

Extended Abstract

Spoken and Written English Chunks

Maryam Ghasemian

M.A, in Department of Foreign Languages and
Linguistics, Shiraz University, Iran
mary.ghasemian98@gmail.com

Naser Rashidi¹ 

Professor of Department of Foreign Languages
and Linguistics, Shiraz University, Iran
naser.rashidi@shirazu.ac.ir

Abstract

Lexical chunks are conventionalized multi-word units stored in the minds of native speakers that convey a single, unified meaning. Native speakers frequently use these chunks in both spoken and written language. The present study aims to examine spoken and written Persian lexical chunks from a grammatical-structural perspective and to analyze their role in creating coherence in speech and text. The data for this research included natural telephone conversations, selected scenes from several television series, Persian magazines, short stories, texts from various websites, and formal messages exchanged on social media platforms. These data were transcribed and compiled into two separate corpora: one spoken and one written. Then, two-word, three-word, and four-word lexical chunks were extracted using AntConc software and transferred to Excel for further analysis. The lexical chunks were manually identified by the researcher and analyzed based on Como's grammatical framework. The quantitative analysis showed structural differences between spoken and written lexical chunks. In spoken two-word chunks, patterns such as noun+verb, adjective+verb, noun+noun, verb+verb, and numeral+noun were more frequent, whereas written two-word chunks followed patterns like preposition+noun, noun+preposition, and pronoun+noun. Three- and four-word spoken lexical chunks often followed the noun+pronoun+noun+verb pattern, while written three-word chunks frequently followed preposition+pronoun+noun and preposition+noun+preposition structures. Written four-word chunks displayed a wider range of grammatical patterns. The qualitative analysis further revealed that these lexical chunks significantly contribute to the coherence and cohesion of both spoken discourse and written texts.

Keywords:

Como framework
Lexical chunks
Spoken Persian
Written Persian

How to cite this article:

Ghasemian, M. & Rashidi, N. (2026). Spoken and Written English Chunks. *Journal of Linguistics & Iranian Languages*, 10(4), 1-17. doi: <https://doi.org/10.22099/jlil.2026.52884.1467>



COPYRIGHTS ©2026 The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY-NC 4.0), which permits unrestricted use, distribution, and reproduction in any medium, as long as the original authors and source are cited. No permission is required from the authors or the publisher.

Introduction

Lexical chunks are groups of words that co-occur and convey a single, unified meaning. These lexical chunks have become entrenched in the mental lexicon of native speakers. Language users frequently employ lexical chunks in both spoken and written discourse. A review of various studies indicates that lexical chunks hold significant importance in different

¹ Corresponding author

languages. For instance, from an educational perspective as well as in terms of communication, spoken lexical chunks enhance fluency and coherence in speech, while written lexical chunks contribute to textual cohesion. Given the differences between spoken and written varieties of language, spoken and written lexical chunks also differ from each other. Distinguishing between spoken and written lexical chunks is possible through comparison, which helps speakers use spoken chunks in speech and written chunks in writing appropriately. Lexical chunks can be applied in letter writing, journalism, speeches, and teaching. In the first part of the present study, a quantitative analysis, spoken and written lexical chunks in Persian were examined in terms of their grammatical structure. In the second part of the study, which is conducted qualitatively, the effect of lexical chunks on the cohesion of spoken and written Persian was investigated.

Materials & Methods

The present study adopted a mixed-methods approach, incorporating both quantitative and qualitative designs. The spoken data of the study consisted of television series and naturally occurring telephone conversations, while the written data included short stories in Persian, texts from various websites, Persian magazines, and formal messages exchanged on social media platforms. In this study, spoken data were compiled into a spoken corpus, and written data into a written corpus. Initially, using the AntConc concordance software, two-word, three-word, and four-word expressions were extracted from both the spoken and written corpora developed by the researcher and transferred to Microsoft Excel. Subsequently, lexical chunks in Persian were identified manually by the researcher and then classified, examined, and analyzed based on grammatical patterns. The grammatical structure of lexical chunks was determined within Como framework. Regarding the present study, a considerable number of investigations have been conducted. Various studies (Asaei & Rahimi, 2021; Mohammadi & Enayati, 2018; Nemati & Mesgarshahr, 2019; Al-Baqami, 2022) have shown that teaching lexical chunks can significantly improve language learners' speaking and writing skills. Lexical chunks can be applied in letter writing, journalism, speeches, and teaching. In the first part of the present study, a quantitative analysis, spoken and written lexical chunks in Persian were examined in terms of their grammatical structure. In the second part of the study, qualitative in nature, the impact of lexical chunks on the cohesion of spoken and written Persian was investigated.

Results & Discussion

In this study, the grammatical pattern of lexical chunks was determined using Como framework. For example, the grammatical pattern of "noun + verb" was frequently observed among two-word lexical chunks. Expressions such as "khashesh mikonam", "sabr kon", "chizi shode", "aqa bebakhshid", "eshkali nadare", and "moshkeli nist" follow the noun + verb pattern, some of which are compound verbs. Expressions such as "khashesh mikonam", "sabr konid", "negah kon", and "ghol midam" are all compound verbs. Compound verbs refer to verbs in which the first element is a noun or adjective and the second element is sometimes an auxiliary verb such as "kardan" or "shodan". The expressions "dastet dard nakone" and "dasteton dard nakone" follow the structural pattern "noun + pronoun + noun + verb".

Qualitative analysis of the data also revealed that lexical chunks, as pre-fabricated units, play an important role in creating textual cohesion. In spoken language varieties, lexical chunks such as "na baba", "khashesh mikonam", "ye lahze", "ghorbonet beram", "dastet dard nakone", "sad dar sad", and "be har hal" were observed, all of which play an important role in discourse cohesion. For instance, the expression "na baba" is used as a response to a question or to express surprise. This expression indicates the speaker's stance toward the question and contributes to

cohesion. The expression “ye lahze” is used when the speaker intends to pause the conversation briefly and take control of the interaction. The expression “ghorbonet beram” is a lexical chunk used to increase intimacy in speech. An expression such as “sad dar sad” was also identified among two-word lexical chunks and is used to confirm the interlocutor’s statement. The expression “be har hal”, which is a three-word lexical chunk, is used for topic shift and conclusion.

In written language, lexical chunks such as “momken ast”, “pas az”, “nesbat be”, and “say konid” were observed. For example, the expression “momken ast” is used to indicate speculation or possibility and helps the writer produce non-absolute statements while creating logical cohesion between academic sentences. The lexical chunk “pas az” contributes to temporal cohesion in sentences and indicates the sequence of events. Among two-word written lexical chunks, the expression “nesbat be” was also observed, which is generally used to express comparison between two entities. An expression such as “say konid” is an imperative form that can also be used in instructional texts and contributes to textual cohesion.

Conclusion

In the present study, spoken and written lexical chunks of the Persian language, extracted and selected from a researcher-developed corpus, were examined based on their grammatical structures. The grammatical structures of spoken and written lexical chunks were identified according to the Como framework. The results of the first part of the study showed that two-word spoken lexical chunks mostly include grammatical patterns such as “noun + verb”, “adjective + verb”, “noun + noun”, “verb + verb”, and “numeral classifier + noun”. In addition, three-word and four-word spoken lexical chunks mostly include the grammatical pattern “noun + pronoun + noun + verb”. Among written two-word lexical chunks, patterns such as “preposition + noun”, “noun + preposition”, and “pronoun + noun” were more frequently observed than other patterns. Also, among written three-word lexical chunks, the patterns “preposition + pronoun + noun” and “preposition + noun + preposition” were identified. Various grammatical patterns were also observed in written four-word lexical chunks. In the second part of the study, it was also shown how spoken and written lexical chunks contribute to the cohesion and coherence of spoken and written discourse.

References

- Albaqami, R. (2022). The role of lexical chunks in promoting English writing competence among foreign language learners in Saudi Arabia. *Arab World English Journal*, 13(2): 91-108.
- Asaei, S. M., & Rahimy, R. (2021). Appropriation-based syllabus and speaking ability: Evidence from Iranian EFL context. *Biannual Journal of Education Experiences*, 4(2): 195-2014.
- Behruz, S. (2022). The use of the approved vocabulary of the academy of Persian Language and Literature: A quantitative and qualitative study of theoretical branch textbooks (2021). *Linguistic Research*, 1-29
- Bijan khan, M., Heshmati, Z., Maleki vika, M., & Veisi, H. (2024). Identification and extraction of Persian collocations using computational methods. *Journal of Information Processing and Management (IranDoc)*, 40(2), 557-604
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Harlow, England: Pearson Education Limited.
- Biber, D., Conrad, S., & Cortes, V. (2004). If you look at.....Lexical bundles in university teaching and textbooks. *Applied Linguistics*, 25(3), 371-405.
<https://doi.org/10.1093/applin/25.3.371>
- Chun-Guang, T. (2014). An empirical research on the corpus-driven lexical chunks instruction. *International Journal of English Language Teaching*, 2(2):1-36.

- Farahani, A., Yazdani, H., Ahmadian, M., & Amarian, M. (2020). Investigating the use of English expressions in applied linguistics research articles: A corpus based study. *Linguistic Research in Foreign Languages*, 2(10), 390-405.
- Ghasemian, M., & Rashidi, N. (2024). An Analysis of Lexical Chunks in Mina 1 and Let's Learn Persian 1 Textbooks: A Comparative Study Based on Lewis Model and Their Degree of Fixedness [In Persian]. *Journal of Studies in Applied Language (JSAL)*, 7(2): 93–111.
- Hakuta, K. (1974). Prefabricated patterns and the emergence of structure in second language acquisition 1. *Language Learning*, 24(2): 287-297.
- Kaveh, M., & Afrashi, A. (2019). Formulaic language and its role in teaching Persian. *Language studies. Institute for Humanities and Cultural Studies*, 99-122.
- Lackman, K. (2011). *Lexical approach activities*. Ken Lackman & Associates Educational Consultants.
- Li, Q. (2014). An empirical study on the application of lexical chunk to college English writing. *Journal of Language Teaching and Research*, 5(3): 682.
- McCarthy, M. (2004). *From corpus to coursebook*. Cambridge: Cambridge University Press
- Mohammadi, M., & Enayati, B. (2018). The effects of lexical chunks teaching on EFL intermediate learners' speaking fluency. *International Journal of Instruction*, 11(3): 179-192.
- Nemati, M., & Mesgarshahr, E. (2019). The effect of explicit teaching of lexical chunks on vocabulary use quality and grammatical accuracy in English writing: An applied approach. *Linguistic Research in Foreign Languages*, 9(3), 959-990.
- Richards, J. C., & Rodgers, T. S. (2014). *Approaches and methods in language teaching*. Cambridge university press.
- Sauri, R., Alderslade, A., & Shapiro, R. (2018). A universal classification of lexical categories and grammatical distinctions for lexicographic and processing purposes. In *Proceedings of the XVII EURALEX International Congress*, 2018, pp.173-185. Oxford University Press.
- Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford University Press.
- Wood, D. 2002. Formulaic language acquisition and production: Implications for teaching. *TESL Canada Journal*, 01-15.

زبان‌شناسی و گویش‌های ایرانی

سال ۱۰، شماره ۴، پیاپی ۱۹ (زمستان ۱۴۰۴) شماره صفحات: ۱ - ۱۷

تکه‌های واژگانی گفتاری و نوشتاری زبان فارسی

مریم قاسمیان^۱، ناصر رشیدی^{۲*}

۱. کارشناس ارشد آموزش زبان فارسی به غیر فارسی زبانان، دانشگاه شیراز، شیراز، ایران

۲. استاد آموزش زبان انگلیسی، دانشگاه شیراز، شیراز، ایران

چکیده

تکه‌های واژگانی، عباراتی نهاده شده در ذهن اهل زبان هستند که از دو واژه یا بیشتر تشکیل شده‌اند، اما معنایی که منتقل می‌کنند یک معنای واحد است. اهل زبان، در گفتار و نوشتار از تکه‌های واژگانی به‌وفور استفاده می‌کنند. هدف پژوهش حاضر بررسی تکه‌های واژگانی گفتاری و نوشتاری زبان فارسی از منظر ساختار دستوری و تحلیل نقش این تکه‌های واژگانی در ایجاد انسجام در گفتار و متن است. داده‌های این پژوهش شامل گفتگوهای تلفنی طبیعی، قسمت‌هایی از چند سریال تلویزیونی، مجله‌های فارسی، داستان‌های کوتاه، متونی از وبسایت‌های مختلف و پیام‌های رسمی ردوبدل شده در شبکه‌های اجتماعی هستند که پس از ترانویسی به دو پیکره مستقل گفتاری و نوشتاری تبدیل شدند. سپس عبارات دو کلمه‌ای، سه کلمه‌ای و چهار کلمه‌ای توسط نرم‌افزار انت-کانک استخراج شده و به اکسل منتقل شدند. تکه‌های واژگانی زبان فارسی توسط محقق به‌طور شمی مشخص گردید و بر اساس چارچوب دستوری کومو بررسی و تحلیل شدند. نتایج حاصل از تحلیل کمی نشان داد که تفاوت‌هایی میان الگوی دستوری تکه‌های واژگانی گفتاری و نوشتاری وجود دارد. تکه‌های واژگانی گفتاری دو کلمه‌ای بیشتر از الگوی دستوری اسم + فعل، صفت + فعل، اسم + اسم، فعل + فعل، تعداد نما + اسم پیروی می‌کنند. درحالی‌که تکه‌های واژگانی نوشتاری دو کلمه‌ای بیشتر از الگوی حرف اضافه + اسم، اسم + حرف اضافه، ضمیر + اسم پیروی می‌کنند. همچنین تکه‌های واژگانی گفتاری سه کلمه‌ای و چهار کلمه‌ای دارای الگوی دستوری اسم + ضمیر + اسم + فعل هستند؛ اما تکه‌های واژگانی نوشتاری سه کلمه‌ای دارای الگوی دستوری حرف اضافه + ضمیر + اسم و حرف اضافه + اسم + حرف اضافه هستند. تکه‌های واژگانی نوشتاری چهار کلمه‌ای نیز انواع مختلفی از الگوهای دستوری را شامل می‌شود. نتایج حاصل از تحلیل کیفی نشان داد که این تکه‌های واژگانی در انسجام گفتار و متن نقش مؤثری ایفا می‌کنند.

واژه‌های کلیدی:

تکه‌های واژگانی
چارچوب کومو
فارسی گفتاری
فارسی نوشتاری

تاریخچه مقاله: دریافت: ۱ اردیبهشت‌ماه ۱۴۰۵

بازنگری: ۱ خردادماه ۱۴۰۵

پذیرش: ۱۲ خردادماه ۱۴۰۵

* آدرس ایمیل نویسنده مسئول: naser.rashidi@shirazu.ac.ir

قاسمیان، مریم و رشیدی، ناصر. (۱۴۰۴). تکه‌های واژگانی گفتاری و نوشتاری زبان فارسی. زبان‌شناسی و گویش‌های ایران، ۱۰(۴)، ۱-۱۷

<https://doi.org/10.22099/jlil.2026.52884.1467>

حق نشر ۲۰۲۶-نویسنده(گان). این مقاله با دسترسی آزاد است که تحت شرایط مجوز Creative Commons Attribution (CC BY-NC 4.0)

منتشر شده است، استفاده، توزیع و تکثیر نامحدود را در هر رسانه‌ای مجاز می‌داند، مشروط بر اینکه به نویسندگان و منبع اصلی اشاره شود. هیچ اجازه‌ای از نویسندگان یا ناشر لازم نیست.

۱. مقدمه

تکه‌های واژگانی^۱، گروهی از واژگان هستند که با یکدیگر به کار رفته و مفهوم واحدی را منتقل می‌کنند. این تکه‌های واژگانی در ذهن اهل زبان نهادینه شده‌اند. اهل زبان در گفتار و نوشتار از تکه‌های واژگانی به‌وفور استفاده می‌کنند. «روشن است که زبان فرمول‌واره‌ای^۲ در فراگیری و تولید زبان نقش بسزایی دارد.» (وود^۳، ۲۰۰۲: ۲).

به عقیده وود (۲۰۰۲)، تکه‌های واژگانی گفتاری به داشتن گفتاری روان و سلیس و تکه‌های واژگانی نوشتاری به انسجام متن کمک می‌کنند. تکه‌های واژگانی در گونه گفتاری و در گونه نوشتاری زبان وجود دارند. تکه‌های واژگانی گفتاری عباراتی هستند که به‌طور روزمره توسط اهل زبان استفاده می‌شوند. تکه‌های واژگانی نوشتاری نیز عباراتی هستند که به‌وفور در گونه نوشتاری زبان، استفاده می‌شوند و به انسجام متن کمک می‌کنند. واژگان تا زمانی که در متن و بافت جای‌گیر نشوند، نه تنها درجه کاربردپذیری و قابلیت‌های زبان‌شناختی و معناشناسی آن‌ها آشکار نمی‌شود، بلکه میزان درک کاربر از آن واژه و سطح دانش زبانی وی پوشیده می‌ماند (بهروز، ۱۴۰۱).

بررسی پژوهش‌های مختلف نشان می‌دهد که تکه‌های واژگانی در زبان‌های مختلف اهمیت قابل‌توجهی دارند. به‌عنوان مثال، از منظر آموزشی و همچنین از منظر برقراری ارتباط تکه‌های واژگانی گفتاری باعث افزایش روانی و انسجام گفتار و تکه‌های واژگانی نوشتاری سبب انسجام متن می‌شوند. در زبان انگلیسی تحقیقاتی در زمینه تکه‌های واژگانی صورت گرفته است و در این تحقیقات الگوی دستوری تکه‌های واژگانی و نوع آن‌ها بررسی می‌شوند. برای نمونه، رویکرد واژگانی، رویکردی برای آموزش واژگان است که نخستین بار توسط لوئیس^۴ (۱۹۹۷) معرفی شد. لوئیس (۱۹۹۷)، تکه‌های واژگانی را به کلمات، چندکلمه‌ای‌ها، هماینها، پاره‌گفتارهای نهادینه‌شده، قالب‌ها و هسته‌های جمله‌ای دسته‌بندی می‌کند (گانگ^۵، ۲۰۱۴).

همچنین این تحقیقات نشان داده‌اند که استفاده از تکه‌های واژگانی به بهبود انسجام در گفتار و متن کمک شایانی می‌کند. با توجه به تفاوت میان گونه‌های گفتاری و نوشتاری زبان، تکه‌های واژگانی گفتاری و نوشتاری نیز با یکدیگر تفاوت دارند. تشخیص تکه‌های واژگانی گفتاری و نوشتاری از یکدیگر توسط گویشور زبان، با مقایسه آن‌ها امکان‌پذیر است و به گویشور کمک می‌کند که در گفتار از تکه‌های واژگانی گفتاری و در نوشتار از تکه‌های واژگانی نوشتاری استفاده کند. تکه‌های واژگانی می‌توانند در نامه‌نگاری، روزنامه‌نگاری، سخنرانی‌ها و تدریس کاربرد داشته باشند؛ در بخش اول پژوهش حاضر که تحلیل کمی است، تکه‌های واژگانی گفتاری و نوشتاری زبان فارسی از نظر ساختار دستوری‌شان مورد بررسی قرار گرفته‌اند و در بخش دوم پژوهش که به صورت کیفی انجام شده به بررسی چگونگی تاثیر تکه‌های واژگانی در ایجاد انسجام گفتار و نوشتار فارسی پرداخته شده است.

۲- پرسش‌های پژوهش

۱- آیا الگوی دستوری تکه‌های واژگانی گفتاری و نوشتاری با یکدیگر تفاوت دارند؟

۲- آیا تکه‌های واژگانی گفتاری و نوشتاری در انسجام گفتار و نوشتار مؤثر هستند؟

¹ Lexical Chunks

² Formulaic Language

³ Wood

⁴ Lewis

⁵ Gaung

۳. مبانی نظری

تکه‌های واژگانی به گروهی از واژگان گفته می‌شود که در ذهن اهل زبان ذخیره شده‌اند و در تولید گفتار و نوشتار منسجم نقشی کلیدی دارند. تاکنون چارچوب‌های نظری مختلفی برای تکه‌های واژگانی ارائه شده است.

در پژوهش حاضر برای بررسی الگوی دستوری تکه‌های واژگانی گفتاری و نوشتاری از چارچوب نظری کومو^۱ استفاده شد؛ زیرا این چارچوب یک چارچوب جهانی است که توسط سائوری^۲ و همکاران (۲۰۱۸) ارائه شده است. هدف کلی چارچوب کومو ایجاد یک چارچوب جهانی برای دسته‌بندی مقوله‌های واژگانی است به طوری که به جای اینکه هر زبان با قواعد سنتی خودش توصیف شود، با یک چارچوب مشترک بتوان همه زبان‌ها را مقایسه و تحلیل کرد. یکی از اصول چارچوب کومو این است که نه تنها برای زبان‌های اروپایی بلکه برای زبان‌های خیلی متفاوت قابل استفاده است. در این رویکرد، تعداد مقوله‌های واژگانی تا حد امکان محدود و کلی در نظر گرفته می‌شود تا چارچوب بتواند برای زبان‌هایی با ساختارهای متفاوت نیز کاربرد داشته باشد. این چارچوب مقوله‌های واژگانی اسم، فعل، صفت، قید، حرف اضافه، ضمیر، حرف ربط^۳، شناسه^۴، وجه نما^۵، میانجی^۶، تعداد نما^۷، پیش مشخص کننده^۸، ادغام شده^۹، شناسه ترکیبی^{۱۰}، آواژه^{۱۱}، عبارت اصطلاحی^{۱۲}، نشانه‌گذاری^{۱۳}، باقیمانده‌ها^{۱۴} و واژک^{۱۵} را شامل می‌شود.

لویس (۱۹۹۷) و ناتینگر و دیکاریو^{۱۶} (۱۹۹۲)، تکه‌های واژگانی را بر اساس نوع آن‌ها دسته‌بندی می‌کنند (گانگ، ۲۰۱۴). ریچاردز و راجرز^{۱۷} (۲۰۱۴) نیز تکه‌های واژگانی را بر همین اساس دسته‌بندی می‌کنند. لویس (۱۹۹۷) این تکه‌های واژگانی را شامل کلمات^{۱۸}، چند کلمه‌ای‌ها^{۱۹}، هماینده‌ها^{۲۰}، پاره گفتارهای نهادینه شده^{۲۱}، قالب‌ها و هسته‌های جمله‌ای^{۲۲} می‌داند؛ درحالی‌که در چارچوب ناتینگر و دیکاریو (۱۹۹۲) دسته‌بندی‌هایی نظیر چند کلمه‌ای‌ها، عبارات نهادینه شده^{۲۳}، قیده‌های عبارتی^{۲۴} و جمله-سازها^{۲۵} مشاهده می‌شوند. در رویکرد ریچاردز و راجرز (۲۰۱۴) نیز تکه‌های واژگانی هماینده‌ها، دو جمله‌ای‌ها^{۲۶}، سه جمله‌ای‌ها^{۲۷}،

¹ Como

² Sauri

³ Conjunction

⁴ Person marker

⁵ Mood marker

⁶ Interjection

⁷ Numeral

⁸ Predeterminer

⁹ Clitic

¹⁰ Compound marker

¹¹ Ideophone

¹² Idiom

¹³ Punctuation

¹⁴ Residual

¹⁵ Token

¹⁶ Nattinger & Decarrio

¹⁷ Richards & Rodgers

¹⁸ Words

¹⁹ Poly-words

²⁰ Collocations

²¹ Institutionalized Utterances

²² Sentence frames and heads

²³ Institutionalized expressions

²⁴ Phrasal constraints

²⁵ Sentence builders

²⁶ Binomials

²⁷ Trinomials

اصطلاحات^۱، تشبیهات^۲، کلمه ربط یا عطف^۳، عبارت‌های روزمره اجتماعی^۴، نشانگر گفتمانی^۵، چند کلمه‌ای‌ها، ضرب‌المثل‌ها^۶ و حرف‌ندا^۷ را شامل می‌شوند. ژو^۸ (۲۰۱۶) به استناد از بکر^۹ (۱۹۷۵) تکه‌های واژگانی را به شش دسته چندکلمه‌ای‌ها قیده‌ای عبارتی، جمله‌سازها^{۱۰}، فرایام‌ها^{۱۱}، پاره‌گفتارهای موقعیتی^{۱۲} و متن‌های کلمه به کلمه^{۱۳} تقسیم می‌کند. لی^{۱۴} (۲۰۱۴) به استناد از مون^{۱۵} (۱۹۹۷) تکه‌های واژگانی را به پنج دسته کلمات ترکیبی^{۱۶}، افعال، اصطلاحات^{۱۷}، عبارات ثابت^{۱۸} و پیش ساخته‌ها^{۱۹} تقسیم می‌کند.

از سوی دیگر سینکلر^{۲۰} (۱۹۹۱) چارچوب گرامر واژگانی را مطرح کرد که بر این اصل استوار است که ساختارهای زبانی بیشتر از آن که بر اساس دستور زبان باشند، بر پایه تکه‌های واژگانی تکرار شونده شکل می‌گیرند و بر اساس این رویکرد، زبان متکی بر باهم‌آیی واژگان و گزینش مشترک واژگان است. بیجن‌خان و همکاران (۱۴۰۳) نیز از چارچوب نظری فرث^{۲۱} (۱۹۵۷) استفاده کردند که مانند چارچوب سینکلر (۱۹۹۱) بر اهمیت هم‌نشینی واژه‌ها در تعیین معنا تأکید دارد و به نقش بافت در شکل‌گیری معنای واژگان می‌پردازد.

برخی پژوهشگران نیز تکه‌های واژگانی را بر اساس ساختار دستوری آن‌ها دسته‌بندی کرده‌اند. برای نمونه چارچوب ناتز^{۲۲} که توسط بایبر^{۲۳} و همکاران (۱۹۹۹) معرفی شد، بر عبارت‌هایی که یک هسته اسمی همراه با وابسته در آن‌ها مشاهده می‌شود تمرکز دارد. این نوع الگوهای دستوری معمولاً در متون گفتاری و نوشتاری پربسامد هستند.

۴. پیشینه پژوهش

در رابطه با پژوهش حاضر تحقیقات زیادی به خصوص در زبان انگلیسی صورت گرفته‌است. هاکوتا^{۲۴} (۱۹۷۴) با تحلیل نمونه گفتار یک کودک پنج ساله ژاپنی که زبان انگلیسی را می‌آموخت نشان داد که الگوهای زبانی پیش ساخته در یادگیری زبان دوم نقش مؤثری دارند. هاکوتا (۱۹۷۴) نمونه گفتار یک کودک پنج ساله ژاپنی را که در حال یادگیری زبان انگلیسی بود، در طول ۶۰ هفته مورد بررسی قرار داد. پژوهشی که هاکوتا انجام داد از جهات زیادی به پژوهش حاضر شبیه نیست؛ زیرا در ابتدا اینکه

¹ Idioms

² Similes

³ Connectives

⁴ Social routine formulae

⁵ Discourse Markers

⁶ Proverbs

⁷ Exclamations

⁸ Xu

⁹ Becker

¹⁰ Sentence frames

¹¹ Meta-messages

¹² Situational-Utterances

¹³ Verbatim texts

¹⁴ Li

¹⁵ Moon

¹⁶ Compounds

¹⁷ Idioms

¹⁸ Fixed phrases

¹⁹ Prefabs

²⁰ Sinclair

²¹ Firth

²² NATS

²³ Biber

²⁴ Hakuta

هاکوتا گفتار کودک را مورد بررسی قرار داده‌است. میان گفتار کودکان و بزرگسالان تفاوت‌های قابل توجهی دیده می‌شود. دوم اینکه، پژوهشی که هاکوتا انجام داد را نمی‌توان به کل کودکان یا جامعه زبانی تعمیم داد؛ زیرا فقط گفتار یک کودک را شامل می‌شود؛ اما پژوهش حاضر پیکره بنیاد است و داده‌های موجود در این پیکره‌ها از منابع مختلفی جمع‌آوری شده‌اند و تا حدودی قابل تعمیم به جامعه زبانی هستند. سوم اینکه هاکوتا (۱۹۷۴) فقط گونه گفتار را مورد بررسی قرار داده است؛ اما در پژوهش حاضر هر دو گونه گفتار و نوشتار مورد بررسی قرار گرفته‌اند؛ بنابراین پژوهش حاضر با پژوهش هاکوتا (۱۹۷۴) همسو نیست.

در پژوهشی دیگر فراهانی و همکاران (۱۳۹۸) در مقاله‌های پژوهشی به مقایسه کاربرد اصطلاحات انگلیسی میان نویسندگان فارسی زبان و انگلیسی زبان پرداختند. نتایج نشان داد که نویسندگان انگلیسی زبان بسیار بیشتر از نویسندگان فارسی زبان، از اصطلاحات انگلیسی استفاده کرده‌اند. این پژوهش نشان می‌دهد که عبارات فرمول‌واره‌ای در ذهن اهل زبان ذخیره شده‌اند و به همین دلیل به کارگیری آن‌ها در سخنرانی‌ها، نوشتن مقالات، نامه‌نگاری، آموزش زبان و غیره حائز اهمیت است.

پژوهش‌های مختلفی مانند (عصائی و رحیمی، ۲۰۲۱؛ محمدی و عنایتی، ۲۰۱۸؛ نعمتی و مسگرشهر، ۱۳۹۸؛ البقामी، ۲۰۲۲) نشان داده‌اند که آموزش تکه‌های واژگانی می‌تواند باعث بهبود چشمگیر مهارت گفتاری و نوشتاری زبان آموزان شود. هر چند این مطالعات بیشتر به اثرات آموزشی تکه‌های واژگانی و عملکرد زبان آموزان پرداخته‌اند و در آن‌ها ساختار دستوری تکه‌های واژگانی بررسی نشده است.

قاسمیان و رشیدی (۱۴۰۳) به مقایسه تکه‌های واژگانی گفتگوهای کتاب‌های آموزش زبان فارسی «مینا ۱» و «فارسی بیاموزیم ۱» با رویکرد پیکره محور در زبان فارسی پرداختند. در این پژوهش گفتگوهای کتاب «مینا ۱» و «فارسی بیاموزیم ۱» ترانویسی شده و به دو پیکره تبدیل شدند. سپس تکه‌های واژگانی دو کلمه‌ای، سه کلمه‌ای و چهار کلمه‌ای بر پایه چارچوب لوپس (۱۹۹۷)، استخراج و تحلیل شدند. یافته‌ها نشان داد که تکه‌های واژگانی موجود در کتاب مینا ۱ از نظر تعداد و درجه ثبات (ثابت یا نیمه ثابت بودن) نسبت به کتاب فارسی بیاموزیم ۱ پیچیده‌تر و متنوع‌تر هستند. به اعتقاد آن‌ها کتاب فارسی بیاموزیم ۱ برای زبان آموزان مبتدی‌تر و کتاب مینا ۱ برای زبان آموزانی که آشنایی اندکی با زبان فارسی دارند مناسب‌تر هستند. در پژوهش حاضر نیز تکه‌های واژگانی گفتاری و نوشتاری با یکدیگر مقایسه شدند؛ بنابراین می‌توان گفت که این مطالعه تا حد قابل توجهی با پژوهش حاضر همسو است.

با پیشرفت زبان‌شناسی پیکره‌ای، شاهد اهمیت پیکره و پژوهش‌های پیکره بنیاد هستیم. همچنین طراحی کتاب‌های آموزشی پیکره بنیاد رو به افزایش است. مک کارتی (۲۰۰۴)، درباره مجموعه کتاب‌های تاج استون (مک کارتی و همکاران)، توضیحاتی را ارائه می‌دهد. بر اساس توضیحات وی، مجموعه کتاب‌های تاج استون بر اساس پیکره کمبریج، بخش انگلیسی آمریکای شمالی گردآوری شده است. نویسندگان مجموعه تاج استون از پیکره استفاده کردند تا پرتکرارترین واژگان و عبارات را شناسایی کنند. داده‌های پژوهش حاضر نیز به پیکره گفتاری و نوشتاری تبدیل شدند و در این پژوهش نیز پرتکرارترین عبارت‌ها شناسایی شد. از این نظر، پژوهش مک کارتی (۲۰۰۴) با پژوهش حاضر همسو است.

کاوه و افراشی (۱۳۹۸)، عبارت‌های فرمول‌واره‌ای را معرفی و توصیف می‌کنند و به تحلیل عبارت‌های فرمول‌واره‌ای در زبان فارسی می‌پردازند. آن‌ها تکه‌های واژگانی را به دو دسته ثابت و زایا تقسیم می‌کنند و داده‌های پژوهش را از پایگاه داده‌های زبان فارسی، وب گردی، سریال‌های تلویزیونی و مشاهدات شخصی گردآوری کرده‌اند. نتایج این پژوهش نشان می‌دهد که عبارات

فرمول‌واره‌ای در زبان فارسی تقریباً در تمامی بافت‌های فرهنگی و اجتماعی به کار می‌روند و همچنین تألیف کتاب‌های آموزش زبان فارسی با تأکید بر عبارات فرمول‌واره‌ای می‌تواند سبب انسجام و روانی گفتار و نوشتار زبان آموزان شود. در پژوهش حاضر نیز داده‌های پژوهش از سریال‌های تلویزیونی و برخی وب‌سایت‌های فارسی زبان گردآوری شده است و سپس به تحلیل این داده‌ها پرداخته شده است. از این نظر پژوهش حاضر با پژوهش کاوه و افراشی (۱۳۹۸) همسو است.

۵. روش پژوهش

هدف پژوهش حاضر بررسی تکه‌های واژگانی گفتاری و نوشتاری زبان فارسی بر اساس الگوی دستوری آن‌ها است. طرح پژوهش حاضر به روش کمی و کیفی است. داده‌های گفتاری پژوهش حاضر، شامل سریال‌های تلویزیونی و گفت‌وگوهای تلفنی طبیعی و داده‌های نوشتاری نیز شامل داستان‌های کوتاه به زبان فارسی، متون وب‌سایت‌های مختلف، مجله‌های فارسی و پیام‌های رسمی رد و بدل شده در شبکه‌های اجتماعی هستند. در این پژوهش ابتدا محقق به ترانویسی تعدادی گفت‌وگوی تلفنی که از قبل ضبط شده بودند، پرداخت. سپس، قسمت‌هایی از چند سریال تلویزیونی توسط محقق انتخاب و ترانویسی شدند. مجله‌های فارسی و داستان‌های کوتاه، متونی از وب‌سایت‌های مختلف و پیام‌های رسمی ردوبدل شده در شبکه‌های اجتماعی جمع‌آوری شدند. داده‌های گفتاری این پژوهش که گفتگوهای تلفنی و سریال‌های تلویزیونی بودند، پس از ترانویسی به پیکره گفتاری و داده‌های نوشتاری این پژوهش که داستان‌های کوتاه، متون وب‌سایت‌های مختلف، مجله‌های فارسی و پیام‌های رسمی ردوبدل شده در شبکه‌های اجتماعی بودند نیز به پیکره نوشتاری تبدیل شدند. پیکره گفتاری پژوهش حاضر ۸۶۷۳۵ واژه و پیکره نوشتاری ۸۴۴۳۲ واژه دارد.

ابتدا، با استفاده از نرم‌افزار واژه‌نمای انت‌کانک^۱ عبارات دوکلمه‌ای، سه‌کلمه‌ای و چهارکلمه‌ای از پیکره گفتاری و پیکره نوشتاری محقق ساخته، استخراج شده و به برنامه اکسل^۲ منتقل شد. سپس تکه‌های واژگانی زبان فارسی توسط محقق به‌طور شمی مشخص گردید و در مرحله بعد، تکه‌های واژگانی بر اساس الگوی دستوری دسته‌بندی، بررسی و تحلیل شدند. ساختار دستوری تکه‌های واژگانی گفتاری و نوشتاری طبق چارچوب کومو مشخص گردید.

۶. یافته‌های پژوهش: بررسی تکه‌های واژگانی بر اساس الگوی دستوری آن‌ها- تحلیل کمی تکه‌های واژگانی

با توجه به تحلیل داده‌ها می‌توان تکه‌های واژگانی زبان فارسی را بر اساس الگوی دستوری آن‌ها بررسی کرد. ابتدا تکه‌های واژگانی گفتاری را بر اساس الگوی دستوری آن‌ها بررسی می‌کنیم. در این پژوهش، الگوی دستوری تکه‌های واژگانی توسط چارچوب کومو مشخص می‌شود.

^۱ Ant-conc

^۲ Excel

جدول ۱

دسته بندی تکه‌های واژگانی گفتاری دو کلمه‌ای بر اساس الگوی دستوری

تکه واژگانی	تعداد زیر پیکره	بسامد	ساختار دستوری
خواهش میکنم	۶	۱۳۲	اسم + فعل
از این	۶	۱۰۱	حرف اضافه + ضمیر اشاره
خیل خب	۶	۹۴	قید + صفت
نه بابا	۶	۹۲	وجه نما + اسم
یعنی چی	۵	۹۱	فعل + ضمیر
خیلی ممنون	۶	۸۸	قید + صفت
یه لحظه	۶	۸۰	تعداد نما + اسم
جناب سروان	۲	۷۶	اسم + اسم
به خاطر	۶	۷۱	حرف اضافه + اسم
هر چی	۶	۶۶	صفت + ضمیر
چی شده	۵	۶۶	ضمیر + فعل
چه جوری	۶	۵۸	ضمیر + اسم
قربونت برم	۶	۵۶	اسم + ضمیر + فعل
ای بابا	۵	۵۴	میانه‌چی + اسم
آقای دکتر	۴	۵۱	اسم + اسم
برای چی	۶	۵۱	حرف اضافه + ضمیر
اون وقت	۶	۴۹	ضمیر + اسم
یه چیزی	۶	۴۹	تعداد نما + اسم
یه خرده	۵	۴۸	تعداد نما + اسم
سلامت باشین	۲	۴۷	صفت + فعل
این همه	۶	۴۴	ضمیر + اسم
یه دقیقه	۶	۴۱	تعداد نما + اسم
آقا چون	۴	۳۹	اسم + اسم
صبر کن	۵	۳۹	اسم + فعل
همه چی	۶	۳۹	اسم + ضمیر
چیکار می‌کنی	۶	۳۹	ضمیر + اسم + فعل
از کجا	۵	۳۸	حرف اضافه + قید
همین الان	۶	۳۸	ضمیر + قید
الو سلام	۵	۳۷	اسم + اسم
زنگ می‌زنم	۵	۳۶	اسم + فعل
تموم شد	۶	۳۵	اسم + فعل
واسه چی	۵	۳۵	حرف اضافه + ضمیر
به به	۵	۳۴	میانه‌چی
ای وای	۶	۳۲	میانه‌چی
چه می‌دونم	۶	۳۲	ضمیر + فعل
اجازه بدین	۵	۳۰	اسم + فعل
اون ور	۶	۳۰	ضمیر + اسم
برای اینکه	۶	۳۰	حرف اضافه + حرف ربط
بیا بریم	۴	۲۹	فعل + فعل

تکه واژگانی	تعداد زیر پیکره	بسامد	ساختار دستوری
چیکار کنم	۵	۲۹	ضمیر + اسم + فعل
قبول شده	۱	۲۷	اسم + فعل
چی شد	۵	۲۷	ضمیر + فعل
یه دونه	۶	۲۷	تعداد نما + اسم
در مورد	۵	۲۶	حرف اضافه + اسم
فکر می‌کنم	۶	۲۶	اسم + فعل
ول کن	۵	۲۶	صفت + فعل
بله قربان	۳	۲۵	اسم + اسم
دمت گرم	۴	۲۵	اسم + ضمیر + صفت
دیگه نه	۶	۲۵	وجه نما + وجه نما
بنده خدا	۴	۲۴	اسم + اسم
بیا بیا	۴	۲۴	فعل + فعل
خوب نیست	۵	۲۴	صفت + فعل
معذرت می‌خوام	۵	۲۴	اسم + فعل
من برم	۵	۲۴	ضمیر + فعل
به سلامتی	۵	۲۳	حرف اضافه + اسم
شب بخیر	۵	۲۳	اسم + صفت
هر کی	۵	۲۳	صفت + ضمیر
آره دیگه	۶	۲۲	اسم + وجه نما
حال شما	۴	۲۲	اسم + ضمیر
دارم میرم	۶	۲۲	فعل + فعل
مهم نیست	۴	۲۲	صفت + فعل
چه خبره	۵	۲۲	ضمیر + فعل
چیکار داری	۵	۲۲	ضمیر + اسم + فعل
اوه اوه	۳	۲۱	میانجی
بابا جون	۴	۲۱	اسم + اسم
زنگ زدم	۳	۲۱	اسم + فعل
هیچ وقت	۵	۲۱	صفت + اسم
اون موقع	۶	۲۰	ضمیر + قید
با اجازتون	۶	۲۰	حرف اضافه + اسم + ضمیر
بهت می‌گم	۴	۲۰	حرف اضافه + ضمیر + فعل
سلام علیکم	۴	۲۰	اسم + حرف اضافه + ضمیر
چیزی شده	۵	۲۰	اسم + فعل
آره بابا	۵	۱۹	اسم + اسم
این چیه	۵	۱۹	ضمیر + فعل
برا چی	۴	۱۹	حرف اضافه + ضمیر
راست می‌گی	۵	۱۹	صفت + فعل
سال دیگه	۵	۱۹	اسم + وجه نما
شما خوبین	۵	۱۹	ضمیر + فعل
نگران نباش	۴	۱۹	صفت + فعل
درست میشه	۴	۱۸	صفت + فعل
راست می‌گه	۵	۱۸	صفت + فعل

تکه واژگانی	تعداد زیر پیکره	بسامد	ساختار دستوری
قول می‌دم	۴	۱۸	اسم + فعل
معلوم نیست	۶	۱۸	صفت + فعل
کار دارم	۵	۱۸	اسم + فعل
آخ آخ	۵	۱۷	میانجی
آقا ببخشید	۳	۱۷	اسم + فعل
اشکالی نداره	۴	۱۷	اسم + فعل
برو برو	۴	۱۷	فعل + فعل
حرف نزن	۴	۱۷	اسم + فعل
منظورم اینه	۵	۱۷	اسم + ضمیر + ضمیر + فعل
وایسا ببینم	۳	۱۷	فعل + فعل
ولش کن	۴	۱۷	صفت + فعل
گفتم که	۴	۱۷	فعل + حرف ربط
یکی دیگه	۵	۱۷	تعداد نما + وجه نما
خب باشه	۵	۱۶	صفت + اسم
سلامت باشی	۲	۱۶	صفت + فعل
نگاه کن	۶	۱۶	اسم + فعل
اشکال نداره	۵	۱۵	اسم + فعل
باید برم	۶	۱۵	وجه نما + فعل
بسه دیگه	۵	۱۵	فعل + وجه نما
تو چی	۵	۱۵	ضمیر + ضمیر
خدمت شما	۳	۱۵	اسم + ضمیر
خیلی خوبه	۶	۱۵	قید + فعل
داری می‌ری	۳	۱۵	فعل + فعل
شما بفرمایید	۴	۱۵	ضمیر + فعل
صبر کنید	۵	۱۵	اسم + فعل
مشکلی نیست	۴	۱۵	اسم + فعل
مواظب باش	۳	۱۵	صفت + فعل
نگران نباشین	۴	۱۵	صفت + فعل

در جدول ۱ دسته‌بندی تکه‌های واژگانی گفتاری دو کلمه‌ای بر اساس ساختار دستوری آن‌ها مشاهده می‌شود. برای نمونه الگوی دستوری «اسم+فعل» با بسامد بالا در میان این تکه‌های واژگانی دو کلمه‌ای دیده می‌شود. عباراتی مانند «خواهش می‌کنم»، «صبر کن»، «چیزی شده»، «آقا ببخشید»، «اشکالی نداره»، «مشکلی نیست» دارای الگوی دستوری اسم+فعل هستند که برخی از آن‌ها فعل مرکب هستند. عباراتی مانند «خواهش می‌کنم»، «صبر کنید»، «نگاه کن»، «قول می‌دم» همگی فعل مرکب هستند. افعال مرکب به افعالی گفته می‌شود که جزء اول آن‌ها اسم یا صفت و جزء دوم آن‌ها گاهی اوقات فعل کمکی مانند «کردن» و «شدن» است. الگوی دستوری «صفت+فعل» نیز در میان تکه‌های واژگانی مشاهده می‌شود که بیشتر آن‌ها فعل مرکب هستند. عباراتی مانند «ول کن»، «درست میشه»، «راست میگه» افعال مرکب در زبان فارسی هستند. الگوی دستوری «اسم+اسم» نیز در میان این تکه‌های واژگانی مشاهده می‌شود. برای نمونه عبارات «بله قربان»، «بنده خدا»، «بابا جون» دارای الگوی دستوری «اسم+اسم» هستند. در میان این تکه‌های واژگانی دو کلمه‌ای الگوی دستوری «فعل+فعل» نیز به چشم می‌خورد.

عبارت‌هایی نظیر «بیا بریم»، «وایسا ببینم»، «داری می‌ری»، «دارم می‌رم»، «برو برو»، «بیا بیا» دارای این الگوی دستوری هستند. عبارت‌هایی مانند «برو برو»، «بیا بیا» از دو فعل امر پشت سرهم تشکیل شده‌اند. عبارت‌های «بیا بریم» و «وایسا ببینم» به ترتیب از دو فعل امر و مضارع التزامی ساخته شده‌اند. عبارت‌های «دارم می‌رم» و «داری می‌ری» افعال حال ناتمام هستند. در میان این تکه‌های واژگانی الگوی دستوری «تعداد نما+ اسم» نیز به چشم می‌خورد. عباراتی مانند «یه چیزی»، «یه خرده»، «یه دقیقه» از این الگوی دستوری پیروی می‌کنند.

جدول ۲

دسته‌بندی تکه‌های واژگانی گفتاری سه کلمه‌ای بر اساس الگوی دستوری

تکه واژگانی	تعداد زیرپیکره	بسامد	ساختار دستوری
دست درد نکنه	۵	۴۲	اسم + ضمیر + اسم+فعل
صد در صد	۶	۲۲	تعداد نما + حرف اضافه + تعداد نما
به هر حال	۵	۲۱	حرف اضافه + صفت + اسم
دستتون درد نکنه	۵	۲۱	اسم + ضمیر + اسم+فعل
بفرمایید خواهش میکنم	۲	۲۰	فعل + اسم + فعل
چی داری می‌گی	۴	۱۶	ضمیر + فعل+ فعل
به خاطر اینکه	۵	۱۵	حرف اضافه + اسم + حرف ربط
من چه می‌دونم	۴	۱۴	ضمیر + صفت + فعل
منظورم اینه که	۵	۱۳	اسم + ضمیر + ضمیر+ فعل + حرف ربط
بسپارش به من	۱	۱۲	فعل + ضمیر + حرف اضافه + ضمیر

در جدول-۲ تکه‌های واژگانی گفتاری سه کلمه‌ای بر اساس الگوی دستوری مشاهده می‌شوند. در این جدول می‌توان تکه‌های واژگانی را با ساختارهای دستوری متفاوت مشاهده کرد. برای نمونه عبارات «دستت درد نکنه» و «دستتون درد نکنه» دارای الگوی ساختاری «اسم+ضمیر+اسم+فعل» هستند. در عبارت «دستت درد نکنه» حرف «ت» نهایی ضمیر پیوسته است که معادل ضمیر جدای «تو» در زبان فارسی است. همچنین در عبارت دستتون درد نکنه «تون» نیز ضمیر پیوسته است که این ضمیر پیوسته معادل ضمیر جدای «شما» است.

جدول ۳

دسته‌بندی تکه‌های واژگانی گفتاری چهارکلمه‌ای بر اساس ساختار دستوری

تکه واژگانی	تعداد زیرپیکره	بسامد	ساختار دستوری
دست شما درد نکنه	۵	۲۰	اسم + ضمیر + اسم + فعل

در جدول ۳ عبارت «دست شما درد نکنه» را مشاهده می‌کنید که ساختار دستوری آن «اسم+ضمیر+اسم+فعل» است. در این عبارت ضمیر شما یک ضمیر جدا است که معادل ضمیر پیوسته «تون» است. ضمیر پیوسته در زبان فارسی در بیشتر مواقع برای گفتگوی سریع‌تر به کار می‌رود. در ادامه به الگوی دستوری تکه‌های واژگانی نوشتاری می‌پردازیم.

جدول ۴

دسته‌بندی تکه‌های واژگانی نوشتاری دوکلمه‌ای بر اساس الگوی دستوری

تکه واژگانی	تعداد زیرپیکره	بسامد	ساختار دستوری
ممکن است	۳	۸۷	اسم + فعل
پس از	۴	۶۷	حرف اضافه + حرف اضافه
نسبت به	۴	۴۹	اسم + حرف اضافه
به طور	۳	۴۵	حرف اضافه + اسم
بسیاری از	۳	۴۲	پیش مشخص کننده + حرف اضافه
این روزها	۱	۴۱	ضمیر + اسم
بهتر است	۴	۴۱	صفت + فعل
سعی کنید	۲	۴۱	اسم + فعل
قبل از	۴	۴۰	اسم + حرف اضافه
پیش از	۳	۳۹	پیش مشخص کننده + حرف اضافه
این موضوع	۴	۳۷	ضمیر + اسم
علاوه بر	۲	۳۶	حرف اضافه + حرف اضافه
این کار	۴	۳۴	ضمیر + اسم
در حال	۳	۳۲	حرف اضافه + اسم
در مورد	۴	۳۲	حرف اضافه + اسم
وجود دارد	۳	۳۲	اسم + فعل
به صورت	۴	۳۱	حرف اضافه + اسم
توجه به	۲	۳۱	اسم + حرف اضافه
برخی از	۱	۳۰	پیش مشخص کننده + حرف اضافه
به خاطر	۳	۳۰	حرف اضافه + اسم
بعضی از	۲	۲۸	پیش مشخص کننده + حرف اضافه
به جای	۳	۲۸	حرف اضافه + اسم
به زودی	۳	۲۷	حرف اضافه + اسم
در زندگی	۳	۲۷	حرف اضافه + اسم
در طول	۳	۲۷	حرف اضافه + اسم
به طرف	۲	۲۶	حرف اضافه + اسم
افزایش اعتماد	۱	۲۵	اسم + اسم
بر روی	۱	۲۵	حرف اضافه + حرف اضافه
در حالی	۳	۲۵	حرف اضافه + اسم
در کنار	۳	۲۵	حرف اضافه + اسم
با او	۳	۲۴	حرف اضافه + ضمیر
به خانه	۳	۲۴	حرف اضافه + اسم
هر روز	۳	۲۴	صفت + اسم

تکه واژگانی	تعداد زیرپیکره	بسامد	ساختار دستوری
استفاده از	۲	۲۳	اسم + حرف اضافه
با توجه	۱	۲۳	حرف اضافه + اسم
برای شما	۳	۲۳	حرف اضافه + ضمیر
شروع به	۳	۲۳	اسم + حرف اضافه
منجر به	۲	۲۳	اسم + حرف اضافه
ورود به	۳	۲۳	اسم + حرف اضافه
به دنبال	۳	۲۲	حرف اضافه + اسم
بخشی از	۳	۲۳	اسم + حرف اضافه
به دست	۳	۲۱	حرف اضافه + اسم
از نظر	۳	۲۰	حرف اضافه + اسم
مربوط به	۲	۲۰	صفت + حرف اضافه
به شدت	۳	۱۹	حرف اضافه + اسم
در خانه	۳	۱۹	حرف اضافه + اسم
به نظر	۲	۱۸	حرف اضافه + اسم
این شغل	۲	۱۷	ضمیر + اسم
این مطلب	۲	۱۷	ضمیر + اسم
این نتیجه	۲	۱۷	ضمیر + اسم
در صورت	۳	۱۷	حرف اضافه + اسم
روی زمین	۲	۱۶	ضمیر + اسم
رژیم غذایی	۲	۱۶	اسم + صفت
سفر با	۱	۱۶	اسم + حرف اضافه
موجود در	۲	۱۶	اسم + حرف اضافه
به همراه	۳	۱۵	حرف اضافه + اسم
بیشتر از	۳	۱۵	قید + حرف اضافه
در مقابل	۳	۱۵	حرف اضافه + اسم
اختلال یادگیری	۱	۱۴	اسم + اسم
در مقایسه	۲	۱۴	حرف اضافه + اسم
زمانی که	۲	۱۴	اسم + حرف ربط
هر چه	۳	۱۴	پیش مشخص کننده + ضمیر
در میان	۳	۱۳	حرف اضافه + اسم
عبارتند از	۱	۱۳	فعل + حرف اضافه
ناشی از	۲	۱۳	صفت + حرف اضافه
برای اینکه	۲	۱۲	حرف اضافه + حرف ربط
به اندازه	۳	۱۲	حرف اضافه + اسم
به سرعت	۳	۱۲	حرف اضافه + اسم
به مدت	۳	۱۲	حرف اضافه + اسم
ترک عادت	۱	۱۲	اسم + اسم
در برابر	۲	۱۲	حرف اضافه + اسم
در شرایط	۳	۱۲	حرف اضافه + اسم
یک روز	۲	۱۲	تعداد نما + اسم

در جدول ۴ انواع الگوی دستوری تکه‌های واژگانی نوشتاری مشاهده می‌شود. الگوی دستوری «حرف اضافه+اسم» در میان این تکه‌های واژگانی دو کلمه‌ای نوشتاری با بسامد بالا دیده می‌شود. برای نمونه عبارات «به جای»، «به زودی»، «در زندگی»، «در طول» از الگوی دستوری «حرف اضافه+اسم» پیروی می‌کنند. الگوی دستوری «اسم+حرف اضافه» نیز در میان تکه‌های واژگانی نوشتاری دو کلمه‌ای مشاهده می‌شود. عباراتی نظیر «استفاده از»، «شروع به»، «منجر به»، «بخشی از» دارای الگوی دستوری «اسم+حرف اضافه» هستند. عبارت‌هایی مانند «این شغل»، «این مطلب»، «این نتیجه» دارای الگوی دستوری «ضمیر+اسم» هستند. البته در اینجا منظور از ضمیر، ضمیر اشاره در زبان فارسی است.

جدول ۵

دسته‌بندی تکه‌های واژگانی نوشتاری سه کلمه‌ای بر اساس نوع تکه واژگانی

تکه واژگانی	تعداد زیرپیکره	بسامد	ساختار دستوری
اعتماد به نفس	۳	۱۳۴	اسم + حرف اضافه + اسم
این است که	۲	۵۶	ضمیر + فعل + حرف ربط
در این شهر	۱	۳۹	حرف اضافه + ضمیر + اسم
با اعتماد به	۲	۲۳	حرف اضافه + اسم + حرف اضافه
با توجه به	۱	۲۳	حرف اضافه + اسم + حرف اضافه
به همین دلیل	۲	۲۰	حرف اضافه + ضمیر + اسم
علاوه بر این	۲	۱۹	حرف اضافه+حرف اضافه+ضمیر
به این نتیجه	۲	۱۷	حرف اضافه + ضمیر + اسم
یکی دیگر از	۲	۱۶	تعداد نما +
در حالی که	۳	۱۵	حرف اضافه + اسم + حرف ربط
به طور مثال	۲	۱۴	حرف اضافه + اسم + اسم
در مقایسه با	۲	۱۴	حرف اضافه + اسم + حرف اضافه
به طور کلی	۲	۱۳	حرف اضافه + اسم + صفت
یکی از بهترین	۲	۱۲	تعداد نما + حرف اضافه + صفت

در جدول ۵ دسته‌بندی تکه‌های واژگانی نوشتاری سه کلمه‌ای بر اساس الگوی دستوری آن‌ها مشاهده می‌شود. در میان این تکه‌های واژگانی الگوی دستوری «حرف اضافه + ضمیر + اسم» با بسامد نسبتاً بالا مشاهده می‌شود. عباراتی مانند «در این شهر»، «به همین دلیل»، «به این نتیجه» از این الگوی دستوری پیروی می‌کنند. همچنین الگوی دستوری «حرف اضافه + اسم + حرف اضافه» نیز مشاهده می‌شود. عبارت‌هایی مانند «با اعتماد به»، «با توجه به»، «در مقایسه با» دارای این الگوی دستوری هستند.

جدول ۶

دسته‌بندی تکه‌های واژگانی نوشتاری چهار کلمه‌ای بر اساس نوع تکه واژگانی

تکه واژگانی	تعداد زیرپیکره	بسامد	ساختار دستوری
افزایش اعتماد به نفس	۱	۲۵	اسم + اسم + حرف اضافه + اسم
با اعتماد به نفس	۲	۲۲	حرف اضافه + اسم + حرف اضافه + اسم

تکه واژگانی	تعداد زیرپیکره	بسامد	ساختار دستوری
مردم چه می‌گویند	۱	۱۴	اسم + ضمیر + فعل
اعتماد به نفس بالا	۱	۱۲	اسم + حرف اضافه + اسم + صفت
به این نتیجه رسیده‌اند	۱	۱۲	حرف اضافه + ضمیر + اسم + فعل

در جدول ۶ دسته‌بندی تکه‌های واژگانی نوشتاری چهارکلمه‌ای را بر اساس ساختار دستوری آن‌ها می‌توان مشاهده کرد. انواع الگوی دستوری را می‌توان در این جدول مشاهده کرد. برای نمونه در عبارت «افزایش اعتماد به نفس» الگوی دستوری «اسم + اسم + حرف اضافه + اسم» را می‌توان مشاهده کرد. در عبارت «مردم چه می‌گویند» الگوی دستوری «اسم + ضمیر + فعل» مشاهده می‌شود.

نتایج این بخش از پژوهش نشان داد که تکه‌های واژگانی گفتاری دو کلمه‌ای بیشتر الگوی دستوری «اسم + فعل»، «صفت + فعل»، «اسم + اسم»، «فعل + فعل»، «تعداد نما + اسم» و تکه‌های واژگانی گفتاری سه کلمه‌ای و چهار کلمه‌ای بیشتر الگوی دستوری «اسم + ضمیر + اسم + فعل» را شامل می‌شوند. در میان تکه‌های واژگانی نوشتاری دو کلمه‌ای نیز الگوی دستوری «حرف اضافه + اسم»، «اسم + حرف اضافه»، «ضمیر + اسم» و در میان تکه‌های واژگانی نوشتاری سه کلمه‌ای الگوی دستوری «حرف اضافه + ضمیر + اسم» و «حرف اضافه + اسم + حرف اضافه» مشاهده شد. در تکه‌های واژگانی نوشتاری چهار کلمه‌ای نیز الگوهای دستوری مختلفی مشاهده شد.

۷- یافته‌های پژوهش: بررسی چگونگی انسجام تکه‌های واژگانی: تحلیل کیفی تکه‌های واژگانی

تحلیل کیفی داده‌ها نشان داد که تکه‌های واژگانی به‌عنوان واحدهایی از پیش ساخته، نقش مهمی در شکل‌گیری انسجام متنی ایجاد می‌کنند. در گونه گفتاری، تکه‌های واژگانی مانند «نه بابا»، «خواهش می‌کنم»، «یه لحظه»، «قربونت برم»، «دست درد نکنه»، «صد در صد» و «به هر حال» مشاهده می‌شوند که در انسجام گفتار نقش مهمی ایفا می‌کنند. برای نمونه، عبارت «نه بابا» به‌عنوان یک پاسخ به یک پرسش یا ابراز تعجب استفاده می‌شود. این عبارت موضع‌گوینده نسبت به پرسشی را نشان می‌دهد و باعث انسجام می‌شود. برای نمونه به گفتگوی زیر توجه کنید:

الف: شنیدی که سارا رفته خارج از کشور؟

ب: نه بابا. جدی؟

در این گفتگو «نه بابا» برای ابراز تعجب به کار رفته است. عبارت «خواهش می‌کنم» در زبان فارسی معمولاً برای پاسخ به تشکر یا بیان درخواست استفاده می‌شود. به گفتگوی زیر توجه کنید:

الف: خیلی ممنونم ازت. لطف کردی.

ب: خواهش می‌کنم.

استفاده از این تکه واژگانی گفتگو را از حالت ناقص و ناپایدار خارج می‌کند. عبارت «یه لحظه» زمانی به کار می‌رود که گوینده می‌خواهد برای لحظه‌ای در گفتگو توقف ایجاد کند و کنترل گفتگو را به دست بگیرد. عبارت «قربونت برم» یک تکه واژگانی است که برای افزایش صمیمیت در گفتار استفاده می‌شود. عبارتی مانند «صد در صد» نیز در میان تکه‌های واژگانی دو کلمه‌ای

مشاهده شد که برای تأیید کردن گفتار طرف مقابل به کار می‌رود. عبارت «به هر حال» که یک تکه واژگانی سه کلمه‌ای است، جهت تغییر موضوع و نتیجه‌گیری به کار می‌رود.

در گونه نوشتاری تکه‌های واژگانی مانند «ممکن است»، «پس از»، «نسبت به»، «سعی کنید» و «در این شهر» مشاهده می‌شوند. برای نمونه عبارت «ممکن است» به عنوان نشانه حدس و گمان یا احتمال به کار می‌رود و به نویسنده متن کمک می‌کند بدون قطعیت متنی را بنویسد و پیوندی منطقی میان جملات علمی ایجاد کند. تکه واژگانی «پس از» به پیوستگی زمانی در جملات کمک می‌کند و ترتیب رویدادها را مشخص می‌کند. در میان تکه‌های واژگانی نوشتاری دو کلمه‌ای، عبارت «نسبت به» نیز دیده می‌شود که معمولاً نشان‌دهنده مقایسه یک چیز با چیز دیگری است. عبارتی مانند «سعی کنید» یک فعل امر است که می‌تواند در متن‌های آموزشی هم به کار رود و باعث انسجام متن شود. در میان تکه‌های واژگانی نوشتاری سه کلمه‌ای عبارتی مانند «در این شهر» مشاهده می‌شود. اگر نویسنده‌ای درباره یک مکان خاص یا شهر خاص صحبت کرده باشد، این تکه واژگانی باعث ادامه بحث درباره آن مکان یا شهر می‌شود.

۸. نتیجه‌گیری

تکه‌های واژگانی گروهی از واژگان هستند که با یکدیگر به کار می‌روند. این عبارات از دو واژه یا بیشتر تشکیل شده، اما معنایی که منتقل می‌کنند یک معنای واحد است. اهل زبان از تکه‌های واژگانی گفتاری در گفتار و از تکه‌های واژگانی نوشتاری در نوشتار خود به‌وفور استفاده می‌کنند. تکه‌های واژگانی در ذهن اهل زبان نهادینه شده‌اند.

در پژوهش حاضر، تکه‌های واژگانی گفتاری و نوشتاری زبان فارسی که از پیکره ساخته شده توسط محقق استخراج و انتخاب شده بودند را بر اساس ساختار دستوری آن‌ها بررسی کردیم. ساختار دستوری تکه‌های واژگانی گفتاری و نوشتاری بر اساس چارچوب کومو مشخص شد. نتایج بخش اول پژوهش نشان داد که تکه‌های واژگانی گفتاری دو کلمه‌ای بیشتر الگوی دستوری «اسم + فعل»، «صفت + فعل»، «اسم + اسم»، «فعل + فعل»، «تعداد نما + اسم» و تکه‌های واژگانی گفتاری سه کلمه‌ای و چهار کلمه‌ای بیشتر الگوی دستوری «اسم + ضمیر + اسم + فعل» را شامل می‌شوند. در میان تکه‌های واژگانی نوشتاری دو کلمه‌ای نیز الگوی دستوری «حرف اضافه + اسم»، «اسم + حرف اضافه»، «ضمیر + اسم» از سایر الگوها بیشتر مشاهده شد. همچنین در میان تکه‌های واژگانی نوشتاری سه کلمه‌ای الگوی دستوری «حرف اضافه + اسم» و «حرف اضافه + اسم + حرف اضافه» مشاهده شد. در تکه‌های واژگانی نوشتاری چهار کلمه‌ای نیز الگوهای دستوری مختلفی مشاهده شد. در بخش دوم پژوهش نیز نشان داده شد که چگونه تکه‌های واژگانی گفتاری و نوشتاری به انسجام و پیوستگی گفتار و نوشتار کمک می‌کنند.

ذکر این نکته مهم است که تفاوت میان گونه گفتاری و نوشتاری زبان در زبان فارسی تا حد زیادی برجسته است. از این رو، تفاوت‌های دستوری قابل توجهی را می‌توان میان تکه‌های واژگانی گفتاری و تکه‌های واژگانی نوشتاری زبان فارسی مشاهده کرد. پژوهشگران در پژوهش حاضر به دنبال پاسخ به پرسش‌های زیر هستند:

۱- آیا الگوی دستوری تکه‌های واژگانی گفتاری و نوشتاری با یکدیگر تفاوت دارند؟

۲- آیا تکه‌های واژگانی گفتاری و نوشتاری در انسجام گفتار و نوشتار موثر هستند؟

در پاسخ به پرسش نخست، نتایج نشان می‌دهد که الگوی دستوری تکه‌های واژگانی گفتاری و نوشتاری از یکدیگر متفاوت هستند.

تکه‌های واژگانی گفتاری دو کلمه‌ای بیشتر از الگوی دستوری اسم + فعل، صفت + فعل، اسم + اسم، فعل + فعل، تعداد نما + اسم پیروی می‌کنند. درحالی‌که تکه‌های واژگانی نوشتاری دو کلمه‌ای بیشتر از الگوی حرف اضافه + اسم، اسم + حرف اضافه، ضمیر + اسم پیروی می‌کنند. همچنین تکه‌های واژگانی گفتاری سه کلمه‌ای و چهار کلمه‌ای دارای الگوی دستوری اسم + ضمیر + اسم + فعل هستند؛ اما تکه‌های واژگانی نوشتاری سه کلمه‌ای دارای الگوی دستوری حرف اضافه + ضمیر + اسم و حرف اضافه + اسم + حرف اضافه هستند. تکه‌های واژگانی نوشتاری چهار کلمه‌ای نیز انواع مختلفی از الگوهای دستوری را شامل می‌شود. در پاسخ به پرسش دوم باید گفت که این تکه‌های واژگانی در انسجام گفتار و متن نقش مؤثری ایفا می‌کنند. به عنوان مثال، عبارت «نه بابا» به عنوان یک پاسخ به یک پرسش یا ابراز تعجب استفاده می‌شود. این عبارت موضع گوینده نسبت به پرسشی را نشان می‌دهد و باعث انسجام می‌شود و همچنین عبارت «ممکن است» به عنوان نشانه حدس و گمان یا احتمال به کار می‌رود و به نویسنده متن کمک می‌کند بدون قطعیت متنی را بنویسد و پیوندی منطقی میان جملات علمی ایجاد کند.

منابع فارسی

- بهروز، س. (۱۴۰۱). کاربرد واژگان مصوب فرهنگستان زبان و ادب فارسی: بررسی کمی و کیفی کتاب‌های درسی در شاخه نظری (۱۴۰۰). *پژوهش‌های زبانی*، ۱-۲۹.
- بیجن‌خان، م.، حشمتی، ز.، ملکی ویکام، م. و ویسی، ه. (۱۴۰۳). شناسایی و استخراج باهمایی‌های زبان فارسی با استفاده از روش‌های رایانشی. *نشریه پردازش و مدیریت اطلاعات (ایرانداک)*، ۴۰(۲)، ۵۵۷-۶۰۴.
- فراهانی، ا.، یزدانی، ه.، احمدیان، م. و عامریان، م. (۱۳۹۹). بررسی کاربرد اصطلاحات انگلیسی در مقاله‌های پژوهشی زبان‌شناسی کاربردی: پژوهشی پیکره محور؛ پژوهش‌های زبان شناختی در زبان‌های خارجی، ۲(۱۰)، ۳۹۰-۴۰۵.
- کاوه، م.، افراشی، آ. (۱۳۹۸). زبان فرمولواره‌ای و نقش آن در آموزش زبان فارسی؛ زبان شناخت. *پژوهشگاه علوم انسانی و مطالعات فرهنگی*، ۹۹-۱۲۲.
- نعمتی، م.، مسگرشهر، ا. (۱۳۹۸). تاثیر تدریس صریح دسته‌های واژگانی بر کیفیت کاربرد واژگان و صحت دستوری در نگارش زبان انگلیسی: یک رویکرد کاربردی. *پژوهش‌های زبان شناختی در زبان‌های خارجی*، ۹(۳)، ۹۵۹-۹۹۰.
- Albaqami, R. (2022). The role of lexical chunks in promoting English writing competence among foreign language learners in Saudi Arabia. *Arab World English Journal*, 13(2): 91-108.
- Asaei, S. M., & Rahimy, R. (2021). Appropriation-based syllabus and speaking ability: Evidence from Iranian EFL context. *Biannual Journal of Education Experiences*, 4(2): 195-2014.
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Harlow, England: Pearson Education Limited.
- Biber, D., Conrad, S., & Cortes, V. (2004). If you look at.....Lexical bundles in university teaching and textbooks. *Applied Linguistics*, 25(3), 371-405.
- <https://doi.org/10.1093/applin/25.3.371>
- Chun-Guang, T. (2014). An empirical research on the corpus-driven lexical chunks instruction. *International Journal of English Language Teaching*, 2(2):1-36.
- Ghasemian, M., & Rashidi, N. (2024). An Analysis of Lexical Chunks in Mina 1 and Let's Learn Persian 1 Textbooks: A Comparative Study Based on Lewis Model and Their Degree of Fixedness [In Persian]. *Journal of Studies in Applied Language (JSAL)*, 7(2): 93-111.
- Hakuta, K. (1974). Prefabricated patterns and the emergence of structure in second language acquisition 1. *Language Learning*, 24(2): 287-297.
- Lackman, K. (2011). *Lexical approach activities*. Ken Lackman & Associates Educational Consultants.

- Li, Q. (2014). An empirical study on the application of lexical chunk to college English writing. *Journal of Language Teaching and Research*, 5(3): 682.
- Mccarthy, M. (2004). *From corpus to coursebook*. Cambridge: Cambridge University Press
- Mohammadi, M., & Enayati, B. (2018). The effects of lexical chunks teaching on EFL intermediate learners' speaking fluency. *International Journal of Instruction*, 11(3): 179-192.
- Richards, J. C., & Rodgers, T. S. (2014). *Approaches and methods in language teaching*. Cambridge university press.
- Sauri, R., Alderslade, A., & Shapiro, R. (2018). A universal classification of lexical categories and grammatical distinctions for lexicographic and processing purposes. In *Proceedings of the XVII EURALEX International Congress*, 2018, pp.173-185. Oxford University Press.
- Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford University Press.
- Wood, D. 2002. Formulaic language acquisition and production: Implications for teaching. *TESL Canada Journal*, 01-15.