

Original Paper

A Corpus-Driven Comparative Study of Lexical Bundles in the Conclusion Sections of Hard and Soft Science International Journal Articles

Yujie Shao

School of Foreign Languages and Cultures, Jilin University, China

Received: June 28, 2026

Accepted: July 29, 2026

Online Published: August 7, 2026

doi:10.22158/selt.v14n3p61

URL: <http://dx.doi.org/10.22158/selt.v14n3p61>

Abstract

Lexical bundles are repeated multi-word units that serve to organise information and convey disciplinary knowledge in academic writing. The use of lexical bundles in the conclusion sections of research papers, particularly across diverse fields, remains underexplored. Four-word lexical bundles in the 400 concluding paragraphs of international journal articles published in 2025 across the four fields of Applied Linguistics, Sociology, Biology, and Electrical & Electronic Engineering have been studied in this paper. AntConc was used to obtain a corpus-based list of lexical bundles and then analysed their structure and function. Lexical bundles are used widely in all fields, but their distribution and function vary among different areas of study. Soft disciplines are more likely to have interpretative and evaluative parts in the research significance and theoretical implications sections of their studies; hard sciences focus on empirical results, research outcomes and future development plans. Therefore, the above results show that lexical bundles contain disciplinary conventions and ways of knowing, and they can offer some references for research on academic writing and discipline-specific writing instruction.

Keywords

lexical bundles, conclusion sections, disciplinary variation, corpus-based analysis

1. Introduction

Scholarly writing is to present and disseminate knowledge in a specific field. Academic papers present research results and are also examples of rhetoric and communication among scholars (Swales, 1990; Hyland, 2008). The conventions of genre and the specific demands of different fields make scholarly papers a rich subject for research. Among the parts, the conclusion is used to summarise the results, point out the achievements, and suggest future research directions (Brett, 1994). Research conclusions are used

to summarise the work and present the findings in the field.

Although they are necessary, the conclusions of the research have not been frequently mentioned. Most previous studies have only covered the Introduction, Abstract and Methods; the Conclusion section has been neglected. Lexical bundles are common sequences of words that have been widely studied for organising academic discourse and showing disciplinary practices (Biber et al., 2004; Hyland, 2008). Research in other fields has shown that the structure and purposes of lexical bundles are not the same. However, the existing studies have done so broadly on academic writing or focused on parts such as the abstract and introduction, leaving lexical bundles in the conclusion section underexplored.

To fill this gap, this paper examines four-word lexical bundles in the conclusion sections of international journal articles from the fields of Applied Linguistics, Sociology, Biology and Electrical & Electronic Engineering. According to Becher and Trowler (2001), the four types of disciplines are soft-applied, soft-pure, hard-pure and hard-applied. Based on a self-built corpus of 400 conclusion sections and a corpus-driven mixed-methods approach, this paper will analyze the distribution, structural patterns and discourse functions of lexical bundles to determine how disciplinary conventions are presented in concluding academic writing.

2. Literature Review

2.1 Genre Analysis and Research Article Conclusions

Genre analysis can help us see how academic writing is a type of social communication. Swales (1990) also proposed that genre studies should consider, in addition to linguistic characteristics, the connections among texts, their purposes of communication and disciplinary norms. In this way, research articles are considered conventional academic genres, and each section has a different rhetorical function. Therefore, the above sections will present the knowledge structure and disciplinary connotations constructed by scholars in language.

Research has shown that the different parts of research papers are organised systematically and have various styles in language use. Research on introductions has shown that writers can create a research environment, point out deficiencies in the research, and introduce the study in light of existing knowledge (Samraj, 2002). Studies of abstracts have also found that this concise type has certain rhetorical patterns and varies in different fields (Pho, 2008). However, conclusion sections have been given less attention than abstracts and introductions. As the end part of research articles, conclusions generally present the research results, explain their importance, point out shortcomings and suggest future research directions (Brett, 1994; Yang & Allison, 2003). Thus, at this time, researchers have accumulated some achievements and begun linking their studies with other areas of research.

Although the previous studies have looked at the rhetorical structure and functions of conclusions, little attention has been given to the linguistic resources used to achieve these functions. According to the above studies, the conclusions of different fields vary in how they summarise knowledge and present the research value (Zhu, 2006). However, there are no organised studies of patterns in conclusion language,

and cross-disciplinary research is even rarer. This gap underscores the need to investigate the specific linguistic resources that realize concluding functions in academic discourse.

2.2 Lexical Bundles in Academic Writing

Lexical bundles have been chosen as the subjects of corpus-based research on academic discourse because they occur frequently in all parts of speech. Biber and others (1999) considered lexical bundles to be frequently occurring sequences of words that could be identified in corpora. Unlike traditional phraseological units, lexical bundles are not always complete grammatical structures or idiomatic expressions; rather, they are relatively frequent and serve to organise discourse. Biber and others (2004), Hyland (2008), and others have shown that lexical bundles are used to organise information, demonstrate the relationship among ideas, and construct an academic style. They can also show the rhetorical preferences of some discourse communities and offer conventional resources for academic communication.

Most research on lexical bundles has focused on their structure and function. Structural analysis studies the formal patterns of lexical bundles. Biber et al. (1999, 2004) put forward the four general types of structures: noun-phrase-based, prepositional-phrase-based, verb-phrase-based, and clause-based bundles, and these have been widely used in the study of academic discourse and further refined by later scholars (Chen & Baker, 2010). Functional Approach studies how lexical bundles help construct discourse. Biber and others (2004) have classified bundles as stance expressions, discourse organisers and referential expressions; Hyland (2008) has put forward a discipline-sensitive system of research-oriented, text-oriented and participant-oriented bundles. The above framework relates word choice to the aim of academic writing and has been applied widely in disciplinary studies.

Recently, some studies have suggested that structure and function should be considered together; thus, lexical bundles with the same form can be used for different purposes in different situations. Combining the above, we have obtained a more all-encompassing view of lexical bundle use. Based on the above, this paper will use the structural framework of Hyland and Jiang (2018) and the functional framework of Hyland (2008) to study lexical bundles in the conclusions of research articles across all fields.

2.3 Lexical Bundles and Disciplinary Variation

Different academic disciplines have different ways of constructing and presenting knowledge, and thus use language in various ways. According to the conventions and epistemological views of the respective disciplines, scholars have conducted various research studies on how to apply lexical bundles in all fields of study (Biber et al., 2004; Hyland, 2008). Lexical bundles can therefore be used to study how disciplinary connotations are expressed in frequent linguistic patterns.

Studies in many areas have found different lexical bundle structures and functions. Research in science and engineering generally favours noun-phrase-based and prepositional-phrase-based bundles that are rich in information and easy to report; in the soft sciences, verbs and clauses are often used for interpretation, evaluation and interaction (Hyland, 2007, 2008). Functionally speaking, the hard sciences are more dependent on research-oriented packages, and the soft sciences use participant-oriented

resources to express evaluations and attitudes (Hyland, 2008). In sum, lexical bundles serve not only as textual resources but also as linguistic reflections of the rhetorical norms prevalent in different academic fields. They influence how scholars construct and convey meaning within their respective discourse communities.

Becher and Trowler (2001) have classified the disciplines and, according to this classification, divided them into hard-pure, soft-pure, hard-applied and soft-applied. Most studies have covered academic writing as a whole or specific parts of it, such as the abstract and introduction, but the conclusion section has been researched relatively little. Few studies have been carried out on the use of lexical bundles in various areas with explicit classification or on the connection between structural patterns and discourse functions in concluding sentences. To address the above problems, this paper investigates four-word lexical bundles in the conclusion sections of international journal articles from Applied Linguistics, Sociology, Biology and Electrical & Electronic Engineering.

3. Methodology

3.1 Research Questions

To explore the distribution patterns, structural features and discourse functions of lexical bundles in the conclusions of research articles across various fields, this paper aims to answer the following three questions: (1) What are the distributional patterns of lexical bundles in different fields of study? (2) How do the structures of lexical bundles vary among these fields? (3) What are the functions of lexical bundles in the conclusions of different fields? By examining the frequency, structural patterns, and discourse functions of lexical bundles, this paper explores how disciplinary conventions affect language selection in the conclusion of academic writing.

3.2 Corpus Construction

A self-built corpus of conclusion sections from research articles in foreign journals will be constructed in this paper. The four areas in the corpus are Applied Linguistics, Sociology, Biology and Electrical and Electronic Engineering, etc. According to the disciplinary framework proposed by Becher and Trowler (2001), the four disciplines are soft-applied, soft-pure, hard-pure and hard-applied knowledge domains. The following are the general rules for building corpora in various fields. First, all the articles were released in 2025 to limit the effect of changes over time. Second, only empirical research papers were included; review papers, editorials, book reviews, letters and other non-research types were excluded. Third, in line with the JCR division of disciplinary classification and journal impact indicators, one journal was chosen. Q1 journals were given more weight; at the same time, to ensure the authority and comparability of the corpus, journal representativeness, international recognition and disciplinary relevance were also considered.

For each subject, ten well-known foreign journals were selected, and ten suitable research articles were obtained from these journals. Therefore, each disciplinary sub-corpus has 100 articles, and the total balanced corpus of research articles is 400. The composition of the corpus is shown in Table 1.

Table 1. Composition of the Corpus

Discipline	Disciplinary Category	Number of Journals	Number of Articles
Applied Linguistics	Soft-applied	10	100
Sociology	Soft-pure	10	100
Biology	Hard-pure	10	100
Electrical & Electronic Engineering	Hard-applied	10	100
Total		40	400

According to disciplinary classification, four sub-corpora were established: the Applied Linguistics Conclusion Corpus, the Sociology Conclusion Corpus, the Biology Conclusion Corpus, and the Electrical & Electronic Engineering Conclusion Corpus. These sub-corpora were further combined into an integrated conclusion corpus to facilitate cross-disciplinary comparison.

3.3 Disciplinary Framework and Corpus Selection Criteria

The disciplinary classification adopted in this study is based on Becher & Trowler's (2001) model of academic tribes and territories. This framework conceptualizes disciplines according to their epistemological orientations and knowledge-construction practices, providing a theoretical basis for interpreting linguistic variation in academic discourse. Table 2 presents the relationship between the selected disciplines and their corresponding knowledge domains.

Table 2. Disciplinary Classification Based on Becher & Trowler (2001)

Knowledge Domain	Discipline Selected in the Present Study	Characteristics
Hard-Pure	Biology	Empirical investigation and systematic explanation of natural phenomena
Hard-Applied	Electrical & Electronic Engineering	Application-oriented knowledge production and technological development
Soft-Pure	Sociology	Interpretive analysis of social phenomena and theoretical development
Soft-Applied	Applied Linguistics	Application of linguistic theories to practical contexts

The selection of journals and articles followed the criteria summarized in Table 3.

Table 3. Corpus Selection Criteria

Criterion	Description
Publication period	Articles published in 2025
Article type	Empirical research articles
Journal source	International journals indexed in JCR
Journal quality	Preference for Q1-ranked journals
Language	English-language publications
Section analyzed	Conclusion sections

3.4 Analytical Framework

The present study establishes a three-level analytical framework incorporating disciplinary, structural, and functional dimensions. The overall framework is summarized in Table 4.

Table 4. Analytical Framework of the Study

Analytical Level	Framework Adopted	Analytical Purpose
Disciplinary level	Becher & Trowler (2001)	Interpret disciplinary variation in academic discourse
Structural level	Hyland and Jiang (2018)	Examine formal patterns of lexical bundles
Functional level	Hyland (2008)	Analyze rhetorical functions of lexical bundles

At the structural level, the study adopts the four-category classification system proposed by Hyland and Jiang (2018). Lexical bundles are categorized into noun phrase-related, prepositional phrase-related, verb phrase-related, and clause-related bundles. The detailed structural classification is presented in Table 5.

Table 5. Structural Classification of Lexical Bundles

Category	Description
NP-related bundles	Bundles involving noun phrase structures
PP-related bundles	Bundles involving prepositional phrase structures
VP-related bundles	Bundles involving verb phrase structures
Clause-related bundles	Bundles involving clausal structures

To provide more detailed structural interpretation, the refined classification criteria proposed by Chen & Baker (2010) are also considered during manual annotation.

3.5 Lexical Bundle Extraction and Selection

The four-word lexical bundle is the unit of analysis in this paper. Based on the above, Hyland (2008) and Chen and Baker (2010) have shown that four-word bundles are relatively frequent and functionally stable in academic language. Lexical bundles were extracted from the conclusion sections of the four disciplinary corpora with AntConc. Both frequency and distribution criteria were used in the previous work on corpora. Bundles should occur at least 30 times per million words and appear in a minimum of three texts in each disciplinary corpus; otherwise, they were considered isolated expressions rather than recurring features.

Automatically extract the sequences and then manually filter out those that cross sentence boundaries, are proper nouns, or lack clear discourse functions. Use the concordance lines in AntConc to find out if each bundle is an important pattern for academic conclusion writing. Finally, the list of lexical bundles was divided into categories according to structure and function.

3.6 Functional Classification and Annotation Procedures

In addition to structural analysis, this paper will also explore the discourse functions of lexical bundles in the conclusion sections. Based on the three categories of functional analysis by Hyland (2008), there are three types of lexical bundles: research-oriented, text-oriented and participant-oriented.

Table 6. Functional Classification of Lexical Bundles Based on Hyland (2008)

Functional Category	Description
Research-oriented bundles	Refer to research activities, procedures, findings, and objects of investigation
Text-oriented bundles	Organize discourse, establish logical relationships, and guide textual development
Participant-oriented bundles	Express authorial stance, evaluation, and interaction with readers

Based on the above, the extracted lexical bundles were manually annotated according to their context and rhetorical functions in the conclusion sections. Given that lexical bundles can have various functions in different contexts, functional annotation was performed based on concordance lines produced by AntConc, not on individual word sequences.

The two types of annotations are structure and function. At the level of structure, the four types of lexical bundles are NP-related, PP-related, VP-related and clause-related bundles. At the level of function, there were three types: research-oriented, text-oriented and participant-oriented. The two types of analysis can show us which language forms are preferred in different subjects and how these forms help to build the rhetoric of the final part of an academic paper.

A Time-Interval Re-annotation Method was employed to improve the accuracy of annotation. After the

first classification, at a later time, I carried out a second round of annotations. The two rounds of coding were compared, and the inconsistent ones were corrected. Thus, the division results will be more uniform and dependable.

3.7 Data Analysis and Statistical Procedures

Based on a corpus, this paper will integrate some quantitative statistics with some qualitative analysis from context. Based on the four disciplinary conclusion corpora, quantitative analysis was carried out to determine the distribution pattern and statistical differences of lexical bundles, and qualitative analysis was also used to explore the rhetorical functions of representative bundles.

Extract the four-word lexical bundles and calculate the normalised frequency in the corpus with AntConc. Compare the distribution of lexical bundles in the various subjects, and then use SPSS to conduct chi-square tests to see if there are any significant differences in structural categories and functional distributions. With the help of concordance lines and context information, it was determined that the representative lexical bundles in the conclusion section serve various rhetorical functions, such as summarizing the research results, evaluating the significance, mentioning limitations and proposing future research directions.

4. Results and Discussion

4.1 Overall Distribution of Lexical Bundles across Disciplines

To investigate the general distribution characteristics of lexical bundles in the conclusion sections of international journal articles, four disciplinary corpora were constructed: Applied Linguistics, Sociology, Biology and Electrical & Electronic Engineering. As shown in Table 4-1, the four sub-corpora have a relatively even number of words, although there are some differences due to disciplinary writing conventions.

Table 7. Distribution of the Four Disciplinary Corpora

Discipline	Tokens	Number of Articles
Applied Linguistics	58,924	100
Sociology	61,407	100
Biology	45,941	100
Electrical & Electronic Engineering	40,291	100
Total	206,563	400

As indicated in Table 7, the total size of the corpus reaches 206,563 words, covering 400 conclusion sections from international journal research articles. Although the Biology and Electrical & Electronic Engineering corpora contain fewer tokens than the two soft disciplines, the difference is mainly related

to disciplinary writing practices. Research articles in social sciences and applied linguistics generally contain longer concluding discussions involving interpretation, implications, and limitations, whereas conclusions in scientific and engineering disciplines tend to be more concise and focused on research findings and future applications.

Following the criteria proposed by Hyland (2008) and Chen & Baker (2010), four-word lexical bundles were extracted using AntConc. The extraction results show that lexical bundles occur frequently across all four disciplines, indicating that recurrent formulaic expressions constitute an important linguistic resource in academic conclusion writing.

Table 8. Representative High-frequency Lexical Bundles across Disciplines

Discipline	Representative lexical bundles	Main discourse purposes
Applied Linguistics	the generalizability of the; it is important to; findings of this study; further research is needed	limitation discussion, implication, evaluation
Sociology	in the context of; in the United States; the extent to which; this study contributes to	contextualization, interpretation, contribution
Biology	in the context of; in this study we; provide valuable research insights into; further research is needed	summary, implication, future direction
Electrical Engineering	state of the art; in this paper we; future work will focus; results show that	research reporting, innovation, future development

There are disciplinary variations. Applied Linguistics and Sociology are social sciences; therefore, they are more likely to use evaluative and interpretative expressions, such as the generalizability of the [study subject], the extent to which, and what this study adds. The above bundles can be used to discuss the research limitations, significance and theoretical implications. Biology and Electrical Engineering are "hard" sciences that generally need to report research results in bundles, so we have developed bundles that present the results and offer some theoretical insights. The above are mainly the empirical results and research contributions of this paper.

Therefore, the various ways lexical bundles are used in the conclusion sections probably reflect different disciplinary epistemological views. In general, the primary concerns of the soft sciences are interpretation, argumentation, and knowledge construction, whereas those of the hard sciences are observation, methodology, empirical data, and application.

4.2 Structural Distribution of Lexical Bundles across Disciplines

To explore the structure of lexical bundles in all the above fields further, the four-word lexical bundles obtained have been organised according to the structural framework presented by Hyland and Jiang (2018). Based on the above framework, the four general structural types of lexical bundles are noun phrase-related bundles, prepositional phrase-related bundles, verb phrase-related bundles and clause-

related bundles. Table 9 shows typical cases of all the structural categories found in the four disciplinary corpora.

Table 9. Structural Classification of Lexical Bundles across Disciplines

Structural category	Representative lexical bundles	Disciplines with frequent occurrence
NP-related bundles	the generalizability of the; the findings of this study; the development of the; the performance of the proposed	Applied Linguistics, Biology, Electrical Engineering
PP-related bundles	in the context of; in the case of; from the perspective of	Sociology, Applied Linguistics
VP-related bundles	has been shown to; provide valuable insights into; contribute to a more	Biology, Applied Linguistics
Clause-related bundles	in this paper we; in this study we; we propose a novel	Electrical Engineering, Sociology

As shown in Table 9, although all four disciplines use lexical bundles from the four structural categories, there are some obvious disciplinary preferences. Noun phrase-related and prepositional phrase-related bundles are generally typical structural resources in the conclusion sections of academic papers, and this is in line with previous research that has found that academic writing frequently uses condensed nominal structures and referential expressions (Biber et al., 1999; Hyland & Jiang, 2018).

Applied Linguistics is distributed relatively evenly among the sections. A considerable number of noun phrase-related bundles appear in this field, such as the generalisation of the, generalisation of the results, and the results of this study. The above expressions are in line with the rhetorical goals of the conclusion section, such as pointing out deficiencies in the research, presenting the results and discussing the contributions to knowledge. At the same time, clause-related and verb phrase-related bundles, such as *it is necessary to*, *future research will investigate*, and *this study contributes to*, are often employed to express suggestions and link the present research with its future progress in academia. The expressions *in the context of*, *from the perspective of* and *through the lens of* appear frequently in the sociology corpus. These bundles provide the background for the research and thus place the results in a broader context of society, theory and culture. Unlike the scientific method, which usually reports the results directly, the conclusions of sociology are often interpretations and theoretical reflections. Therefore, prepositional constructions can be used to build contexts and analytical viewpoints in language.

Biology and Electrical and Electronic Engineering have a relatively high proportion of noun phrase-related and verb phrase-related bundles, respectively. Biology has produced many typical cases, and our understanding of them and their molecular mechanisms demonstrates the focus of biology on scientific objects, processes and causal explanations. The above structures can be used to present the research

results clearly and concisely, as is required in scientific writing, so electrical engineering also employs this type of structure. Bundles such as the performance of the proposed work, the effectiveness of our work, and the development of a are closely related to technological evaluation and methodological innovation. At the same time, the engineering conclusions have a considerable number of clause bundles, and in this paper, we propose a new one; future work will focus on this. The above indicate the research results and future directions of the work.

The general structure of lexical bundle use is different in various fields of study. Hard sciences have a relatively simple structure, so noun phrase bundles are often used to present scientific knowledge and research results; in the soft sciences, prepositional and clause-related bundles are frequently employed for interpretation and argumentation.

4.3 Functional Distribution of Lexical Bundles across Disciplines

Only the structure cannot fully explain the communication function of lexical bundles in academic writing. Therefore, the identified lexical bundles were further classified according to their function according to Hyland (2008). Based on the above framework, lexical bundles in academic discourse can be divided into three general categories according to their functions: research-oriented bundles, text-oriented bundles and participant-oriented bundles.

Table 10. Functional Classification of Lexical Bundles across Disciplines

Functional category	Representative lexical bundles	Typical disciplines
Research-oriented bundles	results show that; has been shown to; provide valuable insights into	Biology, Electrical Engineering
Text-oriented bundles	in the context of; in line with previous; as a result of	Sociology, Applied Linguistics
Participant-oriented bundles	it is important to; future research could explore; we propose a novel	Applied Linguistics, Electrical Engineering

Functional analysis shows that lexical bundles in the conclusion sections have different rhetorical purposes in various fields. Although all four are research-oriented and text-based or participant-based, they have different preferred forms of academic communication.

Research-oriented Bundles are relatively common in Biology and Electrical Engineering. Phrases such as "*it has been shown that*" and "*our results show that*" are frequently used in biology to summarise the results, present evidence and point out new discoveries in research. Electrical Engineering often uses expressions such as *The results show that*, *reported the development of*, and *we have developed a* to emphasize technological progress and methodological achievements. Frequent use of such bundles shows that the hard sciences attach great importance to empirical evidence and research results; at the same time, Sociology and Applied Linguistics are more likely to use text-oriented and participant-

oriented bundles. Sociologists often use the expressions *in light of*, *from the perspective of* and *through the lens of* to link the results with theoretical explanations. Applied Linguistics makes better use of bundles, and in future research, it will explore what is important to study; this paper will present the results, discuss the corresponding implications, and engage with readers. Lexical bundles in the hard sciences are generally used to present research results and empirical evidence, and those in the social sciences are more often applied for interpretation, evaluation and academic discussion.

4.4 Discussion

Therefore, lexical bundles in the conclusions of research papers are closely linked to the knowledge structures and rhetorical patterns of different disciplines. According to Becher and Trowler (2001), the four fields in this study fall under different categories of knowledge: Applied Linguistics is classified as soft-applied, Sociology as soft-pure, Biology as hard-pure, and Electrical Engineering as hard-applied. The Linguistic Patterns in all these fields show different ways of constructing knowledge and communicating it.

Hard sciences, such as biology and electrical engineering, are generally the results of empirical data, research findings and scientific achievements. Bundles have been shown to, the results show that, and we propose a new way to help researchers summarise the results and highlight their achievements and contributions in a relatively concise way. The demand for research-oriented bundles is in line with the evidence-based approach that can be observed in science, and at the same time, soft sciences such as Sociology and Applied Linguistics also pay more attention to interpretation, argumentation and contextualisation. The Conclusions are divided into several parts: Evaluation, Theoretical Discussion, Future Research Directions, etc. "In light of", "Based on the framework of", and "This study adds to the body of knowledge in" are frequently used to show how a researcher's work is connected to previous studies and establish the discipline's value.

Therefore, lexical bundles are not fixed expressions but rather resources that have been formed according to disciplinary demands and communication goals. Although the repeated expressions may seem conventional, they serve different purposes according to academic norms. Thus, we can see how the various sections present their reasons and knowledge. Based on the above results, EAP teachers can improve by paying more attention to disciplinary awareness of academic writing and various effective communication in different academic environments and fields.

5. Conclusion

Lexical bundles in the conclusion sections of international journal articles from four fields have been studied: Applied Linguistics, Sociology, Biology, and Electrical & Electronic Engineering. Based on the self-built corpus of 400 conclusion sections and a corpus-driven mixed-methods approach, this paper studies the distribution, structure and function of lexical bundles in various areas. Lexical bundles are used in all four subjects according to different patterns and rhetorical functions, as shown in the above results. Soft subjects, such as applied linguistics and sociology, prefer groups of content related to

interpretation, context, evaluation and attitude; hard sciences, represented by biology and electrical and electronic engineering, prefer groups of empirical data, research results and methods. Based on the analysis of structure and function, the hard sciences are more research-oriented in form and have condensed structures; the soft sciences are more contextual and argumentative. Therefore, it can be concluded that lexical bundles are associated with disciplinary norms and knowledge construction. Research on academic writing and EAP instruction will benefit from the concept of disciplinary language awareness. Only 2025 research articles from the four fields are in the corpus. The next ones will be added in due course.

References

- Becher, T., & Trowler, P. R. (2001). Academic tribes and territories: Intellectual enquiry and the culture of disciplines (2nd ed., pp.1-70). Philadelphia, PA: SRHE and Open University Press.
- Biber, D., S. Johansson., G. Leech., S. Conrad., & E. Finegan. (1999). *The Longman Grammar of Spoken and Written English*. London: Longman.
- Brett, Paul. (1994). A genre analysis of the results section of sociology articles. *English for Specific Purposes*, 13(1), 47-59. [https://doi.org/10.1016/0889-4906\(94\)90024-8](https://doi.org/10.1016/0889-4906(94)90024-8)
- Deng Liming, & He Ping. (2023). "We may conclude that:" a corpus-based study of stance-taking in conclusion sections of RAs across cultures and disciplines. *Frontiers in psychology*, 14, 1175144-1175144. <https://doi.org/10.3389/fpsyg.2023.1175144>
- Hyland, K. (2004). Disciplinary Discourses: Social Interactions in Academic Writing. *Michigan: University of Michigan Press*, 2004, 25-60.
- Hyland, K. (2007). As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes*, 27(1), 4-21. <https://doi.org/10.1016/j.esp.2007.06.001>
- Hyland, K. (2008). Academic clusters: text patterning in published and postgraduate writing. *International Journal of Applied Linguistics*, 18(1), 41-62. <https://doi.org/10.1111/j.1473-4192.2008.00178.x>
- Hyland, K. (2008). Genre and academic writing in the disciplines. *Language Teaching*, 41(4), 543-562. <https://doi.org/10.1017/S0261444808005235>
- Hyland, K., & Kevin Jiang. (2018). Academic lexical bundles. *International Journal of Corpus Linguistics*, 23(4), 383-407. <https://doi.org/10.1075/ijcl.17080.hyl>
- SWALES, J. M. (1990). Genre analysis: English in academic and research settings. *Cambridge: Cambridge University Press*.
- Yang Ruiying, & Desmond Allison. (2003). Research article in applied linguistics: moving from results to conclusions. *English for Specific Purposes*, 22(4), 365-385. [https://doi.org/10.1016/S0889-4906\(02\)00026-1](https://doi.org/10.1016/S0889-4906(02)00026-1)
- Yu-Hua Chen, & Paul Baker. (2010). Lexical Bundles in L1 and L2 Academic Writing. *Language Learning and Technology*, 14(2), 30. <https://doi.org/10.64152/10125/44213>