

بسترها، چشم‌اندازها، کاربردها و چالش‌های مطالعه‌ی واژگان در متون دانشگاهی: ضرورت توجه بیشتر به زبان فارسی و زبان‌آموزان فارسی‌زبان

رضا غفارنمر^{۱*}، محسن شیرازی‌زاده^۲، غلامرضا کیانی^۳

۱. دانشیار آموزش زبان انگلیسی، دانشگاه تربیت مدرس، تهران، ایران

۲. دانشجوی دکتری آموزش زبان انگلیسی، دانشگاه تربیت مدرس / استادیار آموزش زبان انگلیسی، دانشگاه الزهرا (س)،

تهران، ایران

۳. دانشیار آموزش زبان انگلیسی، دانشگاه تربیت مدرس، تهران، ایران

پذیرش: ۹۲/۷/۲۲

دریافت: ۹۲/۴/۱۳

چکیده

در این مقاله، می‌کوشیم با نگاهی انتقادی کمبود توجه به واژه‌پژوهی در متون دانشگاهی را بررسی کنیم و با بررسی ابعاد مختلف این موضوع، چشم‌انداز روشنی از پیش‌زمینه‌ها، مسیرها و موانع موجود در این حیطه را ارائه کنیم. در این راستا، تمرکز این مقاله هم بر کمبود پژوهش در مورد زبان علمی فارسی است و هم بر کم‌توجهی پژوهش‌ها به زبان‌آموزان فارسی‌زبان. ابتدا تصویری کلی از اهمیت یادگیری و تسلط بر ابعاد مختلف واژگان در زبان دانشگاهی را بررسی می‌کنیم و در ادامه، برخی از پیکره‌ها و لیست‌های واژگانی پراهمیت را معرفی می‌نماییم که در ادبیات تحقیق وجود دارند. پس از تبیین اهمیت و ضرورت پژوهش در این حوزه، کوشیدیم تصویری چندبعدی از زوایای مختلف پژوهشی ارائه کنیم که باید مطرح نظر پژوهشگران این حوزه باشد و مورد بحث قرار گیرد. در این میان، مباحثی چون هم‌نشینی واژگان، زنجیره‌های واژگانی، نوسانات واژگانی درون‌متنی و بین‌متنی و برخی موضوعات دیگر را بررسی کردیم و در نهایت، برخی از چالش‌ها و موانعی را که پیش روی تحقیق در این حوزه قرار دارند معرفی نمودیم.

واژگان کلیدی: واژگان دانشگاهی، متون دانشگاهی، پیکره، زبان‌آموزان فارسی‌زبان، هم‌نشینی، زنجیره‌ی واژگانی.

E-mail: rgsamar@modares.ac.ir

* نویسنده‌ی مسئول مقاله:



۱. مقدمه

افزایش روزافزون رشته‌های دانشگاهی و بالا رفتن شمار پژوهشگران در حوزه‌های مختلف علوم در دو دهه اخیر، با رشد چشمگیر شمار مجلات علمی، پایان‌نامه‌ها و رساله‌ها همراه بوده است. این روند طبیعتاً باعث شده است که پژوهشگران و مؤلفان نیاز به تسلط بر خواندن (وارد، ۲۰۰۹؛ نرگیس، ۲۰۱۳) و نحوه مناسب نگارش متون علمی (بیچنر و باستورکمن، ۲۰۰۶؛ وینگیت، ۲۰۱۲) را بیشتر احساس کنند. از اوایل دهه هشتاد یا نود میلادی، زبان پژوهان نگاه ویژه‌ای به «زبان برای اهداف خاص» داشته‌اند (هاچینسون و واترز، ۱۹۸۷؛ دادلی ایوانز و سینت جان، ۱۹۹۸) و در این میان، «زبان برای اهداف دانشگاهی»^۱ نیز یکی از پرمخاطب‌ترین حوزه‌ها به‌شمار می‌آید؛ به‌طوری که در سال‌های اخیر، مطالعه ویژگی‌های زبان علمی به‌عنوان یک حوزه مستقل و کاربردی تحقیق مطرح شده است (جردن، ۱۹۹۷؛ هایلند، ۲۰۰۶). تاکنون جنبه‌های مختلفی از متون علمی مورد مطالعه فرار گرفته‌اند. برخی از محققان به جنبه‌های دستوری متون علمی توجه کرده‌اند (بایبر، ۲۰۰۶) و برخی دیگر به ابعاد گفتمانی (سمرج، ۲۰۰۸)، فراگفتمانی (لی و وارتنون، ۲۰۱۲) و یا «گونه‌شناختی» (لودرمیلک، ۲۰۰۷) و «حرکت‌شناختی» (هلک و کانر، ۲۰۰۶) آن‌ها. در این میان، شمار قابل توجهی از پژوهشگران نیز وقت خود را صرف مطالعه و بررسی جنبه‌های متعدد واژگانی متون علمی کرده‌اند (کاکسهد، ۲۰۰۰). با نگاهی گذرا به منابع و تحقیقات موجود در حوزه «زبان متون علمی» به‌طور عام و جنبه‌های واژگانی این متون به‌طور خاص، درمی‌یابیم که بیشتر پژوهش‌ها در این زمینه یا درمورد تولیدات زبانی نویسندگان بومی انگلیسی‌زبان صورت گرفته‌اند و یا به تولیدات زبانی آموزندگان زبان انگلیسی به‌عنوان زبان خارجی توجه کرده‌اند. با بررسی مختصر می‌توانیم بگوییم که تاکنون به مطالعه مدون متون علمی فارسی (به‌عنوان یک گونه خاص زبانی) توجه بسیار اندکی شده است. همچنین، با بررسی تحقیقات موجود درمی‌یابیم که حتی مطالعات در رابطه با تولیدات فارسی‌زبانان به زبان انگلیسی یا زبان‌های خارجی دیگر بسیار ناچیز و پراکنده است. در این مقاله، می‌کوشیم ضرورت‌ها، بسترهای لازم و موانع موجود در راه مطالعه متون علمی فارسی و متون علمی انگلیسی یا سایر زبان‌های خارجی را که توسط نویسندگان فارسی‌زبان نوشته می‌شوند تحلیل، نقد و بررسی کنیم؛ به سخن دیگر، می‌کوشیم

توجه زبان‌پژوهان را به این نکته جلب کنیم که در تحقیقات «زبان برای اهداف علمی»، هم به «زبان فارسی» و هم به «زبان‌آموزان فارسی‌زبان» بی‌توجهی شده است. در این مسیر، تأکید ما بر نیاز و چگونگی مطالعه واژگان و عبارتهای خاص متون علمی خواهد بود. به‌طور خاص، در مقاله پیش‌رو، می‌کوشیم به سؤالات اساسی زیر پاسخ دهیم:

۱. پژوهش و مطالعه در حوزه واژگان دانشگاهی از چه ابعادی دارای اهمیت است و این حوزه از پژوهش چه پیشینه‌ای در زبان‌های دیگر، به‌ویژه انگلیسی، دارد؟
 ۲. محققان زبان فارسی به کدام زوایای مختلف پژوهش در این حوزه باید توجه کنند؟
 ۳. یافته‌های این حوزه از پژوهش چه کاربردهایی در زمینه‌های مختلف خواهند داشت؟
 ۴. چالش‌های اصلی بر سر راه محققان در این حوزه مطالعاتی در ایران چیست؟
- گفتنی است که در این مقاله و در راستای تعریف مسئله و بیان ضرورت‌ها، بیشتر به مطالعاتی اشاره خواهیم کرد که در زبان انگلیسی انجام شده‌اند. این موضوع به این معنا نیست که چنین مطالعاتی در زبان‌های دیگر وجود ندارند؛ بلکه تنها به این دلیل که مطالعات صورت‌گرفته در زبان انگلیسی بسیار بیشتر و پرسابقه‌تر هستند، بیشتر این زبان را محور معرفی‌ها و بحث خود قرار داده‌ایم. همچنین، تبیین، تحلیل و نقد مسائل مطرح‌شده در مقاله پیش‌رو بیشتر از نگاه زبان‌شناسی بیکره‌ای صورت می‌گیرد.

۲. اهمیت واژگان در خواندن و نوشتن

یکی از مؤلفه‌های کلیدی تأثیرگذار در مهارت‌های زبانی، از جمله خواندن و نوشتن، میزان تسلط بر واژگان زبان است. این موضوع خود را در خواندن و نوشتن گونه‌های خاص متون به‌طور چشمگیرتری نشان می‌دهد. شاید بتوانیم ادعا کنیم که بدون تسلط بر واژگان خاص یک رشته، امکان نگارش مؤثر و یا درک کامل متون آن وجود ندارد. تاکنون تحقیقات بسیار زیادی در رابطه با تأثیر دانش واژگان بر فهم متون صورت گرفته است (چال، ۱۹۸۵؛ مک اون و همکاران، ۱۹۸۳؛ نیشن، ۲۰۰۱). میزان همبستگی دانش واژگان و فهم متون به‌طور کلی بین ۰٫۳ تا حدود ۰٫۴ گزارش شده است (کارل، ۱۹۹۱؛ برنهاردت، ۲۰۰۵). البته جانستون، ادمز و برونینگ (۱۹۸۵) رابطه بین آشنایی با کلمات تخصصی را مهم‌تر از آشنایی با کلمات عمومی ارزیابی می‌کنند و براساس یافته‌های خود اعلام می‌کنند که ضریب همبستگی بین فهم متن و



آشنایی با کلمات تخصصی ۰,۳۹ است؛ درحالی که رابطه فهم متن با آشنایی با کلمات عمومی ۰,۳۳ است.

نتایج بسیاری از پژوهش‌های دیگر نیز نشان می‌دهند که لازمه فهم مناسب یک متن، آشنایی با میزان قابل توجهی از واژگان آن است؛ به طوری که برخی یافته‌ها نشان می‌دهند که مهم‌ترین عامل تعیین‌کننده دشواری یک متن، میزان دشواری کلمات آن برای خواننده است (استاهل، ۲۰۰۳). اندرسون و فری بادی (۱۹۸۳) مدعی هستند که خواننده‌هایی که میزان بسیار زیادی از واژگان متن را ندانند، ناگزیر خواننده‌های بسیار ضعیفی خواهند بود. لافر و سیم (۱۹۸۵) نیز بیان می‌کنند که مهم‌ترین ضرورت برای موفقیت در خواندن متون علمی، آشنایی با واژگان است. آن‌ها در ادامه می‌افزایند که خواننده باید با حداقل ۶۵ تا ۷۰ درصد واژگان یک متن آشنا باشد. در تحقیقی دیگر، لیو و نیشن (۱۹۸۵) پیشنهاد می‌کنند که برای فهم روان یک متن، خواننده باید با حداقل ۹۵ درصد از واژگان آن آشنا باشد؛ درحالی که هیرش و نیشن (۱۹۹۲) مدعی شده‌اند که خوانندگانی که کمتر از ۹۵ تا ۹۸ درصد از واژگان یک متن را بدانند (چیزی حدود ۵۰۰۰ واژه طبق ادعای آن‌ها)، در خواندن به مشکل برمی‌خورند.

ارتباط بین دانش واژگان و نگارش نیز از موضوعاتی است که بسیاری از پژوهشگران آن را بررسی و تحلیل کرده‌اند. البته گفتنی است که میزان این مطالعات درمقایسه با مطالعات صورت‌گرفته در مورد ارتباط واژگان و خواندن بسیار کمتر است. لکی و کارسون (۱۹۹۴) در مطالعه‌ای، از شرکت‌کنندگان یک دوره زبان انگلیسی برای اهداف دانشگاهی در مورد نیازهای آموزشی آن‌ها پرسیدند. نتایج این تحقیق نشان داد که نیاز به آموزش واژگان برجسته‌ترین موضوعی بود که دانشجویان به آن اشاره کردند. در برخی دیگر از مطالعات، پژوهشگران متون نوشته‌شده توسط نویسندگان غیربومی انگلیسی را تجزیه و تحلیل کرده‌اند. در این راستا، پژوهشگران بسیاری (هارلی و کینگ، ۱۹۸۹؛ لافر و نیشن، ۱۹۹۵) نتیجه گرفته‌اند که کاربران توانا تر زبان دوم در نوشته‌های خود دامنه وسیع‌تری از واژگان را به کار می‌برند و همچنین واژگان و عباراتی که آن‌ها استفاده می‌کنند از سطح پیچیدگی بیشتری برخوردار است. تحقیقات دیگر نیز نشان می‌دهند که نوشته‌هایی که دارای دامنه واژگان وسیع‌تر هستند، همواره در مجموع نمره بهتری را کسب می‌کنند. در مطالعه جذاب دیگری، رینولدز (۱۹۹۵) نوشته‌های انگلیسی نویسندگان بومی و غیربومی را با هم مقایسه کرد. بخشی از نتایج این پژوهش نشان

دادند که نویسندگان غیربومی برای بیان معانی مشابه از گستره واژگانی وسیع‌تری استفاده می‌کنند؛ درحالی که نویسندگان غیربومی تنها چند واژه را تکرار می‌کنند. طبق نظر مؤلف، این ناتوانی نویسندگان غیربومی که به کم شدن انسجام متنشان منجر می‌شود، در دانش ضعیف واژگانی آن‌ها ریشه دارد. در پژوهش دیگری، کودا (۱۹۹۳) نوشته‌های دانشجویانی را بررسی می‌کند که درحال یادگیری زبان ژاپنی، به‌عنوان زبان دوم هستند. یافته‌های این مطالعه هم نشان داد که همبستگی بسیار زیادی ($R=0.7$) بین دانش واژگانی و کیفیت کلی نگارش این دانشجویان وجود داشت. تحلیل رگرسیون داده‌های این تحقیق نیز نشان داد که میزان گستردگی واژگان تنها مؤلفه پیش‌بینی‌کننده عملکرد نگارش است؛ به‌طوری که این مؤلفه نیمی از واریانس موجود در عملکرد نگارش را توجیه می‌کند.

دانستیم که شناخت و آشنایی با واژگان نقش بسیار پررنگی در توانایی خواندن و نوشتن و حتی شنیدن و صحبت کردن ایفا می‌کند. در متون علمی، این نقش با توجه به وجود واژگان خاص و همچنین وجود معانی خاص برای برخی واژه‌ها براساس نوع متنی که در آن آمده‌اند، ممکن است بسیار پررنگ‌تر نیز باشد. این اهمیت باعث شده است که از مدت‌ها پیش، پژوهشگران بسیاری به شناسایی و استخراج واژگان پراهمیت انواع مختلف متون توجه کنند. این تلاش‌ها در برخی از موارد به تهیه لیست‌های واژگان منجر شده است. در موارد دیگر نیز برخی از گروه‌های تحقیقاتی پیکره‌های منسجم و شکلی را از متون علمی تهیه کرده‌اند. این پیکره‌ها به نرم‌افزارهایی مجهز هستند که به کاربر خود امکان انجام انواع مختلف تحلیل واژگان را می‌دهند. در ادامه، برخی از این لیست‌ها و پیکره‌ها را معرفی می‌کنیم. بدیهی است که آنچه معرفی می‌کنیم، تنها لیست‌ها و پیکره‌هایی هستند که یا با نگاه به متون دانشگاهی تهیه شده‌اند و یا توجه خاصی به تولیدات زبانی دانشجویان غیرانگلیسی‌زبانی داشته‌اند که به زبان انگلیسی نوشته‌اند.

۲-۱. لیست واژگان دانشگاهی

شاید لیست واژگان دانشگاهی^۴ (AWL) که حاصل پروژه کارشناسی ارشد اوریل کاکسهد بود و در سال ۲۰۰۰ منتشر شد، جدی‌ترین و پرمخاطب‌ترین کوششی باشد که تاکنون در زمینه شناسایی و دسته‌بندی واژگان علمی در زبان انگلیسی صورت گرفته است. پیش از انتشار این



لیست، تنها لیست واژگان عمومی وست^۵ (۱۹۵۳) بر مبنای تحقیقی مدون منتشر شده بود و لیست‌ها و واژه‌نامه‌های دیگر که تعدادشان کم نیز نبود، پشتوانه پژوهشی و محاسباتی محکمی نداشتند. از سوی دیگر، لیست واژگان عمومی نیز جدای از انتقادات متعددی که به آن وارد بود (ریچاردز، ۱۹۷۴)، فقط پرکاربردترین واژگان عمومی انگلیسی را در خود گنجانده بود و نمی‌توانست در مطالعه متون علمی زبان انگلیسی، نقش پررنگی را ایفا کند. به هر حال، لیست کاکسهد توجه بسیار زیادی را به خود جلب کرد؛ به طوری که حتی مشهورترین لغت‌نامه‌های انگلیسی، از جمله لغت‌نامه آکسفورد و لانگمن نیز واژه‌های دانشگاهی این لیست را در ویرایش‌های جدید خود نشانه‌گذاری کردند.

این لیست که ۵۷۰ خانواده لغت را دربر گرفته، حاصل یک پژوهش پیکره‌شناختی نسبتاً جامع است. پیکره به‌کاررفته برای استخراج واژگان دانشگاهی شامل ۴۱۴ متن با طول نسبتاً یکسان بود که در مجموع ۳٫۵ میلیون واژه داشتند. این پیکره چهار گروه هنر، تجارت، حقوق و علوم را دربر می‌گرفت. کتب درسی، مقالات، کتاب‌ها و گزارشات آزمایشگاهی از جمله متون مختلف گنجانده‌شده در این پیکره هستند. کاکسهد (۲۰۰۰) چهار معیار را برای گزینش واژگان از پیکره جمع‌آوری‌شده به‌کار برد. اولین معیار این بود که واژگان انتخابی نباید جزو واژگان لیست واژگان عمومی وست (۱۹۵۳) باشند. معیار بعدی برای گزینش واژگان، بسامد آن‌ها بود. براساس معیار بسامد، تنها واژگانی می‌توانستند انتخاب شوند که حداقل ۱۰۰ بار در متون هریک از چهار رشته مورد بررسی به‌کار رفته باشند. سومین معیار، پراکنندگی بود که براساس آن، واژگان منتخب باید حداقل ۱۰ بار در ۱۵ رشته از ۲۸ رشته پیکره به‌کار می‌رفتند. براساس چهارمین معیار، یعنی معیار یکپارچگی، اعضای هریک از خانواده لغت باید حداقل ۱۰ بار در هر چهار بخش پیکره به‌کار رفته باشند. لیست واژگان دانشگاهی شامل ده لیست خرد است که اولین لیست خرد شامل پرکاربردترین خانواده‌های لغت و آخرین لیست خرد شامل کم‌کاربردترین آن‌ها است. براساس یافته‌های پژوهش کاکسهد (۲۰۰۰)، این لیست چیزی حدود ۱۰ درصد از واژگان متون علمی را پوشش می‌دهد؛ در نتیجه، می‌توانیم ادعا کنیم که ترکیب این لیست با لیست لغات عمومی وست (۱۹۵۳) می‌تواند چیزی حدود ۸۶ درصد از واژگان این متون را پوشش دهد که رقم قابل توجهی به‌نظر می‌رسد.

۲-۲. لیست زنجیره‌های واژگانی دانشگاهی (AFL)

یکی از حیطه‌های جدید مطالعه واژگان، بررسی عبارات چندکلمه‌ای است. عبارات چندکلمه‌ای یا زنجیره‌های واژگانی به مجموعه‌های متوالی از واژگان گفته می‌شود که به کنار هم بودن تمایل بسیار زیادی نشان می‌دهند؛ به طوری که احتمال دارد که این تمایل به با هم بودن نشان‌دهنده نوعی مابه‌ازای ذهنی در ساختار حافظه زبانی گویشوران است (الیس، ۱۹۹۶). اخیراً پژوهش‌های متعددی در این حوزه انجام شده است که بیشتر آن‌ها ترکیبی از زبان‌شناسی پیکره‌ای و روان‌شناسی زبان هستند. در این میان، یکی از کاربردی‌ترین مطالعات، پژوهش سیمپسون-والچ و الیس (۲۰۱۰) است. این محققان کوشیدند که با تکیه بر اصول محاسباتی زبان‌شناسی پیکره‌ای، لیستی از زنجیره‌های واژگانی خاص زبان علمی را استخراج و طبقه‌بندی کنند. آن‌ها با ترکیب دو شاخص «اطلاعات دوسویه^۷» و «میزان بسامد^۸»، شاخص جدیدی را با عنوان «شاخص ارزش تدریسی زنجیره‌ها^۹» ابداع کردند و آن را معیار گزینش خود قرار دادند. این پژوهشگران پس از تجزیه و تحلیل پیکره‌های بزرگ از زبان علمی کتبی و شفاهی توانستند سه لیست مدون از زنجیره‌های واژگان علمی را براساس معیارهای مشخص به دست آورند. لیست اول شامل ۹۷۹ زنجیره واژگانی مختص زبان شفاهی علمی است، لیست دوم شامل ۷۱۲ زنجیره واژگانی پرکاربرد مختص به متون علمی است و لیست سوم این مجموعه شامل ۲۰۷ مورد از زنجیره‌هایی است که بین گفتار و نوشتار علمی مشترک هستند. در ادامه این مطالعه، محققان لیست تهیه‌شده را براساس نقش زبانی عبارات طبقه‌بندی کردند تا برای استفاده در مطالب درسی و کلاس‌های آموزش زبان، کارآمدتر و کاربردی‌تر شود.

۲-۳. پیکره انگلیسی گفتار دانشگاهی میشیگان (MICASE^{۱۰})

یکی از پیکره‌هایی که به زبان علمی اختصاص یافته، پیکره انگلیسی گفتار دانشگاهی میشیگان (MICASE) است. این پیکره که به کوشش گروهی از استادان و پژوهشگران مرکز زبان انگلیسی دانشگاه میشیگان، واقع در آن آرپور (سیمپسون و دیگران، ۲۰۰۲) جمع‌آوری شده است، تنها به گفتار دانشگاهی اختصاص دارد. این مجموعه شامل حدود یک میلیون و هشتصد هزار کلمه است و به صورت اتفاقی از بین مجموعه گسترده‌ای از رشته‌های دانشگاهی



گردآوری شده است. جمع‌آوری آن به این صورت بوده است که رویدادهای زبانی متفاوت (سخنرانی‌های کلاسی، بحث‌های علمی و جلسات مشاوره‌ای) در دانشگاه میشیگان توسط گروه پژوهشی ضبط، با دقت پیاده‌سازی و به متن تبدیل شده‌اند. امکان جست‌وجوی بسیار کاملی در اینترنت^{۱۱} در این پیکره وجود دارد که نشان می‌دهد که این پروژه یک پروژه حرفه‌ای و حساب‌شده بوده است. برای نمونه، می‌توان جست‌وجوی یک واژه یا عبارت را با توجه به مسائلی، همچون جنسیت گوینده، رشته دانشگاهی، استاد یا دانشجو بودن گوینده و حتی نوع رویداد زبانی، محدود کرد.

۴-۲. پیکره پروژه‌های دانشجویان سطح بالای میشیگان (MICUSP^{۱۲})

پیکره پروژه‌های دانشجویان سطح بالای میشیگان نیز مجموعه ارزشمند دیگری است که بیش از ۸۳۰ مقاله کلاسی دانشجویان رشته‌های مختلف (زیست‌شناسی، تاریخ، روان‌شناسی، پرستاری، مهندسی مکانیک و ...) را که در حد عالی ارزیابی شده‌اند، در خود جای داده است. این پیکره شامل حدود دومیلیون و ششصد هزار واژه است. انواع مختلفی از پروژه‌ها، از جمله گزارش‌های کلاسی، پروپوزال‌های علمی، مقالات انتقادی و تحقیقات میدانی، در این پیکره گنجانده شده‌اند. این پیکره طوری طراحی و تدوین شده است که امکان جست‌وجو در بخش‌های خاص مقالات، مثل مقدمه یا نتیجه‌گیری را نیز به کاربر خود می‌دهد. از ویژگی‌های دیگر این مجموعه تفکیک بین نویسندگان بومی و غیربومی انگلیسی است. از مزیت‌های این ویژگی این است که با این کار، امکان مقایسه ویژگی‌های زبانی این دو گروه نیز برای کاربران فراهم می‌شود. گفتنی است که این پیکره پروژه‌های دانشجویان را براساس تعداد سال‌های اشتغال به تحصیل آن‌ها طبقه‌بندی کرده است. نسخه برخط این پیکره^{۱۳} نیز از لحاظ قابلیت‌های تحلیلی و ارائه نمودارهای مختلف قابل توجه است.

۵-۲. پیکره انگلیسی آمریکایی معاصر (COCA^{۱۴})

پیکره انگلیسی آمریکایی معاصر که با بیش از ۴۵۰ میلیون واژه، بزرگ‌ترین پیکره موجود زبان انگلیسی به‌شمار می‌آید، از معدود پیکره‌هایی است که هر ساله به‌روز می‌شود. این پیکره

همچنین از لحاظ ساختاری، پیکره‌ای بسیار متعادل به‌شمار می‌رود؛ به‌طوری که بخش‌های تشکیل‌دهنده آن (بخش گفتاری، شنیداری، روزنامه، داستان و ...) اندازه نسبتاً یکسانی دارند تا پیکره نسبت به یک گونه زبانی سویه‌مند نشود. یکی از بخش‌های پراهمیت این پیکره که شاید بتوان گفت که خود یک پیکره مستقل است، بخش دانشگاهی یا علمی است. این بخش از پیکره چیزی حدود ۸۵ میلیون واژه از مقالات علمی آمریکا در رشته‌های مختلف را در خود جای داده است. در این بخش از پیکره، به رشته‌های تعلیم و تربیت، تاریخ، پزشکی، حقوق، بازرگانی و ... اشاره شده است. یکی از نکات قابل توجه در مورد این مجموعه متن دانشگاهی این است که نرم‌افزار طراحی‌شده برخط^{۱۰} موجود روی آن اجازه جست‌وجو و تحلیل نسبتاً پیچیده‌ای را در رابطه با واژگان، عبارتهای چندکلمه‌ای و همنشین‌های واژگانی می‌دهد. نکته دیگر که بسیاری از محققان را مجذوب می‌کند این است که تمام واژه‌های این پیکره از لحاظ نقش زبانی برچسب‌گذاری شده‌اند؛ در نتیجه در آن، امکان جست‌وجوی نقش‌های دستوری نیز وجود دارد. برای روشن‌تر شدن موضوع، فرض می‌کنیم که یک محقق علاقه‌مند است که قیدهایی را بررسی کند که با واژه analyze در بخش دانشگاهی این پیکره آمده است. قابلیت‌های نرم‌افزاری این پیکره این امکان را به وی می‌دهد که لیستی از این قیدها را به ترتیب میزان بسامد و یا مؤلفه‌های دیگر -مانند شاخص وابستگی دوطرفه- ایجاد کند. حتی می‌توان برای نرم‌افزار تعریف کرد که تا چند واژه از سمت چپ و راست واژه analyze را جست‌وجو کند.

۶-۲. پیکره‌های مقالات انگلیسی زبان آموزان غیرانگلیسی زبان

نوع دیگری از پیکره‌ها که در سال‌های اخیر بسیار مورد توجه قرار گرفته‌اند پیکره‌هایی هستند که تولیدات زبانی دانشجویان غیرانگلیسی‌زبان به زبان انگلیسی را در خود جای داده‌اند. این نوع پیکره‌ها معمولاً از لحاظ اندازه کوچک‌تر از پیکره‌های عمومی‌تر هستند و بیشتر به مقالات تحلیلی موضوع‌محور کلاسی توجه کرده‌اند که در حدود سیصد کلمه هستند. مرکز زبان‌شناسی پیکره‌ای دانشگاه کاتولیک لوویان یکی از مراکزی است که به جمع‌آوری چنین پیکره‌هایی از زبان‌آموزان انگلیسی با زبان‌های مادری مختلف توجه زیادی کرده است. این مرکز مجموعه این پیکره‌ها را با عنوان «پیکره بین‌المللی آموزندگان زبان انگلیسی» (ICLE^{۱۱}) در قالب لوح فشرده منتشر کرده و نرم‌افزار درونی مجهزی را برای تحلیل پیکره روی آن تعیین نموده است.



تاکنون پیکره‌هایی با محتوای تولیدات زبان انگلیسی زبان‌آموزان با زبان‌های مادری مختلف (عربی، فرانسوی، آلمانی، ژاپنی، اسپانیایی، چینی، پرتغالی، ایتالیایی، مجاری، سوئدی، کاتالان، یونانی و ...) در مراکز مختلف تهیه شده‌اند. لیست کامل و مفصلی از این پیکره‌ها و جزئیاتشان در وبسایت مرکز زبان‌شناسی پیکره‌ای دانشگاه کاتولیک لوویان^{۱۷} پیدا می‌شود. پس از شناخت و بررسی این پیکره‌ها، آنچه بیش از همه توجه یک زبان‌پژوه فارسی‌زبان را جلب می‌کند، غفلت آشکار نسبت به زبان‌آموزان انگلیسی‌فارسی‌زبان است.

پروژه دیگری نیز با عنوان «پیکره بین‌المللی زبان انگلیسی» (ICE^{۱۸}) وجود دارد که گونه‌های مختلف زبان انگلیسی به‌عنوان زبان اول و دوم را در خود جای داده است. انگلیسی سریلانکا، اوگاندا، ایرلند، آمریکا، بریتانیا، هند، هنگ‌کنگ، غنا، فیلیپین، جامائیکا و کانادا^{۱۹} از گونه‌های موجود در این مجموعه هستند.

۳. ضرورت گسترده‌ی چشم‌انداز در مطالعه‌ی واژگان علمی

با نگاهی به آثار و ادبیات موجود در حوزه‌های واژه‌پژوهی و واژه‌نگاری در زبان فارسی، زبان‌آموزان فارسی‌زبان و آموزندگان زبان فارسی، درمی‌یابیم که مطالعه‌ی واژگان به‌صورت نظام‌مند و مدون در ایران دوران کودکی خود را سپری می‌کند. تحقیقات جامع و محکم در این حوزه نیازمند وجود بسترهای علمی متعددی است که هنوز فراهم نشده‌اند. ممکن است که یکی از این بسترها وجود آثاری باشد که بتوانند زمینه‌ی تحقیق را فراهم کنند و علاقه‌مندان به پژوهش را از جنبه‌های متعدد، مهم و درعین‌حال فراموش‌شده‌ی این حوزه آگاه سازند که این موضوع به‌نوعی هدف اصلی نگارش مقاله پیش‌رو است. در هر حال، واضح است که قدم گذاشتن در مسیر هر پژوهشی نیازمند داشتن نگاه جامع به ابعاد مختلف آن موضوع است. درزمینه‌ی پژوهش واژگان به‌طور عام و مطالعه‌ی واژگان علمی به‌طور خاص نیز باید پژوهشگر نگاهی بسیط و در عین حال جزئی‌نگر داشته باشد؛ وگرنه محصول پژوهشی وی نارس و یا کم‌اثر خواهد شد. در ادامه، به برخی از دیدگاه‌ها و نکاتی اشاره می‌کنیم که یک پژوهشگر واژگان باید در پژوهش خود به آن‌ها توجه کند.

۳-۱. همنشینی در واژگان علمی

یکی از نکاتی که محققان زبان از گذشته به آن توجه می‌کردند، باهم‌آیی یا همنشینی واژگان است. البته گفتنی است که در دهه گذشته و به دلیل ایجاد بسترهای لازم یارانه‌ای و الکترونیکی شدن متون، تحقیق در مورد این موضوع بیشتر شده است. تسلط بر عبارات چندکلمه‌ای و نحوه همنشینی دو یا چند واژه یکی از برجسته‌ترین مهارت‌هایی است که آموزندگان زبان دوم باید کسب کنند. چنین دانشی یکی از عناصر اصلی توانش زبانی به‌شمار می‌آید که به‌طور خاص برای رسیدن به سطوح عالی توانش زبانی ضروری به‌نظر می‌رسد (لافر و والدمن، ۲۰۱۱؛ پاولی و سایدر، ۱۹۸۳؛ ریچاردز، ۱۹۷۶؛ اشمیت، ۲۰۰۴؛ سینکلیر، ۱۹۹۱؛ رای، ۲۰۰۲). تسلط بر این دانش بخش مهمی از توانش واژگانی آموزندگان زبان دوم است و به‌نوعی از مؤلفه‌های اصلی کسب روانی کلام و نگارش در زبان خارجی به‌شمار می‌رود. همان‌طور که گلدهیل (۲۰۰۰) بیان می‌کند، بدون تسلط بر عبارت‌شناسی و آشنایی با الگوهای همنشینی، نوشتن به‌صورت حرفه‌ای نیز تقریباً غیرممکن به‌نظر می‌رسد. هیل (۱۹۹۹) در رابطه با ضرورت آشنایی با عبارات چندکلمه‌ای می‌گوید که دانشجویانی که گاه ایده‌های بسیار خوبی نیز برای نوشتن دارند، تنها به این دلیل که همنشین‌های واژگان کلیدی موضوعی را که راجع به آن می‌نویسند نمی‌دانند، ناگزیر نمره زیادی را از دست می‌دهند. این تسلط نداشتن بر همنشین‌ها زبان‌آموزان را مجبور می‌کند که برای بیان مفهوم خود، به استفاده از جملات طولانی‌تری متوسل شوند که به‌نوبه خود احتمال ارتکاب خطاهای زبانی را در متنی که تولید خواهند کرد بالا خواهند برد. به این ترتیب، توجه به شناسایی، طبقه‌بندی و در نظر گرفتن همنشین واژگان در مطالب درسی برای واژه‌پژوهان ضروری به‌نظر می‌رسد. با نگاهی به لغت‌نامه‌ها و منابع واژگانی دیگر در زبان‌های مختلف و به‌ویژه زبان انگلیسی، درمی‌یابیم که امروزه تمامی این منابع به موضوع همنشینی نیز توجه کرده‌اند. بدیهی است که با انجام تحقیقات لازم در این حوزه، بستر لازم برای تهیه لغت‌نامه‌های عمومی و تخصصی به‌روز زبان فارسی فراهم می‌شود. همچنین، نگاه به همنشینی و مطالعه در این زمینه، در تهیه مطالب درسی برای آموزش زبان فارسی به غیرفارسی‌زبانان کاربردهای فراوانی خواهد داشت.



۲-۳. عبارات چندکلمه‌ای یا زنجیره‌های واژگانی

نگاه به زنجیره‌های واژگانی یکی دیگر از ابعاد نو و جذاب واژه‌پژوهی است که در اصل همان همنشینی واژگان در معنای کلان‌تر و با نگاه پدیدارانه‌تر است. زنجیره‌های واژگانی بیشتر به باهم‌آیی بیش از دو واژه اطلاق می‌شود و البته با توجه به نوع زبان مورد تحقیق، شکل و ماهیت ویژه‌ای به خود می‌گیرد. عباراتی همچون «به‌طور کلی»، «در این صورت»، «با توجه به»، «شایان ذکر است که» و حتی «می‌توان این‌طور نتیجه گرفت که» نمونه‌های احتمالی از زنجیره‌های واژگانی هستند؛ زیرا هنوز پژوهشی مدون در زبان فارسی ثابت نکرده است که این مثال‌ها صلاحیت‌های لازم (از لحاظ ساختاری و محاسباتی) را برای شناخته شدن به‌عنوان زنجیره‌های واژگانی دارند. آنچه این زنجیره‌ها را به موضوع ارزشمندی برای پژوهش تبدیل کرده، فقط هویت مستقل متنی و پیکره‌ای آن‌ها نیست؛ بلکه اهمیت چشمگیر این عبارات در پردازش زبانی در ذهن (رای و پرکینز، ۲۰۰۰)، کاربرد قابل توجه شناخت و طبقه‌بندی آن‌ها برای آموزش زبان عمومی و دانشگاهی و طراحی مطالب درسی باعث اهمیت بررسی این زنجیره‌ها شده است.

۲-۳. نوسانات واژگانی در محصولات زبانی دانشگاهی

بدیهی است که هر رشته دانشگاهی ویژگی‌های منحصر به فرد زبانی خود را دارد و قطعاً این موضوع در مورد بعد واژگانی متون نیز صدق می‌کند. بسیاری از پژوهش‌های صورت‌گرفته در سایر زبان‌ها و به‌ویژه زبان انگلیسی، نشان می‌دهند که نوع نوسانات در واژگان به‌کاررفته در متون رشته‌های مختلف علاوه بر اینکه به رشته دانشگاهی وابسته است که متن به آن تعلق دارد، از مؤلفه‌های جزئی دیگری نیز تأثیر می‌پذیرد. داگلاس بایبر (۲۰۰۶) که یکی از پژوهشگران برجسته زبان دانشگاهی است، در پژوهش نسبتاً مفصل خود در رابطه با ویژگی‌های خاص زبان دانشگاهی در انگلیسی، به این موضوع اشاره می‌کند که در مجموع رشته‌های مرتبط با بازرگانی و مهندسی در مقایسه با رشته‌های علوم طبیعی، علوم اجتماعی و مخصوصاً علوم انسانی، تنوع واژگانی کمتری دارند و این موضوع خود را در کتب درسی-درمقایسه با تدریس‌های شفاهی کلاس- بیشتر نشان می‌دهد. بایبر (۲۰۰۶) همچنین واژگان خاص رشته‌های مختلف را بررسی می‌کند و نتیجه می‌گیرد که هر چند رشته‌های مختلف از واژگان ویژه خود برای بیان منظور استفاده می‌کنند، رشته‌های مرتبط با علوم طبیعی و

مهندسی واژگان خاصی را به کار می‌برند که به‌سختی برای عموم گویشوران قابل توضیح است؛ درحالی که واژگان تخصصی علوم انسانی را می‌توان با یک عبارت به افراد غیرمتخصص توضیح داد. گفتنی است که جزئیات نتایج پژوهش بایبر (۲۰۰۶) نشان می‌دهند که برخی از یافته‌ها پس از تفکیک گفتار دانشگاهی و نوشتار دانشگاهی و یا تفکیک زبان دانشگاهی به زیرگونه‌های مختلف (کتاب درسی، جزوه، ارائه کلاسی و ...)، شکل دیگری به خود می‌گیرند. نوسان بین‌رشته‌ای در سطح واژگانی به کلمات محدود نیست و برخی از تحقیقات اخیر نشان می‌دهند که حتی زنجیره‌های واژگانی نیز از حیث بسامد و محتوا در رشته‌های مختلف به گونه‌های مختلف به کار می‌روند. هایلند (۲۰۰۸ a, b) در دو پژوهش مستقل، به‌خوبی نشان می‌دهند که زنجیره‌های واژگانی در انواع متون دانشگاهی مربوط به رشته‌های مختلف، هم از حیث میزان شیوع و هم از حیث کاربرد، تفاوت‌های قابل توجهی دارند؛ به سخن دیگر، نویسندگان دانشگاهی در رشته‌های مختلف، از الگوهای خاص خود در به‌کارگیری زنجیره‌های واژگانی استفاده می‌کنند. نوسانات بین‌متنی در رشته‌های مختلف، در حوزه معنایی نیز دیده می‌شود. بایبر (۲۰۰۶) معتقد است که حجم قابل توجهی از اسامی به‌کاررفته در متون علمی (حدود ۵۰ درصد) را اسم‌هایی با معانی انتزاعی تشکیل می‌دهند. وی در ادامه می‌گوید که واژگانی که معانی مرتبط با مسائل ذهنی دارند (تصمیم گرفتن یا نتیجه گرفتن) در متون درسی بازرگانی و علوم انسانی درمقایسه با رشته‌های دیگر بیشتر وجود دارند. واژگانی با معانی بسیار انتزاعی و یا پروسه‌ای نیز در متون مهندسی و بازرگانی فراوان هستند. همچنین، متون مهندسی درمقایسه با متون رشته‌های دیگر بسیار بیشتر از واژگان ملموس و درعین‌حال غیرجاندار استفاده می‌کنند. یافته‌های بایبر (۲۰۰۶) نشان می‌دهند که تفاوت ملموسی بین متون رشته‌های مختلف در میزان و نوع استفاده از کلمات و قیود ربط وجود دارد.

کلیت این نوع تحقیقات به ما می‌آموزد که تغییرات و نوسانات زبانی در سطوح مختلف بین متون رشته‌های مختلف موضوعی نسبتاً مشخص و قابل تأمل است؛ در نتیجه علاقه‌مندان به پژوهش در زبان فارسی باید به این بعد نیز در تحقیقات خود درمورد متون علمی به‌طور عام و واژگان علمی به‌طور خاص توجه کنند. بررسی تولیدات زبانی فارسی‌زبانان به زبان انگلیسی یا زبان‌های خارجی دیگر می‌تواند فصل نوینی در پژوهش باشد؛ زیرا ممکن است که تولیدات زبانی این افراد در یک زبان خارجی ویژگی‌ها و الگوهای منحصربه‌فرد خود را داشته باشد و با



تولیدات زبانی کاربران بومی‌زبان و حتی کاربرانی با زبان‌های مادری دیگر متفاوت باشند (صفا و زمانی، ۱۳۹۱).

۴. کاربردهای پژوهش در واژگان دانشگاهی

یکی از مهم‌ترین دستاوردهای پژوهش در حوزه‌های مختلف کاربرد یافته‌های آن برای پیشبرد تحقیقات دیگر و یا برای جامعه بیرونی است. هرچند ممکن است بسیاری از پژوهش‌ها در حوزه شناخت و بررسی ابعاد زبان، ماهیتی بنیادی داشته باشند و نتوانیم کاربردی مستقیم و گسترده برای آن‌ها بیابیم، یکی از ویژگی‌های واژه‌پژوهی در متون علمی، کاربردهای ملموس و نسبتاً چشمگیری است که یافته‌های این حوزه از تحقیقات برای محققان، مدرسان، دانشجویان و اقشار دیگر جامعه دارند (کاکسهد، ۲۰۱۱). در ادامه، به بخشی از کاربردهای ممکن این حوزه از تحقیقات اشاره می‌کنیم.

۴-۱ کاربرد در تهیه و تدوین مطالب درسی و آزمون

تحقیقات مؤثر در مورد واژگان دانشگاهی به لیست‌های قابل استفاده‌ای از واژگان منجر می‌شود؛ از جمله لیست واژگان علمی فارسی، لیست همنشین واژگان علمی، لیست زنجیره‌های واژگانی در متون علمی و ... بدیهی است که این مجموعه‌ها و یافته‌های دیگر تحقیقاتی که این مجموعه‌ها را همراهی می‌کنند، زیرساخت بسیار مناسبی برای تهیه و تدوین مطالب درسی پژوهش‌محور هستند. تحقیقات گسترده در زمینه زبان علمی که البته به بعد واژگان محدود نمی‌شود، علاوه بر اینکه بسترهای لازم برای تهیه کتب آموزشی (در راستای آموزش زبان علم) را مهیا خواهند کرد، می‌توانند پیش‌زمینه‌ای برای تهیه و تدوین آزمون‌های استاندارد زبان دانشگاهی باشند. یکی از اهداف دو آزمون توانش زبان انگلیسی آیلتس^۲ و تافل^۱، سنجش میزان تسلط آزمون‌وران بر جنبه‌های مختلف زبان علمی است؛ در نتیجه، بسیاری از دانشگاه‌های معتبر می‌کوشند که با لازم دانستن نمرات این آزمون‌ها برای ثبت‌نام، از ورود افرادی که با زبان علمی انگلیسی آشنایی کافی ندارند و طبیعتاً در ادامه تحصیل با مشکل مواجه خواهند شد، جلوگیری کنند. در نمونه‌های کنونی این دو آزمون، سنجش دانش واژگان آزمون‌وران که اغلب در بطن مهارت‌های زبانی گنجانده می‌شود، یکی از پایه‌های اصلی سنجش توانایی زبانی به‌شمار می‌رود. واضح است که در آینده،

در صورت نیاز به چنین آزمون‌هایی در زبان فارسی، پیشینه پژوهشی غنی در رابطه با ابعاد مختلف زبان فارسی دانشگاهی ضروری است.

از سوی دیگر، تحقیق روی زبان‌آموزانی که زبان اولشان فارسی است، ابعاد مختلفی از ویژگی‌های لغوی این زبان‌آموزان را روشن می‌کند. نتایج چنین تحقیقاتی می‌تواند پیش‌زمینه‌ای برای تهیه مطالب درسی باشد که فقط برای زبان‌آموزان فارسی‌زبانی تألیف شده‌اند که می‌کوشند در یک زبان خارجی، فعالیت علمی و پژوهشی کنند. بدیهی است که در این مطالب درسی، باید نگاه ویژه‌ای به خطاهای متداول فارسی‌زبانان در گزینش واژگان و یا نحوه و مکان به‌کارگیری واژگان توسط این گروه از زبان‌آموزان وجود داشته باشد که این موضوع نیازمند انجام مطالعات تطبیقی است که الگوهای تولید زبان خارجی را با الگوهای تولید زبان علمی توسط گویشوران بومی مقایسه کرده‌اند.

به‌طور خلاصه می‌توان گفت که نتایج حاصل از مطالعات درمورد واژگان دانشگاهی در زبان فارسی و در تولیدات فارسی‌زبانان در یک زبان خارجی، در تهیه مطالب درسی برای سه گروه خاص مؤثر است. گروه اول دانشجویان و پژوهشگران فارسی‌زبان مبتدی هستند که آشنایی و تسلط کافی به متون علمی فارسی ندارند و در خواندن و به‌ویژه نوشتن مقاله و پایان‌نامه با مشکل روبه‌رو می‌شوند. دسته دوم پژوهشگران فارسی‌زبان هستند که به خواندن مقالات و یا انتشار آثار خود در مجلات بین‌المللی علاقه‌مند هستند؛ ولی نداشتن آشنایی کافی با نوع متون دانشگاهی در زبان خارجی باعث ایجاد مشکلات زیادی برای آن‌ها می‌شود. نکته بسیار قابل توجه در رابطه با این گروه این است که بسیاری از این پژوهشگران تسلط بسیار خوبی بر جنبه‌های علمی تحقیقات خود دارند و حتی می‌توانند به راحتی منابع رشته خود را در یک زبان خارجی بخوانند؛ اما نداشتن تسلط بر چگونگی نگارش که یکی از ابعاد اساسی آن تسلط بر انواع واژگان علمی است، آن‌ها را از انتشار آثارشان در مجلات بین‌المللی باز می‌دارد. گروه سومی که می‌توانند از مطالب درسی پژوهش‌محور بالا بهره ببرند، غیرفارسی‌زبانانی هستند که به فعالیت پژوهشی در رشته تحصیلی خود به زبان فارسی علاقه‌مند هستند. وجود تحقیقات جامع در زمینه الگوهای زبان دانشگاهی در زبان فارسی به‌طور عام و در حوزه واژگان به‌طور خاص، به حل دشواری‌های پیش‌روی آموزش زبان فارسی به غیرفارسی‌زبانان نیز کمک شایانی خواهد کرد.



۲-۴. کاربرد در ترجمه

کاربرد دیگر تحقیقات درمورد واژگان در متون علمی، کمک به معادل‌یابی دقیق برای کلمات است. پژوهش‌های جامع درمورد واژگان علمی، به ما اطلاعات بسیاری را در رابطه با ابعاد مختلف زندگی یک واژه می‌دهد. چنین پژوهش‌هایی میزان بسامد واژه در انواع متون را روشن خواهند کرد. رفتار واژه در همنشین‌گزینی در متون مختلف نیز از جمله اطلاعات دیگری است که از یافته‌های پژوهش‌ها به دست می‌آید. میزان کلیدی بودن واژه در انواع متون، ساختارهای خاص دستوری که واژه به آن‌ها تمایل بیشتری نشان می‌دهد (گریس و استفانویچ، ۲۰۰۴)، نقش‌های مختلفی که ریشه واژه در گونه‌های متنی مختلف به خود می‌گیرد، تغییرات بسامدی و دیگر تغییرات رفتاری که ممکن است واژه در بخش‌های مختلف یک متن علمی (مقدمه، روش‌شناسی، نتیجه‌گیری) از خود نشان می‌دهد، جزئی از اطلاعاتی هستند که نتایج این حوزه از پژوهش به ما ارائه خواهند کرد. چنین یافته‌هایی ممکن است که به ایجاد نوعی شناسنامه برای هر واژه در متون علمی منجر شوند. چنین شناسنامه‌ای که اطلاعات بسیار مفصلی (هم کمی و هم کیفی) از رفتار واژه در متون علمی را در خود جای داده است، ممکن است در گزینش معادل فارسی برای واژگان غیرفارسی و بالعکس بسیار رهگشا باشد. بدیهی است که وجود چنین شناسنامه‌ای این امکان را به مترجمان خواهد داد که با نگرشی بسیار جامع و با توجه به هر رشته تحصیلی، معادلی مناسب و پژوهش‌محور برای واژگان مختلف بیابند. گفتنی است که معادل‌یابی برای واژگان در متون علمی نتیجه کار گروهی بین زبان‌شناسان و متخصصان رشته‌های تحصیلی است؛ ولی در هر حال، نتایج واژه‌پژوهی برای هر دو گروه رهگشا و سودمند خواهد بود.

۳-۴. کاربرد در تهیه نرم‌افزارهای مختلف

کاربرد دیگر یافته‌های واژه‌پژوهی در متون علمی، ثمرات آن‌ها برای تهیه و تدوین نرم‌افزارهای آموزشی و پژوهشی است (فوتوس و برون، ۲۰۰۴). یکی از بارزترین کاربردهای چنین پژوهش‌هایی، تهیه نرم‌افزارهای لغت‌نامه جامع و چندرسانه‌ای است. پژوهش‌های مبتنی بر پروژه‌های پیکره‌شناسی زیرساخت بسیار مناسبی برای ایجاد لغت‌نامه‌های بسیار متفاوت و کارآمدتر از لغت‌نامه‌های سنتی در قالب نرم‌افزار هستند. نرم‌افزارهای حاصل از چنین

پژوهش‌هایی ویژگی‌هایی دارند که باعث می‌شوند که لغت‌نامه‌های سنتی کمتر به‌کار روند. این نرم‌افزارها با توجه به اینکه بر پایهٔ پیکره‌های جامع و مدون متون علمی و حتی متون عمومی بنا شده‌اند، امکانات بسیار قابل توجهی را برای کاربران خود فراهم می‌کنند. برای نمونه، این نرم‌افزارها علاوه بر معانی متفاوتی که برای هر واژه در رشته‌های مختلف ارائه می‌دهند، این امکان را به کاربر می‌دهند که تمام جمالاتی (از پیکره‌ای که لغت‌نامه روی آن سوار شده است) را که آن واژه در آن‌ها به‌کار رفته است، به‌عنوان مثال‌هایی از چگونگی کاربرد آن واژه در واقعیت ببیند. چنین محصولی به کاربر امکان جست‌وجوی هم‌نشین‌ها، مترادف‌ها، متضادها و شکل‌های لغوی متفاوت (صرف فعل، شکل جمع، کلمات مرتبط از لحاظ معنایی و ...) واژه مورد نظر را نیز می‌دهد. چنین نرم‌افزاری که بی‌شک نیازمند انجام پروژه‌های پیکره‌شناختی و تحقیقات گستردهٔ واژه‌پژوهی در متون علمی است، نقش یک دایره‌المعارف الکترونیک کم‌حجم و بسیار کاربرپسند را برای محققان و دانشجویان رشته‌های مختلف دانشگاهی ایفا می‌کند.

نتایج پروژه‌های گسترده و تحقیقات در زمینهٔ واژگان علمی، بخشی از زیرساخت‌های لازم برای تهیهٔ نرم‌افزارهای همیار در نگارش را نیز فراهم می‌کند. چنین نرم‌افزارهایی می‌توانند به‌صورت افزونه^{۲۲} روی برنامه‌های تایپ نصب شوند و با توجه به رشته‌ای که نگارنده قصد نگارش در آن را دارد، پیشنهادهای زبانی مختلفی را ارائه کنند و یا متن نوشته‌شده را از نظر زبانی ارزیابی کنند و اشتباهات زبانی و یا ناهماهنگی‌های شیوهٔ نگارش با گونهٔ متنی را متذکر شوند. به این ترتیب، نگارنده می‌تواند درحین نگارش یک متن علمی، از همیاری این افزونه‌ها بهره‌برد و حتی برخی از ابعاد نگارش خود را توسط چنین نرم‌افزارهایی ارزیابی کند.

از کاربردهای مفروض دیگر برای یافته‌های واژه‌پژوهی در متون علمی، کمک شایان به تهیهٔ نرم‌افزارهای آموزشی برای دانشجویان تحصیلات تکمیلی و محققان تازه‌کار است. شناخت ابعاد مختلف واژگانی متون علمی چه در زبان فارسی و چه در زبان‌های خارجی که فارسی‌زبانان به انتشار آثار خود در آن‌ها علاقه‌مند هستند، بستر لازم (نه کافی) را برای تهیهٔ نرم‌افزارهای خودآموز زبان علمی فراهم می‌کند. طبقه‌بندی واژگان براساس میزان بسامد و دشواری و ارائهٔ نمونه‌های مختلف از چگونگی کاربرد واژگان در رشته‌های خاص نمونه‌هایی از ابعاد آموزشی هستند که باید به آن‌ها در چنین نرم‌افزارهایی توجه شود. بدیهی است که این موارد و موارد مشابه آن نیازمند تحقیق در این حوزه هستند.



۵. موانع پیش‌روی پژوهش درباره‌ی واژگان دانشگاهی در ایران

تاکنون کوشیدیم که تصویری نسبتاً جامع از ضرورت، چگونگی و همچنین کاربردهای واژه‌پژوهی در زبان علمی ارائه کنیم. لازم است یادآوری کنیم که بحث ما در این مقاله از دو جنبه واکاوی شده است. از یک سو کوشیدیم ابعاد مختلف بحث را بر زبان فارسی علمی متمرکز کنیم و از سوی دیگر تلاش کردیم به تولیدات زبانی پژوهشگران فارسی‌زبان در زبان‌های خارجی نیز توجه کنیم. در این بخش، کاستی‌ها و موانع عمده‌ای را بررسی می‌کنیم که باعث شده‌اند روند تحقیق در واژگان دانشگاهی در زبان فارسی و یا تولیدات زبانی فارسی‌زبانان به زبان‌های دیگر بسیار کند و ناچیز باشد. این در حالی است که این حوزه از پژوهش در زبان‌های دیگر بسیار گسترده و چشمگیر است.

۵-۱. نبودن پیکره‌های زبانی لازم

واضح است که اساس و بنیان هر گونه تحقیق جامع توصیفی، تحلیلی و یا مقایسه‌ای در واژگان یک زبان و یا یک گونه زبانی، وجود پیکره‌های مدون، همگن و متعادل است (گریس، ۲۰۰۹). بیشتر تحقیقاتی که در زبان‌های دیگر درمبحث واژگان انجام شده‌اند، به پشتوانه پیکره‌هایی بوده‌اند که در بخش ابتدایی مقاله پیش‌رو به برخی از آن‌ها اشاره کردیم. با انجام جست‌وجویی مختصر درمی‌یابیم که تعداد معدودی پیکره در زبان فارسی تهیه شده‌اند (دادگان زبان فارسی به کوشش عاصی^۳؛ پیکره بی‌جن خان و همکاران، ۲۰۰۹؛ پیکره وابستگی نحوی زبان فارسی^۴). ازجمله مشکلات موجود برای همه این پیکره‌ها تنوع و حجم نسبتاً کم متون آن‌ها در حال حاضر است؛ به‌طوری که از یک پیکره جامع ملی فاصله دارند. همچنین، به‌نظر می‌رسد که برخی از این پیکره‌ها باوجود زحمات پدیدآورندگان در جمع‌آوری و حتی نشانه‌گذاری‌های دقیق (پیکره بی‌جن خان)، فقط برای اهداف تحقیقاتی خاص ایجاد شده‌اند؛ درنتیجه در دسترس مخاطبان خاصی قرار دارند و در حال حاضر امکان دسترسی عمومی به آن‌ها ازسوی کاربران اینترنتی وجود ندارد. البته پیکره نسبتاً جامع‌تر زبان فارسی که با عنوان «دادگان زبان فارسی» به‌کوشش دکتر مصطفی عاصی گردآوری شده است، نسخه برخطی در اینترنت دارد؛ اما جست‌وجویی مختصر در این پیکره و مقایسه سطحی آن با پیکره‌های سایر زبان‌های مطرح

دنیا نشان می‌دهد که این نسخه نیز با وجود امکانات و تلاش بسیاری که در تهیه آن صرف شده، از لحاظ عملکرد دچار اختلال‌های مکرر و جدی است و کاربران در استفاده از آن با مشکل مواجه می‌شوند. به نظر می‌رسد که نحوه طبقه‌بندی متون در این پیکره، میزان کاربرپسند بودن جست‌وجو و همچنین تنوع ارائه برون‌داد برای جست‌وجوها همچنان نیازمند بررسی و تغییرات است و مجری این پیکره این نقاط ضعف را تأیید می‌کند (عاصی، ۱۳۹۲: گفتگوی شخصی).

نکته مهم دیگر این است که تاکنون برای جمع‌آوری پیکره‌ای تخصصی از متون علمی فارسی تلاشی نشده است و کمبود این زیرساخت کمبود تحقیقات در این موضوع را نیز به همراه داشته است. پیکره‌ای مدون از متون علمی فارسی با طبقه‌بندی متون براساس نوع زبان (گفتار و نوشتار)، کلان‌رشته‌ها (علوم انسانی، مهندسی و ...)، ریزرشته‌ها (تاریخ، جغرافی، مهندسی مخابرات و ...)، تاریخ انتشار، نوع متن (کتاب، مقاله، پایان‌نامه، گزارش کلاسی و ...)، دانشگاه محل خدمت نویسندگان (تربیت مدرس، صنعتی شریف، شیراز و ...) و بسیاری از موارد دیگر بی‌شک گنجینه‌ای ارزشمند و زیرساختی گرانبها برای پژوهش در زبان علمی فارسی است.

مسئله دیگر در رابطه با کاستی‌های زیرساختی تحقیقات واژه‌پژوهی نبودن هرگونه پیکره (خرد یا کلان) از متون یا گفتار فارسی‌زبانان در یک زبان خارجی است. همان‌طور که پیش‌تر گفتیم، بسیاری از کشورهای دنیا علاوه بر اینکه پیکره‌ای بسیار جامع و کلان از زبان رسمی کشور خود را تهیه کرده‌اند، پیکره‌هایی از تولیدات گویشوران زبان کشورشان به زبان‌های خارجی (به‌ویژه انگلیسی) دارند. مشخص است که وجود پیکره‌ای شکل از تولیدات زبانی فارسی‌زبانان در هر زبان خارجی، پیش‌زمینه انجام تحقیقات و تحلیل‌های متنوعی در رابطه با ویژگی‌های خاص این گروه از زبان‌آموزان است و نتایج چنین تحقیقاتی برای آموزش آن زبان خارجی به فارسی‌زبانان بسیار مفید خواهد بود.

نوع سوم از پیکره‌هایی که تاکنون هیچ تلاشی برای گردآوری آن‌ها صورت نگرفته است و به دلیل تازگی مباحث این حوزه از مطالعات چندان نیز عجیب نیست، پیکره‌ای از تولیدات فارسی زبان‌آموزان غیرفارسی‌زبان است. با شکل‌گیری رشته آموزش زبان فارسی به غیرفارسی‌زبانان در دانشگاه‌های مختلف کشور و گسترش مباحث و تحقیقات در این حوزه،



باید پیش‌زمینه‌هایی برای جمع‌آوری و ساماندهی چنین پیکره‌ای نیز فراهم شود تا با تکیه بر یافته‌های مستند به آن، گام‌های مؤثری در زمینه آموزش کارآمدتر و اختصاصی‌تر زبان فارسی به زبان‌آموزانی با زبان‌های مادری متفاوت برداشته شود.

۲-۵. کمبود ادبیات و پیشینه تحقیق

از کاستی‌های دیگر موجود در حوزه واژه‌پژوهی در متون علمی به‌طور خاص و واژه‌پژوهی به‌طور عام در کشور ما، کمبود منابع و تحقیقات کافی در این زمینه است. برای نمونه، جست‌وجوی هم‌زمان دو واژه «پیکره» و «زبان» در همه بخش‌های قابل جست‌وجوی مقالات در پایگاه استنادی علوم جهان اسلام^{۲۰}، از بین بیش از ۳۵۰ هزار مقاله تنها به سه مقاله منتج شد که این موضوع حتی با احتساب میزانی از خطا در جست‌وجو، نشان‌دهنده کمبود تحقیقات در حوزه زبان‌شناسی پیکره‌ای است. مسلم است که وجود منابع کافی به زبان فارسی و انجام تحقیقات مختلف در این حوزه نقطه آغازی برای چنین پژوهش‌هایی است و پژوهش‌های اولیه در قالب تحقیق، تألیف و ترجمه، منابع و ادبیات مناسبی برای محققان دیگر به‌شمار می‌روند. این مسئله یکی از اهداف عمده و از جنبه‌های نو در مقاله پیش‌رو است. باید بگوییم که یکی از ضروری‌ترین اقدامات پژوهشی در زمینه واژه‌پژوهی در متون علمی، ترجمه منابع کلاسیکی است که در حوزه زبان‌شناسی پیکره‌ای، واژه‌پژوهی و زبان برای اهداف دانشگاهی نگاشته شده‌اند (سینکلیر، ۱۹۹۱؛ میر، ۲۰۰۲؛ کندی، ۱۹۹۸). ترجمه برخی از این آثار به زبان فارسی باعث می‌شود که محققان تازه‌کاری که با زبان‌های خارجی آشنایی زیادی ندارند، با این مباحث آشنا شوند و قدم در راه پژوهش در حوزه زبان‌شناسی پیکره‌ای، واژگان و متون دانشگاهی بگذارند. بدیهی است که گستره این تحقیقات هم شامل زبان فارسی به‌عنوان زبان مادری، هم شامل آموزش زبان فارسی به غیرفارسی‌زبانان و هم شامل آموزش زبان‌های خارجی به فارسی‌زبانان خواهد شد؛ در نتیجه کاربردهای متنوع و گسترده‌ای در بخش‌های مختلف آموزش، ترجمه، تهیه مطالب درسی و ... خواهد داشت.

۳-۵. توجه نکردن مسئولان ذی ربط

شاید یکی از کاستی‌های زیربنایی و عمده که باعث کم‌توجهی به پژوهش‌های پیکره‌شناختی در حوزه واژگان و متون دانشگاهی شده است، کم‌توجهی برخی از تصمیم‌گیران و مسئولان باشد. نهادهایی همچون فرهنگستان ادب فارسی^{۲۶}، پژوهشگاه علوم و فناوری اطلاعات ایران (ایران داک)^{۲۷}، پایگاه اطلاعات علمی جهاد دانشگاهی^{۲۸}، بانک اطلاعات نشریات کشور^{۲۹}، کتابخانه‌های دانشگاهی و عمومی بزرگ و بسیاری از نهادهای دیگر می‌توانند با ایجاد هماهنگی‌های لازم، زمینه ایجاد پیکره و تحقیقات وسیع در زمینه واژگان و متون علمی را فراهم کنند. پژوهشگاه علوم و فناوری اطلاعات ایران مجموعه بسیار بزرگی از پایان‌نامه‌های کارشناسی ارشد و دکتری نوشته‌شده در سال‌های اخیر را به‌صورت دیجیتال در اختیار دارد. بانک اطلاعات نشریات کشور نیز متن الکترونیک تعداد بسیار زیادی از مجلات علمی پژوهشی کشور را در خود جای داده است. منابع موجود در همین دو سازمان، گنجینه ارزشمندی از متون مورد نیاز برای ایجاد پیکره متون دانشگاهی فارسی را فراهم می‌کند. اختصاص بودجه‌ها و گرانت‌های پژوهشی از سوی سازمان‌های نام‌برده و نهادهای مرتبط دیگر نیز مشوق بسیار خوبی برای ایجاد بسترهای لازم و انجام تحقیقات متنوع در ابعاد مختلف متون فارسی و غیرفارسی دانشگاهی خواهد بود. آگاه کردن مسئولان مرتبط از طرح‌های تحقیقاتی و انجام رایزنی‌های لازم فضای بسیار بهتری را برای تحقیق در حوزه واژه‌پژوهی و متون دانشگاهی ایجاد می‌کند (این موضوع یکی از اهداف مقاله پیش‌رو است).

۶. نتیجه‌گیری و سخن آخر

مطالعه محاسباتی و آماری ابعاد مختلف متون، یکی از حوزه‌های درحال‌رشد زبان‌شناسی است. برخی از ابعاد این مطالعات با عنوان «زبان‌شناسی محاسباتی» و برخی دیگر با عنوان «زبان‌شناسی پیکره‌ای» شناخته می‌شوند. آنچه این مطالعات را از برخی دیگر از حوزه‌های بررسی زبان، مانند بررسی نحوی (چامسکی، ۲۰۰۲)، متمایز می‌کند، بنیان داده‌محور آن‌ها است. مطالعات پیکره‌محور بر زبان رخ داده^{۲۰} بنا شده‌اند، نه زبان ممکن‌الوقوع^{۲۱} (ویدوسون، ۲۰۰۳) و از این جهت، نتایج چنین مطالعاتی مبتنی بر واقعیت روزمره زبان است، نه ذهنیت گویشوران بومی. در این میان، مطالعه و بررسی واژگان و رفتار متنی آن‌ها حوزه‌ای گریا و گسترده است. تلفیق مطالعه پیکره‌محور واژگان با مطالعه متون دانشگاهی مولد پرسش‌ها و جستارهای گوناگون و عمیق است. در این مقاله، کوشیدیم اهمیت تحقیق در واژگان دانشگاهی را برای خوانندگان روشن و کاربردهای آن را بازگو کنیم. در خلال بحث، همواره به‌طور مستقیم و غیرمستقیم به این موضوع اشاره کردیم که تاکنون در ایران، پژوهشگران به مطالعه واژگان دانشگاهی، آن‌گونه که لازم است، توجه نکرده‌اند. ما دلیل این کم‌توجهی را به دو بخش کلی تقسیم کردیم و کوشیدیم در مباحث خود همواره به ضرورت توجه بیشتر به آن‌ها اشاره کنیم. دلیل نخست این است که محققان ایرانی تاکنون آن‌گونه که لازم است، واژگان علمی در زبان فارسی را مطالعه نکرده‌اند. در اینجا، باید به این نکته اشاره کنیم که مطالعه واژگان علمی در زبان فارسی مبحث کلانی است و ریزبخش‌هایی چون مطالعه واژگان دانشگاهی در پایان‌نامه‌ها، کتاب‌ها و مقالات یا مطالعه این واژگان در رشته‌های مختلف و یا در متون نگاشته‌شده به فارسی توسط غیرفارسی‌زبانان و انواع دیگر متون را دربر می‌گیرد. بخش دیگری از کم‌توجهی به واژگان دانشگاهی در ایران به بررسی واژگانی در متون غیرفارسی مربوط می‌شود که توسط فارسی‌زبانان نوشته شده‌اند. با وجود تحقیقات گسترده‌ای که در زبان‌های مختلف، به‌ویژه زبان انگلیسی، در رابطه با ویژگی‌های واژگانی نویسندگان غیربومی با زبان‌های مادری مختلف انجام شده است، تحقیقات در مورد متون نوشته‌شده در زبان‌های خارجی توسط فارسی‌زبانان بسیار ناچیز است.

در بخش دیگری از مقاله، کوشیدیم نمایی کلی‌نگرانه نسبت به حیطه‌های مختلف واژه‌پژوهی



ارائه کنیم تا به این ترتیب، خوانندگان با زمینه‌های کلی و بخشی از پرسش‌های رایج در این حوزه آشنا شوند. از موضوعات دیگری که در این مقاله بررسی کردیم، چالش‌ها و موانع پیش‌روی مطالعه واکژگان دانشگاهی بود؛ موانعی از قبیل کمبود پیکره‌های لازم، نبودن مشوق‌های پژوهشی، کمبود ادبیات و پیشینه تحقیق به زبان فارسی و کم‌توجهی مسئولان ذی‌ربط. در مجموع کوشیدیم در این مقاله، از ابعاد مختلف مطالعه واکژگان در متون دانشگاهی پرده برداریم و با بیان چالش‌ها، چشم‌اندازها، بسترهای لازم و کاربردهای این حوزه از مطالعات، افق روشنی از این مسیر پژوهشی ناپیموده را ارائه کنیم. در شکل ۱، خلاصه‌ای از بخش‌های بررسی‌شده در این مقاله را ارائه کردیم که نوعی نقشه راه برای علاقه‌مندان به پژوهش در این حوزه به‌شمار می‌رود. امید است که بخش‌های مطرح‌شده در این مقاله، ضرورت توجه بیشتر به واژه‌پژوهی با تمرکز بر متون دانشگاهی را روشن و علاقه‌مندان به پیمودن این مسیر را تاحدی از موانع موجود در آن، پیچ‌وخم‌های آن، نقشه راه و افق روشن پیش‌رو آگاه کرده باشند.

۷. پی‌نوشت‌ها

1. English for academic purposes
2. genre analysis
3. move analysis
4. academic word list
5. general service list
6. academic formula list
7. mutual information
8. frequency
9. formula teaching worth
10. Michigan corpus of academic spoken English
11. <http://quod.lib.umich.edu/m/micase/>
12. Michigan corpus of upper-level student papers
13. <http://micusp.elicorpora.info/>
14. corpus of contemporary American English
15. <http://corpus.byu.edu/coca/>
16. international corpus of learner English
17. <https://www.uclouvain.be/en-cecl.html>
18. international corpus of English
19. <http://ice-corpora.net/ice/avail.htm>

20. IELTS
21. TOEFL
22. add-on
23. <http://pldb.ihcs.ac.ir/>
24. <http://www.dadegan.ir/perdt>
25. <http://www.isc.gov.ir/Default.aspx?lan=fa>
26. <http://www.persianacademy.ir/>
27. <http://www.irandoc.ac.ir/>
28. <http://www.sid.ir/fa/>
29. <http://www.magiran.com/>
30. attested language
31. possible language

۸. منابع

- صفا، پریش و مینا زمانی (۱۳۹۱). «بررسی مقایسه‌ای باهم‌آیی‌ها در کنش زمانی بیان شدت در فرانسه و فارسی». پژوهش‌های زبان و ادبیات تطبیقی (جستارهای زبانی). د ۳. ش ۳. صص ۱۲۳-۱۴۱.

References:

- Biber, D. (2006). *University Language: A Corpus-based Study of Spoken and Written Registers*. Vol. 23. Amsterdam: John Benjamins Publishing Company.
- Bijankhan, M.; J. Sheykhzadegan; M. Bahrani & M. Ghayoomi (2011). "Lessons from building a Persian written corpus: Peykare". *Language Resources and Evaluation*. 45(2). pp. 143- 164.
- Bitchener, J. & H. Basturkmen (2006). "Perceptions of the difficulties of postgraduate L2 thesis students writing the discussion section". *Journal of English for Academic Purposes*. 5(1). pp. 4- 18.
- Carrell, P.L. (1991). "Second language reading: Reading ability or language proficiency?". *Applied Linguistics*. 12(2). pp. 159- 179.
- Chall, J., & S. Stahl (1985). "Reading comprehension research in the past decade: Implications for educational publishing". *Book Research Quarterly*. 1(1). pp. 95-



102.

- Chomsky, N. (2002). *Syntactic Structures*. Berlin: De Gruyter Mouton.
- Conner Loudermilk, B. (2007). "Occluded academic genres: an analysis of the mba thought essay". *Journal of English for Academic Purposes*, 6(3), pp. 190- 205.
- Coxhead, A. (2000). "A new academic word list". *TESOL Quarterly*, 34(2), pp. 213-238.
- -----, (2011). "The academic word list 10 years on: Research and teaching implications". *TESOL Quarterly*, 45(2), pp. 355- 362.
- Dudley-Evans, T. & M.J. Saint John (1998). *Developments in English for Specific Purposes*. Cambridge: Cambridge University Press
- Ellis, N. (1996). "Sequencing in SLA". *Studies in Second Language Acquisition*, 18(1), pp. 91-126.
- Fotos, S. & C.M. Browne (2004). *New Perspectives on CALL for Second Language Classrooms*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Gledhill, C.J. (2000). *Collocations in Science Writing*. Vol. 22. Tübingen: Gunter Narr Verlag.
- Gries, S.T. (2009). *Quantitative Corpus Linguistics with R: A Practical Introduction*. New York: Routledge.
- Gries, S.T. & A. Stefanowitsch (2004). "Extending collocation analysis: A corpus-based perspective on `alternations`". *International Journal of Corpus Linguistics*, 9(1), pp. 97- 129.
- Halleck, G. B. & U.M. Connor (2006). "Rhetorical moves in TESOL conference proposals". *Journal of English for Academic Purposes*, 5(1), PP. 70- 86.
- Harley, B. & M.L. King (1989). "Verb lexis in the written compositions of young L2 learners". *Studies in Second Language Acquisition*, 11(4), PP. 415- 439.
- Hill, J. (1999). *Collocational Competence*. *ETP English Teaching Professional*, Issue 11.
- Hirsh, D. & P. Nation (1992). "What vocabulary size is needed to read unsimplified

texts for pleasure?”. *Reading in a Foreign Language*. 8(2). pp. 689- 689.

- Hutchinson, T. & A. Waters (1987). *English for Specific Purposes*. Cambridge: Cambridge University Press.
- Hyland, K. (2006). *English for Academic Purposes: An Advanced Resource Book*. New York: Routledge.
- ----- (2008a). “Academic clusters: Text patterning in published and postgraduate writing”. *International Journal of Applied Linguistics*. 18(1). pp. 41-62.
- ----- (2008b). “As can be seen: Lexical bundles and disciplinary variation”. *English for Specific Purposes*. 27(1). pp. 4- 21.
- Johnson, C. W., Adams, M., & Bruning, R. (1985). Keywords and vocabulary acquisition: Some words of caution about words of assistance. *ECTJ*. 33(2). PP. 125-138.
- Jordan, R.R. (1997). *English for Academic Purposes: A Guide and Resource Book for Teachers*. Cambridge: Cambridge University Press.
- Kennedy, G. (1998). *An Introduction to Corpus Linguistics*. London: Longman.
- Koda, K. (1993). “Task-induced variability in FL composition: Language-specific perspectives”. *Foreign Language Annals*. 26(3). pp. 332- 346.
- Laufer, B. & I. Nation (1995). Lexical richness in L2 written production: Can it be measured. *Applied linguistics*. 16(3). pp. 307- 322.
- Laufer, B. & D.D. Sim (1985). *Taking the Easy Way out: Non-use and Misuse of Clues in EFL Reading*. Paper Presented at the English Teaching Forum.
- Laufer, B. & T. Waldman (2011). “Verb-noun collocations in second language writing: A corpus analysis of learners’ English”. *Language Learning*. 61(2). pp. 647-672.
- Leki, I. & J.G. Carson (1994). “Students’ perceptions of EAP writing instruction and writing needs across the disciplines”. *Tesol Quarterly*. 28(1). pp. 81- 101.
- Li, T. & S. Wharton (2012). “Metadiscourse repertoire of L1 Mandarin



undergraduates writing in English: A cross-contextual, cross-disciplinary study". *Journal of English for Academic Purposes*. 11(4), pp. 345- 356.

- Liu, N., & Nation, I. (1985). Factors affecting guessing vocabulary in context. *RELC journal*. 16(1), pp. 33- 42.
- McKeown, M.G.; I.L. Beck; R.C.; Omanson & C.A. Perfetti (1983). "The effects of long-term vocabulary instruction on reading comprehension: A replication". *Journal of Literacy Research*. 15(1), pp. 3- 18.
- Meyer, C.F. (2002). *English Corpus Linguistics: An Introduction*. Cambridge: Cambridge University Press.
- Nation, I.S.P. (2001). *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.
- Nergis, A. (2013). "Exploring the factors that affect reading comprehension of EAP learners". *Journal of English for Academic Purposes*. 12(1), pp. 1- 9.
- Pawley, A. & F.H. Syder (1983). "Two puzzles for linguistic theory: Nativelike selection and nativelike fluency". In J.C. Richards and R.W. Schmidt (Eds.). *Language and Communication*. New York: Longman.
- Reynolds, D.W. (1995). "Repetition in nonnative speaker writing". *Studies in Second Language Acquisition*. 17(02), pp. 185- 209.
- Richards, J.C. (1974). "Word lists: Problems and prospects". *RELC Journal*. 5(2), pp. 69-84.
- -----, (1976). "The role of vocabulary teaching". *TESOL Quarterly*. 10(1), pp. 77- 89.
- Safa, P. & M. Zamani (2012). "A Comparison study of the collocations in linguistic action in Persian and French language". *Language Related Research*. Vol. 3. No. 3. pp. 123-141 [In Persian].
- Samraj, B. & L. Monk (2008). "The statement of purpose in graduate program applications: Genre structure and disciplinary variation". *English for Specific Purposes*. 27(2), pp. 193- 211.

- Schmitt, N. & R. Carter (2004). "Formulaic sequences in action". *Formulaic Sequences: Acquisition, Processing and use*. pp. 1- 22.
- Simpson-Vlach, R. & N.C. Ellis (2010). "An academic formulas list: New methods in phraseology research". *Applied Linguistics*. 31(4). pp. 487- 512.
- Sinclair, J. (1991). *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- Stahl, S.A. (2003). "Vocabulary and readability: How knowing word meanings affects Comprehension". *Topics in Language Disorders*. 23(3). pp. 241- 247.
- Ward, J. (2009). "EAP reading and lexis for Thai engineering undergraduates". *Journal of English for Academic Purposes*. 8(4). pp. 294- 301.
- West, M.P. (1953). *A General Service List of English Words*. London: Longman.
- Wingate, U. (2012). "Using academic literacies and genre-based models for academic writing instruction: A 'literacy' journey". *Journal of English for Academic Purposes*. 11(1). pp. 26- 37.
- Wray, A. (2002). *Formulaic Language and the Lexicon*. Cambridge: Cambridge University Press.
- Wray, A. & M.R. Perkins (2000). "The functions of formulaic language: An integrated model". *Language & Communication*. 20(1). pp. 1- 28.