Employing keywords and lexical bundles within figure captions in earth science research articles

by Tsaqifa

Submission date: 16-Apr-2024 05:02PM (UTC+0700)
Submission ID: 2298933971
File name: 5_Tsaqifa_1_1.pdf (423.84K)
Word count: 5392
Character count: 29692

LITE: Jurnal Bahasa, Sastra, dan Budaya
ISSN online: 2548-9588 ISSN print.: 1907-3607
http://publikasi.dinus.ac.id/index.php/lite
Published by Universitas Dian Nuswantoro, Semarang

Employing keywords and lexical bundles within figure captions in earth science research articles

Tsabita Intan Tsaqifa[1*], Ni Gusti Ayu Roselani [1]
[1]*Universitas Gadjah Mada, Jl.Bulak Sumur, Yogyakarta, Indonesia*

Article History

Submitted date:
2024-01-15
Accepted date:
2024-03-29
Published date:
2024-03-31

Keywords:

figure legends; keywords; lexical bundles; corpus linguistics; AntConc

Abstract

The presentation of data is considered an essential part of academic writing, especially in research articles, and it could not stand alone without a short text to describe the visual data, namely figure legends. This research aims to identify the keywords and lexical bundles frequently used in figure legends. In addition, this research also examines the functions conveyed by these linguistic features. Based on the identification of the corpus, there are 334 keywords found with the 20 highest frequencies belonging to the terminologies frequently used in earth science (fault, seismic, quartz, and formation) and words related to the visual representation of the figure legends (*line/lines*, *map*, *area*, *figure* and *legend*). Furthermore, noun-related bundles are found to be the greatest number for lexical bundles, with 165, followed by verb-related bundles with 79, clause-related bundles with 54, and preposition-related bundles with 17. These lexical bundles perform research-oriented, text-oriented, and participant-oriented functions. These findings suggest that lexical bundles play a role in explaining the research condition, connecting visual data and research results discussed in research articles, and displaying information about visual data or specific parts of research articles to readers.

Kata Kunci:

legenda gambar; kata kunci; gugus leksikal; linguistik korpus; AntConc

Abstrak

Penggunaan kata kunci dan gugus leksikal pada legenda gambar di artikel riset ilmu kebumian

Penyajian data dianggap sebagai bagian penting dalam penulisan akademis, terutama dalam artikel penelitian, dan tidak dapat berdiri sendiri tanpa adanya teks singkat untuk mendeskripsikan data visual, yaitu legenda gambar. Penelitian ini bertujuan untuk mengidentifikasi kata kunci dan gugus leksikal yang sering digunakan dalam legenda gambar. Selain itu, penelitian ini juga meneliti fungsi yang disampaikan oleh fitur-fitur linguistik tersebut. Berdasarkan hasil identifikasi terhadap korpus, ditemukan 334 kata kunci dengan 20 frekuensi tertinggi merupakan terminologi yang sering digunakan dalam ilmu kebumian (patahan, seismik, kuarsa, dan formasi) dan kata-kata yang berhubungan dengan representasi visual dari legenda gambar (garis, peta, area, figur, dan legenda). Selanjutnya, untuk gugus leksikal, gugus terkait kata benda ditemukan dalam jumlah terbanyak, yaitu 165, diikuti oleh gugus terkait kata kerja sejumlah 79, gugus terkait klausa sejumlah 54, dan gugus terkait preposisi sejumlah 17. Gugus leksikal ini menjalankan fungsi yang berorientasi pada penelitian, berorientasi pada teks, dan berorientasi pada peserta. Temuan ini menunjukkan bahwa gugus leksikal berperan dalam menjelaskan kondisi penelitian, menghubungkan data visual dan hasil penelitian yang dibahas dalam artikel penelitian, dan juga menampilkan informasi tentang data visual atau bagian tertentu dari artikel penelitian kepada pembaca.

* Corresponding author:
tsabita.i.t@mail.ugm.ac.id

1 Introduction

The presentation of data is considered an essential part of academic writing, especially in research articles, since most of the scientific data are represented in a visual way (Hyland, 2006). Visual data representation needs the accompaniment of figure legends to describe the figures in research articles, as figure legends "can be really useful when it comes to interpreting data and looking for patterns" (Humphreys & Ruxton, 2022, p. 1). However, there has been little research on these important short texts in research articles.

Figure legends are claimed to have great importance in scientific texts, as they construct meanings aside from accompanying academic texts (Moghaddasi et al., 2019). A figure legend is necessary to guarantee that the figure is understandable without having to refer to the entire paper, describing what is represented in the figure and providing readers with every pertinent information to interpret it (Aliotta, 2018).

In addition, Cargill and O'Connor (2011) propose that figure legends in academic texts consist of five parts as its general form:

(1) A succinct heading that encapsulates the essence of the figure.
(2) An overview of the outcomes, representations, or models depicted within or related to the figure.
(3) Further clarification on the figure's elements, the methodologies employed, or key information that underscores the figure's role in the overall findings.
(4) An outline of the measurement units or statistical symbols present in the figure.
(5) A breakdown of any additional symbols or notations utilized in the figure.

Figure legends are clearly visible and easy to notice as they are usually written as a part of the visual representation of the data in research articles, as shown in Figure 1.

Figure 1. An example of a figure and its legend.
[Source: Adapted from Supendi et al. (2023). A conjugate fault was revealed by the destructive Mw 5.6 (November 21, 2022) Cianjur earthquake, West Java, Indonesia. *Journal of Asian Earth Sciences*, 257, p.8.]

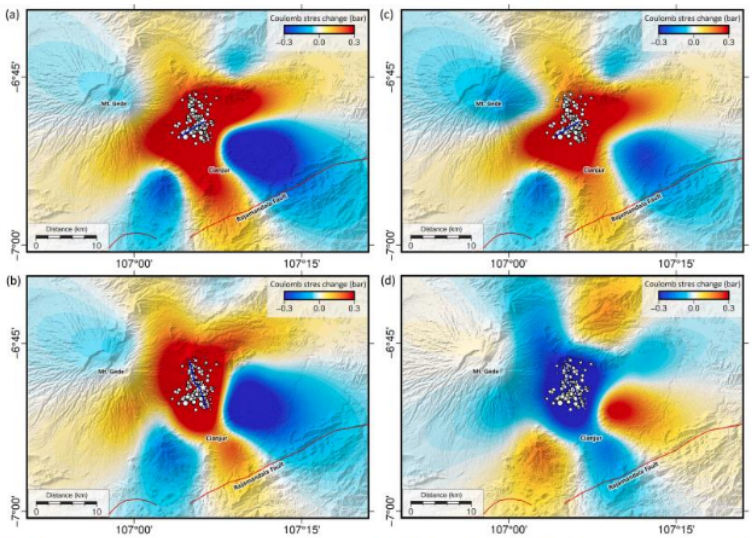

Fig. 5. Modeled Coulomb stress change for four different receiver faults (a) strike-slip fault corresponding to WSW-ENE fault; (b) strike-slip fault corresponding to NNW-SSE fault; (c) strike-slip fault corresponding to the Rajamandala Fault; (d) strike-slip fault corresponding to the Rajamandala Fault. The blue line in each plot denotes the source fault, which corresponds to either the NNW-SSE fault or its conjugate pair. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Current corpus linguistics has developed massively to the point that identifying linguistics expressions could be done in an effective way. It is even possible to count repeated combinations of one or more words, i.e., keywords and lexical bundles in a corpus.

Keywords are words that have an exceptionally high frequency of occurrence "in a corpus of specialized texts against the frequencies of the same wordforms in a larger and more general reference corpus" (Groom, 2010, p. 60). Keywords consist of single words that disclose the 'aboutness' (Scott & Tribble, 2006) and identify the key lexical items that are associated to the discourse functionally (Gilmore & Millar, 2018; Grabowski, 2015; Hyland & Jiang, 2020).

Lexical bundles are, according to statistics, the most frequent "recurring sequences" of words found in a group of texts: extended collocations that occur in a given set of texts more frequently than would be predicted by chance (Biber et al., 1999, p. 990). Lexical bundles serve to support effective communication in pragmatic term and, in academic discourse, to reduce readers' processing times by providing a structure to the discourse, guiding readers through it or connecting ideas (Biber et al., 2004; Cortes, 2013; Hyland, 2004).

Previous studies have identified keywords in academic texts (Malá, 2020) and research articles published in journals (Pípalová, 2019; Soler & Wang, 2019). Furthermore, some notable studies have conducted to examine the lexical bundles in academic writing, such as in argumentative essays (Nguyen, 2023; Oktavianti & Prayogi, 2022), theses and dissertations (Narkprom & Phoocharoensil, 2022; Öztürk & Taşçi, 2023), academic textbooks (Ardi et al., 2023; Grabowski, 2015), academic lectures (Liu & Chen, 2020), academic texts (Ruan, 2017) and research articles (Kurniawan & Haerunisa, 2023; Shahmoradi et al., 2021; Shahriari, 2017; Varghaei & Khodadadi, 2022). However, there are also previous studies of figure legends that aimed to investigate the rhetorical move (L. Liu et al., 2023; Moghaddasi et al., 2019) and linguistics structures and functions (Du et al., 2021).

Based on the previous studies, there have been only a few studies in figure legends of research articles. This essential part of research articles needs to be examined more in-depth since it will be useful for the improvement of academic writing. Thus, this research attempts to eliminate the gap by identifying the keywords and lexical bundles frequently found in figure legends. Additionally, this research also examines the functions conveyed by these linguistic features. In achieving these objectives, this research adapts the methods and analyses suggested by Du et al. (2021).

2 Methods

2.1 Corpus of the study

Figure legends are commonly used in hard sciences. The current study investigates one of the fields of hard science, specifically *earth science,* which belongs to the group of natural science. Data were taken from 30 research articles published in the year 2023 from the *Journal of Asian Earth Science*. This journal has been chosen as a data source as it is Scopus-indexed in Q1. Therefore, a corpus was successfully compiled, including a total of 323 figure legends with a total of 21,377 words.

2.2 Keyword calculation and lexical bundles identification

Keywords were calculated with *AntConc* version 4.2.4 (Anthony, 2023). The corpus was compared with the academic genre of British National Corpus (BNC) as a reference corpus. Lexical bundles were classified as three-word chunks because they are more frequently encountered than four- or five-word chunks and have a greater variety of structures and functions than two-word chunks (Du et al., 2021). Lexical bundles were calculated by using the N-Gram feature, set up to a minimal frequency of 5 since the corpus is under one million words and a minimal range of 3. The results were then listed to an Excel file to ease the analysis of the functions by following the categorization proposed by Biber et al. (1999) and Hyland and Jiang (2018).

3 Results and Discussion

3.1 Keywords in figure legends

Based on the calculation by employing *AntConc* with a comparison referring to the academic genre of BNC, there is a total of 334 keywords. However, not all keywords found are meeting the criteria since some unidentified characters were included in the calculation results. These characters were opted out, resulting in 20 keywords with the highest keyness value.

The majority of 20 keywords occurred in the data belong to the terminologies frequently used in earth science (fault, seismic, quartz, and formation) and words related to the visual representation of the figure legends (line/lines, map, area, figure, and legend). This result shows that analyzing research articles in a particular field would result in the highest keywords mostly related to the field. In addition, the result shows that keywords related to graphic shapes were also used in the figure legends, in accordance with when analyzing research articles in more than one field of study, resulting in the highest keywords related to colors and graphic shapes as "textual presentations of figures, making intelligible what graphic illustration means" (Du et al., 2021, p. 4).

Table 1: The 20 highest keyness value keywords.
[Source: Author's data]

keyword	frequency	keyness likelihood	effect size
fault	108	175.009	0.010
line	105	170.140	0.010
map	84	136.066	0.008
sample	76	123.091	0.007
showing	72	116.605	0.007
seismic	69	111.741	0.006
area	67	108.499	0.006
lines	64	103.636	0.006
interpretation	62	100.394	0.006
formation	61	98.773	0.006
figure	60	97.152	0.006
data	59	95.532	0.006
along	59	95.532	0.006
legend	56	90.670	0.005
references	56	90.670	0.005
version	55	89.049	0.005
quartz	55	89.049	0.005
referred	54	87.429	0.005
article	54	87.429	0.005
reader	54	87.429	0.005

One of the terminologies in the earth science field, *fault*, has the highest keyness value and frequency. This term is considered to be an essential one in earth science since geologists do research by studying fault formations. As a result, *fault* has frequently been used in the figure legends in (1) and (2).

(1) Geology and current **fault** system in southeast Aceh and its surroundings. (Earth Science 30)
(2) The red lines in each plot represent the T70 reflector, which separates the upper and lower **fault** systems. (Earth Science 9)

Furthermore, *seismic*, *quartz*, and formation, similar to a *fault*, are typically used in earth science to examine geological phenomena (3) and (4). However, it is interesting that *quartz* is included in the highest keywords result since quartz is a type of rock mineral. This suggests that the authors of the RAs mainly discuss this term in their research, as in (5).

(3) **Seismic** attribute maps are made along the top of the high amplitude reflection section with time windows of 30 ms downward. (Earth Science 24)
(4) The hornblende phenocrysts from the Pleistocene Bato Dacite Porphyry shows higher pressure and temperature of **formation** coupled by anomalous higher H2O…. (Earth Science 8)
(5) BSE images confirm the bending and fracturing of **quartz** grains along the kink planes that serve as a proxy for the identification of kink plane migration during progressive deformation. (Earth Science 15)

Line and *lines* in figure legends are used to represent the shapes. Additionally, *line* and *lines* make clear of the description in the graphics as in (6), and in the case of the earth science field, point out the geological formation and structure as a guide for the reader and avoid misconceptions as in (7).

(6) Black dotted **line** marks the initial velocity model, while solid black marks the final modeled velocity model. (Earth Science 25)
(7) An orange solid **line** marks the edge of the olivine grain. (Earth Science 16)

Other graphic shapes such as *maps, areas, figures, and legends are frequently found with high keyness values*. *Map*, as an essential part of the field of earth science, gives information regarding to enrich the background knowledge of the readers about the location of the study as in (8). Along with *the map, the area* is used to point out the particular region in Figure (9).

(8) Geological **map** of the study area, showing the distribution of Cambrian outcrops and the location of the Danangou section. (Earth Science 11)
(9) The grey **area** in the map of the Asian part of the Russian Federation (inset) shows the Siberian craton. (Earth Science 16)

Turning to *sample*, *interpretation*, and data, which have a typically high frequency in the overall corpus, show the process, methodology, or result of the research. For instance, see (10), (11), and (12).

(10) Distribution of the pore density of the loess **sample** before and after the ring tests. (Earth Science 3)
(11) …methodology routinely applied to carbonate **interpretation**, in which the timeline overlaps the smooth, convex, high-relief, aggradational platforms in a deep water environment… (Earth Science 24)
(12) Age compositions of concordant **data** of rims of zircons from the studied samples. (Earth Science 23)

Moreover, *showing* and *referring* are the two verb-related words with the highest keyness values in the overall figure legends. *Showing* is used to help science writers arrange as much information as possible into a short pa of text as legends (Du et al., 2021) as in (13). Meanwhile, referred is used to guide the reader to a particular section of the article, as in (14).

(13) Photographs **showing** the strata of the Pakchon Formation and the structural elements in the Pakchon (equal–angle, lower hemisphere projections). (Earth Science 26)
(14) For interpretation of the references to color in this figure legend, the reader is **referred** to the web version of this article. (Earth Science 1)

3.2 Lexical bundles in figure legends

3.2.1 Structural classification of lexical bundles

A total of 73 types of three-word lexical bundles were identified in the overall corpus. However, not all of them can be categorized according to their structure. Therefore, only a total of 13 three-

word lexical bundles were successfully categorized into each of the 12 types of lexical bundle structures suggested by Biber et al. (1999) and Hyland and Jiang (2018), which are summarized in a table adapted from previous research conducted by Du et al. (2021). The result of a structural classification of lexical bundles found in the study is presented in Table 2, and the frequency of each lexical bundle is shown in Table 3.

Table 2: Classification of lexical bundles in academic writing based on the structure. [Source: Author's data, adapted from Du et al., 2021]

verb phrase-related bundles	• passive verb (*is referred to*, *are shown in*, *is shown in*) • copular be (-) • imperative (*note that the*)
clause-related bundles	• abstract subject (*the reader is*) • *as*-fragments (-) • *conj*-fragments (-) • *adj*-fragments (-) • *wh*-fragments (-)
noun-related bundles	• noun phrase with of-phrase fragment (*web version of*, *the location of*, *geological map of*, *the development of*) • noun phrase with other post-modifier fragment (*the references to*, *the relationship between*)
preposition-related bundles	• prepositional phrase expressions (*with respect to*) • comparative expressions (*before and after*)

Table 3: Categories of lexical bundles with each frequency based on the structure. [Source: Author's data]

structural categories	frequency
verb-related bundles	**79**
passive verb	71
copular be	0
imperative	8
clause-related bundles	**54**
abstract subject	54
as-fragments	0
conj-fragments	0
adj-fragments	0
wh-fragments	0
noun-related bundles	**165**
noun phrase with of-phrase fragment	103
noun phrase with other post-modifier fragment	62
preposition-related bundles	**17**
prepositional phrase expressions	12
comparative expressions	5

Noun-related bundles are found in the highest number, which is 165, followed by verb-related bundles with a total of 79, clause-related bundles with a total of 54, and preposition-related bundles with a total of 17. The results of noun-related bundles with the highest frequency are in line with the results of the study conducted by Du et al. (2021). In contrast, preposition-related bundles have a total number that is less than that of this previous research.

Verb-related bundles type consists of passive verbs, copular be, and imperative. Passive verbs that were identified include *is referred to*, *are shown in*, and is shown in as in examples (15), (16), and (17). Turning to imperative, the three-word bundles found are noted as in example (18). Meanwhile, the copular be was not found in the figure legends.

(15) For interpretation of the references to color in this figure legend, the reader **is referred to** the web version of this article. (Earth Science 1)
(16) Rose diagrams of fast directions for each station **are shown in** blue and yellow bars are the average fast direction. (Earth Science 27)
(17) Information on the concentration-distance profiles along lines A–B and C–D **is shown in** Fig. 4. (Earth Science 16)
(18) **Note that the** spacing of contour lines in the upper and lower fault systems is different. (Earth Science 9)

Turning to clause-related bundles, this type consists of an abstract subject, as-fragments, conj-fragments, adj-fragments, and wh-fragments. The abstract subject identified in the corpus is *the reader* as in example (19). However, no three-word bundles fall into the category of *as*-fragments, *conj*-fragments, *adj*-fragments, or *wh*-fragments at all.

(19) For interpretation of the references to color in this figure legend, **the reader is** referred to the web version of this article. (Earth Science 1)

Additionally, the noun-related bundle's type consists of two categories, namely, noun phrases with of-phrase fragments and noun phrases with other post-modifier fragments. Several three-word bundles identified to be in the category of the noun phrases with of-phrase fragments include a web version of, the location of, a *geological map of,* and *the development of* as in (20), (21), (22), and (23). Furthermore, the three-word bundles that belong to the category of noun phrases with other post-modifier fragments are *the references to* and *the relationship between* as in (20) and (24).

(20) For interpretation of **the references to** color in this figure legend, the reader is referred to the **web version of** this article. (Earth Science 1)
(21) Filled violet triangles mark **the location of** broadband stations, while small filled red circles mark the earthquake relocations obtained from simultaneous inversion. (Earth Science 25)
(22) **Geological map of** the area around field stations along the Song Ca Fault (B) and the Rao Nay Fault (C). (Earth Science 29)
(23) Geological structures and **the development of** fissures and caves in landslide-prone areas. (Earth Science 3)
(24) **The relationship between** M and longitude during the Early Oligocene. (Earth Science 7)

Moreover, preposition-related bundle types are categorized into prepositional phrase expressions and comparative expressions. Three-word bundles that belong to the category of prepositional phrase expressions include *with respect to* as in example (25). Furthermore, the comparative expressions category consists of one three-word bundle that is before and after, as in (26).

(25) Distribution of selected records **with respect to** surface wave magnitude and hypocentral distance used in this study. (Earth Science 10)
(26) Three seismic property models **before and after** inversion beneath line A, line B and line C. (Earth Science 22)

3.2.2 Functional categories of lexical bundles

Regardless of their structure, three-word bundles can also be categorized into several function categories of lexical bundles according to the bundle grouping proposed by Hyland and Jiang (2018). Table 4 shows the function categories and lexical bundles included in each category. Additionally, there are some three-word bundles that serve more than one function.

Furthermore, the research-oriented bundle's function consists of location (time and place), procedure, quantification, description of tangible attributes, and description of intangible attributes.

The function of these lexical bundles is to explain the research since figure legends "specify the exact conditions used in the experiment" (Du et al., 2021), as in (27) and (28).

(27) Comparison of the seismic sections **before and after** the spectral enhancement. (Earth Science 22)

(28) Schematic diagrams depicting the superposed buckling model for explaining **the development of** folds in the Daling Group of rocks in the Ramgarh Thrust (RT) sheet in the Darjeeling-Sikkim Himalaya (DSH). (Earth Science 15)

Table 4: Functional categories of lexical bundles in academic writing.
[Source: Author's data, adapted from Du et al., 2021]

research-oriented bundles	• location (*before and after*, *the location of*) • procedure (*geological map of*) • quantification (-) • description of tangible attributes (*the location of*) • description of intangible attributes (*the development of*)
text-oriented bundles	• transition signals (*the relationship between*) • resultative signals (*with respect to*) • structuring signals (*are shown in*, *is shown in*, *is referred to*) • framing signals (*with respect to*, *the references to*)
participant-oriented bundles	• stance features (-) • engagement features (*note that the*, *the reader is*, *is referred to*, *web version of*)

In addition, the function of text-oriented bundles includes transition signals, resultative signals, structuring signals, and framing signals. Text-oriented bundles are commonly found in figure legends due to their purpose of connecting visual data and research results discussed in research articles to make it easier for readers to receive information conveyed by the authors (Cargill & O'Connor, 2011; Du et al., 2021). For instance, see (29) and (30).

(29) Residual distribution of all data points and (b–d) residual distribution **with respect to** ML, frequency, and hypocentral distance. (Earth Science 12)

(30) The boundary faults and some larger faults within the LF15 fault zone **are shown in** red. (Earth Science 9)

Moreover, participant-oriented bundles are mainly composed of stance features and engagement features. The function of lexical bundles is to connect the authors and readers. The authors try to show the readers information about visual data or certain parts of the research article with figure legends, as in (31) and (32).

(31) **Note that the** compositional trends of EPR tholeiitic MORB cross in the 'calc-alkalic' field whereas the trends of the Mount Shasta calc-alkalic lavas cross in the 'calcic' field. (Earth Science 19)

(32) For interpretation of the references to color in this figure legend, **the reader is** referred to the web version of this article. (Earth Science 1)

4 Conclusion

Figure legends are essential in data visualization as they construct meanings and accompany figures in scientific texts. Based on the corpus analysis, there are 334 keywords found, with the 20 highest keywords belonging to the terminologies frequently used in earth science (fault, seismic, quartz, and formation) and words related to the visual representation of the figure legends (line/lines, map, area, figure, and legend). In addition, there are also verb-related words identified in the corpus, such as showing and referred.

Furthermore, a total of 73 types of three-word lexical bundles were identified in the overall corpus, with 13 three-word bundles matching the categorization of the structures. Noun-related bundles are found in the highest number, which is 165, followed by verb-related bundles with a total of 79, clause-related bundles with a total of 54, and preposition-related bundles with a total of 17.

These lexical bundles serve the functions of research-oriented bundles, text-oriented bundles, and participant-oriented bundles. These results indicate that lexical bundles convey the role of explaining the research condition, connecting visual data and research results discussed in research articles, and showing the readers information about visual data or certain parts of research articles.

References

Aliotta, M. (2018). *Mastering academic writing in the sciences: A step-by-step guide*. CRC Press.

Anthony, L. (2023). *AntConc*. Tokyo, Japan: Waseda University. Retrieved from http://www.antlab.sci.waseda.ac.jp/

Ardi, P., Oktafiani, Y. D., Widianingtyas, N., Dekhnich, O. V., & Widiati, U. (2023). Lexical bundles in Indonesian EFL textbooks: A corpus analysis. *Journal of Language and Education*, 9(2), 25–39. https://doi.org/10.17323/jle.2023.16305

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

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Pearson Education Limited.

Cargill, M., & O'Connor, P. (2011). *Writing scientific research articles: Strategy and steps*. Wiley.

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. https://doi.org/10.1016/j.jeap.2012.11.002

Du, Z., Jiang, F., & Liu, L. (2021). Profiling figure legends in scientific research articles: A corpus-driven approach. *Journal of English for Academic Purposes*, 54. https://doi.org/10.1016/j.jeap.2021.101054

Gilmore, A., & Millar, N. (2018). The language of civil engineering research articles: A corpus-based approach. *English for Specific Purposes*, 51, 1–17. https://doi.org/10.1016/j.esp.2018.02.002

Grabowski, Ł. (2015). Keywords and lexical bundles within English pharmaceutical discourse: A corpus-driven description. *English for Specific Purposes*, 38, 23–33. https://doi.org/10.1016/j.esp.2014.10.004

Groom, N. (2010). *Closed-class keywords and corpus-driven discourse analysis* (pp. 59–78). https://doi.org/10.1075/scl.41.05gro

Humphreys, R. K., & Ruxton, G. D. (2022). *Presenting scientific data in R: Creating effective graphs and figures*. Oxford University Press.

Hyland, K. (2004). *Disciplinary discourses*. Michigan Classics Ed. University of Michigan Press/ELT. https://doi.org/10.3998/mpub.6719

Hyland, K. (2006). *English for academic purposes: An advanced resource book*. Routledge.

Hyland, K., & Jiang, F. (Kevin). (2020). "This work is antithetical to the spirit of research": An anatomy of harsh peer reviews. *Journal of English for Academic Purposes*, 46, 100867. https://doi.org/10.1016/j.jeap.2020.100867

Hyland, K., & Jiang, K. (2018). Academic lexical bundles: How are they changing? *International Journal of Corpus Linguistics*, 24(3), 383–407. https://doi.org/10.1075/ijcl.17080.hyl

Kurniawan, E., & Haerunisa, Z. F. (2023). A comparative study of lexical bundles in accepted and rejected applied linguistic research article introductions. *Studies in English Language and Education*, 10(2), 628–648. https://doi.org/10.24815/siele.v10i2.28119

Liu, C. Y., & Chen, H. J. H. (2020). Analyzing the functions of lexical bundles in undergraduate academic lectures for pedagogical use. *English for Specific Purposes*, 58, 122–137. https://doi.org/10.1016/j.esp.2019.12.003

Liu, L., Jiang, F. (Kevin), & Du, Z. (2023). Figure legends of scientific research articles: Rhetorical moves and phrase frames. *English for Specific Purposes*, 70, 86–100. https://doi.org/10.1016/j.esp.2022.11.005

Malá, M. (2020). Phraseology in learner academic English: Corpus-driven approaches. *Discourse and Interaction*, 13(2), 75–88. https://doi.org/10.5817/DI2020-2-75

Moghaddasi, S., Graves, H. A. B., Graves, R., & Gutierrez, X. (2019). "See Figure 1": Visual moves in discrete mathematics research articles. *English for Specific Purposes*, 56, 50–67. https://doi.org/10.1016/j.esp.2019.08.001

Narkprom, N., & Phoocharoensil, S. (2022). Lexical bundles in native English speakers' and Thai writers' dissertations. *GEMA Online Journal of Language Studies*, 22(3), 43–62. https://doi.org/10.17576/gema-2022-2203-03

Nguyen, N. X. (2023). Structures and functions of lexical bundles in student and expert timed argumentative writing: A corpus-driven study. *Pertanika Journal of Social Sciences and Humanities*, 31(3). https://doi.org/10.47836/pjssh.31.3.06

Oktavianti, I. N., & Prayogi, I. (2022). Discourse functions of lexical bundles in Indonesian EFL learners' argumentative essays: A corpus study. *Studies in English Language and Education*, 9(2), 761–783. https://doi.org/10.24815/SIELE.V9I2.23995

Öztürk, Y., & Taşçi, S. (2022). A corpus-based analysis of lexical bundles in non-native postgraduate academic writing and a potential L1 influence. *In rEFLections* (Vol. 30, Issue 2). https://doi.org/10.61508/refl.v30i2.267463

Pípalová, R. (2019). Constructing the global from the local: On the FSP status of keywords in academic discourse. *Linguistica Pragensia*, 29(2), 192–212. https://doi.org/10.14712/18059635.2019.2.5

Ruan, Z. (2017). Lexical bundles in chinese undergraduate academic writing at an English Medium University. *RELC Journal*, 48(3), 327–340. https://doi.org/10.1177/0033688216631218

Scott, M., & Tribble, C. (2006). *Textual patterns* (Vol. 22). John Benjamins Publishing Company. https://doi.org/10.1075/scl.22

Shahmoradi, N., Jalali, H., & Ghadiri, M. (2021). Lexical bundles in the abstract and conclusion sections: The case of applied linguistics and information technology. *Applied Research on English Language*, 10(3), 47–76. https://doi.org/10.22108/are.2021.128024.1703

Shahriari, H. (2017). Comparing lexical bundles across the introduction, method and results sections of the research article. *Corpora*, 12(1), 1–22. https://doi.org/10.3366/cor.2017.0107

Soler, J., & Wang, Y. (2019). Linguistic differences between well-established and predatory journals: A keyword analysis of two journals in political science. *Learned Publishing*, 32(3), 259–269. https://doi.org/10.1002/leap.1244

Varghaei, E., & Khodadadi, G. (2022). Comparing lexical bundles in medical research article abstracts of Iranian and foreign journals. *GEMA Online Journal of Language Studies*, 22(3), 86–102. https://doi.org/10.17576/gema-2022-2203-05

Employing keywords and lexical bundles within figure captions in earth science research articles

ORIGINALITY REPORT

19%	9%	18%	6%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

MATCH ALL SOURCES (ONLY SELECTED SOURCE PRINTED)

7%

★ Zhongquan Du, Feng Kevin Jiang, Luda Liu. "Profiling figure legends in scientific research articles: A corpus-driven approach", Journal of English for Academic Purposes, 2021

Publication

Exclude quotes	On	Exclude matches	< 1%
Exclude bibliography	On		

Employing keywords and lexical bundles within figure captions in earth science research articles

GRADEMARK REPORT

FINAL GRADE

/0

GENERAL COMMENTS

PAGE 1

PAGE 2

PAGE 3

PAGE 4

PAGE 5

PAGE 6

PAGE 7

PAGE 8

PAGE 9

PAGE 10

PAGE 11