد
فارسی	mC	۷	۳.۹۱	۵	۱۷.۳۴	۳	۲۵	۲	۲۸.۵۷
	Cm	۱۷۱	۹۶.۰۹	۲۴	۸۲.۷۶	۹	۷۵	۵	۷۱.۴۳
	mC	۱۲۸	۲۵.۶	۴۴	۳۵.۷۷	۲۴	۴۶.۱۵	۱۳	۴۸.۱۵
وام‌واژه	Cm	۲۷۲	۷۴.۴	۷۹	۶۴.۲۳	۲۸	۵۳.۸۵	۱۴	۵۱.۸۵
	nC	۱۲۷۵	۹۸.۲۳	۸۸	۹۴.۶۲	۱۴	۷۷.۷۸	۶	۶۶.۶۷
	Cn	۲۳	۱.۷۷	۵	۵.۳۸	۴	۲۲.۲۲	۳	۳۳.۳۳
وام‌واژه	nC	۲۵۳	۶۶.۳۳	۱۲۰	۷۱.۰۱	۴۲	۶۳.۶۴	۱۵	۵۵.۵۶
	Cn	۱۲۹	۳۳.۷۷	۴۹	۲۸.۹۹	۲۴	۳۶.۳۶	۱۲	۴۴.۴۴





ستون‌ها از چپ به راست، به‌ترتیب: ستون اول کل هجا، ستون دوم هجای منحصراً، ستون سوم قافیه منحصراً، و ستون چهارم خوشه منحصراً

نمودار (۲): بسامد واج‌های خیشومی در خوشهٔ همخوانی به تفکیک نوع واج و نوع واژه

بررسی همه‌جانبهٔ این ۱۱ واژه نشان می‌دهد که ماهیت این واژه‌ها چه از نظر ریشه‌شناختی و چه از نظر کاربردی نوع تلفظ (استاندارد و گویشی) و ویژگی‌های ساختارهای مورد تشکیل قرار است و می‌توان این فرضیه را مطرح کرد که تلفظ امروزی برخی از این واژه‌ها تحت‌تأثیر ورود وام‌واژه‌ها است.

از نظر تاریخی، واژه گوزن در اوستایی به‌صورت *ǰəvesnə* تلفظ می‌شود (Bartholomae, 2012: 510). این واژه با دو تلفظ *gæwæzn* (مکنزی، ۱۳۷۳: ۷۸) و *gævazæn* (فروه‌وشی، ۱۳۷۱؛ ۴۳۹) در زبان پهلوی ثبت شده است که نوع توزیع خیشومی‌ها در خوشهٔ همخوانی که در این پژوهش به تصویر کشیدیم، می‌تواند از تلفظ *gævazæn* حمایت کند. علی‌رغم اینکه در لغتنامه‌ها برای واژه

«شمش» اطلاعات ریشه‌شناختی ثبت نشده است، شواهد نشان می‌دهد که این واژه دارای ریشه سامی (:  $\text{š.m.š}$  به معنای درخشیدن) است و واژه شمس عربی نیز از همین ریشه مشتق شده است. ما در ادامه این واژه را به عنوان یک وام‌واژه مدنظر قرار خواهیم داد.

از بین ۶ واژه‌ای که در آنها واج  $m$  در جایگاه  $C_1$  ظاهر شده است (رمب، گرمب/گرمب، دمب، تلمب، دالمب و دولومب، و شمش)، دو واژه رَمب و غرمب/گرمب، واژه‌هایی هستند که تحت تأثیر قاعده همگونی محل تولید اینگونه تلفظ می‌شوند. به عبارت دیگر، این دو واژه در اصل رنب و غرنب/گرنب بوده‌اند که براساس قاعده همگونی محل تولید، مجاورت واج  $n$  با واج دولبی  $b$  موجب شده که محل تولید آوای  $n$  از جایگاه لثوی به جایگاه دولبی تغییر کرده و واج  $n$  به آوای  $m$  تبدیل شود.

باین حال، چنین تفسیری هرچند از نظر ریشه‌شناختی وجود واج  $m$  در این دو واژه را رد می‌کند، اما عدم حضور واج  $m$  در جایگاه  $C_1$  را رد نخواهد کرد و محدودیتی مطلق به همراه نخواهد داشت، زیرا قاعده همگونی محل تولید نشان می‌دهد که توالی واجی  $mb/p$  می‌تواند به‌طور بالقوه، یک توالی مجاز در زبان فارسی باشد، کما اینکه برخی وام‌واژه‌های زبان فارسی (بمب، پمپ، کمپ، و استمپ) از چنین مجوزی بهره برده و در زبان فارسی به کار می‌روند. باین حال، می‌توان سطحی انتزاعی‌تر از قواعد را در نظر گرفت که براساس آن، از نظر ریشه‌شناختی، توالی  $mb/p$  توالی غیرمجازی در زبان فارسی بوده و هیچ واژه‌ای با چنین توالی‌ای در زبان فارسی وجود نداشته است و واژه‌هایی که دارای چنین توالی‌ای هستند، صورت‌های تغییر یافته توالی  $nb/p$  هستند که براساس قاعده همگونی محل تولید به  $mb/p$  تغییر می‌یابند. به عبارت دیگر، قاعده‌ای واجی مانع از حضور مدخل با خوشه  $mb/p$

می‌شود، ولی در عین حال قاعده‌ای آواشناختی، بدون دخالت قاعده قبلی توالی‌های nb/p را به توالی mb/p تبدیل می‌کند. براین اساس، چنین قاعده‌ای هرچند مانع از حضور واژه‌هایی با توالی mb/p می‌شود، ولی در اعمال قواعد آواشناختی‌ای که nb/p را به mb/p تبدیل می‌کند، تأثیر نمی‌گذارد. باید توجه داشت که فرض چنین قاعده‌ای را تنها برای مقطعی تاریخی از تاریخ زبان فارسی می‌توان مفروض شمرد، زیرا حضور وام‌واژه‌های ناقض این رابطه، وجود چنین قاعده‌ای را نفی می‌کند.

توجیه آواشناختی‌ای که در مورد دو واژه رمب و گرمب/غرمب ارائه‌شد، به نظر می‌رسد در مورد واژه دمب، تلمب، و دالامب و دولومب نیز صادق است. واژه تلمب تنها در لغتنامه سخن و براساس حضور آن در ترکیب آهن و تلمب ضبط شده‌است و برای آن مدخلی مجزا تعیین شده است. این ترکیب در لغتنامه‌های دیگر، مانند لغتنامه دهخدا، به صورت آهن و تلپ ضبط شده است. نکته قابل توجه دیگری هم که در وام‌واژه‌ها می‌توان مشاهده کرد این است که تمام وام‌واژه‌هایی که به توالی mb/p ختم شده‌اند به انگلیسی/فرانسوی تعلق دارند، مانند بمب، پمپ، کمپ، و چنین توالی‌ای در وام‌واژه‌های عربی مشاهده نمی‌شود. از سوی دیگر، در واژه‌هایی که واج n در جایگاه C<sub>2</sub> خوشه همخوانی ظاهر می‌شود، در گفتار طبیعی شاهد حذف واج خیشومی هستیم که گاهی با کشش واکه‌ای همراه است، مانند جشن و پهن. این واج در صورت‌های مشتق و مرکب هم، زمانی که در مرز بین واژه و وند، و یا در مرز دو واژه قرار می‌گیرد، عموماً با حذف همراه است، مانند جشنواره و گشنگیری. اگر بپذیریم که واژه‌های ناقض رابطه توزیع تکمیلی تحت تأثیر وام‌واژه‌ها در گویش معیار این گونه تلفظ می‌شوند، تلفظ این واژه‌ها در گویش‌های محلی، که کمتر تحت تأثیر وام‌واژه‌ها هستند، می‌تواند به تلفظ تاریخی این واژه‌ها نزدیک‌تر باشد. نوع تلفظ این واژه‌ها در برخی از گویش‌ها نشان می‌دهد که برای ممانعت از

حضور این دو واج در دو جایگاه مذکور، در خوشه همخوانی، تلفظ این واژه‌ها متفاوت از تلفظ آن در گویش معیار است.<sup>۱</sup> چنین تمایزی یا با حذف یکی از واج‌ها و یا با افزودن واکه و شکسته شدن هجای سنگین صورت می‌پذیرد. برای نمونه، واژه پهن در گویش شیرازی با حذف واج h و کشش جبرانی واکه æ به صورت pæ:n تلفظ می‌شود. همین واژه در گویش اناری بدون دو واج h و n، و به صورت pæ: بدون واج n تلفظ می‌شود. در گویش سبزواری دو واژه پهن و جشن به صورت dzæ:fen و pæ:hen تلفظ می‌شوند.

از نظر ساختاری، از میان ۶ واژه حاوی خوشه mC، ۲ واژه گرمب/گرمب، و دالامب و دولومب، به مقوله نام‌آواها تعلق دارند. تفاوت نام‌آواها با دیگر مقولات زبانی در این است که در نام‌آواها رابطه میان دال و مدلول، رابطه‌ای تصویری<sup>۲</sup> است، در حالی که رابطه میان دال و مدلول در دیگر مقولات زبانی رابطه‌ای قراردادی است. رابطه قراردادی میان دال و مدلول موجب می‌شود که ترکیب واجی واژه‌ها از ساختار واج‌شناختی زبان پیروی کند، این در حالیست که در نام‌آواها، ترکیب واجی تقلیدی از آوایی است که واژه به آن دلالت دارد. البته این موضوع به این معنی نیست که نام‌آواها از ساختار واجی زبان پیروی نمی‌کنند، کما اینکه تفاوت نام‌آوایی که در زبان‌های مختلف به پدیده‌ای مشابه دلالت دارند، مؤید نوعی تبعیت از ساختار واجی زبان مورد نظر است و نام‌آواها علاوه بر تبعیت از ساختار واجی زبان، تا حد امکان متعهد به بازنمایی آوایی مشابه به پدیده مورد

۱. این بخش از شواهد براساس صحبت با گویشوران این گویش‌ها و گزارش آنها از نوع تلفظ این واژه‌ها گردآوری شده است.

2. iconic

نظر هستند. بنابراین ۲ واژه مورد بحث ممکن است به دلیل ماهیت تصویری شان ناقص رابطه توزیع تکمیلی در واژه‌های فارسی باشند.

علاوه بر این، برخی از واژه‌های فارسی‌ای که ذکر شد (رمب، گرمب/گرمب)، در واقع ریشه افعال رمبیدن، گرمبیدن/گرمبیدن هستند که تصریف شخص و شمار و همینطور تصریف گذشته، موجب شکسته شدن هجای CVCC در این افعال می‌شود و عملاً تضمین‌کننده عدم حضور توالی mb در این واژه‌ها است (rom.bi.dæm). علاوه بر این، مدلول برخی از این واژه‌ها می‌تواند بر محلی بودن آنها دلالت داشته باشد. برای نمونه از آنجا که زیستگاه حیوانی مثل گوزن به مناطق جغرافیایی خاصی محدود است، می‌توان واژه گوزن را دارای خاستگاهی محلی و یا متأثر از تلفظی محلی دانست. گ برای واژه جشن نیز، از آنجا که مفهوم جشن از ابتدای پیدایش آن به مراسم‌های مذهبی اطلاق می‌شده است، این واژه نیز ممکن است از آئین مشخصی با خاستگاهی محلی و تلفظی محلی گسترش یافته باشد.

همانطور که تا به اینجا در مورد واژه‌های فارسی مشاهده کردیم و در ادامه هم خواهیم دید، به دلیل محدودیت‌هایی که در توالی واج‌ها در هر زبانی وجود دارد، واژه‌های هر زبانی حول توالی‌های واجی مشخصی تجمع دارند، باین حال ۱۰ واژه‌ای که شکل مطلق رابطه توزیع تکمیلی را نقض می‌کنند (واژه شمش از این فهرست حذف شده است)، حاوی ۴ خوشه منحصر به فردی هستند که تنها در این ۱۰ واژه دیده می‌شوند (خوشه mb برای واج m و خوشه‌های fn (:جشن، گشن، رشن)، zn (:گوزن)، و hn (:پهن) برای واج n). همانطور که در ادامه خواهیم دید، از بین این ۴ خوشه، خوشه fn تنها در این سه واژه مشاهده شده و چنین خوشه‌ای در هیچ وام‌واژه‌ای حضور ندارد. یکی از فرضیاتی که می‌تواند پاسخگوی چنین توزیعی باشد، ماهیت ریشه‌شناختی یا تلفظ تاریخی متفاوت این واژه‌های است.

شواهدی که تا به اینجا ذکر شد می‌تواند ماهیت ۱۱ واژه‌ای را که ناقض شکل مطلق رابطه توزیع تکمیلی در واژه‌ها فارسی هستند، از جوانب مختلف مورد تشکیک قرار دهد. بدیهی است نوع توزیع واج‌های خیشومی این واژه‌ها که در این پژوهش ارائه شده است، می‌تواند در مطالعه ریشه‌شناختی این واژه‌ها، از فرضیه غیرفارسی‌بودن یا تلفظ تاریخی متفاوت این واژه‌ها حمایت کند. با این حال، می‌توان گفت که وجود وام‌واژه‌هایی که ناقض چنین رابطه‌ای هستند، چنین توالی‌ای را در خوشه همخوانی فارسی معاصر تثبیت کرده است.

#### ۲-۴. توزیع عضو غیرخیشومی خوشه دوهمخوانی

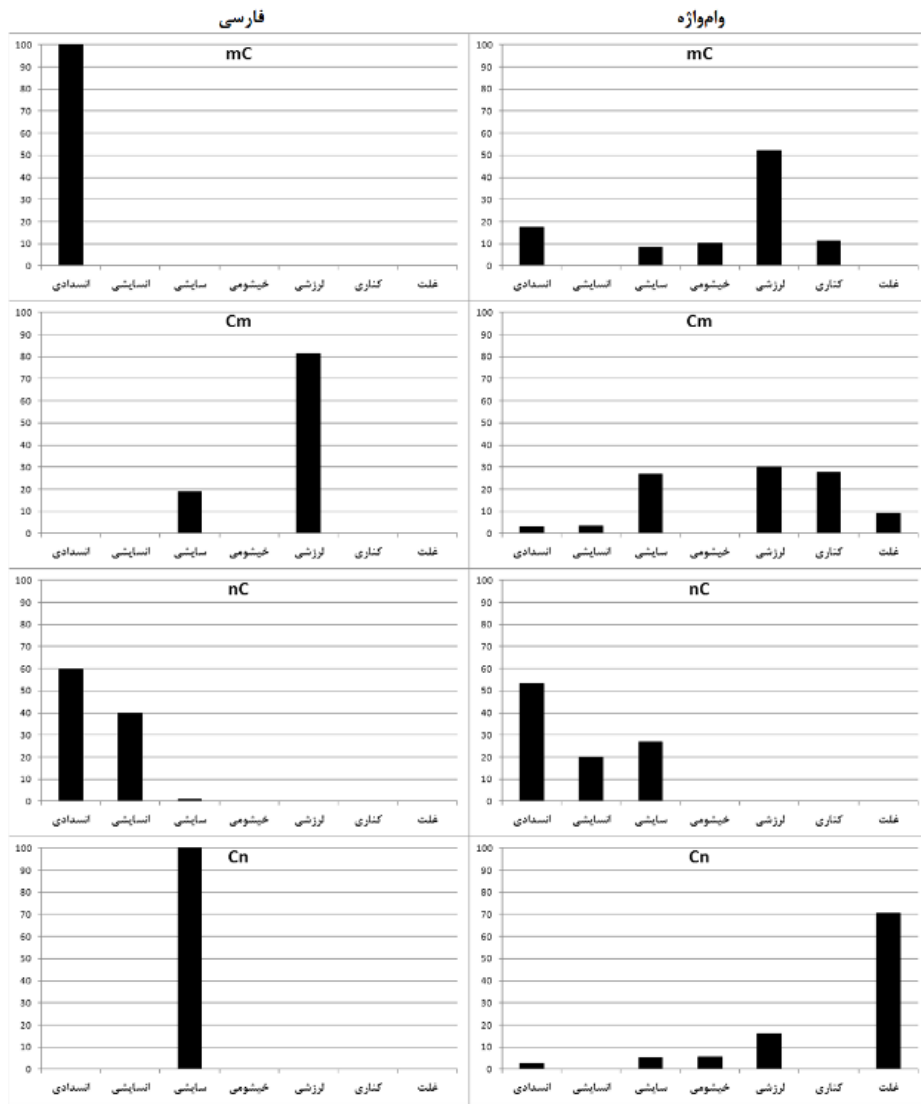
در این بخش به بررسی توزیع مختصه‌های واجی عضو غیرخیشومی خوشه همخوانی در سه سطح شیوه تولید، جایگاه تولید، و وضعیت واک می‌پردازیم. در هر سه سطح، بسامد مختصه‌های واجی براساس مشارکت‌شان در کل هجاها محاسبه شده است.

#### ۲-۴-۱. توزیع مختصه‌های واجی در خوشه همخوانی براساس شیوه تولید

جدول و نمودار ۳ درصد نسبی توزیع مختصه‌های واجی عضو غیرخیشومی خوشه در کل هجاها را براساس شیوه تولیدشان، و به تفکیک نوع واژه نشان می‌دهد. سنجش توازن و پراکندگی مختصه‌های واجی واژه‌های فارسی و

جدول (۳): درصد نسبی توزیع مختصه‌های واجی عضو غیر خیشومی در خوشه‌های همخوانی براساس شیوه تولید

شیوه تولید	فارسی				وام‌واژه			
	Cn	nC	Cm	mC	Cn	nC	Cm	mC
انسدادی	۰	۵۹.۵	۰	۱۷.۶۱	۲۶۳	۵۳.۰۲	۳.۰۳	۱۷.۶۱
انسایشی	۰	۳۹.۶۷	۰	۰	۰	۲۰.۱۳	۳.۳۵	۰
بایشی	۱۰۰	۰.۸۳	۱۸.۷۸	۸.۵۲	۵.۱۱	۲۶.۸۵	۲۶.۶۷	۸.۵۲
خیشومی	۰	۰	۰	۱۰.۲۳	۵.۵۷	۰	۰	۱۰.۲۳
لرزشی	۰	۰	۸۱.۲۲	۵۲.۲۷	۱۶.۱	۰	۳۰.۱۳	۵۲.۲۷
کناری	۰	۰	۰	۱۱.۳۶	۰	۰	۲۷.۶۲	۱۱.۳۶
غلت	۰	۰	۰	۰	۷۰.۵۹	۰	۹.۲۱	۰
دامنه	۱۰۰	۵۸.۶۷	۶۲.۴۴	۴۳.۷۵	۶۷.۹۶	۳۲.۸۹	۳۷.۰۹	۴۳.۷۵
گستره پراکندگی	-	۲	۲	۵	۶	۲	۶	۵
میزان پراکندگی	۱	۳	۲	۵	۵	۳	۶	۵



نمودار (۳): درصد نسبی توزیع مختصه‌های واجی غیرخیشومی در خوشه‌های همخوانی براساس شیوه تولید



وامواژه‌ها را می‌توان با مقایسه دامنه، گستره پراکندگی و میزان پراکندگی در این دو نوع واژه مشاهده کرد. مقایسه این سه مولفه در واژه‌های فارسی با متناظرشان در وامواژه‌ها در راستای پیش‌بینی‌های فرضیه دوم این پژوهش بوده و نشان می‌دهد که وامواژه‌ها نسبت به واژه‌های فارسی از توزیع متوازن‌تر و پراکنده‌تری در سطح شیوه تولید برخوردار هستند. تنها موردی که این رابطه را نقض می‌کند گستره پراکندگی و میزان پراکندگی در خوشه nC است که در بخش ۱، ۳، ۴ در مورد این عدم تمایز توضیح داده خواهد شد.

از آنجا که گستره پراکندگی براساس میزان گرفتگی بست در تولید واج ارزش‌گذاری شده است، برتری گستره پراکندگی مختصه‌های واجی در سطح شیوه تولید وامواژه‌ها نشان می‌دهد که وامواژه‌ها از نظر نوع گرفتگی از تنوع بیشتری برخوردارند. در حالی که در واژه‌های فارسی عضو غیرخیشومی سه خوشه mC، nC، و Cn تنها با واج‌های گرفته (انسدادی‌ها، انسایشی‌ها، و سایشی‌ها) پر شده است و تنها در خوشه Cm واژه‌های فارسی است که چنین گرایش دیده نمی‌شود. از سوی دیگر، از آنجا که در جدول و نمودار ۳، چینش مختصه‌های واجی شیوه تولید براساس میزان گرفتگی مرتب شده است، اطلاعات جدول و نمودار ۳ نشان می‌دهد که در خوشه mC و nC واژه‌های فارسی به‌طور مطلق، و در خوشه Cm واژه‌های فارسی گرایش غالب، پیروی از اصل رسایی است. تنها خوشه Cn در واژه‌های فارسی از اصل رسایی پیروی نمی‌کند که البته این موضوع را نیز می‌توان به‌عنوان شاهدهی در جهت تشکیک در ماهیت فارسی این واژه‌ها در نظر گرفت. در مقابل، در وامواژه‌ها تنها در خوشه nC شاهد پیروی از اصل رسایی هستیم و در سه ترکیب دیگر گرایش غالب در نقض اصل رسایی است.

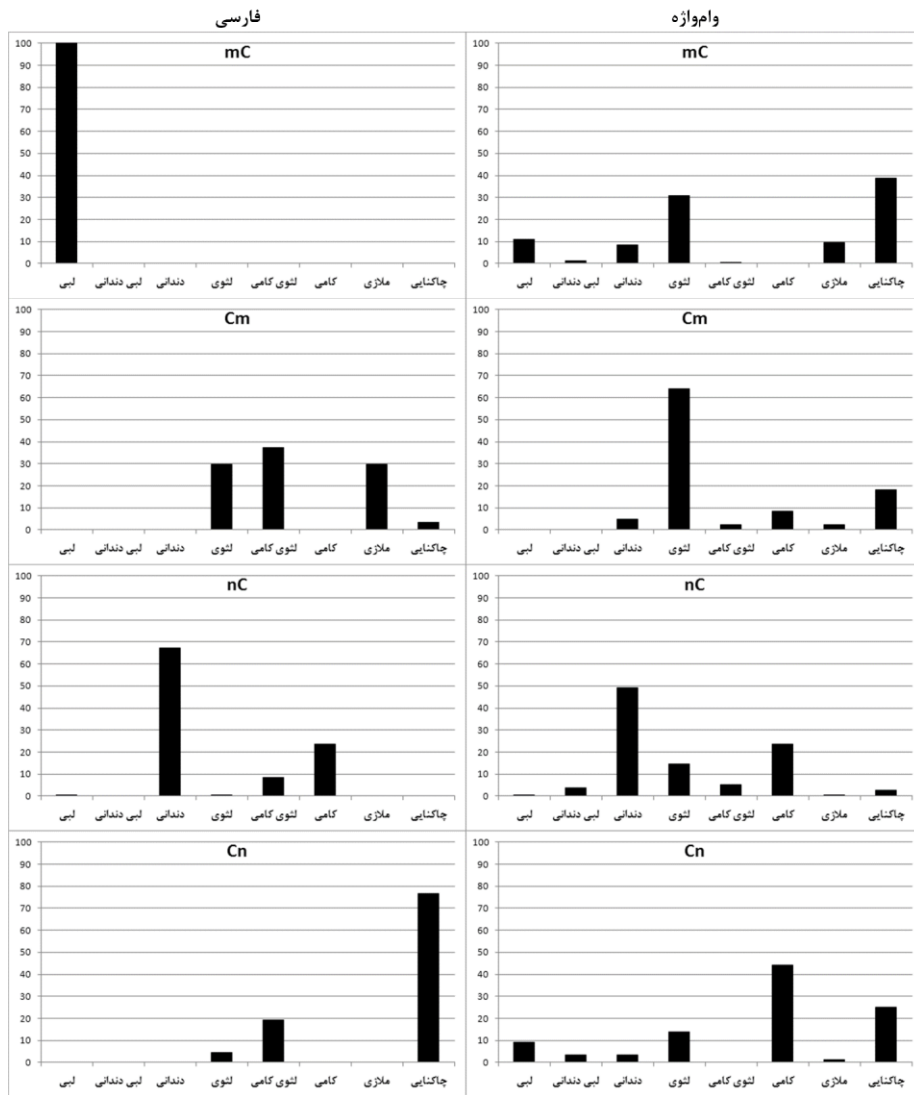
دو گروه واژه‌های فارسی دارای خوشه  $mC$  و  $Cn$ ، که پیشتر ماهیت‌شان را از جوانب مختلف مورد تشکیک قرار دادیم، در این دو نمودار نیز توزیعی ناهنجار یا به عبارت بهتر، تجمعی مطلق را در یک گروه نشان می‌دهند: جایگاه  $C2$  در خوشه  $mC$  را واج‌های انسدادی به‌طور کامل اشغال کرده‌اند و جایگاه  $C1$  در خوشه  $Cn$  را واج‌های سایشی به‌طور کامل اشغال کرده‌اند. بنابراین، نقض اصل رسایی و تجمع در دو مختصه انسدادی و سایشی، به ترتیب، در دو خوشه  $mC$  و  $Cn$  را می‌توان به‌عنوان شاهد دیگری در جهت تشکیک ماهیت این واژه‌ها به‌شمار آورد.

#### ۴-۲-۲. توزیع مختصه‌های واجی در خوشه همخوانی براساس جایگاه تولید

جدول و نمودار ۴ درصد نسبی توزیع مختصه‌های واجی عضو غیرخیشومی خوشه در کل هجاها را براساس جایگاه تولیدشان، و به تفکیک نوع واژه نشان می‌دهند. در توزیع واج‌ها، براساس جایگاه تولید نیز شاهد رفتار توزیعی متمایز مختصه‌های واجی در خوشه‌های واژه‌های فارسی و وام‌واژه‌ها هستیم: مقایسه سه مولفه دامنه، گستره پراکندگی، و میزان پراکندگی در واژه‌های فارسی با وام‌واژه‌ها در راستای پیش‌بینی‌های فرضیه دوم این پژوهش بوده و نشان می‌دهد که وام‌واژه‌ها از توزیع متوازن‌تر و پراکنده‌تری نسبت به واژه‌های فارسی برخوردار هستند. تنها خوشه  $Cm$  در وام‌واژه‌ها دامنه تغییرات بیشتری را نسبت به متناظر خود در واژه‌های فارسی نشان می‌دهد که درباره این مورد نقض در بخش ۴-۳ به‌طور مفصل بحث خواهد شد.

جدول (۴): درصد نسبی توزیع مختصه‌های واجی عضو غیرخیشومی در خوشه‌های همخوانی براساس جایگاه تولید

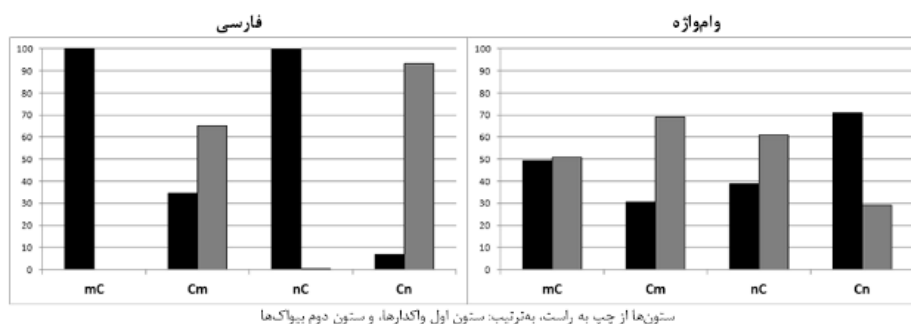
وام‌واژه			فارسی			جایگاه تولید		
Cn	nC	Cm	mC	Cn	nC		Cm	mC
۹.۱۳	۰.۳۷	۰	۱۱	۰	۰.۱۳	۰	۱۰۰	لبی
۳.۴۲	۳.۸۴	۰	۱.۱۸	۰	۰	۰	۰	لبی دندانی
۳.۴۲	۴۹.۳۹	۴.۹۱	۸.۲۵	۰	۶۷.۳۳	۰	۰	دندانی
۱۳.۶۹	۱۴.۴۹	۶۴.۱۱	۳۰.۶۵	۴.۳۷	۰.۵۷	۲۹.۷۹	۰	لثوی
۰	۵.۲۱	۲.۴۵	۰.۵۹	۱۹.۱۳	۸.۵۶	۳۷.۳۳	۰	لثوی کامی
۴۴.۱۱	۲۳.۴۱	۸.۳۶	۰	۰	۲۳.۵۲	۰	۰	کامی
۱.۱۴	۰.۵۵	۲.۱۸	۹.۴۳	۰	۰	۲۹.۷۹	۰	ملازی
۲۵.۱۰	۲.۷۴	۱۷.۹۹	۳۸.۹۰	۷۶.۵۰	۰	۳.۱۹	۰	چاکانی
۴۲.۹۷	۴۹.۰۲	۶۱.۹۳	۳۸.۳۱	۷۲.۱۳	۶۷.۱	۳۴.۰۴	۱۰۰	دامنه
۷	۷	۵	۷	۴	۵	۴	-	گستره پراکندگی
۷	۸	۶	۷	۲	۵	۴	۱	میزان پراکندگی



نمودار (۴): درصد نسبی توزیع مختصه‌های واجی عضو غیرخیشومی در خوشه همخوانی براساس جایگاه تولید

### ۴-۲-۳. توزیع مختصه‌های واجی در خوشه همخوانی براساس وضعیت واک

جدول و نمودار ۵ توزیع مختصه‌های واجی عضو غیرخیشومی خوشه‌های چهارگانه را در کل هجاها، براساس وضعیت واکداری نشان می‌دهد. همانطور که در نمودار ۵ مشاهده می‌کنید، در واژه‌های فارسی جایگاه  $C_1$  گرایش غالب به بیواکی، و جایگاه  $C_2$  گرایش به واکداری دارد. این درحالیست که در وام‌واژه‌ها شاهد توزیع متوازن تری هستیم: مقایسه دامنه تغییرات در هر خوشه‌ واژه‌های فارسی، با خوشه متناظر وام‌واژه نشان می‌دهد که توزیع واج‌ها از نظر وضعیت واکداری در تمامی خوشه‌های وام‌واژه‌ها، به جز خوشه  $C_m$ ، دارای توزیع متوازن تری نسبت به خوشه‌ واژه‌های فارسی است که چنین نتیجه‌ای از فرضیه دوم این پژوهش حمایت می‌کند. علت عدم پیروی خوشه  $C_m$  از این فرضیه در ادامه تشریح خواهد شد.



نمودار (۵): درصد نسبی توزیع مختصه‌های واجی عضو غیرخیشومی در خوشه همخوانی براساس وضعیت واک

جدول (۵): درصد نسبی توزیع مختمه‌های واجی عضو غیرخیشومی در خوشه همخوانی براساس وضعیت واج

دامنه	وام‌واژه		دامنه	فارسی		ترکیب خوشه
	بی‌واک	واکدار		بی‌واک	واکدار	
۱۴۴	۵۰.۷۲	۴۹.۲۸	۱۰۰	۰	۱۰۰	mC
۳۸۵۷	۶۹.۲۹	۳۰.۷۱	۳۰.۴۵	۶۵.۲۳	۳۴.۷۷	Cm
۲۱۸۸	۶۰.۹۴	۳۹.۰۶	۹۹.۸	۰.۱	۹۹.۹	nC
۴۱.۷۲	۲۹.۱۴	۷۰.۸۶	۸۶.۳۵	۹۳.۱۷	۶.۸۳	Cn

#### ۴-۲-۴. علل توزیع پراکنده‌تر و متوازن‌تر مختصه‌های واجی در وام‌واژه‌ها

تا به اینجا نشان دادیم که مختصه‌های واجی وام‌واژه‌ها، نسبت به واژه‌های فارسی، از توزیع پراکنده‌تر و متوازن‌تری برخوردار هستند. سوالی که در این میان می‌توان مطرح کرد این است که علت پراکندگی و توزیع متوازن مختصه‌های واجی در وام‌واژه‌های زبان چیست؟ آرایش واجی واژه‌های هر زبان حاوی قواعد و محدودیت‌هایی است که موجب می‌شود شاهد تجمع و تمرکز مختصه‌های واجی در آن زبان باشیم. این درحالیست که در وام‌واژه‌ها به دو دلیل شاهد توزیع پراکنده‌تر و متوازن‌تر مختصه‌های واجی هستیم: تصادفی بودن وام‌گیری واژه و مبادی مختلف ورود وام‌واژه.

تصادفی بودن وام‌گیری به این نکته اشاره دارد که وام‌گیری واژه‌ها براساس نیازهای ارتباطی صورت می‌گیرد و این نیازهای ارتباطی بر ساختار واج‌شناختی زبان مقصد تقدم دارند. به عبارت ساده‌تر، تقدم نیازهای ارتباطی بر محدودیت‌های واج‌شناختی زبان مقصد در فرایند وام‌گیری موجب می‌شود که واژه ابتدا براساس نیازهای ارتباطی وارد زبان مقصد شده و سپس، محدودیت‌های واج‌شناختی زبان مقصد جرح و تعدیل‌هایی بر واژه اعمال کند. بنابراین، محدودیت‌های واج‌شناختی زبان مقصد نمی‌تواند مانع از ورود واژه‌ای شود که براساس نیاز ارتباطی وارد زبان شده است. در چنین حالتی ممکن است واژه‌هایی وارد زبان شود که از نظر ساختار واجی از ویژگی‌های واج‌شناختی کانونی زبان مقصد فاصله زیادی داشته باشد، با این حال، جرح و تعدیل‌هایی که در واژه صورت می‌گیرد تنها تا جایی می‌تواند اعمال شود که هویت وام‌واژه تغییر نکند. در چنین حالتی اگر برای محدودیت‌های واج‌شناختی زبان مقصد ساختی سلسله‌مراتبی در نظر بگیریم، تنها جرح و تعدیل‌هایی حداقلی بر واژه اعمال می‌شود که در این سلسله مراتب از اهمیت

بیشتری برخوردار است. می‌توان تصور کرد که چنین محدودیت‌هایی تنها در قالب برخی از قواعد کلی، نظیر تعدیل نوع هجا (مانند *sprei / ʔes.pe.rej*) و جایگزین کردن واج‌هایی که در زبان مبدأ حاضر و در زبان مقصد غایب است (مانند *lystɪ / lus.ter*)، می‌توانند در وام‌واژه‌ها دخل و تصرف داشته باشند. ذکر این نکته ضروری است که توصیف فرایند وام‌گیری به صفت تصادفی از منظری واج-شناختی صورت گرفته و به این معنی است که عاملی غیرزبانی در ورود واژه دخیل است، وگرنه از منظری جامعه‌شناختی، وام‌گیری را نمی‌توان تصادفی دانست و می‌توان برای آن دلایل اجتماعی، سیاسی و اقتصادی مختلفی را برشمرد.

از سوی دیگر، این فرایند تصادفی از زبان‌های مبدأ مختلفی صورت می‌گیرد که هرکدام از ساختار واج‌شناختی مشخصی پیروی می‌کنند. تفاوت این ساختار واج-شناختی براساس اصل قراردادی بودن صورت - معنا موجب می‌شود که تناظر یک‌به‌یکی بین این ساختارهای واج‌شناختی وجود نداشته باشد. به این ترتیب، تصادفی بودن فرایند وام‌گیری و مبادی مختلف ورود واژه موجب می‌شود شاهد تنوعی از ویژگی‌های واج‌شناختی باشیم که در ساختار زبان مقصد به‌طور متوازن پراکنده شده‌اند. به عبارت دیگر، پراکندگی در توزیع، به توزیع متوازن ناگزیر ویژگی‌های واج‌شناختی منجر می‌شود. براین اساس، توزیع متوازن بیش از اینکه مسأله‌ای آماری باشد، مسأله‌ای احتمالاتی است.

#### ۳-۴. بسامد خوشه‌های حاوی واج خیشومی

بررسی ۵۴ خوشه منحصراً حاوی واج خیشومی نشان می‌دهد که از میان این ۵۴ خوشه، تنها ۱ خوشه (۱,۸۵ درصد) منحصراً در واژه‌های فارسی حضور دارد. ۱۴ خوشه (۲۵,۹۳ درصد) میان واژه‌های فارسی و وام‌واژه‌ها مشترک بوده و ۳۹



خوشه (۷۲,۲۲ درصد) تنها در وام‌واژه‌ها حضور دارند. این آمار نشان می‌دهد که خوشه‌های مستخرج از وام‌واژه‌ها نسبت به خوشه‌های مستخرج از واژه‌های فارسی از تنوع بیشتری برخوردارند. نکته جالب توجهی که در این آمار مشاهده می‌شود این است که تنها خوشه‌ای که منحصراً در واژه‌های فارسی ظاهر شده است، خوشه  $Jn$  است که در واژه‌های فارسی جشن، رشن، و گشن ظاهر شده است و ماهیت ریشه‌شناختی و نوع تلفظ این واژه‌ها را پیشتر مورد تشکیک قرار دادیم. انحصاری بودن این خوشه در این واژه‌ها را نیز می‌توان به عنوان شاهدهی بر چنین تشکیکی در نظر گرفت.

جدول ۶ حضور خوشه‌های بالفعل حاوی واج خیشومی را به ترتیب نزولی نمایش می‌دهد. از بین پنج خوشه نخست این جدول، سه خوشه دارای ساختار  $nC$  هستند که نشان‌دهنده حضور فعال واج  $n$  در موضع  $C_1$  است. در واقع این سه خوشه (:  $nd$ ,  $nj$  و  $ndʒ$ ) جزو سه خوشه پربسامد زبان فارسی هستند: در فهرست بسامدی‌ای که علم‌الهدی (Alamolhoda, 2000) از خوشه‌های زبان فارسی ارائه کرده است، این سه خوشه جزء ۱۶ خوشه نخست زبان فارسی و جزء ۱۰ خوشه نخست واژه‌های به‌لحاظ ریشه‌شناختی فارسی هستند.

جدول (۶): فهرست بسامدی خوشه‌های منحصر همخوانی در کل هجانه‌ها

درصد	تعداد	خوشه	رتبه	درصد	تعداد	خوشه	رتبه	درصد	تعداد	خوشه	رتبه
۰.۱۷	۴	nʃ	۳۷	۰.۷۳	۱۷	zn	۱۹	۳۱.۹۳	۷۵۰	nd	۱
۰.۱۷	۴	nb	۳۸	۰.۶۴	۱۵	mz	۲۰	۱۷.۷۱	۴۱۶	nʃ	۲
۰.۱۳	۳	bn	۳۹	۰.۵۵	۱۳	ms	۲۱	۸.۱۳	۱۹۱	ndʒ	۳
۰.۱۳	۳	Gm	۴۰	۰.۵۵	۱۳	rn	۲۲	۷.۷۵	۱۸۲	sm	۴
۰.۱۳	۳	nʒ	۴۱	۰.۵۱	۱۲	mib	۲۳	۴.۰۹	۹۶	rm	۵
۰.۱۳	۳	nʔ	۴۲	۰.۴۷	۱۱	jm	۲۴	۳.۰۲	۷۱	ʃm	۶
۰.۱۳	۳	tn	۴۳	۰.۴۳	۱۰	cm	۲۵	۲.۴۳	۵۷	jn	۷
۰.۱۳	۳	fn	۴۴	۰.۳۸	۹	tm	۲۶	۲.۳۴	۵۵	ns	۸
۰.۰۹	۲	ʃm	۴۵	۰.۳۸	۹	mn	۲۷	۲.۲۶	۵۳	zm	۹
۰.۰۹	۲	nh	۴۶	۰.۳۸	۹	ʔn	۲۸	۲.۰۰	۴۷	nt	۱۰
۰.۰۹	۲	sn	۴۷	۰.۳۴	۸	dʒm	۲۹	۱.۴۰	۳۳	lm	۱۱
۰.۰۹	۲	mt	۴۸	۰.۳۴	۸	mG	۳۰	۱.۴۰	۳۳	mʔ	۱۲
۰.۰۴	۱	cn	۴۹	۰.۳۰	۷	mip	۳۱	۱.۳۳	۳۱	hm	۱۳
۰.۰۴	۱	Gn	۵۰	۰.۳۰	۷	ʃn	۳۲	۱.۲۳	۲۹	xm	۱۴
۰.۰۴	۱	mʃ	۵۱	۰.۳۰	۷	nf	۳۳	۱.۱۵	۲۷	hn	۱۵
۰.۰۴	۱	mʃ	۵۲	۰.۲۱	۵	ml	۳۴	۱.۱۱	۲۶	nz	۱۶
۰.۰۴	۱	nʃ	۵۳	۰.۲۱	۵	md	۳۵	۰.۹۸	۲۳	mr	۱۷
۰.۰۴	۱	nx	۵۴	۰.۲۱	۵	ʔm	۳۶	۰.۸۱	۱۹	nc	۱۸

نکته مهمی که جدول ۹ به همراه دارد، پیروی آن از یکی از اصول خانواده قوانین قدرت<sup>۱</sup> با نام اصل پارتو<sup>۲</sup> است. قانون قدرت رابطه‌ای نسبی میان دو کمیت است که براساس آن تغییر در یک کمیت، تغییر نسبی در کمیت دیگر را به همراه دارد. براساس اصل پارتو، که توسط ویلفردو پارتو<sup>۳</sup>، اقتصاددان ایتالیایی، کشف شده است، چیزها دارای توزیع یکسان نیستند، بلکه ۸۰ درصد از توزیع متعلق به ۲۰ درصد از داده‌ها، و ۲۰ درصد از توزیع متعلق به ۸۰ درصد از داده‌هاست. توزیع خوشه‌های منحصر در کل هجاها، که در جدول ۶ آمده است، نشان می‌دهد که توزیع این داده‌ها از اصل پارتو تبعیت می‌کند: ۲۰ درصد از ۵۴ خوشه منحصر (~۱۱ خوشه منحصر (۸، ۱۰)، ۸۳،۰۶ درصد از بسامد کل هجاها را تشکیل می‌دهند. در نمودار ۸ این توزیع به تصویر کشیده شده است.<sup>۴</sup>

پیروی داده‌های این پژوهش از این اصل، از این نظر حائز اهمیت است که این اصل برای توصیف داده‌های طبیعی به کار می‌رود. این درحالیست که داده‌هایی که در این پژوهش مطابقت آنها با این اصل مورد آزمایش قرار گرفته است، هرچند از فهرستی (تقریباً) پیکره‌بنیاد به دست آمده است، با این حال کل داده‌های پیکره مورد سنجش قرار نگرفته، بلکه «تنها» طبقه خیشومی‌ها، و توزیع آنها «تنها» در خوشه‌های همخوانی مورد سنجش قرار گرفته است. بنابراین، تبعیت این داده‌ها از این اصل نشان‌دهنده طبیعی بودن رفتار توزیعی طبقه خیشومی‌ها در خوشه‌های همخوانی است که از آن می‌توان طبیعی بودن طبقه واج‌های خیشومی در زبان

1. power law
2. Pareto principle
3. Vilfredo Pareto

۴. ذکر این نکته ضروری است که توزیع ۲۰/۸۰ داده‌ها یک نسبت تقریبی است و در واقع هیچ‌گاه در داده‌های طبیعی شاهد توزیعی با چنین دقتی نیستیم؛ بلکه توزیع داده‌های طبیعی به چنین نسبتی نزدیک هستند.

فارسی را نتیجه گرفت. به عبارت دیگر، رفتار توزیعی واج‌های خیشومی در خوشه همخوانی و تبعیت آن از اصل پارتو، طبیعی بودن طبقه واج‌های خیشومی را در زبان فارسی تأیید می‌کند.

جدول ۷ توزیع خوشه‌های منحصر در کل هجاها را براساس نوع ترکیب خوشه نشان داده است. جدول ۸ توزیع خوشه‌های منحصر در کل هجاها را به تفکیک واژه‌های فارسی و وام‌واژه‌ها نشان می‌دهد. در بخش ۴-۲-۲ دیدیم که توزیع عضو غیرخیشومی در خوشه Cm وام‌واژه‌ها، نسبت به خوشه متناظر در واژه‌های فارسی، از نظر جایگاه تولید، دامنه تغییرات بیشتری نشان می‌دهند و در نتیجه، توزیع نامتوازنی نسبت به واژه‌های فارسی دارد. همین موضوع را در بخش ۴-۳-۲ در مورد همین خوشه از نظر وضعیت واک شاهد بودیم. سؤالی که در این میان به وجود می‌آید این است که چرا خوشه Cm از این فرضیه تبعیت نمی‌کند؟ پاسخ این پرسش بدین قرار است که وام‌گیری گسترده واژه‌هایی که در زبان مبدأ به طبقه ساختوازی مشخصی تعلق دارد، موجب می‌شود که شاهد تجمع برخی ویژگی‌های واج‌شناختی در زبان مقصد باشیم که نتیجه این تجمع، عدم توازن بوده و در نتیجه، تمایز دامنه تغییرات در توزیع مختصه‌های واجی در واژه‌های فارسی و وام‌واژه‌ها را تضعیف می‌کند.



جدول (۷): فهرست بسامدی خوشه‌های همخوانی به تفکیک نوع ترکیب خوش

رتبه	mC	تعداد	درصد	Cm	تعداد	درصد	nC	تعداد	درصد	Cn	تعداد	درصد
۱	m?	۳۳	۲۴٫۶۳	sm	۱۸۲	۳۳٫۵۳	nd	۷۵۰	۴۹٫۰۵	jn	۵۷	۳۷٫۵۰
۲	mr	۲۳	۱۷٫۱۶	rm	۹۶	۱۷٫۶۸	nj	۴۱۶	۲۷٫۲۱	hn	۲۷	۱۷٫۷۶
۳	mz	۱۵	۱۱٫۱۹	jm	۷۱	۱۳٫۰۸	nds	۱۹۱	۱۲٫۴۹	zn	۱۷	۱۱٫۱۸
۴	ms	۱۳	۹٫۷۰	zm	۵۳	۹٫۷۶	ns	۵۵	۳٫۶۰	rn	۱۳	۸٫۵۵
۵	mb	۱۲	۸٫۹۶	lm	۳۳	۶٫۰۸	nt	۴۷	۳٫۰۷	mn	۹	۵٫۹۲
۶	mn	۹	۶٫۷۲	hm	۳۱	۵٫۷۱	nz	۲۶	۱٫۷۰	Ꞥn	۹	۵٫۹۲
۷	mG	۸	۵٫۹۷	xm	۲۹	۵٫۳۴	nc	۱۹	۱٫۲۴	fn	۷	۴٫۶۱
۸	mp	۷	۵٫۲۲	jm	۱۱	۲٫۰۳	nf	۷	۰٫۴۶	tn	۳	۱٫۹۷
۹	ml	۵	۳٫۷۳	cm	۱۰	۱٫۸۴	nb	۴	۰٫۲۶	fn	۳	۱٫۹۷
۱۰	md	۵	۳٫۷۳	tm	۹	۱٫۶۶	nj	۴	۰٫۲۶	bn	۳	۱٫۹۷
۱۱	mt	۲	۱٫۴۹	dsm	۸	۱٫۴۷	n3	۳	۰٫۲۰	sn	۲	۱٫۳۳
۱۲	mf	۱	۰٫۷۵	Ꞥm	۵	۰٫۹۲	n?	۳	۰٫۲۰	Gn	۱	۰٫۶۶
۱۳	mj	۱	۰٫۷۵	Gm	۳	۰٫۵۵	nh	۲	۰٫۱۳	cn	۱	۰٫۶۶
۱۴				fm	۲	۰٫۳۷	nj	۱	۰٫۰۷			
۱۵							nx	۱	۰٫۰۷			

جدول (۸): فهرست بسامدی خوشه‌های همخوانی به تفکیک نوع وازه (فارسی) / واموازه

درصد	تعداد	خوشه	درصد	تعداد	خوشه	درصد	تعداد	خوشه	درصد	تعداد	خوشه
۰.۳۴	۳	nɜ	۱.۲۶	۱۱	ndʒ	۲۰.۸۲	۱۸۲	sm	۴۷.۸۶	۷۰۶	nd
۰.۳۴	۳	nʔ	۱.۲۶	۱۱	nz	۶.۵۲	۵۷	jn	۲۵.۱۵	۳۷۱	nʃ
۰.۳۴	۳	Gm	۱.۲۶	۱۱	jm	۶.۲۹	۵۵	ns	۱۲.۲	۱۸۰	ndʒ
۰.۳۴	۳	fn	۱.۱۴	۱۰	cm	۵.۲۶	۴۶	nt	۴.۷۵	۷۰	ʃm
۰.۲۳	۲	nh	۱.۰۳	۹	tm	۵.۱۵	۴۵	nʃ	۴.۰۷	۶۰	rm
۰.۲۳	۲	ʃm	۱.۰۳	۹	mn	۵.۰۳	۴۴	nd	۱.۹	۲۸	xm
۰.۲۳	۲	sn	۱.۰۳	۹	ʔn	۴.۹۲	۴۳	zm	۱.۰۲	۱۵	nz
۰.۲۳	۲	nb	۰.۹۲	۸	dʒm	۴.۱۲	۳۶	rm	۰.۹۵	۱۴	hn
۰.۲۳	۲	mt	۰.۹۲	۸	mG	۳.۷۸	۳۳	lm	۰.۶۸	۱۰	zm
۰.۱۱	۱	nʃ	۰.۸	۷	mp	۳.۷۸	۳۳	mʔ	۰.۴۷	۷	ʃn
۰.۱۱	۱	ʃm	۰.۸	۷	nf	۳.۲	۲۸	hm	۰.۴۱	۶	mb
۰.۱۱	۱	Gn	۰.۶۹	۶	mb	۲.۶۳	۲۳	mr	۰.۲	۳	hm
۰.۱۱	۱	mʃ	۰.۵۷	۵	ml	۲.۱۷	۱۹	nc	۰.۱۴	۲	nb
۰.۱۱	۱	cn	۰.۵۷	۵	ʔm	۱.۷۲	۱۵	zn	۰.۱۴	۲	zn
۰.۱۱	۱	nx	۰.۵۷	۵	md	۱.۷۲	۱۵	mz	۰.۰۷	۱	nt
۰.۱۱	۱	xm	۰.۴۶	۴	nʃ	۱.۴۹	۱۳	ms			
۰.۱۱	۱	mʃ	۰.۳۴	۳	bn	۱.۴۹	۱۳	hn			
			۰.۳۴	۳	tn	۱.۴۹	۱۳	rn			

نگاهی به جدول ۷ نشان می‌دهد که در خوشه Cm، بیشترین بسامد به خوشه sm تعلق دارد. نگاهی به جدول ۸ هم نشان می‌دهد که خوشه sm منحصرأ به وام‌واژه‌ها تعلق دارد و از طرفی، پربسامدترین خوشه در وام‌واژه‌هاست. علاوه بر این، نگاهی به جدول ۶ نشان می‌دهد که خوشه sm چهارمین خوشه در بین کل خوشه‌های مورد مطالعه این پژوهش است. نگاهی به وام‌واژه‌های حاوی این خوشه نشان می‌دهد که اکثریت غالب این وام‌واژه‌ها به طبقه ساختوازی خاصی تعلق دارند: از بین ۳۷۱ وام‌واژه‌ای که واج m در جایگاه C<sub>2</sub> خوشه آنها حضور دارد، ۱۸۲ واژه (۴۹,۰۶ درصد) به توالی sm ختم شده‌اند که ۸۶,۲۶ درصد از این ۱۸۲ واژه (۱۵۷) از ترکیب پسوند اسم‌ساز انگلیسی ism تشکیل شده‌اند. این شواهد نشان می‌دهد وام‌گیری گسترده واژه‌هایی که در زبان مبدأ به طبقه ساختوازی خاصی تعلق دارند، منجر به تجمع مختصه‌های واجی آن طبقه ساختوازی در زبان مقصد خواهند شد. این موضوع تنها در مورد وام‌واژه‌ها صادق نیست. در واقع، تجمع و عدم پراکندگی متوازن مختصه‌های واجی در واژه‌های بومی زبان نیز به دلیل استفاده از ترکیبات ساختوازی مشخص در آن زبان است، چنانکه در داده‌های این پژوهش خوشه nd، که پربسامدترین خوشه فارسی است، حاصل ترکیب وندهای اشتقاقی‌ای نظیر مند، ووند، و یا ترکیب واژه‌هایی نظیر بند (ستاک حال بستن) و چند است. واج‌های خیشومی دارای دو ویژگی هستند که آنها را از دیگر واج‌های فارسی متمایز می‌کند. ویژگی نخست به کانالی که جریان هوا از آن خارج می‌شود (خیشوم) باز می‌گردد. ویژگی دوم واج‌های خیشومی در این است که در تولید این واج‌ها شاهد ایجاد بست انسدادی در دهان هستیم. ممکن است ذکر ویژگی دوم اینگونه مورد اعتراض قرار بگیرد که این ویژگی واج‌های خیشومی را از انسدادی‌ها متمایز نمی‌کند. باین حال، نوع تولید واج‌های خیشومی از انسدادی‌ها متمایز است. تولید



آوا از سه مرحله گرایش، گیرش و رهش تشکیل شده است که دو مرحله آخر شیوه تجمع و خروج هوا را تعیین می‌کنند. نکته‌ای که موجب تمایز واج‌های خیشومی می‌شود این است که بست انسدادی ایجادشده در تولید این دو آوا نقشی در شیوه خروج جریان هوا ندارد و همین امر موجب شده این ویژگی در توصیف این واج نقشی نداشته باشد. در تولید آواهای خیشومی پایین آمدن ملاز تعیین‌کننده خروج جریان هواست و بست انسدادی تنها طنین مورد نیاز برای تولید آوا را فراهم می‌کند. با این مقدمه در ادامه به علل پربسامدبودن واج  $n$  در موضع  $C_1$  خوشه همخوانی  $nC$  می‌پردازیم.

جدول ۶ نشان می‌دهد که سه خوشه  $nd$ ،  $nj$  و  $ndʒ$  ۵۷٫۷۷ درصد از کل بسامد خوشه‌های همخوانی حاوی واج خیشومی را تشکیل می‌دهند. به همین دلیل باید دید که واج  $n$  چه امکاناتی را در جایگاه  $C_1$  فراهم می‌کند که موجب پربسامدشدن این واج در این موضع می‌شود. واج  $n$  از اتصال منعطف‌ترین و پرکاربردترین اندام تولیدی (نوک زبان) با لثه حاصل می‌شود. چنین کاربردی با اصل مرز اجباری تحدید و با قاعده همگونی محل تولید توسعه می‌یابد. براساس اصل مرز اجباری هیچ دو عنصر مشابهی نباید در مجاورت یکدیگر قرار بگیرند (McCarthy, 1988؛ نقل شده در بی‌جن‌خان (۱۳۹۲)). براساس این اصل، از آنجا که واج  $n$  دارای بست انسدادی است، مجاورت آن با واج‌های گرفته (انسدادی‌ها، انسایشی‌ها، و سایشی‌ها) غیرمجاز خواهد بود. با این حال، این مجاورت غیرمجاز با عملکرد قاعده همگونی محل تولید مرتفع می‌شود: هر واج گرفته‌ای که پس از واج  $n$  قرار بگیرد، واج  $n$  از نظر جایگاه تولید با آن واج همگون خواهد شد. با این وجود، تمام واج‌های گرفته دارای چنین قابلیت نیستند و واج  $n$  تنها با واج‌هایی می‌تواند همگون شود که جایگاه تولید آنها پس از انشعاب خیشوم واقع شده باشد؛ یعنی

تمام واج‌های گرفته‌ای که در فضای دهان تولید می‌شوند. به عبارت بهتر، تمام واج‌هایی که بست آنها با زبان شکل می‌گیرد. به این ترتیب، واج‌هایی که می‌توانند پس از واج  $n$  در خوشه همخوانی ظاهر شوند و جایگاه واج  $n$  با جایگاه این واج‌ها همگون شود، واج‌هایی هستند که دارای مشخصه [+ گرفته] و [- حنجره‌ای] هستند. از بین ۲۳ همخوان زبان فارسی ۱۸ همخوان گرفته وجود دارد که از میان این ۱۸ همخوان، تنها دو همخوان گرفته در فضای دهان تولید نمی‌شوند ( $h$  و  $ʔ$ ). این مسأله موجب می‌شود که واج  $n$  در ترکیب خوشه  $nC$  بتواند از نظر جایگاه تولید با این ۱۶ واج همگون شود. اگر همگونی جایگاه تولید را در خوشه  $nC$  نوعی اقتصاد زبانی در تولید آوا به‌شمار آوریم، ترکیب واج  $n$  با این ۱۶ واج در خوشه  $nC$ ، خوشه بهینه را برای زبان به‌همراه خواهد داشت. داده‌های به دست آمده از این پژوهش نیز از چنین تفسیری حمایت می‌کنند: از بین تمام خوشه‌های دارای الگوی  $nC$  که در این پژوهش مورد مطالعه گرفت (اعم از فارسی و وام‌واژه) ۱۵ خوشه منحصر  $nC$  به دست آمده است که در همه این خوشه‌ها عضو غیرخیشومی دارای مشخصه [+ گرفته] است. از بین این ۱۵ خوشه، دو خوشه وجود دارد که دارای مشخصه [+ حنجره‌ای] هستند ( $nh$  و  $nʔ$ ). ۱۳ خوشه باقیمانده همگی دارای مشخصه [+ گرفته، - حنجره‌ای] هستند. سه خوشه‌ای که در این داده‌ها دیده نمی‌شود، عبارتند از:  $np$ ،  $nv$  و  $nG$ ؛ با این حال عدم حضور این سه خوشه به دلیل عدم همگونی واج  $n$  با سه واج  $p$ ،  $v$  و  $G$  نیست، چراکه در تولید طبیعی واژه‌هایی نظیر انوار، جان‌پناه و عنقریب شاهد همگونی واج  $n$  با این سه واج هستیم. بنابراین، عدم حضور این خوشه را می‌توان تصادفی دانست. حضور دو خوشه  $nh$  و  $nʔ$  خدشه‌ای به تفسیر ما وارد نمی‌کند، زیرا این دو خوشه تنها ۰٫۳۳ درصد از کل خوشه‌ها را

تشکیل می‌دهند (هر دو خوشه در وام‌واژه‌ها مشاهده شده‌اند) و ۹۹,۶۷ درصد خوشه‌ها دارای واجی با مشخصه [+گرفته، - حنجره‌ای] هستند و این دو مشخصه است که موجب بهینه‌بودن خوشه nC می‌شود. تفسیر ما از علل پربسامدبودن نیز به این مشخصه‌ها اشاره دارد.

از طرفی ترکیب واج n با واج‌هایی با مشخصه [+گرفته] در خوشه nC ترکیباتی را به همراه خواهد داشت که از اصل رسایی پیروی خواهند کرد. در خوشه nC هم، از آنجا که واج‌های گرفته از درجه رسایی کمتری نسبت به واج n (: خیشومی‌ها) برخوردارند، قرارگرفتن این واج‌ها در موضع دوم خوشه، پیروی از این اصل را به همراه خواهد داشت. علاوه‌براین، توزیع واج‌های گرفته در این موضع نیز نشان می‌دهد گرفتگی واج‌ها در این موضع با میزان توزیع‌شان همبستگی مثبت دارد: واج‌های گرفته از نظر میزان گرفتگی به ترتیب، عبارتند از انسدادی‌ها، انسایشی‌ها، و سایشی‌ها که توزیع این سه ویژگی در موضع C<sub>2</sub> خوشه nC، به ترتیب، ۵۲,۷۵، ۸,۳، و ۴,۰۴ درصد است. چنین رابطه‌ای نه تنها به پیروی اکید خوشه nC از اصل رسایی اشاره دارد، بلکه نشان‌دهنده ارتباط میان بست انسدادی واج n و میزان گرفتگی عضو غیرخیشومی خوشه nC است. بنابراین، با توجه به شیب نزولی توزیع واج‌های گرفته، هرچه گرفتگی بست عضو غیرخیشومی بیشتر باشد، میزان مشارکت آن در خوشه nC افزایش یافته و بالعکس، هرچه میزان گرفتگی بست عضو غیرخیشومی کمتر باشد، از میزان مشارکت واج‌ها در خوشه nC کاسته می‌شود.

بنابراین می‌توان گفت که پربسامدبودن الگوی خوشه nC ناشی از بست انسدادی واج n است که موجب می‌شود این واج بتواند با کمک قاعده همگونی محل تولید، با ۱۶ واج [+گرفته، - حنجره‌ای] همگون شود. از چنین منظری اصل مرز اجباری علی‌رغم محدودیت‌هایی که در همنشینی واج‌ها به وجود می‌آورد،

موجب عملکرد قاعده‌ای (قاعده همگونی محل تولید) می‌شود که این قاعده ظرفیت همنشینی واج  $n$  با دیگر واج‌ها در خوشه  $nC$  را افزایش می‌دهد. علاوه بر این، حضور واج‌های گرفته در موضع  $C_2$  خوشه  $nC$  پیروی از اصل رسایی را در خوشه‌های حاصل از چنین ترکیبی به همراه خواهد داشت که خوشه‌هایی بهینه را برای زبان فراهم می‌کنند.

به نظر می‌رسد که گرفتگی عضو غیرخیشومی در خوشه  $nC$  یک محدودیت مطلق در زبان فارسی است، زیرا تنها ویژگی واج‌شناختی‌ای که در آن پراکندگی مختصه‌های واجی (چه از نظر گستره پراکندگی و چه از نظر میزان پراکندگی) در خوشه‌های چهارگانه واژه‌های فارسی و وام‌واژه‌ها یکسان است، در خوشه  $nC$  مشاهده شده است. با این حال، نگاهی چندجانبه به خوشه  $nC$  واژه‌های فارسی و وام‌واژه‌ها تمایزهایی را در توزیع مختصه‌های واجی عضو غیرخیشومی این دو نوع واژه آشکار می‌کند.

عضو غیرخیشومی خوشه  $nC$  واژه‌های فارسی دارای مشخصه  $[+]$  گرفته، - حنجره‌ای، + واکدار است. تنها واژه‌ای که این رابطه را نقض می‌کند، واژه سوشیانت است که در آن عضو غیرخیشومی بیواک است. واژه سوشیانت در آیین زرتشتی به منجی موعود اطلاق می‌شود که خوشه  $nt$  در آن بازمانده پسوند  $-ant$  است که در زبان‌های دوره باستان برای ساخت صفت فاعلی معلوم به کار می‌رفته است. این واژه تنها واژه‌ای است که این پسوند در آن به شکل باستانی آن باقی مانده است که به نظر می‌رسد به دلیل اسم خاص بودن آن و ملاحظات مذهبی تغییری در تلفظ آن ایجاد نشده است. سایر واژه‌هایی که در دوره باستان با این پسوند ساخته شده‌اند، در در فارسی میانه، واج  $t$  در آنها واکدار و به  $d$  تبدیل شده است. از بین دیگر

واج‌های گرفته هم واج‌های سایشی تنها در خوشه nZ به کار رفته‌اند که این خوشه تنها در دو واژه پانزده و شانزده و ترکیبات حاصل از این دو واژه حضور دارد. درمقابل عضو غیرخیشومی خوشه nC در وام‌واژه‌ها دارای مشخصه [+گرفته] بوده و نسبت به دو مشخصه دیگر (حجره‌ای بودن و وضعیت واک) بی تفاوت است. محدودیت خوشه‌های فارسی به سه مشخصه [+گرفته، - حجره‌ای، + واکدار] و محدودیت خوشه وام‌واژه‌ها به مشخصه [+گرفته] نشان می‌دهد که گرفتگی واج‌ها در موضع دوم خوشه nC محدودیتی اکید در زبان فارسی است. بنابراین، هرچند گستره پراکندگی و میزان پراکندگی خوشه nC در واژه‌های فارسی و وام‌واژه‌ها یکسان است، در یک نگاه کلی‌تر، خوشه وام‌واژه‌ها نسبت به دو مشخصه [-حجره‌ای، + واکدار] بی تفاوت است که بی تفاوتی نسبت به این دو مشخصه را می‌توان به پراکنده‌تر بودن مختصه‌های واجی در خوشه nC وام‌واژه‌ها تفسیر کرد. علاوه بر این، اشتراک خوشه واژه‌های فارسی و وام‌واژه‌ها در مشخصه گرفتگی، از ادعای ما مبنی بر پربسامد بودن ترکیب nC، به علت انسدادی بودن بست واج n، حمایت می‌کند. توضیحات ارائه شده نشان از اهمیت بست دهانی در خیشومی‌ها دارد که غالباً در توصیف واج‌های خیشومی نادیده گرفته می‌شود.

## ۵. نتیجه‌گیری

شاید بتوان یکی از مهمترین نتایج این پژوهش را اهمیت نقش بست انسدادی در توصیف واج‌های خیشومی دانست. در آموزه‌های ساختگرایی مختصه‌های واجی براساس نقش ممیزشان تعریف می‌شوند. در چنین نگرشی بست انسدادی واج‌های خیشومی به دلیل اینکه فاقد نقش ممیز است، در توصیف این واج در زبان فارسی

نقش ندارد. نتایج به دست آمده از این پژوهش شواهدی را بر عدم بسندگی چنین نگرشی فراهم می‌کند. انسدادی بودن بست واج n موجب می‌شود که جایگاه تولید این واج قادر به همگونی با واج‌های دارای مشخصه [+گرفته، - حنجره‌ای] شود. چنین امکانی موجب می‌شود که این واج بتواند به طور بالقوه با ۱۶ واج زبان فارسی همگون شود. علاوه بر این، حضور واج‌های گرفته در این موضع، پیروی از اصل رسایی را نیز برای خوشه nC به همراه خواهد داشت. علاوه بر واژه‌های فارسی، نشان دادیم که مختصه [+گرفته] محدودیتی در خوشه nC است که وام‌واژه‌ها هم باید از آن پیروی کنند (۳-۴).

واج‌های خیشومی در واژه‌های فارسی گرایش شدیدی به توزیع تکمیلی دارند. ما فرضیاتی (آواشناختی، ساختواژی، گویشی، تاریخی، و ناهنجاری‌های توزیعی) را نیز در جهت تشکیک در ماهیت واژه‌هایی که شکل مطلق این رابطه را نقض می‌کنند، مطرح کردیم (۱-۴).

در بخش ۴-۲-۴ نشان دادیم که پراکندگی متوازن مختصه‌های واج‌شناختی در وام‌واژه‌ها به دلیل تصادفی بودن فرایند وام‌گیری واژه‌ها و مبادی مختلف ورود وام‌واژه‌هاست. با این حال، اگر طبقه ساختواژی مشخصی از زبانی به طور گسترده وارد زبان مقصد شود، موجب تجمع و توزیع نامتوازن مختصه‌های ساختواژی خواهد شد. در واقع، علت تجمع و پراکندگی نامتوازن مشخصه‌های واج‌شناختی واژه‌های بومی هم وجود سازوکارهای ساختواژی مشخص در زبان است. نتایج به دست آمده در این بخش به ویژگی‌های توزیعی-واج‌شناختی‌ای اشاره دارد که می‌تواند به درک روانشناختی - واج‌شناختی ما از واژه‌های یک زبان در برابر وام‌واژه‌های آن دلالت داشته باشد.

از نتایج جانبی به‌دست‌آمده از این پژوهش، رفتار توزیعی مشابه واج‌های خیشومی در خوشه همخوانی بود که این شواهد از طبیعی بودن طبقه واج‌های خیشومی حمایت می‌کند. علاوه بر این، در توزیع خوشه‌های حاوی واج‌خیشومی نیز شاهد بودیم که توزیع این خوشه‌ها از اصل پارتو حمایت می‌کند که این موضوع نیز مؤید طبیعی بودن طبقه واج‌های خیشومی در زبان فارسی است (۳-۴).

## منابع

- اسلامی، محرم. (۱۳۸۷). «واژه‌گزینی و ملاحظات آوایی». *پردازش علائم و داده‌ها*. ۹/۱: ۷۳-۸۴.
- اسلامی، محرم، و محمود بی‌جن خان. (۱۳۸۴). «واژه‌گزینی و واج‌آرایی». *مجموعه مقالات دومین هم‌اندیشی واژه‌گزینی*. تهران: فرهنگستان ادب و زبان فارسی. ۴۳۰-۴۴۸.
- اسلامی، محرم، رحیمی، افشین و سودابه اسلامی. (۱۳۹۲). «نظام آوایی زبان فارسی در آینه آمار». *زبان و زبانشناسی*. ۹(۱۸): ۶۵-۹۰.
- اسلامی، محرم، شریفی آتشگاه، مسعود، و علیزاده‌لمجیری، صدیقه و طاهره زندی. (۱۳۸۳). «واژگان زبانی زبان فارسی». *مجموعه مقالات اولین کارگاه پژوهشی زبان فارسی و رایانه*.
- بی‌جن خان، محمود. (۱۳۹۲). *واجشناسی: نظریه بهینگی*. تهران: انتشارات سمت.
- کردزعفرانلو کامبوزیا، عالیه و رضا خیرآبادی. (۱۳۹۱). «بررسی رابطه ساخت هجایی با میزان زبایی وندهای اشتقاقی فعل‌افزای زبان فارسی». *پژوهش‌های زبان و ادبیات تطبیقی*. ۳(۳): ۱۴۳-۱۵۶.
- کردزعفرانلو کامبوزیا، عالیه، تاج‌آبادی، فرزانه، اسماعیلی متین، زهرا، و سارا خوردبین. (۱۳۹۵). «واج آرایی کلمات فارسی با ساخت هجایی CV.CVCC». *جستارهای زبانی*. ۷۵(۱): ۸۱-۱۰۷.
- فروهوشی، بهرام (۱۳۷۱). *فرهنگ فارسی به پهلوی*. دانشگاه تهران.

مکنزی، دیوید نیل (۱۳۷۳)؛ **فرهنگ کوچک زبان پهلوی**. برگردان مهشید میرفخرایی. پژوهشگاه علوم انسانی و مطالعات فرهنگی.

- Alamolhoda, M. (2000). *Phonostatistics and Phonotactics of the Syllable in Modern Persian*. Helsinki: The Finnish Oriental Society.
- Bartholomae, C. (2012). *Altiranisches Wörterbuch. In Altiranisches Wörterbuch*. De Gruyter Mouton.
- Jusczyk, P. W., Friederici, A. D., Wessels, J. M., Svenkerud, V., & A. M. Jusczyk. (1993). "Infants' sensitivity to the sound patterns of native language words". *Journal of Memory and Language*. 32(3): 402-420.
- Kramsky, J. (1939). "A Study in the Phonology of Persian". *Archiv Orientalni* (11): 66-83.
- Kramsky, J. (1948). "A Phonological Analysis of Persian Monosyllables". *Archiv Orientalni*. (16): 103-134.
- Liberman, I. Y., Shankweiler, D., Fischer, F. W., & B. Carter. (1974). "Explicit Syllable and Phoneme Segmentation in the Young Child". *Journal of Experimental Child Psychology*. 18: 201-212.
- McCarthy, J. (1988). "Feature Geometry and Dependency: A Review". *Phonetica*. 43: 84-108.
- Nourai, A. (1999). *An Etymological Dictionary of Persian, English, and other Indo-European Languages*. (n.p.).
- Pike, K. L. (1947). *Phonemics: A Technique for Reducing Languages to Writing*. Michigan: The University of Michigan Press.
- Zolfaghari Serish, M. (2004). *Phonological analysis of consonant clusters in Persian*. Tehran: Tarbiat Modares University.