


The Use of Lexical Bundles by Native English Authors in Applied Linguistics: A Corpus-Driven Study

Vol. 13, No. 6, Tome 72
pp. 541-569
January & February
2023

Mohammad Hassanzadeh^{1*}  & Hadith Tamleh² 

Abstract

Lexical bundle research has recently come to the forefront of corpus-driven studies. Previous corpus studies have documented conflicting results regarding the frequency and function of lexical bundles (LBs) in academic prose. To date, however, no study has exclusively investigated LBs in the "discussion" sections of research articles generated by professional native English authors. The current study addressed this gap by examining the frequency, structure, and function of the most frequent four-word LBs. The corpus was composed of the *discussions* of published research papers authored by native (L1) writers. The data were extracted from five reputable international journals in the field of applied linguistics, consisting of over 300,000 words. Using AntConc, all the lexical sequences were retrieved with a frequency of 10 and a range of 5. The results revealed that LBs were predominantly used by English writers. Structurally, it was found that phrasal bundles were the most frequent in our corpus. The findings also demonstrated that functionally, referential bundles were extensively employed. In addition, stance bundles and discourse organizing bundles were the most prevalent after referential bundles. Finally, the findings are discussed in terms of the implications for non-native writers regarding the use of LBs in academic prose.

Keywords: AntConc; corpus; discussion; lexical bundles; native English writers.

Received: 20 February 2021
Received in revised form: 15 May 2021
Accepted: 7 May 2021

1. Assistant Professor of Applied Linguistics, Vali Asr University of Rafsanjan, Rafsanjan, Iran ; E-mail: mhassanzadeh@vru.ac.ir, ORCID ID: <https://orcid.org/0000-0003-1510-1149> ,
2. PhD Candidate of Linguistics, Leiden University, Netherlands;
ORCID ID: <https://orcid.org/0000-0003-4217-3623>

1. Introduction

Since research articles (RAs) are an indispensable part of academia, writing a highly qualified paper entails the competent deployment of linguistic features. The current study investigated a particular type of morphological feature dubbed “lexical bundles” (LBs), which refer to frequently-occurring word combinations. With the growing interest in this area, some corpus-driven studies have examined LBs across different academic genres (Biber, Conrad, & Cortes, 2004), academic registers (Biber, & Barbieri, 2007), disciplines (Cortes, 2006; Durrant, 2017), expertise levels (Staples, Egbert, Biber, & McClair, 2013), L1 versus L2 writing (Ädel & Erman, 2012; Esfandiari, & Barbary, 2017), and rhetorical moves (Alamri, 2020). The findings of prior research on L1 and L2 writing have illustrated inconclusive results concerning the function and frequency of LBs. For instance, Ädel and Erman (2012) observed that native English writers relied on LBs to a greater extent than non-native writers. However, there have been corpus-based studies indicating that non-native writers utilized LBs with a higher frequency than their English counterparts (Bychkovska & Lee, 2017; Pan, Reppen, & Biber, 2016). By the same token, the frequency of functional patterns of LBs has been found to vary in a number of previous corpus studies (e.g., Ädel & Erman, 2012; Bychkovska & Lee, 2017). This study was set out to contribute to this path of inquiry by investigating the frequency, structure, and function of the most frequent four-word LBs in a corpus of 'discussion' sections of RAs written by native English academic writers in applied linguistics.

1.1. Research Questions

1. What are the frequently used four-word lexical bundles in research articles' discussions written by native English academic authors in applied linguistics?
2. What are the structural and functional properties of these frequently used four-word lexical bundles?

2. Literature Review

2.1 Frequency of LBs

Frequency is the most basic attribute of LBs since a multi-word sequence ought to have the requisite frequency threshold to be considered as a bundle. Depending on the size of a corpus, the frequency threshold might vary from 10 (Biber et al., 1999) to 40 times per million words (pmw) (Pan et al., 2016). A variety of occurrences have been identified to be used by authors in preceding bundle studies in L1 and L2 writing. As an example, Esfandiary and Barbary (2017) observed that English academic authors used significantly more LBs than Persian writers. Conversely, Bychkovska and Lee (2017) found that Chinese undergraduate students used more LBs in their essays than English students did.

2.2 Range of LBs

Range or dispersion is another criterion for identifying LBs. Similar to frequency, the range threshold varies depending on the corpus size. For instance, Adel and Erman (2012) set the low dispersion threshold of 'three' thanks to the size of the corpus, while for a corpus of 176 texts, the range threshold was set at 20 by Biber and Barbieri (2007).

2.3 Structure and Function of LBs

LBs fall into different structural and functional patterns. Following Biber et al.'s (2004) functional and structural taxonomies, LBs were structurally classified into three categories: NP/PP based bundles (phrasal bundles), VP-based bundles, and Dependent clause bundles. Functionally, they serve three primary functions, namely stance bundles, discourse organizing bundles, and referential bundles. Previous research has shown varying results regarding the frequency of stance expressions and discourse organizers in L1 writing.

3. Methodology

3.1 Corpus

The present study used a corpus of research article discussions produced by native English academic writers in applied linguistics. The RAs were extracted from five highly-ranked international journals (*Language Learning*, *Applied Linguistics*, *TESOL Quarterly*, *Studies in Second Language Acquisition*, and *Second Language Writing*). The corpus was composed of 243 discussion sections published between 2005 and 2019.

3.2 Bundle identification procedure

In the initial stage, the discussions were removed from all non-textual content (i.e. plain texts). Using AntConc (3.5.8.0), a list of four-word LBs with a frequency of 10 and a range of 5 were retrieved. Then, the LBs were structurally and functionally analyzed based on Biber et al.'s (2004) structural and functional taxonomy.

4. Results

After retrieval, 142 types and 2,637 tokens of LBs were found to be used in the discussions, suggesting the prevalence of LBs in the academic prose of native English writers. The most frequent LBs found in the corpus were *in the present study*, *in the current study*, *in the case of*, *it is possible that*, *the results of the*, and *on the other hand*, which occurred over 50 times across the corpus. Structurally, most LBs were phrasal bundles consisting of NP-based and PP-based bundles. The functional analysis revealed that referential bundles accounted for 60.6% of all LBs.



دوماهنامه بین‌المللی

۱۳، ش ۶ (پیاپی ۷۲)، بهمن و اسفند ۱۴۰۱، صص ۵۵۱-۵۶۹

مقاله پژوهشی

<http://dorl.net/dor/20.1001.1.23223081.1401.0.0.92.4>

چگونگی به‌کارگیری دسته‌های واژگانی توسط نویسندگان

انگلیسی‌زبان در رشته زبان‌شناسی کاربردی:

پژوهشی پیکره‌ای

محمد حسن زاده^{۱*}، حدیث تمله^۲

۱. استادیار رشته زبان‌شناسی کاربردی، دانشگاه ولی عصر رفسنجان، رفسنجان، ایران.

۲. دانشجوی دکتری رشته زبان‌شناسی، دانشگاه لایدن، هلند.

تاریخ پذیرش: ۱۴۰۰/۰۲/۱۷

تاریخ دریافت: ۱۳۹۹/۱۲/۰۲

چکیده

پژوهش‌های مرتبط با دسته‌های واژگانی اخیراً در رأس مطالعات پیکره‌ای قرار گرفته‌اند. پژوهش‌های پیشین پیکره‌ای در پیوند با بررسی دسته‌های واژگانی در نوشتار نویسندگان انگلیسی‌زبان، گاه به نتایج متناقضی درباره فرآوانی و تقسیم‌بندی کارکرد این توالی‌ها دست یافته‌اند. با این حال، هیچ مطالعه‌ای تا به امروز، به بررسی دسته‌های واژگانی در بخش «بحث» مقالاتی که صرفاً نویسندگان انگلیسی‌زبان به نگارش درآورده‌اند، نپرداخته است. پژوهش کنونی با بررسی فرآوانی، ساختار و کارکرد توالی‌های واژگانی چهارکلمه‌ای تلاش کرده تا این شکاف را برطرف کند. داده مورد بررسی در این پژوهش از بخش «بحث» مقالات نویسندگان انگلیسی‌زبان به دست آمد. تمامی این مقالات از پنج مجله ممتاز بین‌المللی در رشته زبان‌شناسی کاربردی استخراج شد. درنهایت، پیکره‌ای متشکل از بیش از سیصد هزار واژه گردآوری شد. با بهره‌گیری از نرم‌افزار AntConc همه دسته‌های چهارکلمه‌ای، با فرآوانی ده و دامنه پنج استخراج شد. یافته‌های به دست آمده نشان داد که دسته‌های واژگانی در پیکره بررسی‌شده فرآوانی بالایی داشتند. در رابطه با تقسیم‌بندی ساختاری دسته‌های واژگانی، نتایج حاکی از آن بود که دسته‌های عبارتی فرآوانی بیشتری نسبت به دسته‌های ساختاری داشته‌اند. همچنین، یافته‌ها نشان داد که به‌لحاظ کارکرد از دسته‌های ارجاعی به‌وفور بهره‌گیری شده است. افزون‌براین، دسته‌های حالتی و گفتمان‌پردازها پس از دسته‌های ارجاعی، به‌ترتیب بیشترین فرآوانی را داشتند. در پایان آموزه‌های این پژوهش برای نویسندگان غیرانگلیسی‌زبان و دوره‌های تحصیلات تکمیلی دانشگاه‌ها در پیوند با به‌کارگیری دسته‌های واژگانی در نگارش دانشگاهی بیان شده است.

واژه‌های کلیدی: دسته‌های واژگانی، پیکره، بحث، نگارش دانشگاهی، نویسندگان انگلیسی‌زبان.

E-mail: mhassanzadeh@vru.ac.ir

* نویسنده مسئول مقاله:

۱. مقدمه

کنکاش و پژوهش علمی یکی از بهترین ابزارهای پیشرفت و انتقال دانش در دنیای امروز به‌شمار می‌رود و به رشته‌ی تحریر درآوردن یک پروژه‌ی علمی مستلزم آگاهی نویسندگان از سبک و سیاق نگارش دانشگاهی (آکادمیک) است. گونه‌های مختلف گفتمان، نیازمند به‌کارگیری واژگان ویژه‌ای هم هستند که از نظر ساختار و مفهوم، هماهنگ با آن متن یا گفتمان باشند. مقالات دانشگاهی نیز از این قاعده مستثنی نیستند. در همین راستا، زمل^۱ (1998) اظهار داشته که نگارش دانشگاهی ویژگی‌های متمایز خود را دارد؛ بدین صورت که یکی از پیش‌شرط‌های آن، به‌کارگیری نوعی زبان با واژگان و استانداردهای مخصوص خود است. با نگاه به اهمیت والای نگارش دانشگاهی، این نوشتار از جنبه‌های گوناگونی بررسی شده است و پیشینه‌ی مطالعات زبانی سرشار از پژوهش‌هایی است که بر روی ویژگی‌های بارز آن متمرکز شده‌اند. برای نمونه، در طول دو دهه‌ی گذشته، پژوهش‌های پیکره‌ای^۲ گسترده‌ای برای بررسی توالی‌های چندکلمه‌ای^۳ صورت گرفته است. نقش محوری و اساسی توالی‌های چندکلمه‌ای زیر عنوان دو مقوله «عبارت‌شناسی»^۴ (Granger & Meunier, 2008) و «زبان الگویی»^۵ (Wray, 2002) همواره مورد تأکید قرار گرفته است، زیرا آشنایی با این توالی‌ها به فراگیری روان‌تر زبان دوم کمک می‌کند (Staples et al., 2013).

بسیاری از مطالعات پیکره‌ای به بررسی توالی‌های خاص چندکلمه‌ای در متون دانشگاهی پرداخته‌اند. این توالی‌های چندکلمه‌ای که دسته‌های واژگانی^۶ نامیده می‌شوند به‌منزله‌ی پرتکرارترین واحدهای واژگانی سه یا چند کلمه‌ای تعریف شده‌اند که در یک گونه‌ی خاص کارکردی^۷ به‌کار می‌روند (Biber et al., 1999). به‌کارگیری دسته‌های واژگانی از ویژگی‌های بارز هر نوشتار شیوا و روان به‌شمار می‌آید که به معنابخشی و انسجام در متن کمک می‌کند و عدم به‌کارگیری آن‌ها می‌تواند بیانگر ناپختگی و ناآشنایی نویسنده با استانداردهای نگارش بین‌المللی شمرده شود.

گفتنی است میان دسته‌های واژگانی و دیگر توالی‌های چندکلمه‌ای تفاوت‌هایی وجود دارد. برای مثال، معنای اصطلاحات^۸ را که گروه دیگری از توالی‌های چندکلمه‌ای هستند نمی‌توان از تک تک کلمات آن توالی استنباط کرد، در صورتی که دسته‌های واژگانی معنای واضحی دارند

و می‌توان معنای کلی هر دسته واژگانی را از عناصر موجود در آن استنباط کرد. ویژگی شاخص دیگر دسته‌های واژگانی، فراوانی^۹ بالای آن‌هاست، بدین صورت که دسته‌های واژگانی نسبت به دیگر توالی‌ها، فراوانی بالاتری دارند. برای مثال کورتز^{۱۰} (2004) خاطر نشان کرد که فراوانی اصطلاح «kick the bucket» در یک میلیون کلمه ۰.۵ است، در صورتی که یک دسته واژگانی در پیکره‌ای متشکل از یک میلیون کلمه باید بیست بار تکرار شده باشد تا به منزله دسته واژگانی تلقی شود. علاوه بر فراوانی، دسته‌های واژگانی باید دامنه^{۱۱} معینی نیز داشته باشند که نشان‌دهنده تکرار این توالی در متون مختلف باشد. همچنین، این توالی‌ها را می‌توان بر اساس الگوهای ساختاری و کارکردی تقسیم‌بندی کرد. در بخش پیشینه پژوهش، توضیحات تکمیلی این ویژگی‌ها ارائه خواهد شد.

این پژوهش با توجه به لزوم کنکاش بیشتر راجع به چگونگی به‌کارگیری واژگان توسط نویسندگان حرفه‌ای و چیره‌دست سعی دارد به بررسی رایج‌ترین دسته‌های واژگانی چهارکلمه‌ای در پیکره‌ای متشکل از بخش «بحث» مقالات حوزه زبان‌شناسی کاربردی که نویسندگان انگلیسی‌زبان به نگارش درآورده‌اند بپردازد. افزون‌براین، این پژوهش به تجزیه و تحلیل ساختاری و کارکردی دسته‌های واژگانی نیز می‌پردازد تا از این راه به شناخت ژرف‌تری از کاربرد این توالی‌ها در میان نویسندگان انگلیسی‌زبان دست یابیم. در راستای اهداف مذکور، پرسش‌های زیر مورد بررسی قرار گرفته است:

۱. متداول‌ترین دسته‌های واژگانی چهارکلمه‌ای در مقالات زبان‌شناسی کاربردی که نویسندگان انگلیسی‌زبان به نگارش درآورده‌اند، کدام‌ها هستند؟
۲. طبقه‌بندی ساختاری این نوع از توالی‌های واژگانی به چه صورت است؟
۳. طبقه‌بندی کارکردی این نوع از توالی‌های واژگانی به چه صورت است؟

۲. پیشینه پژوهش

در این بخش با بررسی مطالعات قبلی مرتبط با توالی‌های واژگانی، به توصیف، تجزیه و تحلیل فراوانی، دامنه، ساختار و کارکرد این توالی‌ها در مطالعات پیکره‌ای می‌پردازیم. همچنین، با تحلیل پژوهش‌هایی که دسته‌های واژگانی را در نوشتار نویسندگان انگلیسی‌زبان بررسی

کرده‌اند، آموزه‌های نظری و تجربی متعلق به پژوهش حاضر را ارائه خواهیم کرد. دسته‌های واژگانی به دلیل نقش مهمی که در یادگیری/آموزش زبان ایفا می‌کنند همواره در سال‌های اخیر مورد توجه پژوهشگران قرار گرفته‌اند. گروهی از مطالعات پیکره‌ای پیشین که به بررسی دسته‌های واژگانی به‌کاررفته توسط نویسندگان انگلیسی‌زبان پرداخته‌اند نتایج متفاوتی را در پیوند با فراوانی و تقسیم‌بندی کارکردی این توالی‌ها نشان می‌دهند، Adel & Erman (2012; Bychkovska & Lee, 2017; Esfandiari & Barbary, 2017). این تفاوت‌ها سزاوار توجه و بررسی بیشتری هستند تا به داده‌های ژرف‌تری در این باره دست یابیم. افزون‌براین، پژوهشگران در پیوند با کاربرد دسته‌های واژگانی در میان نویسندگان انگلیسی‌زبان و غیرانگلیسی‌زبان اختلاف‌نظر دارند (Adel & Erman, 2012; Bychkovska & Lee 2017). برخی از مطالعات پیکره‌ای که به بررسی دسته‌های واژگانی در نوشتار دانشگاهی مربوط به زبان اول و زبان دوم پرداخته‌اند به این نتیجه رسیدند که نویسندگان انگلیسی‌زبان از دسته‌های واژگانی بیشتری استفاده می‌کنند (Adel & Erman, 2012; Esfandiari & Barbary 2017). این مسئله می‌تواند نشان‌دهنده عدم توجه و یا عدم دانش کافی نویسندگان غیرانگلیسی‌زبان از نحوه به‌کارگیری دسته‌های واژگانی و اهمیت آن‌ها باشد. از طرفی، دسته دیگری از این پژوهش‌ها نشان داده‌اند که فراوانی توالی‌های واژگانی در نوشتار دانشگاهی نویسندگان غیرانگلیسی‌زبان بیشتر است (Bychkovska & Lee 2017; Pan et al., 2016). از این نکته نیز نمی‌توان غافل شد که پژوهشگران غیرانگلیسی‌زبان هم تمایل دارند که نوشته‌ها و پژوهش‌های خود را در مجلات معتبر منتشر کنند تا بر اهمیت کار خود بیفزایند و جایگاه علمی خود را ارتقا ببخشند. قطعاً چنین امری در جهان رقابتی امروز، نوشتاری حرفه‌ای براساس استانداردهای بالای نگارش دانشگاهی را می‌طلبد. از آنجایی که توالی‌های واژگانی سهم قابل‌اعتنایی از نوشتار دانشگاهی را به خود اختصاص می‌دهند، بررسی دسته‌های به‌کار برده شده از سوی نویسندگان انگلیسی‌زبان می‌تواند به درک بهتر متون دانشگاهی کمک کند و سطح آگاهی نویسندگان غیرانگلیسی‌زبان را در این زمینه افزایش دهد. درواقع با واکاوی آثار پژوهشگران چیره‌دست می‌توان الگوی شفاف‌تری به‌ویژه پیش پای نویسندگان و پژوهندگان تازه‌کار و غیرتازه‌کار فارسی‌زبان نهاد تا به نگرش و جمع‌بندی معقول‌تری نسبت به شیوه بهره‌گیری از دسته‌های واژگانی در نوشته‌های خویش دست یابند.

برخی پژوهش‌ها کوشیده‌اند تا دسته‌های واژگانی را از نظر ساختار و کارکرد در مقالات رشته‌ی زبان‌شناسی کاربردی بررسی کنند (e.g., Nasrabady et al., 2020; Ren, 2021; Shirazizadeh & Amirfazlian, 2021)، اما در گردآوری پیکره این مطالعات تمایزی میان نویسندگان انگلیسی‌زبان و غیرانگلیسی‌زبان قائل نشده‌اند. همچنین این مطالعات، دسته‌های واژگانی را عموماً در تمامی بخش‌های مقالات سنجیده‌اند. درحالی‌که در این بین، بخش «بحث» مقالات شاید بیش از سایر قسمت‌ها به تبحر نویسنده در قلم‌فرسایی بستگی دارد، زیرا میزان استنادها و وام‌گیری از سخنان دیگران در این بخش محدودتر و از سوی دیگر، توانمندی نویسنده در تفسیر، ارزیابی، و مقایسه یافته‌ها، عیان‌تر می‌شود و بدین سان پختگی زبانی اهمیتی دوچندان پیدا می‌کند.

۱-۲. دسته‌های واژگانی: فراوانی، دامنه، ساختار، و کارکرد

همانطور که پیش‌تر اشاره شد دسته‌های واژگانی، فراوانی بالایی دارند که به‌منزله مهم‌ترین معیار برای شناسایی این توالی‌ها از آن‌ها یاد می‌شود (Cortes, 2015). مطالعات پیشین آستانه فراوانی گوناگونی را در نظر گرفته‌اند. برای نمونه، در مطالعه بایبر^{۱۲} و همکاران (1999) آستانه فراوانی ۱۰، در مطالعه ادل و ارمن^{۱۳} (2012) ۲۵، و در مطالعه پن^{۱۴} و همکاران (2016) ۴۰ در هر یک میلیون کلمه در نظر گرفته شد.

یکی دیگر از معیارهای مهم برای شناسایی دسته‌های واژگانی تعیین دامنه یا پراکندگی^{۱۵} است. پراکندگی، تکرار دسته‌های واژگانی را در متون مختلف نشان می‌دهد. تعیین دامنه اهمیت بسیاری دارد، چون نشان می‌دهد که دسته واژگانی موردنظر از ویژگی‌های قلم نویسنده یا پژوهشگر نیست، بلکه ویژگی اصلی پیکره است (Pan et al., 2016). در مطالعات پیشین، آستانه پراکندگی دسته‌های واژگانی با توجه به اندازه پیکره متغیر بوده است؛ به این صورت که با افزایش تعداد متون پیکره، آستانه دامنه نیز بالا می‌رفته است. در مطالعه ادل و ارمن برای ۸۲ متن موجود در پیکره، آستانه پراکندگی سه و در پژوهش بایبر و باربیری^{۱۶} (2007) برای ۱۷۶ متن، آستانه پراکندگی بیست تعیین شد.

دسته‌های واژگانی معمولاً از نظر ساختار و کارکرد، برای تجزیه و تحلیل بیشتر دسته‌بندی

می‌شوند. بایبر و همکاران (1999) به ترتیب نه و هشت الگوی ساختاری را در مکالمه و نگارش دانشگاهی برای تحلیل ساختار دسته‌های واژگانی شناسایی کردند. بایبر و همکاران (2004) بعدها این الگوها را به سه دسته اصلی ساختاری تقسیم کردند. براساس این مدل، دسته‌های واژگانی به لحاظ ساختاری به سه گروه اصلی، دسته‌های اسمی/گزاره‌ای^{۱۷} (دسته‌های عبارتی^{۱۸})، دسته‌های فعلی^{۱۹}، و دسته‌های جمله‌واره‌ای وابسته^{۲۰} تقسیم می‌شوند. گروه دوم و سوم بیشتر از دسته‌های جمله‌واره‌ای^{۲۱} تشکیل شده‌اند.

در مطالعات پیشین تقسیم‌بندی کارکرد دسته‌های واژگانی براساس مدل‌های بایبر و همکاران (2004) و هایلند^{۲۲} (2008b) صورت می‌گرفت. در مدل بایبر و همکاران دسته‌های واژگانی به سه گروه کارکردی تقسیم می‌شوند. (۱) دسته‌های حالتی^{۲۳} به توالی‌های واژگانی اطلاق می‌شود که نگرش، ارزیابی، و دیدگاه نویسنده را از نظر قطعیت یا عدم اطمینان بیان می‌کند. (۲) گفتمان‌پردازها^{۲۴} رابطه گفتمان‌های مختلف در یک متن را نشان می‌دهد. و در آخر (۳) دسته‌های ارجاعی^{۲۵} که در متون دانشگاهی، فراوانی بالایی داشته و بر مفاهیمی چون تعداد، مقدار، اندازه، زمان یا مکان دلالت دارد و یا این‌که ویژگی مشخصی را نشان می‌دهد. در مدل پیشنهادی هایلند (2008b) نیز دسته‌های واژگانی به لحاظ کارکردی به سه گروه عمده تقسیم می‌شوند: (۱) دسته‌های پژوهش - محور^{۲۶} که نویسندگان از طریق این توالی‌ها به تجربیات خود نسبت به دنیای واقعی ساختار می‌بخشند؛ (۲) دسته‌های واژگانی وابسته به متن^{۲۷} که با معنا و انسجام متن ارتباط دارند؛ و (۳) دسته‌های مخاطب - محور^{۲۸} که بر نویسنده یا خواننده متن متمرکز می‌شوند.

۲-۲. دسته‌های واژگانی در گونه^{۲۹}‌های گفتاری و نوشتاری

در دهه‌های اخیر، بررسی و مطالعه دسته‌های واژگانی افزایش چشمگیری داشته است. بایبر و همکاران (2004) دسته‌های واژگانی را در انواع مختلفی از گونه‌های نوشتاری و گفتاری همچون مکالمه، کلاس درس، کتب و نگارش دانشگاهی بررسی کردند. نتایج این مطالعه نشان داد که گونه‌های گفتاری از دسته‌های واژگانی بیشتری تشکیل شده‌اند. همچنین این پژوهش نشان داد که به لحاظ ساختار و کارکرد دسته‌های واژگانی، میان گونه‌های گفتاری و نوشتاری

نیز تفاوت وجود دارد؛ برای مثال گونه گفتاری بیشتر از دسته‌های فعلی تشکیل شده بود، هرچند در گونه‌های نوشتاری دسته‌های اسمی و گزاره‌ای، فراوانی بیشتری داشتند. به‌لحاظ کارکرد، نتایج همچنین حاکی از آن بود که گونه گفتاری بیشتر وابسته به دسته‌های حالتی است، اما کتب درسی و نگارش دانشگاهی در بردارنده تعداد زیادی از دسته‌های ارجاعی هستند.

۳-۲. دسته‌های واژگانی در نوشتار نویسندگان با سطوح مهارت زبانی متفاوت

گروه دیگری از مطالعات پیکره‌ای دسته‌های واژگانی را در سطوح مختلفی از مهارت زبانی نویسندگان بررسی کرده است؛ برای مثال کورتز (2004) دریافت که دسته‌های واژگانی بیشترین فراوانی را در نوشتار نویسندگان حرفه‌ای در رشته‌های تاریخ و زبان‌شناسی دارند. از طرفی دسته‌های واژگانی یافت شده به‌ندرت در نوشتار دانشجویان به‌کار گرفته شده بود. در راستای همین نتایج، استاپلز^{۳۰} و همکاران (2013) نیز در مطالعه خود به این نتیجه دست یافتند که فراوانی دسته‌های واژگانی با افزایش سطح مهارت نویسنده افزایش می‌یابد. با این حال، به لحاظ کارکرد، تفاوت زیادی میان نویسندگان با سطوح گوناگون مهارت زبانی یافت نشد، اما در پیوند با ساختار دسته‌های واژگانی، چن و بیکر^{۳۱} (2016) خاطرنشان کردند که نویسندگان با سطح مهارت زبانی پایین‌تر بیشتر از دسته‌های فعلی استفاده می‌کنند، در حالی که نویسندگان مسلط‌تر به مقدار زیاد از دسته‌های اسمی و گزاره‌ای استفاده می‌کنند.

۴-۲. دسته‌های واژگانی در متون گوناگون دانشگاهی

برای بررسی دسته‌های واژگانی در نوشته‌های گوناگون، هایلند (2008a) در پژوهش پیکره‌ای خود این نوع از توالی‌ها را در مقالات، رساله‌های دکتری و پایان‌نامه‌ها در چهار رشته مختلف واکاوی کرد. نتایج حاکی از آن بود که پایان‌نامه‌ها دسته‌های واژگانی بیشتری نسبت به مقاله‌ها و رساله‌ها داشتند. همچنین به‌لحاظ کارکرد، دسته‌های ارجاعی بیشتری در پایان‌نامه‌ها یافت شد. از طرفی گفتمان‌پردازها در مقاله‌ها و رساله‌ها به فراوانی یافت شد، اما در مقاله‌ها، دسته‌های حالتی، فراوانی بیشتری نسبت به رساله‌های دکتری داشتند.

۵.۲. دسته‌های واژگانی در نوشتار نویسندگان زبان اول و دوم

شمار دیگری از پژوهش‌های پیکره‌ای به مقایسه دسته‌های واژگانی در نوشتار نویسندگان انگلیسی‌زبان و نوشتار نویسندگان غیرانگلیسی‌زبان پرداخته‌اند. در پیوند با تعدد و فراوانی دسته‌های واژگانی، نتایج برخی از مطالعات پیکره‌ای حاکی از آن است که نویسندگان انگلیسی‌زبان به دسته‌های واژگانی پرتعدادی در نوشتار خود متکی هستند؛ برای مثال، اسفندیاری و باربری^{۳۲} (2017) در پژوهش خود دریافته‌اند که نویسندگان انگلیسی‌زبان در مقایسه با نویسندگان فارسی‌زبان دسته‌های واژگانی بیشتری را در مقالات خود به‌کار برده‌اند. از طرفی دیگر، مطالعات دیگر نتایج متناقضی ارائه می‌دهد، به این صورت که نویسندگان انگلیسی‌زبان نسبت به نویسندگان غیرانگلیسی‌زبان از توالی‌های کم‌تری استفاده می‌کنند (Bychkovska & Lee, 2017; Pan et al., 2016; Pérez-Llantada, 2014).

با در نظر گرفتن ساختار دسته‌های واژگانی، نتایج مطالعه پیکره‌ای که توسط پان و همکاران (2016) انجام شد نشان داد که نویسندگان انگلیسی‌زبان در نوشتار مقالات خود به لحاظ ساختاری از دسته‌های اسمی و گزاره‌ای به وفور استفاده می‌کنند. در همین راستا، لی و بایچکوسکا^{۳۳} (2017) نیز در مطالعه خود به این نتیجه دست یافتند که از نظر ساختاری، بیشتر توالی‌ها به صورت دسته‌های عبارتی به‌کار رفته‌اند. لو و دنگ^{۳۴} (2019) دسته‌های واژگانی به‌کار برده شده در رساله‌های دکتری دانشجویان انگلیسی‌زبان و چینی‌زبان را بررسی کردند. دانشجویان چینی در تمام گروه‌های اصلی ساختاری به‌طور قابل توجهی بیشتر از دسته‌های واژگانی استفاده کرده بودند. براساس یافته‌ها، نویسندگان چینی تقریباً دو برابر نویسندگان انگلیسی‌زبان از دسته‌های فعلی استفاده می‌کنند. شیرازی‌زاده و امیرفضلیان در توضیح این نتایج عنوان می‌کنند که این تفاوت بارز و چشمگیر می‌تواند به این دلیل باشد که دانشجویان چینی هنوز دانش لازم در رابطه با استانداردهای مقاله‌نویسی را کسب نکرده‌اند. گروهی از مطالعات پیکره‌ای به بررسی کارکرد دسته‌های واژگانی در نوشتار نویسندگان انگلیسی‌زبان و نویسندگان غیرانگلیسی‌زبان پرداخته‌اند. نتایج به‌دست‌آمده این پژوهش‌ها نشان داده که دسته‌های ارجاعی در مقایسه با دو گروه دیگر یعنی دسته‌های حالتی و گفتمان‌پردازها، با فراوانی بیشتری همراه بوده‌اند (Ädel & Erman, 2012; Biber et al., 1999; Biber et al., 1999).

(al., 2004; Pan et al., 2016).

باوجوداین، برخی از مطالعات پیکره‌ای، در بررسی دسته‌های حالتی و گفتمان‌پردازها به نتایج ناهمگونی دست یافته‌اند. با تمرکز بر نوشتار دانشگاهی، ادل و ارمن (2012) به بررسی دسته‌های واژگانی در پیکره‌ای متشکل از نوشتار دانشجویان کارشناسی زبان‌شناسی پرداختند. نتایج به‌دست‌آمده در رابطه با کارکرد دسته‌های نشان داد که دانشجویان انگلیسی‌زبان از دسته‌های حالتی به وفور استفاده می‌کنند. با این حال، گفتمان‌پردازها کمتر مورد استفاده قرار گرفته بودند. اما تحقیق لی و بایچکوسکا (2017) نشان داد که نویسندگان انگلیسی‌زبان از گفتمان‌پردازها بیشتر از دسته‌های حالتی استفاده می‌کنند.

۶-۲. دسته‌های واژگانی در بخش‌های گوناگون مقاله

تعداد کمی از مطالعات پیکره‌محور به بررسی دسته‌های واژگانی تنها در یک بخش از نوشتار دانشگاهی پرداخته‌اند. برای نمونه‌ای کمیاب، کورتز (2013) در مطالعه خود به بررسی دسته‌های واژگانی در پیکره‌ای متشکل از بخش «مقدمه» مقالات در رشته‌های مختلف پرداخت. نتایج مطالعه کورتز نشان داد که به‌لحاظ ساختاری، دسته‌های عبارتی و به‌لحاظ کارکردی، دسته‌های ارجاعی بیشترین فراوانی را در پیکره داشته‌اند. در نمونه‌ای دیگر، میرزایی (2019) دسته‌های واژگانی را در بخش بحث مقالاتی که روش پژوهش آن‌ها بر پایه رویکرد کیفی، کمی، و ترکیبی بود تحلیل کرد. میرزایی به این نتیجه دست یافت که دسته‌های واژگانی اسمی و گزاره‌ای بیشترین فراوانی را دارند. بررسی کارکرد دسته‌های واژگانی بر پایه الگوی هابلند (2008b) صورت گرفت. نتایج حاکی از آن بود که دسته‌های پژوهش - محور، دسته‌های وابسته به متن، و دسته‌های مخاطب - محور به‌ترتیب بیشترین فراوانی را داشته‌اند. همانطور که عنوان شد، پیشینه پژوهش، یافته‌های متفاوتی را در رابطه با فراوانی و کارکرد دسته‌های واژگانی در نگارش دانشگاهی نویسندگان انگلیسی‌زبان نشان می‌دهد. پژوهش کنونی اما با بررسی دسته‌های واژگانی در قسمت بحث مقالات نویسندگان انگلیسی‌زبان نگاهی ژرف‌تر به فراوانی و کارکرد این توالی‌ها کرده است.

۳. روش و رویکرد تحقیق

۳-۱. گردآوری پیکره

پیکره موردنیاز شامل قسمت «بحث» مجموعه مقالاتی است که توسط پژوهشگران انگلیسی‌زبان در رشته زبان‌شناسی کاربردی نوشته شده و به چاپ رسیده است. همگی مقالات گردآوری شده شامل بخش جداگانه‌ای زیرعنوان «بحث» بود و مقالاتی که در آن‌ها بخش بحث با سایر بخش‌ها از جمله یافته‌ها یا نتیجه‌گیری به هم پیوسته بود لحاظ نشد.

جدول ۱: لیست مجله‌ها و تعداد مقاله‌های برگرفته

Table 1: List of Journals and the Number of Extracted Articles

English Journals	No. of RAs	No. of Words	Year
Language Learning	47	47,974	2005-2019
Applied Linguistics	21	25,395	2005-2019
TESOL Quarterly	60	76,366	2005-2019
Studies in Second Language Acquisition	83	142,285	2005-2019
Second Language Writing	32	50,222	2005-2019
Total	243	342,242	

معیار گزینش مجلات براساس شاخص ضریب تأثیر^{۳۰} که یکی از شاخصه‌های اصلی اعتبار مجلات است، صورت گرفت. در نهایت، پنج مجله بین‌المللی ممتاز و سرشناس (Language Learning, Applied Linguistics, TESOL Quarterly, Studies in Second Language Acquisition, and Second Language Writing) در رشته زبان‌شناسی کاربردی که همگی دارای ضریب تأثیر بالای ۲۰۰ بودند، برای پژوهش حاضر تعیین شدند که تاریخ انتشار تمامی مقاله‌های انتخاب‌شده متعلق به بازه زمانی ۲۰۰۵ تا ۲۰۱۹ بود (جدول ۱). همچنین، برای شناسایی نویسندگان انگلیسی‌زبان از غیرانگلیسی‌زبان، تمام کسانی که نام و نام خانوادگی

انگلیسی داشته و وابسته به یک نهاد یا مؤسسه دانشگاهی در یک کشور انگلیسی‌زبان بودند، برگزیده شدند. گفتنی است هنگامی که مقاله‌ای حاصل همکاری چندین نویسنده می‌بود، انگلیسی‌زبان بودن تمامی نویسندگان بر پایه معیارهای مذکور مورد بررسی قرار می‌گرفت و پس از جست‌وجوهای صورت‌گرفته روی پیشینه، تبار و محل زندگی افراد در درگاه گوگل و همچنین صفحه شخصی آن افراد، در صورتی که انگلیسی‌زبان نبودن حتی یکی از آنها به نحوی احراز می‌شد مقاله مورد نظر از دستور کار خارج می‌شد.

۲-۳. شناسایی دسته‌های واژگانی

پس از استخراج بخش «بحث» مقالات و ذخیره آن‌ها به صورت فایل‌های جداگانه، از نرم‌افزار AntConc (آخرین نسخه موجود، 3.5.8.0) برای تحلیل پیکره استفاده شد. نخستین بار لورنس آنتونی^{۳۶} (2007) نرم‌افزار AntConc را برای تجزیه و تحلیل پیکره طراحی کرد. این نرم‌افزار را می‌توان به صورت رایگان از درگاه «<http://www.laurenceanthony.net>» بارگیری کرد. برای استفاده از این نرم‌افزار تمامی متون از محتوای غیرممتنی مانند عنوان‌ها، پانویس‌ها، زیرنویس‌ها، جدول‌ها، شکل‌ها، نمونه‌ها، منابع، و یا نقل‌قول‌ها پاک شدند.

بر اساس مطالعه بایبر و همکاران (1999)، فراوانی دسته‌های واژگانی در این مطالعه «ده» و دامنه آن‌ها «پنج» در نظر گرفته شد. از آنجایی که تعداد واژگان پیکره کم‌تر از یک میلیون بود، اما میزان بالایی از متون در آن قرار داشت (جدول ۱)، فراوانی، پایین در نظر گرفته شد، ولی آستانه دامنه بالایی برای تشخیص دسته‌های واژگانی لحاظ شد. بنابراین، این اطمینان به دست آمد که همگی عبارت‌های چند کلمه‌ای به دست آمده دسته‌های واژگانی هستند. پس از تعیین فراوانی و دامنه، با بهره‌گیری از نرم‌افزار AntConc فهرستی از توالی‌های چهارکلمه‌ای از پیکره استخراج شد. دسته‌های واژگانی چهار کلمه‌ای فراوانی بالاتری از توالی‌های پنج کلمه‌ای دارند و نسبت به دسته‌های واژگانی سه کلمه‌ای نیز تقسیم‌بندی ساختاری و کارکردی واضح‌تری دارند (Hyland, 2008b). به همین دلیل، دسته‌های واژگانی چهار کلمه‌ای برای این پژوهش برگزیده شدند.

۳-۳. تقسیم‌بندی ساختاری و کارکردی دسته‌های واژگانی

پس از استخراج فهرستی از دسته‌های واژگانی و مشخص کردن تعداد و فراوانی دسته‌های واژگانی از طریق نرم‌افزار AntConc، بررسی ساختار و کارکرد آن‌ها براساس الگوی بایبر و همکاران (2004) صورت گرفت. این الگو به دلیل این‌که به گونه‌ای گسترده در مطالعات پیشین مورد بهره‌گیری قرار گرفته (e.g. Ädel & Erman, 2012; Bychkovska & Lee, 2017; Cortes 2013; Shin, 2019) و همچنین براساس طیف گسترده‌ای از متون نوشتاری طراحی شده است، برای پژوهش کنونی در نظر گرفته شد. نرم‌افزار AntConc این قابلیت را به پژوهشگر می‌دهد که هر یک از دسته‌های واژگانی را در متن به‌کار برده شده بررسی کند. بنابراین، با بهره‌گیری از گزینه Concordancing Tool در نرم‌افزار، ساختار و کارکرد هر کدام از دسته‌های واژگانی به‌طور جداگانه و بر پایه الگوی برگزیده در متن تجزیه و تحلیل شد.

۳-۴. پایایی^{۳۷}

به‌منظور انجام تحلیلی پایا، ساختار و کارکرد تمام دسته‌های واژگانی توسط پژوهشگر دوم مجدداً بررسی شد. نتایج پایایی مابین تحلیل‌گران^{۳۸} نشان داد که ضریب توافق کاپا^{۳۹} بین دو پژوهشگر ۰/۹۰ بود. در بررسی ساختاری دسته‌های واژگانی تفاوت چندانی بین دو پژوهشگر مشاهده نشد، اما اختلاف اصلی در تقسیم‌بندی کارکردی دسته‌های واژگانی بود. به این دلیل که برخی از دسته‌های واژگانی چندین کاربرد داشتند و هر کدام از تحلیلگران کارکردی متفاوت را در نظر گرفته بودند. اختلاف جزئی به‌وجودآمده توسط دو پژوهشگر از طریق بحث و گفت‌وگو و بررسی مجدد دسته‌های واژگانی برطرف شد. گفتنی است که قرار دادن هر یک از دسته‌های واژگانی در گروه کارکردی خاص به معنای به‌کارگیری اختصاصی از آن دسته واژگانی در آن گروه کارکردی نیست. دسته‌های واژگانی بر پایه بیشترین کارکردی که در هر گروه داشته‌اند تقسیم‌بندی شدند (Biber et al., 2004).

۴. یافته‌ها

۴-۱. فراوانی دسته‌های واژگانی در پیکره

نتایج این بخش در راستای نخستین پرسش این پژوهش است، به این مضمون که تعداد و فراوانی دسته‌های واژگانی در بخش «بحث» مقاله‌های پژوهشی که توسط نویسندگان انگلیسی‌زبان در رشته‌ی زبان‌شناسی کاربردی نوشته شده و به چاپ رسیده، چگونه است.

جدول ۲: تعداد و فراوانی دسته‌های واژگانی در پیکره

Table 2: Frequency of bundles in the Corpus

Bundle Types	Bundle Tokens	Corpus Size
142	2,637	324,242

پس از استخراج دسته‌های واژگانی چهارکلمه‌ای از پیکره با استفاده از نرم‌افزار AntConc، نتایج نشان داد که نویسندگان انگلیسی‌زبان از ۱۴۲ توالی واژگانی با فراوانی کلی ۲۶۳۷ استفاده کرده بودند (جدول ۲). دسته‌های واژگانی «in the present study, in the current study, in the case of, it is possible that, the results of the, and on the other hand» بیشترین فراوانی را داشتند (پیوست ۱).

۴-۲. تجزیه و تحلیل ساختاری دسته‌های واژگانی

همانطور که قبلاً اشاره شد، در طبقه‌بندی ساختاری دسته‌های واژگانی از مدل بایبر و همکاران (2004) پیروی شده است.

جدول ۳: تعداد و فراوانی دسته‌های واژگانی به‌لحاظ ساختاری

Table 3: Number of Structural Types and Tokens

Categories	Types		Tokens	
	Frequency	Percent	Frequency	Percent
NP/PP-based bundles	95	66.9%	1994	73.7%
VP-based bundles	37	26.1%	569	21.6%
Dependent clause bundles	10	7%	124	4.7%

پس از تقسیم‌بندی ساختاری، یافته‌های به دست آمده حاکی از آن بود که فراوانی دسته‌های عبارتی (دسته‌های گزاره‌ای /اسمی) بیشتر از دسته‌های فعلی و دسته‌های جمله‌واره‌ای است. همانطور که جدول ۳ نشان می‌دهد ۹۸ دسته گزاره‌ای/اسمی یافت‌شده فراوانی‌ای بالغ بر ۱۹۰۰ داشتند و در مجموع ۷۵ درصد از کل فراوانی را به خود اختصاص دادند. مثال ۱ و ۲ نمونه‌هایی از دسته‌های اسمی/گزاره‌ای را نشان می‌دهد. تعداد دسته‌های فعلی و دسته‌های جمله‌واره‌ای وابسته یافت‌شده به ترتیب ۳۵ و ۹ است که ۲۰ و ۵ درصد از فراوانی کل دسته‌های واژگانی را داشته‌است (مثال ۳ و ۴).

Example 1. *In the present study*, the main-point effect was significant for both groups ,regardless of the distance between the contradicted precondition and the anomaly. **(PP-based bundle)**

Example 2. *The results of the* current study, however, showed an increase in coordinate clauses followed by a decline. **(NP-based bundle)**

Example 3. Whilst considering the benefits that forms of elaborated processing might bring to the acquisition of L2 lexical items, *it is important to* also consider the role of repetition. **(VP-based bundle)**

Example 4. The findings indicate *that there is a* relationship between these stages of development and word frequency. **(Dependent clausal bundle)**

۳-۴. تجزیه و تحلیل کارکرد دسته‌های واژگانی

براساس تقسیم‌بندی کارکردی بایبر و همکاران (2004)، نویسندگان انگلیسی‌زبان به لحاظ معنایی انواع بیشتری از دسته‌های ارجاعی را به کار گرفته‌اند که ۶۰.۱ درصد از فراوانی دسته‌های واژگانی را به خود اختصاص می‌دهد (مثال ۵).

Example 5. They demonstrated the *degree to which* the exposure to English as an additional language was part of their broader experience of acquiring a repertoire of discourses in the home, community, and school. **(Referential Bundle)**

جدول ۴: تعداد و فراوانی دسته‌های واژگانی به‌لحاظ کارکرد

Table 4: Number of Structural Types and Tokens

Categories	Types		Tokens	
	Frequency	Percent	Frequency	Percent
Stance bundles	26	18.3%	503	19.1%
Discourse organizing bundles	30	21.1%	473	17.9%
Referential bundles	86	60.6%	1661	63%

همانطور که در جدول ۴ مشاهده می‌شود، تعداد دسته‌های حالتی و گفتمان‌پردازانه‌های یافت‌شده تقریباً یکسان است، اما به‌لحاظ فراوانی، ۲۰.۲ درصد دسته‌های حالتی و ۱۹.۱ درصد گفتمان‌پردازانه‌ها در بخش «بحث» مقالات پژوهشگران انگلیسی‌زبان دارای فراوانی بودند. مثال ۶ و ۷ نمونه‌هایی از دسته‌های حالتی و گفتمان‌پردازانه‌ها را در پیکره نشان می‌دهد.

Example 6. *The fact that the* increase in dependent clause use is evident in both the raw data and the ratio data in this study may be a reflection of *the fact that the* proficiency levels in this study are more distinct from each other. (Stance Bundle)

Example 7. In the present study, both groups showed evidence of efficient L1-based structure building. In the L2, *on the other hand*, there were significant differences in structure building between the groups. (Discourse organizer)

۵. بحث و نتیجه‌گیری

هدف اصلی این پژوهش بررسی چگونگی به‌کارگیری دسته‌های چهارکلمه‌ای در بخش «بحث» مقالاتی بود که پژوهشگران انگلیسی‌زبان نوشته بودند. پس از جمع‌آوری پیکره و استخراج دسته‌های واژگانی، تجزیه و تحلیل ساختاری و کارکردی دسته‌هایی که بیشترین فراوانی را داشته‌اند صورت گرفت.

بررسی تعداد و فراوانی کلی دسته‌های چهارکلمه‌ای نشان داد که این گونه توالی‌ها گستردگی بالایی در بخش بحث مقالاتی که توسط نویسندگان انگلیسی‌زبان به نگارش درآمده است دارند. بنابراین، می‌توان ادعا کرد که یکی از ویژگی‌های بارز و مهم نگارش دانشگاهی به‌ویژه در بخش بحث، به‌کارگیری دسته‌های واژگانی است. نتیجه به‌دست‌آمده در راستای برخی از پژوهش‌های پیشین است که نقش و اهمیت بسزای دسته‌های واژگانی را در نگارش

متون دانشگاهی دریافته بودند (Cortes, 2006; Erman & Warren, 2000; Li & Schmitt, 2009; Pérez-(Llantada, 2014). افزون بر این، نتایج این پژوهش، تمایز ساختاری میان دسته‌های عبارتی و جمله‌وارهای را که توسط بایبر و همکاران (2004) مطرح شده بود تأیید می‌کند. براساس مدل ساختاری بایبر و همکاران، دسته‌های عبارتی شامل عبارت‌های اسمی و گزاره‌ای هستند. از سوی دیگر، دسته‌های جمله‌وارهای از دسته‌های فعلی و جمله‌وارهای وابسته تشکیل شده‌اند. در این پژوهش نیز تمامی دسته‌های واژگانی عبارتی شامل عبارت‌های اسمی و گزاره‌ای بودند. افزون‌براین، دسته‌های جمله‌وارهای نیز فقط متشکل از دسته‌های فعلی و جمله‌وارهای وابسته بودند.

نتایج تقسیم‌بندی ساختاری دسته‌های واژگانی حاکی از آن بود که فراوانی دسته‌های عبارتی به‌کار برده‌شده توسط نویسندگان انگلیسی‌زبان از دسته‌های جمله‌وارهای در بیکره بسیار بیشتر بود. نتایج پژوهش‌های پیشین نیز (Chen & Baker, 2010; Cortes, 2004; Esfandiari & Barbary, 2017; Shirazizadeh & Amirfazlian, 2021) بر این امر دلالت دارد که در نگارش دانشگاهی نویسندگان انگلیسی‌زبان، فراوانی دسته‌های عبارتی از جمله‌وارهای بیشتر است. براساس گفته بایبر و همکاران (1999)، 60 درصد نگارش دانشگاهی از دسته‌های واژگانی عبارتی به‌ویژه اسمی و گزاره‌ای تشکیل شده است. یک دلیل برای توضیح این نتیجه را می‌توان به این مسئله ربط داد که یکی از مشخصه‌های بارز بخش بحث مقالات در انتقال داده‌ها، به‌کارگیری دسته‌های واژگانی اسمی و گزاره‌ای است. گفتنی است که نویسندگان برای انتقال بهتر معنا و مفهوم پژوهش خود به‌ویژه در بخش بحث، لازم است با انواع گوناگونی از دسته‌های واژگانی اسمی و گزاره‌ای آشنا باشند. در رابطه با فراوانی ساختاری دسته‌های واژگانی در بخش بحث مقالات، میرزایی (2019) نیز مشاهده کرد که دسته‌های واژگانی اسمی و گزاره‌ای بیشترین فراوانی را دارند.

جدا از فراوانی و ساختار دسته‌های واژگانی، شناخت نقش کارکرد ایفا شده توسط دسته‌های واژگانی نیز در فصاحت و رسایی متون دانشگاهی اهمیت بسزایی دارد. یافته‌های پژوهش‌های پیشین (e.g., Biber & Barbieri, 2007; Conrad & Biber, 2005) حاکی از آن است که دسته‌های ارجاعی نسبت به دو گروه دیگر یعنی دسته‌های حالتی و گفتمان‌پردازها، گستردگی بیشتری دارند. در راستای همین یافته‌ها، پژوهش حاضر نیز نشان داد که

متداول‌ترین نقش کارکردی که در نوشتار بخش بحث مقالات توسط دسته‌های واژگانی ایفا می‌شود، دسته‌های ارجاعی هستند. هرچند یافته‌های برخی از پژوهش‌های پیشین در رابطه با کارکرد دسته‌های واژگانی، با نتایج مطالعه حاضر در تضاد است. برای مثال در مطالعه چن و بیکر (2010)، دانشجویان انگلیسی‌زبان در نوشتار خود از گفتمان‌پردازهای بیشتر و از دسته‌های ارجاعی کمتری استفاده کرده بودند. پس از بررسی دقیق‌تر دسته‌های واژگانی توسط نویسندگان انگلیسی‌زبان در این پژوهش، می‌توان ادعا کرد که متداول‌ترین کارکردی که دسته‌های واژگانی در بخش بحث مقالات ایفا می‌کنند، نقش ارجاعی است. ادل و ارمن (2012) نیز در پژوهش خود مشاهده کردند دسته‌های ارجاعی در نوشتار نویسندگان انگلیسی‌زبان نسبت به گفتمان‌پردازها و دسته‌های حالتی، فراوانی بالاتری دارد.

یافته‌های به‌دست آمده در این پژوهش می‌تواند سطح آگاهی نویسندگان غیرانگلیسی‌زبان را در پیوند با چگونگی به‌کارگیری دسته‌های واژگانی افزایش دهد. در نتیجه، نویسندگان غیربومی از جمله پژوهشگران ایرانی می‌توانند با آگاهی از این یافته‌ها، مسیر آگاهانه‌تری را بپیمایند. یک روش کارساز در یادگیری دسته‌های واژگانی نگاه دقیق به کارکرد این توالی‌هاست (Cortes, 2004). از آنجایی که به‌کارگیری دسته‌های واژگانی در نوشتار دانشگاهی اهمیت بالایی دارد، تقسیم‌بندی ساختاری و کارکردی دسته‌های واژگانی به‌کاررفته از سوی نویسندگان انگلیسی‌زبان می‌تواند برای طراحی تمرین‌های درسی در کلاس‌های دوره‌های کارشناسی و همچنین تحصیلات تکمیلی در دانشگاه‌ها به‌کار گرفته شود که به موجب آن استادان دانشجویان را به‌شکل پویاتر و هوشمندانه‌تری با دسته‌های واژگانی آشنا سازند؛ به‌ویژه در درس نگارش دانشگاهی، نگارش پیشرفته، و روش تحقیق، زبان‌آموزان نیاز مبرمی به فراگیری و تمرین بر روی موارد ذکرشده دارند.

این پژوهش فقط به بررسی آن دسته از دسته‌های واژگانی پرداخته که نویسندگان انگلیسی‌زبان به‌کار برده بودند. از آنجایی که مطالعات پیشین در پیوند با به‌کارگیری دسته‌های واژگانی میان نویسندگان انگلیسی‌زبان و غیرانگلیسی‌زبان در بردارنده نتایج بعضاً تناقض‌آمیزی بوده است، پژوهشگران علاقه‌مند به چنین موضوعاتی می‌توانند به بررسی مقایسه‌ای به‌کارگیری این گونه از دسته‌های واژگانی میان نویسندگان انگلیسی‌زبان و غیرانگلیسی‌زبان بپردازند. قطعاً نتایج چنین مطالعاتی در رفع ناهمگونی‌های موجود در بررسی‌های پیشین و درک بهتر از کاربرد

این توالی‌ها توسط نویسندگان زبان اول و دوم تأثیر بسزایی خواهد داشت. از آنجایی که انگلیسی، زبان اول چندین کشور است بایسته است تفاوت‌های بین‌فرهنگی در آثار نویسندگان انگلیسی‌زبان که می‌تواند بر نگارش آن‌ها تأثیر بگذارد نیز در آینده بررسی شود. افزون‌براین، این پژوهش تنها بر روی بخش «بحث» مقالات متمرکز شده بود. با توجه به اهمیت دسته‌های واژگانی در متون دانشگاهی به‌ویژه مقالات علمی، به‌کارگیری این توالی‌ها در سایر بخش‌های مقالات نیز مستلزم بررسی بیشتر است. این‌گونه پژوهش‌ها به نویسندگان فارسی‌زبان کمک می‌کند که مقالات و دستاوردهای علمی خود را براساس عرف‌های بین‌المللی و با استانداردهای بالا به رشته‌ تحریر درآورند. همچنین، پیکره‌ این تحقیق فقط متشکل از مقالات رشته‌ زبان‌شناسی کاربردی بود. از آنجایی که به‌کارگیری دسته‌های واژگانی در رشته‌های مختلف متغیر است (Cortes, 2008b, 2008a, Hyland 2008, Durrant, 2017, 2004) بنابراین، یافته‌های این پژوهش را لزوماً نمی‌توان به سایر رشته‌ها تعمیم داد و سزاوار است پژوهش‌هایی از این دست در پیوند با رشته‌های دیگر نیز در آینده به انجام برسد. همچنین در مطالعات پیکره‌ای معمولاً اندازه استاندارد برای هر پیکره یک میلیون واژه است. از آنجایی که تعداد واژگان در پژوهش کنونی از یک میلیون کم‌تر بوده، پژوهشگران در آینده می‌توانند با بررسی دسته‌های واژگانی در بخش بحث مقالات نویسندگان انگلیسی‌زبان از پیکره‌ای بزرگ‌تر بهره‌گیرند تا شاید به نتایج جامع‌تری در این خصوص دست یابند.

۶. پی‌نوشت‌ها

1. Zamel
2. corpus-driven studies
3. multi-word sequences
4. phraseology
5. formulaic language
6. lexical bundles (LBs)
7. functional
8. idioms
9. frequency
10. Cortes
11. range
12. Biber

13. Ädel & Erman
14. Pan
15. dispersion
16. Biber & Barbieri
17. NP/PP-based bundles
18. phrasal bundles
19. VP-based bundles
20. dependent clause bundles
21. clausal bundles
22. Hyland
23. stance bundles
24. discourse organizing bundles
25. referential bundles
26. research-oriented bundles
27. text-oriented bundles
28. participant-oriented bundles
29. genre
30. Staples
31. Chen & Baker
32. Esfandiari & Barbary
33. Bychkovska & Lee
34. Lu & Deng
35. impact factor
36. Laurence Anthony
37. Reliability
38. inter-rater reliability
39. measure of Kappa agreement

7. References

- Ädel, A., & Erman, B. (2012). Recurrent word combinations in academic writing by native and non-native speakers of English: A lexical bundles approach. *English for Specific Purposes*, 31(2), 81–92.
- Ahmadi, H. S., Ghonsooly, B., & Fatemi, A. H. (2013). An analysis of lexical bundles in research article abstracts by Iranian and native English-speaking authors of applied linguistics articles. *The Asian ESP Journal*, 9(1), 5–25.
- Alamri, B. (2020). A comparative study of Saudi and international journals of Applied Linguistics: The move–bundle connection approach. *Journal of Language*

and Education, 6(2), 9–30.

- Biber, D., & Barbieri, F. (2007). Lexical bundles in university spoken and written registers. *English for Specific Purposes*, 26, 263–286.
- Biber, D., Conrad, S., & Cortes, V. (2004). If you look at ...: Lexical bundles in university teaching and textbooks. *Applied Linguistics*, 25, 371–405.
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Pearson.
- Bychkovska, T., & Lee, J. J. (2017). At the same time: Lexical bundles in L1 and L2 university student argumentative writing. *Journal of English for Academic Purposes*, 30, 38–52.
- Chen, Y. H., & Baker, P. (2010). Lexical bundles in L1 and L2 academic writing. *Language Learning & Technology*, 14(2), 30–49.
- Conrad, S. M., & Biber, D. (2005). The frequency and use of lexical bundles in conversation and academic prose. *Lexicographica*, 20, 56–71.
- Cortes, V. (2002). Lexical bundles in Freshman composition. *Using corpora to explore linguistic variation*, 9, 131–145.
- Cortes, V. (2004). Lexical bundles in published and student disciplinary writing: Examples from history and biology. *English for Specific Purposes*, 23, 397–423.
- Cortes, V. (2006). Teaching lexical bundles in the disciplines: An example from a writing intensive history class. *Linguistics and Education*, 17, 391–406.
- Cortes, V. (2013). The purpose of this study is to: Connecting lexical bundles and moves in research article introductions. *Journal of English for Academic Purposes*, 12(1), 33–43.
- Cortes, V. (2015). Situating lexical bundles in the formulaic language spectrum: Origins and functional analysis developments. *In Corpus-based research in applied linguistics* (197–216). John Benjamins.
- Erman, B., & Warren, B. (2000). The idiom principle and the open choice

- principle. *Text- Interdisciplinary Journal for the Study of Discourse*, 20(1), 29–62.
- Esfandiari, R., & Barbary, F. (2017). A contrastive corpus-driven study of lexical bundles between English writers and Persian writers in psychology research articles. *Journal of English for Academic Purposes*, 29, 21–42.
 - Granger, S., & Meunier, F. (Eds.). (2008). *Phraseology: An interdisciplinary perspective*. John Benjamins.
 - Hyland, K. (2008a). As can be seen: lexical bundles and disciplinary variation. *English for Specific Purposes*, 27, 4–21.
 - Hyland, K. (2008b). As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes*, 27(1), 4–21.
 - Li, J., & Schmitt, N. (2009). The acquisition of lexical phrases in academic writing: A longitudinal case study. *Journal of Second Language Writing*, 18(2), 85–102.
 - Lu, X., & Deng, J. (2019). With the rapid development: A contrastive analysis of lexical bundles in dissertation abstracts by Chinese and L1 English doctoral students. *Journal of English for Academic Purposes*, 39, 21–36.
 - Mirzai, M. (2019). *Lexical Bundles in the Discussion Section of Research Articles in Applied Linguistics: A Contrastive Analysis of Quantitative, Qualitative, and Mixed Methods Research*. (Unpublished doctoral dissertation). The University of Zabol, Zabol.
 - Nasrabad, P., Elahi Shirvan, M., & Golparvar, S. E. (2020). Exploring lexical bundles in recent published papers in the field of applied linguistics. *Journal of World Languages*, 6(3), 175–197.
 - Pan, F., Reppen, R., & Biber, D. (2016). Comparing patterns of L1 versus L2 English academic professionals: Lexical bundles in Telecommunications research journals. *Journal of English for Academic Purposes*, 21, 60–71.
 - Perez-Llantada, C. (2014). Formulaic language in L1 and L2 expert academic

writing: Convergent and divergent usage. *Journal of English for Academic Purposes*, 14(2), 84–94.

- Ren, J. (2021). Variability and functions of lexical bundles in research articles of applied linguistics and pharmaceutical sciences. *Journal of English for Academic Purposes*, 50, 100968.
- Shirazizadeh, M., & Amirfazlian, R. (2021). Lexical bundles in theses, articles and textbooks of applied linguistics: Investigating intradisciplinary uniformity and variation. *Journal of English for Academic Purposes*, 49. <https://doi.org/10.1016/j.jeap.2020.100946>.
- Shin, Y. K. (2019). Do native writers always have a head start over nonnative writers? The use of lexical bundles in college students' essays. *Journal of English for Academic Purposes*, 40, 1-14.
- Staples, S., Egbert, J., Biber, D., & McClair, A. (2013). Formulaic sequences and EAP writing development: Lexical bundles in the TOEFL iBT writing section. *Journal of English for Academic Purposes*, 12(3), 214–225.
- Wei, Y., & Lei, L. (2011). Lexical bundles in the academic writing of advanced Chinese EFL learners. *RELC Journal*, 42(2), 155–166.
- Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge University Press.
- Zamel, V (1998). Questioning Academic Discourse. In V. Zamel and R. Spack (Eds.), *Negotiating academic literacies: Teaching and learning across languages and cultures* (187–197). Erlbaum.

۸ پیوست ۱

Appendix A: The Retrieved Four-Word LBs in the Corpus

Rank	Cluster	Frequency	Rank	Rank	Cluster	Frequency	Rank
1	in the present study	80	41	72	of this study was	14	13
2	in the current study	78	37	73	the beginning of the	14	9
3	in the case of	60	41	74	the difference between the	14	10
4	it is possible that	58	46	75	the two types of	14	9
5	the results of the	57	38	76	used in this study	14	13
6	on the other hand	51	41	77	despite the fact that	13	12
7	it is important to	50	42	78	in addition to the	13	11
8	the fact that the	45	34	79	in contrast to the	13	11
9	at the same time	41	31	80	in relation to the	13	13
10	of the current study	41	25	81	is likely to be	13	9
11	the results of this	37	27	82	it may be the	13	11
12	in terms of the	36	31	83	may be the case	13	11
13	results of this study	36	27	84	of the relationship between	13	12
14	as well as the	34	27	85	should be noted that	13	12
15	to the fact that	34	25	86	a great deal of	12	10
16	in the context of	32	25	87	are more likely to	12	10
17	on the basis of	32	21	88	can be seen in	12	11
18	the extent to which	31	26	89	in light of the	12	10
19	the end of the	29	19	90	in the same way	12	10
20	as a result of	28	25	91	in the target language	12	9
21	the L1 and L2	27	10	92	may be due to	12	10
22	the nature of the	27	19	93	of English as a	12	10
23	the use of the	27	20	94	on the part of	12	11
24	of the present study	26	17	95	that the use of	12	9
25	at the end of	24	17	96	the course of the	12	12
26	is important to note	24	22	97	the learners in the	12	8

Rank	Cluster	Frequency	Rank	Rank	Cluster	Frequency	Rank
27	the case of the	24	16	98	the learners in this	12	7
28	with respect to the	24	16	99	the way in which	12	9
29	learners in this study	23	15	100	to be able to	12	12
30	at the level of	22	16	101	to the use of	12	10
31	is possible that the	21	19	102	a large number of	11	10
32	the first research question	21	18	103	can be used to	11	8
33	due to the fact	20	16	104	in the number of	11	9
34	important to note that	20	19	105	in this study was	11	10
35	in other words the	20	17	106	L1 and L2 writers	11	5
36	on the one hand	20	18	107	learners are able to	11	9
37	the findings of this	20	13	108	of the use of	11	9
38	be the case that	19	15	109	that there is a	11	11
39	it may be that	19	16	110	the importance of the	11	7
40	in a variety of	18	13	111	the participants in this	11	10
41	in the united states	18	14	112	the present study was	11	8
42	in this study the	18	17	113	this was not the	11	10
43	the degree to which	18	14	114	to note that the	11	10
44	the other two groups	18	12	115	was not the case	11	10
45	be due to the	17	9	116	a result of the	10	8
46	findings of this study	17	11	117	and the fact that	10	9
47	in the use of	17	14	118	as was the case	10	10
48	it is likely that	17	16	119	could be argued that	10	9
49	of the target language	17	15	120	for the most part	10	10
50	over the course of	17	14	121	found in this study	10	7
51	results of the current	17	13	122	in a way that	10	10
52	the second research question	17	16	123	in line with the	10	10

Rank	Cluster	Frequency	Range	Rank	Cluster	Frequency	Range
53	a wide range of	16	15	124	in terms of their	10	10
54	as a function of	16	13	125	in the process of	10	9
55	both L1 and L2	16	9	126	is also important to	10	10
56	it is clear that	16	15	127	it could be argued	10	9
57	participants in this study	16	14	128	it is interesting to	10	10
58	the participants in the	16	13	129	it is not surprising	10	10
59	this is consistent with	16	12	130	of the target words	10	9
60	with regard to the	16	13	131	on the use of	10	9
61	at the beginning of	15	11	132	our results show that	10	5
62	between L1 and L2	15	9	133	plays a role in	10	7
63	by the fact that	15	13	134	the basis of the	10	9
64	in the form of	15	13	135	the findings of the	10	6
65	it should be noted	15	14	136	the majority of the	10	9
66	the results of experiment	15	6	137	the meaning of the	10	8
67	the strength of the	15	12	138	the present study the	10	9
68	this study was to	15	14	139	the results indicate that	10	7
69	at the time of	14	9	140	the ways in which	10	8
70	in the absence of	14	14	141	to be the case	10	10
71	in this study were	14	12	142	used in the current	10	7