It can be said (that): Lexical Bundles in English Extended Abstracts of Locally Published Turkish Research Articles

Fevzi Umut Özçelik[1] Betül Bal Gezegin[2]

ORCID: [1] 0000-0001-6198-8292, [2] 0000-0001-9818-9347

[1]*Ordu University, School of Foreign Languages, 52000, Ordu, Türkiye*

[2]*Ondokuz Mayıs University, Department of Education, 55100, Samsun, Türkiye*

[1]umutozcelik@odu.edu.tr [2]betul.balgezegin@omu.edu.tr

(Received 10 September 2024; Accepted 2 April 2025)

ABSTRACT: Building on existing literature on phraseology, this study explores how Turkish scholars employ lexical bundles in their extended abstracts. To this end, a corpus of the extended abstracts from five disciplines was collected. The structures and functions of the four-word lexical bundles identified in the extended abstracts were analyzed based on the previously structured two taxonomies (Biber et al., 1999; Hyland, 2008a). The analysis revealed 116 four-word lexical bundles. The findings indicate that, in contrast to previous studies using the same taxonomies, most of the identified lexical bundles in this study are structurally categorized as verb-constructed and functionally classified as text-oriented. We conclude with the potential benefits of this parallel lingual environment created by extended abstracts, which may act as a catalyst to overcome the possible linguistic challenges faced by English as an Additional Language scholars.

Keywords: lexical bundles, extended abstracts, Turkish scholars

It can be said (that): İngilizce Genişletilmiş Özetlerde Türk Araştırmacıların Kullandıkları Sözcüksel Öbekler

ÖZ: Deyimbilim üzerine mevcut literatürü temel alan bu çalışma, Türk akademisyenlerin genişletilmiş özetlerinde sözcüksel demetleri ne ölçüde ve nasıl kullandıklarını araştırmaktadır. Bu amaçla, beş disiplin üzerine makalelerin genişletilmiş özetlerinden yarım milyonluk bir derlem toplanmıştır. Ardından, daha önce yapılandırılmış iki taksonomiden (Biber vd.,

http://doi.org/10.18492/dad.1546802

Dilbilim Araştırmaları Dergisi,
2025 Dilbilim Derneği, Ankara.

1999; Hyland, 2008a) yararlanılarak, dört kelimelik sözcük öbeklerinin yapıları ve işlevleri analiz edilmiştir. Bulgular 116 adet dört kelimelik sözcük öbeği ortaya çıkarmıştır. Bu sonuç aynı taksonomileri kullanan önceki çalışmaların aksine, bu çalışmada tespit edilen sözcüksel öbeklerinin çoğunun yapısal olarak fiil yapılı ve işlevsel olarak metin odaklı olarak sınıflandırıldığını göstermektedir. Çalışmanın sonucu olarak, genişletilmiş özetlerin yarattığı bu paralel dilsel ortamın potansiyel faydaları ve İngilizceyi ek bir dil olarak kullanan akademisyenlere sağlayabileceği kolaylıklar vurgulanmaktadır.

Anahtar kelimeler: sözcüksel demetler, genişletilmiş özetler, Türk akademisyenler

1 Introduction

Although coined some hundred years ago to denote the challenging aspect of engaging in academic pursuits, "publish or perish" as a notion has gained greater importance in recent years (Flowerdew & Habibie, 2021). The process of publishing an article has been proven to encompass both discursive and non-discursive aspects. Scholars aiming to publish in English face various issues stemming from their self-motivation or pressure from the local academic authorities (Cargill, 2020). These aspects have been argued to be legitimately challenging (Flowerdew, 2019), whereas others deny that they constitute an impediment (Hyland, 2016).

It is undeniable that English has been the dominant language in the academic world. English for Academic Purposes (EAP) has expanded branches due to the exponential increase in the number of researchers and journals (Flowerdew & Habibie, 2021). In line with this, English for Research Publication Purposes (ERPP), coined as a sub-branch of EAP, has been serving well to address the concerns experienced by English as an Additional Language (EAL) scholars and novice researchers (Cargill & Burgess, 2008). The rhetorical components both on the lexico-grammatical level (e.g., lexical bundles) and on the discursive level (e.g., IMRAD moves for Introductions), which primarily constitute the pedagogical backbone (e.g., Swales & Feak, 2012) of English as Lingua Academia (ELA, Philipson, 2008), are the linguacultural outcomes of the so-called Tyrannosaurus Rex of English (Swales, 1997). This phenomenon leads to at least two significant undesirable consequences. On the one hand, it puts pressure on periphery EAL scholars, creating linguistic injustice (Soler, 2021); on the other hand, it diminishes linguistic and rhetorical diversity (Canagarajah, 2002; Salager-Meyer, 1997).

In this study, we analyze one of the lexico-grammatical features, i.e., lexical bundles (Biber et al., 1999), deployed in extended abstracts of locally published research articles in Türkiye. The results show that lexical bundles are widely used

in extended abstracts. However, their functional and structural analyses suggest that these choices are more characteristic of non-native speakers. We argue that extended abstracts offer EAL scholars a practical bridge to transnational audiences. To serve this function effectively, such abstracts must align with the linguacultural conventions of English as Lingua Academia. Thanks to the parallel linguistic environment created by extended abstracts, EAL scholars around the world who use English for publication purposes, such as those in Türkiye (i) will be able to overcome the limiting effects of T-rex English (Swales, 1997) by eliminating the linguacultural constraints (Canagarajah, 2002), (ii) may not have to confront the epistemological differences that even scholars in European countries have encountered (see Belcher & Yang, 2020), and finally (iii) can avoid the phenomenon of domain loss (Ferguson, 2007), which has the potential to affect scholars within plurilingual discourse communities.

1.1 Abstracts and Extended Abstracts

Research articles are the building blocks of knowledge in academia. They serve as the fundamental means of communication through which researchers disseminate knowledge and gain a reputation, thus placing their existence in an academic community (Swales, 2004). Genre-based approaches have revealed that research articles generally consist of sections, such as abstract, introduction, literature review, method, analysis, findings, discussion, and conclusion, though there may be discipline-specific variations. Theoretical and empirical studies conducted to better grasp the structure of these building blocks of scholarly writing have revealed that not only does each section have its distinctive rhetorical moves (see Swales, 2004 for Introductions; Hyland, 2000 for Abstracts), but phraseological approaches have also revealed that certain lexical formulas are employed to fulfil the communicative functions of those distinctive moves (e.g., Cortes, 2013).

As a section of a research article, abstracts fulfil multiple tasks. These include acting as 'mini texts' that summarize the study, as 'screening devices' that help readers determine to read the rest of the article, and as 'previews' that orient the reader to the article (Huckin, 2001, p. 93). Hyland (2004) suggests that "the abstract is generally the readers' first encounter with a text and often points at which they decide whether to continue and give the accompanying article further attention or to ignore it" (p.63). These key functions of abstracts have provided the foundation for research to explore the macro-structure and lexico-grammatical characteristics through the application of cross-cultural, cross-disciplinary, and cross-sectional approaches.

The rhetorical structures of abstracts have been the focus of many studies. Building on Hyland's (2000) classification of the rhetorical structures, i.e.,

introduction, purpose, method, product (result), and conclusion (IMRD), several studies have explored research article abstracts with particular attention to specific disciplines, including applied linguistics (Can et al., 2016; Lorés, 2004; Pho, 2008; Ren & Li, 2011; Tseng, 2011), protozoology (Cross & Oppenheim, 2006), English language teaching (Kaya & Yağız, 2020), nanoscience and nanotechnology (Hwang et al., 2017), and conservation biology and wildlife behavior (Samraj, 2005).

This body of research often follows a comparative or contrastive approach. Its primary purpose generally involves detecting the common practices and identifying the possible effects stemming from discipline, language, experience, or a combination of these variables. To this end, for instance, Kaya and Yağız (2020) compared abstracts written in English by Turkish and non-Turkish scholars living in Anglophone countries in the field of English Language Teaching. In a broader scope, Çandarlı (2012) compared abstracts written in English by native Turkish and English researchers in the field of education. Additionally, Van Bonn and Swales (2007) conducted a study comparing abstracts written in French and English. Elsewhere, Martin-Martin (2002) sought differing rhetorical move preferences in the abstracts of English and Spanish writers in the field of social sciences, namely phonetics and psychology.

The extended abstract, an academic genre that is still relatively new but increasingly widespread, is the subject of investigation in this paper. In contrast to the typically perceived abstracts in conference proceedings, which are intended for potential participants of those events (Sidek et al., 2016; Martin & Burgess, 2023), the extended abstracts discussed in this paper refer to the relatively longer versions of standard English abstracts provided for articles published in Turkish.

Several journals available in the database of DergiPark[1] require authors to write an extended version (e.g., 500 – 1000 words) of the short English abstracts upon acceptance for publication. When studies are published online or in print, extended abstracts typically appear after the 'references' section, although some journals place them immediately following the abstract section.

With all these in mind, we argue that extended abstracts can prove to be a significant building block in the course of the establishment of a research community that is more recognized by Anglophone and other non-Anglophone scholars. ‘Extended Abstracts’ carries the potential of becoming a well-established genre through which scholars spread their results to transnational academic communities. Furthermore, it creates a space where scholars can

[1] a local journal management and storage system serving under the administration of Scientific and Technological Research Council of Türkiye- TUBITAK

create, process, and publish without facing the difficulties of publishing in English (e.g., Flowerdew, 1999; Phillipson, 2008). We also recognize the importance of understanding and incorporating both the discursive and non-discursive features of scientific genres to fully reap the benefits of each. In this context, it is crucial to explore and contrast one of the key discursive linguistic choices, namely lexical bundles, with the established practices of scientific communities. Improper use of these bundles can lead to even academically qualified researchers being perceived as outsiders (Hyland, 2008b) or as novice members of the academic community.

The extended abstract as a genre is a promising for several reasons. Firstly, recent studies on scholarly writing in L2 contexts have highlighted that English has become the dominant language for academic and scientific purposes, including research publications and conferences (e.g., Flowerdew, 2000; Phillipson, 2008; Tardy, 2004). The term 'lingua academia' (Phillipson, 2008, p.250) has emerged due to this widespread use. While English as a common language for science facilitates efficient information storage and dissemination (Flowerdew, 1999; Grabe, 1988; Wood, 2001), it also raises concerns of 'academic imperialism' (Fewer, 1997), reducing linguistic and rhetorical diversity (Canagarajah, 2002; Salager-Meyer, 1997) and limiting the reach of scientific communities outside Inner Circle countries (Kachru, 1985). It is important to note the significance of extended abstracts in this context. EAL communities can utilize extended abstracts, which may assist scholars who are either reluctant to use English or not fully proficient in it for academic writing, provided these abstracts adhere to global academic writing conventions.

Additionally, while the parallel linguistic environment may enhance the dissemination of the results to wider transnational academic communities through well-structured extended English abstracts, scholars can write in their mother tongue without facing the challenge of epistemological differences, which are directly related to specific discursive and stylistic features inherent in the language itself (see Belcher & Yang, 2020). Another reason why extended abstracts should be regarded as a significant genre to be integrated is related to the phenomenon of domain loss (Ferguson, 2007). It is likely to be experienced in countries where English has become the dominantly preferred academic language, and we believe that extended abstracts can play a significant role in overcoming domain loss in the academic world.

Fulfilling their overarching objectives of (i) publishing their studies and (ii) reaching wider transnational academic communities in a "fair" environment (Flowerdew, 2020, emphasis added), EAL scholars like those in Türkiye may prefer utilizing the emerging genre of Extended Abstracts by reporting their research in their local language. This might establish an academic environment where the local language will also be preserved for scientific purposes.

1.2 Lexical Bundles

First coined by Biber et al. (1999), lexical bundles are defined as "recurrent expressions, regardless of their idiomaticity, and regardless of their structural status" (p.990). Identifying lexical bundles involves finding "the most frequently recurring sequences of words" (Biber, 2009, p.282) in a given discourse. To be considered as a lexical bundle, an item should meet two criteria: (i) frequency of usage (e.g., 40 times per million words) and (ii) range of usage (e.g., must occur in at least five different texts). While the former pertains to the frequency of occurrence in an entire corpus, the latter impedes any idiosyncratic use by a particular writer or speaker. Although there is no one-for-all formula for the lexical bundle studies, recent studies (e.g., Güngör & Uysal, 2020; Hyland, 2008b) have applied what they call 'a conservative approach' to their corpora to retrieve lexical bundles (e.g., minimum 20 times per million words and an occurrence in at least 10% of the texts).

The results of the existing research on lexical bundles show that they are quite common compared to other 'grammar patterns' (Biber, 2009, emphasis original). These word strings generally comprise noun phrases (e.g., one of the most), prepositional phrases (e.g., on the basis of), anticipatory 'it' (e.g., it can be seen that), and passive constructions (e.g., is based on the) (Biber et al., 1999). They are transparent in nature; that is, their meanings can be deduced from the constituents composing them (Conrad & Biber, 2005). Although they are not complete structural units (see Biber et al., 1999), they provide "a kind of pragmatic 'head' for larger phrases and clauses where they function as discourse frames for the expression of new information" (Biber & Barbieri, 2007, p.270, emphasis in original). While spoken registers mainly consist of clausal bundles, academic prose involves phrasal (i.e., noun and prepositional) word sequences (Biber et al., 2004).

Observations on the relative incompleteness but noticeably strong associations with certain grammatical structures paved the way for the development of structural categorizations. Biber et al. (1999) composed a taxonomy to group lexical bundles according to their structural features under twelve categories. This taxonomy is employed in this study to categorize identified lexical bundles structurally. The following list illustrates each category and examples for them.

- Noun phrase with of-phrase fragment (*one of the most*)
- Noun phrase with other post-modifier fragment (*an important role in*)
- Prepositional phrase with embedded of-phrase fragment (*at the end of*)
- Other prepositional phrase fragment (*as in the case*)
- Anticipatory it + verb phrase/adjective phrase (*it is possible to*)
- Passive verb + prepositional phrase fragment (*is based on the*)

- Copula be + noun phrase/ adjective phrase (*is one of the*)
- (verb phrase +) that-clause fragment (*has been shown that*)
- (verb/adjective +) to-clause fragment (*are likely to be*)
- Adverbial clause fragment (*as shown in figure*)
- Pronoun/noun phrase + be (+…) (*there was no significant*)
- Other expressions (*as well as the*)

Employing the constructed framework, previous studies (e.g., Biber et al., 1999; Biber, 2009; Hyland, 2008a) analyzed the structural patterns of these formulaic word sequences. They revealed that such bundles often function as a bridge between two subsequent units, as the final lexical item frequently forms part of the following structure in a sentence, and that 15% of bundles in conversation were complete structural units. At the same time, this proportion was only 5% for bundles in academic prose (Biber et al., 1999). Additionally, the lexical bundles in written prose were found to have internal variable slots in contrast to bundles in conversation, which are patterns of highly fixed elements (Biber, 2009).

Hyland (2008b) associated disciplinary and register variations with the diverse purposes and target audiences for which the lexical bundles are used. Initially, Biber et al. (2004) identified the primary discursive functions of lexical bundles as (i) stance bundles, (ii) discourse organizers, and (iii) referential expressions. In his attempt to identify frequency, forms, and functions of lexical bundles by utilizing a relatively large corpus composed of research articles, master's theses, and Ph.D. dissertations, Hyland (2008a) modified the framework of Biber et al. (2004). His categorization of lexical bundles was divided into three main functions, further into different sub-functional categories. This study adopts a modified version of Hyland's (2008a) taxonomy, as revised by Salazar (2014) (see Appendix).

Investigating lexical bundles within a specific register or genre has been shown to be empirically grounded, as the extraction process relies entirely on objective criteria (Salazar, 2014). Moreover, their frequent use in a given discourse reflects disciplinary affiliation, highlighting the widely adopted and highly distinctive phraseological practices of that particular discursive domain (Hyland, 2012).

The recognition and appropriate use of these lexical bundles, particularly by students, is essential, as insufficient or incorrect usage may result in unconventional expressions, thereby reducing the likelihood of acceptance by experts in a field (cf. Scott & Tribble, 2006). Meanwhile, appropriate use of lexical bundles may help writers compose a text that meets readers' expectations in academic communities (Coxhead & Byrd, 2007).

Generally, research on lexical bundles has followed a similar methodology with different research objectives. Researchers composed special corpora either to compare them with a reference corpus or to delve into the features of those corpora. However, this was not the case for this study due to the unavailability of a similar corpus (Section 2.2) to be used as a reference corpus. Regarding research objectives, researchers sought to identify lexical bundle usage differences stemming from register, discipline, and language background variables. The relationship between the first language and the command of using English lexical bundles has been the subject of numerous studies. The methodological path researchers generally follow in this type of inquiry goes through composing a non-native corpus, which includes texts written by learners of English, and contrasting the results with those obtained from the native English speaker corpus. Potential derivations, arguably stemming from first language (L1) Swedish (Ädel & Erman, 2012), L1 Chinese (Bychkovska & Lee, 2017; Kim & Kessler, 2022; Liu & Lu, 2019; Lu & Deng, 2019; Lyu & Gee, 2020; Pan and et al. 2016; Wei & Lei, 2011), L1 Spanish (Pérez-Llantada, 2014) have been investigated.

The overall findings of these studies can be summarized with Pérez-Llantada's (2014) word 'hybrid'. It means that the lexical bundles used by L2 users generally, but not always, differ compared to the lexical bundles employed by L1 users.

There are a few studies that particularly focus on specific genres in the Turkish context. Bal-Gezegin (2019) investigated research articles by Turkish scholars. Öztürk (2014) compared Turkish and native English postgraduate students' research papers with those of expert native English scholars' lexical bundle usage. The analyses revealed that although the lexical bundle usage for Turkish postgraduate students is relatively high, Turkish students use lexical bundles quite repetitively. Güngör and Uysal (2020) sought to identify lexical bundles that are distinctive to Turkish scholars. Their analyses uncovered 54 distinctive lexical bundles for Turkish scholars, thereby contributing to cross-linguistics influence on lexical bundle acquisition.

Despite their significance, extended abstracts have not yet been examined from a phraseological perspective. This study aims to address this gap by investigating the following research questions through an analysis of extended abstracts written by Turkish scholars for articles published in the context of Türkiye.

- What are the most common four-word lexical bundles used by Turkish scholars in their extended abstracts in English?
- What are the functional and structural features of the identified lexical bundles?

- Do the lexical bundles demonstrate convergence/divergence from the most frequent bundles identified in the published taxonomies?

2 Corpus and Method

2.1 Corpus of Turkish Scholars Extended Abstracts

Since this study seeks to uncover the lexical bundles employed by Turkish scholars in the extended abstracts of their studies, a specialized corpus comprising half a million words was compiled from the extended abstracts of different knowledge domains. Table 2 below presents the corpus of 572 extended abstracts published in fifteen different journals. Compared to previous research on lexical bundles, the number of texts is relatively high because of the average text length of the extended abstracts (e.g., 600-1500 words). The following criteria were specified before the data collection process to reach a balanced corpus:

a. there must be at least two different journals from a single discipline.
b. papers must be authored by Turkish scholars.
c. papers must be written in the local language (Turkish).

It should be noted that the current number of journals that require an extended abstract is limited. Thus, applying the above-mentioned criteria resulted in the selection of fifteen journals published under five main publishing bodies. (see Appendix for a complete list of journals and the number of articles). We adopted the hard and soft sciences distinction used in previous studies to categorize the journals by discipline (e.g., Sun & Crosthwaite, 2022). This classification organizes fields of specialization based on the characteristics of their subject matter, such as Religion, Communication, and more (see Biglan, 1973).

Furthermore, two social sciences journals required additional consideration. These two journals have inclusive publication policies that encompass the disciplines covered by other journals in the corpus that require extended abstracts (see Table 2). To maintain disciplinary balance, articles from these overlapping disciplines -namely education, economics, religion and communication- published in the two social sciences journals were systematically excluded from the corpus. As a result, articles that were drawn from these social science journals represent the disciplines of art, archeology, humanities, linguistics, history, radio, and television.

Regarding the size of this specialized corpus, Koester's (2010) comprehensive discussion on 'building a small corpus' is considered. Biber (2006, p. 251) suggests that a corpus "must be large enough" to adequately represent the topic or construction under investigation. Koester (2010), drawing

on Flowerdew L.'s (2004) notion of a 'contextualization' parameter—which includes specific context, participants, and objectives—argues that specialized corpora, are more likely to reliably represent a particular register or genre than general corpora because they are carefully targeted (p. 69).

Aligned with the objectives of this research, which is set out to explore to what extent and how L1 Turkish researchers employ English lexical bundles in their extended abstracts, a total of 501,337-word corpus was compiled.

Table 2. Turkish Scholars Extended Abstract Corpus (TSEAC)

Disciplines	#of words	#of abstracts	#of journals
Social Sciences	100,045	123	2
Education	100,460	94	4
Economics	100,482	123	5
Religion	100,171	107	2
Communication	100,179	125	2
TOTAL	501,337	572	15

2.2 Identification of Lexical Bundles

Formulaic sequences must meet specific frequency criteria to be counted as lexical bundles. The cut-off frequency point varies in studies on lexical bundles. The conservative approach taken by previous scholars (Hyland, 2008a, b) is also applied in this study to ensure the reliability of the research. In line with Bestgen's (2018) recommendations, we adopted a normalized conservative frequency cut-off point (i.e., 20 occurrences). Regarding the dispersion criterion, we followed Bao and Liu (2022), who argue that dispersion is particularly crucial in the analysis of mini texts such as abstracts. To minimize the impact of idiosyncratic usage and ensure more representative results, a 2% cut-off point was adopted for identifying lexical bundles in this study.

The analysis was conducted using AntConc (Anthony, 2022), a free corpus analysis tool, in accordance with the stated criteria. The extracted lexical bundles were then transferred to an Excel file and coded based on their structural and functional characteristics. The two authors independently coded ten percent of the examples using the presented taxonomies. Any disagreements were resolved through discussion until full consensus was achieved.

To enhance the quality of tertium comparisons, we applied the same treatments as those used in previous studies. In cases of 'complete overlap' or 'complete subsumption' (Chen & Baker, 2010, p. 33), the bundles were analyzed through concordance analysis, and their frequencies were adjusted to exclude instances where they were part of longer bundles. Additionally, topic-specific bundles (e.g., Ministry of National Education) were manually excluded from the analysis.

3 Findings & Discussion

3.1 Overview

Implementing the specified criteria on the TSEA corpus revealed 116 four-word lexical bundles identified in 5657 tokens. This number constitutes nearly 5% of all the words. Bundles were individually observed at least 1727 times.

To assess the consistency of the findings, the identified lexical bundles were compared with those from previous studies that focused on slightly different academic genres. Of particular relevance to the current research are studies that provide a list of identified lexical bundles based on data produced by Turkish researchers (i.e., Bal-Gezegin, 2019; Güngör & Uysal, 2020; Öztürk, 2014). The cross-examination revealed similar patterns. In at least one of these studies, 62 of 116 bundles were also identified before. The high resemblance ratio of 53% reflects important aspects of the TSEA corpus. It suggests that most of the identified bundles have been used in other genres and disciplines by both novice and expert writers. This ratio also supports the validity of the study's findings.

3.2 Structural Analysis of Lexical Bundles

The application of the suggested taxonomy to the identified lexical bundles revealed results that align with the findings of most of the previous studies. As shown in Table 3, verb-constructed forms (i.e., VP-based and clausal bundles) constitute nearly half of the total number of identified lexical bundles.

Table 3. Structural Analysis of Lexical Bundles

Category	Subcategory	Types (% of all structures)	Tokens (% of all cases)
NP-based	NP with of-phrase fragment	23 (19,82%)	717
	NP with other post-modifier fragment	6 (5,17%)	175
	Other NP	2 (1,72%)	109
	Subtotal	31 (26,72%)	1001 (17,69%)
PP-based	PP with embedded of phrase	13 (11,20%)	821
	Other PP fragment	15 (12,93%)	908
	Subtotal	28 (24,13%)	1729 (30,56%)
VP-based	copula be + NP /adjective phrase	1 (0,86%)	152
	VP with active verb	-	-
	VP with infinitive verb	1 (0,86%)	20
	VP with passive verb	8 (6,89%)	229
	Lexical bundle beginning with past participle	3 (2,58%)	65
	Other	2 (1,72%)	55
	Subtotal	15 (12,93%)	521 (9,20%)
Clause-based	PP + copula be	2 (1,72%)	134
	NP + copula be	5 (4,31%)	134
	Anticipatory it + copula be +adjective phrase	4 (3,44%)	228
	Anticipatory it + passive verb + that	17 (14,65%)	1365
	NP / Complementizer + passive verb	-	-
	NP + active verb	2 (1,72%)	51
	NP + active verb + that	-	-
	That fragment	4 (3,44%)	169
	Wh-fragment	1 (0,86%)	82
	There fragment	2 (1,65%)	47
	Other	1 (0,86%)	28
	Subtotal	38 (32,75%)	2238 (39,56%)
Conjunctions		4 (3,44%)	168 (2,96%)
Total		116	5657

Table 3 illustrates the overall type and token frequency of lexical bundles. Percentages in the 'types' column denote the percentage of that type when compared to the total of 116 four-word bundles. Similarly, percentages in the token column represent the percentage of a total of 5657 tokens of four-word bundles. The initial analysis reveals that clause-based lexical bundles are the most frequently used structure (32,75%). The following most frequent bundles are NP-based ones (26,72%). The least frequent category is conjunctions, with 4 (3,44%) lexical bundles. Regarding the token frequency, clause-based bundles constitute 39,56% of total tokens (5657). This is followed by prepositional phrase-based (PP-based) (30,56%), noun phrase-based (NP-based) (17,69%), VP-based (9,20%), and conjunctions (2,96%).

The total percentage of NP-based and PP-based bundles (50,86%) constitutes slightly more than half of all the identified bundle types. This relatively high proportion is in line with the previous literature. Considering the features of this genre, the results are evident, as bundles in academic prose typically form parts of noun or prepositional phrases (Biber & Conrad, 2019; Biber et al., 1999; Hyland, 2008a). The use of noun phrases also reflects the "nominal style" (Lan et al., 2022, p.6) of academic writing since they are beneficial means for designating aspects of information (Conrad & Biber, 2005).

However, the equal distribution of clause-based and VP-based (i.e., verb-constructed) bundles compared to NP- and PP-based bundles requires critical attention. This finding contrasts with some previous research (e.g., Hyland, 2008a) while supporting others (e.g., Hyland, 2008b; Pérez-Llantada, 2014). It was found that 45,68% of the identified bundles belong to these categories. Additionally, the overall token frequency of these categories is 48,77%. Taken together, these results imply that Turkish authors frequently employed a bundle that comprises a verb. This contradicts the characteristic features of the research article register, which is typically marked by extended noun phrases, nominalizations, and the use of passive voice (see Biber & Conrad, 2019). According to Conrad and Biber (2005), PP and NP-based bundles constituted approximately 60% of the identified items in their academic corpus. Similarly, Bal-Gezegin (2019) reported that PP and NP-based bundle rates were relatively higher in her corpus of research articles written by Turkish scholars. The findings of this study suggest that while the authors of these extended abstracts appear to be aware of the linguistic resources that characterize academic prose, they may not yet have achieved full competence in academic writing. This is evidenced by their use of lexical bundles, which diverges from expert academic writing, where NP and PP-based bundles are notably more prominent (Biber et al., 2004; Chen & Baker, 2010). In addition, the prevalence of verb-constructed lexical bundles implies the low proportion of noun phrases, prepositional and

adjectival phrases, which was claimed to “verbose” the style and make it “unqualified” (Pérez-Llantada, 2014, p.92).

Many of these bundles belong to passive fragments (6,89%) and anticipatory-it passive verb fragments (14,65%). The employment of anticipatory it-structures and passive verb + prepositional phrase fragments has been construed as the "depersonalized mode" within academic discourse (Cortes, 2004, p.408). Similarly, Hyland (2008b) asserted that, syntactically, these patterns introduce extraposed structures, and functionally, they convey evaluation without explicitly identifying the source. It was also suggested that Turkish scholars disguise their authorial presence since they feel incompetent in presenting arguments (Güngör & Uysal, 2020).

The primary source for this verbose style is the set of anticipatory-it-passive verbs (e.g., *it can be seen that*, 1365 token frequency) and passive-verb constructions (e.g., *was found in the*, 229 token frequency). This phenomenon can be attributed to direct L1 transfer and a preference for covert authorial presence. Corpus studies on Turkish academic writing across different genres, such as archeology, education, and engineering have shown that passive structures like *edil-* (been), *yapıl-* (be made), *bulun-* (be found), *görül-*(be seen) are in the top twenty frequent academic words list in numerous disciplines (Tüfekçioğlu & Albayrak, 2022). Given these structural similarities, it is reasonable to suggest that Turkish scholars transfer their Turkish expressions to their academic English (also see Güngör & Uysal, 2020). Moreover, the preference for a covert authorial presence (Işıktaş, 2018) may contribute to the frequent use of verbs. Avoiding overt self-reference often necessitates the use of inanimate subjects (e.g., study, research, analysis, tools), which in turn promotes the use of passive constructions.

3.3 Functional Analysis of Lexical Bundles

Previous literature has demonstrated that lexical bundles carry out discourse functions, which vary based on the register (Biber et al., 1999, 2004). In line with this, the functional categorization of identified lexical bundles is quite significant. To achieve this, the lexical bundles identified in this study were analyzed according to the taxonomy offered by Hyland (2008a) and modified by Salazar (2014). The results point to the underlying functions of extended abstracts that authors try to fulfil by employing them.

Table 4. Functional Categorization of Lexical Bundles

Function	Structure Type					
	PP-based	NP-based	Clause-Based	VP-based	Conjunction	Total (% of all functions)
Research-Oriented Bundles						44 (38%)
Location	2	1	1	-	-	4
Procedure	2	13	8	3	-	26
Quantification	1	1	1	1	-	4
Description	1	6	1	2	-	10
Grouping	-	-	-	-	-	-
Referential	-	-	-	-	-	-
Text-Oriented Bundles						60 (51%)
Additive	2	-	-	-	1	3
Comparative	-	-	-	-	-	-
Inferential	4	5	16	4	3	32
Causative	-	-	-	-	-	
Structuring	2	-	-	-	-	2
Framing	13	-	-	2	-	15
Citation	-	-	-	-	-	-
Generalization	-	-	3	-	-	3
Objective	1	-	1	3	-	5
Participant-Oriented Bundles						12 (9%)
Stance	-	5	5	-	-	10
Engagement	-	-	2	-	-	2

Table 4 shows that text-oriented lexical bundles were the dominant category (51%) in terms of the functions of lexical bundles. The research-oriented (38%) bundles follow this category, and the least frequently used bundles belong to the participant-oriented category (9%). These functions are further discussed below.

3.3.1 Text-Oriented Bundles

Text-oriented bundles pertain to the structural organization of the text, and semantically, they are concerned with the interpretation of arguments (Hyland,

2008a). The biggest difference between text-oriented bundles lies in their use of inferential bundles. Inferential bundles were used nearly two times as often as framing bundles, which is the second highest category for text-oriented bundles. A close examination of the data revealed that authors mainly used inferential bundles in the form of *it (is)(was) (determined)(observed)(seen)(found) that.* The results indicate that authors predominantly utilize extended abstracts to present the findings of their studies.

From a rhetorical perspective, this result may suggest that extended abstracts could not be utilized to achieve their primary purpose. These types of extracts are requested from authors to increase the recognizability of their studies. Although there is a place for results and implications in the moves of abstracts (Swales, 2004), neglecting other moves, such as introduction, may violate academic conventions. From a register perspective, the relatively lower frequency of NP-based and PP-based (9) compared to verb-constructed (VP-based and Clausal, 20) utilized to signal conclusions drawn from the research (i.e., inferential bundles) may imply the need for a pedagogical intervention to equip them with diverse expressions to prevent divergence from academic norms (see Li & Schmitt, 2009). Additionally, drawing on the findings in her cross-linguistic study (L1 Spanish – L2 English), Pérez-Llantada (2014) states that "L2 writers tend to use these bundles to construct reason–result argumentative discourse" (p.92). Pérez-Llantada further suggests that the lack of stance bundles as a result of the overuse of verb-based inferential bundles partially deteriorates the academic prose since it leads to an increase in verb usage (see Pérez-Llantada, 2014).

3.3.2 Research-Oriented Bundles

Contrary to the findings in this study, research-oriented bundles have been found to be a dominant category in scientific discourse (Hyland, 2008b). Table 4 shows that authors used more procedural and descriptive bundles, and they did not use either grouping or referential bundles. The absence of referential bundles may be attributed to the genre conventions of abstracts, whose main purpose is to provide a concise overview of the research without delving into extensive details (see Hyland, 2004).

Close examination of the data revealed that there is an overlap around the function of describing the purpose of the research in the forms of *the (main) purpose (aim) of the (study)(research)* and *(the) aim of the study*. Research-oriented description bundles were used to give details about the research itself in the form of *the scope of the, the subject of the,* and *the development of the*. These results align with the findings of Hyland and Jiang's (2022) study on an EAL corpus. They showcased that bundles in EAL scholar articles diverge from

those of expert researchers. Moreover, they use more research-oriented bundles, thereby emphasizing the aspects related to research instead of constructing academic text with an interactional discourse (Hyland & Jiang, 2022).

3.3.4 Participant-Oriented Bundles

The least frequently used functional category was participant-oriented bundles. While stance features indicate the position of the writers through, for example, hedging bundles, engagement bundles concern the writer's intention to engage readers in the meaning-making process.

The majority of the stance bundles are in the form of *anticipatory it + is + adj + to*. On the other hand, a large proportion of the engagement markers are in the form of *(it) can be said (that).* The high frequency of this bundle was linked to the linguistic influence of L1 Japanese learners (Allen, 2009). Similarly, Güngör and Uysal (2020) identified this bundle as distinctive to Turkish scholars.

The relatively low frequency of participant-oriented bundles and their utilization with bundles distinctive to Turkish scholars can be attributed to language transfer (Güngör & Uysal, 2020), contextual influence of academic writing (Işık-Taş, 2018), or inadequate language proficiency (Hu & Cao, 2011). Güngör and Uysal (2020) illustrated that the 'it can be said' engagement marker is distinctive to Turkish scholars. Işık-Taş (2018) showcased in her cross-linguistic study that, contrary to internationally oriented colleagues, locally oriented Turkish scholars avoid using personal pronouns, and they avoid "overt authorial presence" (p.26) in their publications to make their text appear objective by not being involved in the texts. Since stance-making involves the "ways in which writers present themselves and express their judgments, opinions, and commitments" (Hyland, 2005a, p. 176), it serves as a useful lens for interpreting authorial choices. Drawing on Işık-Taş's (2018) findings, we argue that our extended abstracts corpus reflects characteristics commonly found in papers by locally oriented Turkish authors, among whom avoiding overt authorial presence is a frequent practice. Performing the actions through participant-oriented bundles requires high-level and well-established knowledge of lexico-grammatical and semantic features (Hyland, 2005b); therefore, the low frequency of stance features was also attributed to low English competence by Hu and Cao (2011).

4 Conclusion

Three research questions guided this research paper. In line with the first research question, the high number of lexical bundles implies that Turkish writers tend to use lexical bundles quite frequently (Hyland, 2008a; Öztürk, 2014). The relatively high percentage for the overall lexical bundle usage (about 5% of all the words in the corpus) may imply that extended abstracts in our corpus bear a resemblance to what is called novice writing. It has been showcased that novice academic writing is more phrasal than published articles, and these writers tend to rely increasingly on prefabricated strings to convey their arguments (Hyland, 2008a). Likewise, relatively more reliance on lexical bundles by less proficient L2 English writers has also been documented (Paquot & Granger, 2012).

The following research questions were related to the functional and structural analyses of the lexical bundles. A close examination of token frequency revealed that verb-constructed bundles (i.e., verb-based and clause-based) are as frequent as the total of PP- and NP-based bundles (48%). It can be suggested that texts in the Extended Abstract corpus share more commonalities with the corpora collected from non-native authors, which has been claimed to be written in a verbosed (Pérez-Llantada, 2014) style as a result of the high verb usage similar to this study.

The following significant finding was the low rate of participant-oriented bundles. As was implied, the low frequency might stem from the tendency of Turkish authors to write with the "depersonalized mode" (Cortes, 2004, p.208), thereby disguising their presence.

Previous literature has warned that mis-, over-, or under-use of multi-unit sequences might indicate a lack of competence in academic writing and decrease the likelihood of being accepted by a scientific community (Ädel & Erman, 2012; Hyland, 2008a; Scott & Tribble, 2006). Considering this and the results obtained in this study, despite the need to view them with caution due to the lack of a comparable corpus, increasing awareness about the conventional way of using lexical bundles and integrating them into academic writing lessons constitute an essential importance to increase the recognition of studies conducted in Türkiye (c.f. Uysal, 2014).

The findings of this study align with Hyland's (2008) assertion that lexical bundles vary in their occurrence and usage across different disciplinary contexts, as seen in the extended abstracts targeted in this study. The frequent occurrence of lexical bundles signals that adopting a discipline-specific approach to teaching lexical bundles is suggested. EAP course designers and instructors should integrate these discipline-specific bundles into their instructional materials, and EAP practitioners could provide explicit instruction focused on these disciplinary bundles. In addition to explicit instruction, systematic exposure to these formulaic expressions, particularly genre-specific

exposure, can lead to better understanding and mastery of writing academic texts in general and specific sections of an academic paper in particular.

Author Contributions: Both authors equally contributed to all the stages of this research including conceptualization, analysis, draft preparation, writing, revisions and editing. Both authors have read and agreed to the final version of the article.
Submission statement and verification: This study has not been previously published elsewhere. It is not under review in another journal. Publication of the study has been approved, either implicitly or explicitly, by all authors and the responsible authorities at the university/research center where the study was conducted. If the study is accepted for publication, it will not be published in the same form in another printed or electronic medium in Turkish or any other language without the written permission of the Journal of Linguistic Research.
Conflict of Interest Statement: The authors declare that there are no financial or academic conflicts of interest between themselves or with other institutions, organizations or individuals that may affect this study.
Data Use: The corpus for this study was derived from academic journals that publish both articles and their accompanying extended abstracts. The dataset is accessible for review upon request.
Ethical Approval/Participant Consent: There is no need for ethical approval in the study.
Financial Support: No financial support was received for the study.

References

Ädel, A., & Erman, B. (2012). Recurrent word combinations in academic writing by native and non-native speakers of English: A lexical-bundles approach. *English for Specific Purposes, 31*(2), 81–92. https://doi.org/10.1016/j.esp.2011.08.004

Allen, D. (2009). Lexical bundles in learner writing: An analysis of formulaic language in the ALESS learner corpus. *Komaba Journal of English Education, 1*(4), 105–127.

Anthony, L. (2022). *AntConc* (Version 4.1.4) [Computer software]. Waseda University. https://www.laurenceanthony.net/software

Bal-Gezegin, B. (2019). Lexical bundles in published research articles: A corpus-based study. *Journal of Language and Linguistic Studies, 15*(2), 520–534. https://doi.org/10.17263/jlls.586188

Bao, K., & Liu, M. (2022). A corpus study of lexical bundles used differently in dissertation abstracts produced by Chinese and American PhD students of linguistics. *Frontiers in Psychology, 13*, Article 893773. https://doi.org/10.3389/fpsyg.2022.893773

Belcher, D., & Yang, H. S. (2020). Global perspectives on linguacultural variation in academic publishing. *Journal of English for Research Publication Purposes, 1*(1), 28–50. https://doi.org/10.1075/jerpp.19009.bel

Bestgen, Y. (2018). Evaluating the frequency threshold for selecting lexical bundles by means of an extension of Fisher's exact test. *Corpora, 13*(2), 205–228. https://doi.org/10.3366/cor.2018.0144

Biber, D. (2006). *University language: A corpus-based study of spoken and written registers*. John Benjamins. https://doi.org/10.1075/scl.23

Biber, D. (2009). A corpus-driven approach to formulaic language in English: Multi-word patterns in speech and writing. *International Journal of Corpus Linguistics, 14*(3), 275–311. https://doi.org/10.1075/ijcl.14.3.08bib

Biber, D., & Barbieri, F. (2007). Lexical bundles in university spoken and written registers. *English for Specific Purposes, 26*(3), 263–286. https://doi.org/10.1016/j.esp.2006.08.003

Biber, D., & Conrad, S. (2019). *Register, genre, and style* (2nd ed.). Cambridge University Press. https://doi.org/10.1017/9781108686136

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*. Longman.

Biglan, A. (1973). The characteristics of subject matter in different academic areas. *Journal of Applied Psychology, 57*(3), 195–203. https://doi.org/10.1037/h0034701

Bychkovska, T., & Lee, J. J. (2017). At the same time: Lexical bundles in L1 and L2 university student argumentative writing. *Journal of English for Academic Purposes, 30*, 38–52. https://doi.org/10.1016/j.jeap.2017.10.008

Can, S., Karabacak, E., & Qin, J. (2016). Structure of moves in research article abstracts in applied linguistics. *Publications, 4*(3), https://doi.org/10.3390/publications4030023

Canagarajah, A. S. (2002). *The geopolitics of academic writing*. University of Pittsburgh Press. https://doi.org/10.2307/j.ctt5hjn6c

Cargill, M. (2020). Teaching ERPP to undergraduate STEM students in Chinese universities? Addressing contextual realities in an EFL academic environment. *Journal of English for Research Publication Purposes, 1*(1), 66–78. https://doi.org/10.1075/jerpp.19011.car

Cargill, M., & Burgess, S. (2008). Introduction to the special issue: English for research publication purposes. *Journal of English for Academic Purposes, 7*(2), 75–76. https://doi.org/10.1016/j.jeap.2008.02.006

Çandarlı, D. (2012). A cross-cultural investigation of English and Turkish research article abstracts in educational sciences. *Studies About Languages, 20*, 12–16. https://doi.org/10.5755/j01.sal.0.20.1770

Chen, Y.-H., & Baker, P. (2010). Lexical bundles in L1 and L2 academic writing. *Language Learning & Technology, 14*(2), 30–49. http://doi.org/10125/44213

Conrad, S., & Biber, D. (2005). The frequency and use of lexical bundles in conversation and academic prose. *Lexicographica, 20*, 56–71. https://doi.org/10.1515/9783484604674.56

Cortes, V. (2004). Lexical bundles in published and student disciplinary writing: Examples from history and biology. *English for Specific Purposes, 23*(4), 397–423. https://doi.org/10.1016/j.esp.2003.12.001

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

Coxhead, A., & Byrd, P. (2007). Preparing writing teachers to teach the vocabulary and grammar of academic prose. *Journal of Second Language Writing, 16*(3), 129–147. https://doi.org/10.1016/j.jslw.2007.07.002

Cross, C., & Oppenheim, C. (2006). A genre analysis of scientific abstracts. *Journal of Documentation, 62*(4), 428–446. https://doi.org/10.1108/00220410610700953

Ferguson, G. (2007). The global spread of English, scientific communication and ESP: Questions of equity, access and domain loss. *Ibérica, 13*, 7–38.

Fewer, G. (1997, February 27). Beyond the language barrier [Letter to the editor]. *Nature, 385*(6619), 764. https://doi.org/10.1038/385764c0

Flowerdew, J. (1999). Writing for scholarly publication in English: The case of Hong Kong. *Journal of Second Language Writing, 8*(2), 123–145. https://doi.org/10.1016/s1060-3743(99)80125-8

Flowerdew, J. (2000). Discourse community, legitimate peripheral participation, and the non-native English-speaking scholar. *TESOL Quarterly, 34*(1), 127–150. https://doi.org/10.2307/3588099

Flowerdew, L. (2004). The argument for using English specialised corpora to understand academic and professional language. In U. Connor & T. A. Upton (Eds.), *Discourse in the professions: Perspectives from corpus linguistics* (pp. 11–33). John Benjamins. https://doi.org/10.1075/scl.16.02flo

Flowerdew, J. (2020). English for research and publication purposes: A personal view. *Journal of English for Research Publication Purposes, 1*(2), 170–183. https://doi.org/10.1075/jerpp.20013.flo

Flowerdew, J., & Habibie, P. (2021). *Introducing English for research publication purposes*. Routledge. https://doi.org/10.4324/9780429317798

Grabe, W. (1988). English, information management, and technology transfer: A rationale for English as an international language. *World Englishes, 7*(1), 63–72. https://doi.org/10.1111/j.1467-971x.1988.tb00215.x

Güngör, F., & Uysal, H. H. (2020). Lexical bundle use and cross-linguistic influence in academic texts. *Lingua, 242*, 102859. https://doi.org/10.1016/j.lingua.2020.102859

Hu, G., & Cao, F. (2011). Hedging and boosting in abstracts of applied linguistics articles: A comparative study of English- and Chinese-medium journals. *Journal of Pragmatics, 43*(11), 2795–2809. https://doi.org/10.1016/j.pragma.2011.04.007

Huckin, T. (2001). Abstracting from abstracts. In M. Hewings (Ed.), *Academic writing in context* (pp. 93–103). University of Michigan Press.

Hwang, J. V., Nguyen, T.-H., & Su, T.-J. (2017). Move analysis for scientific abstract sections: A study of nanoscience and nanotechnology research article abstracts. *World Transactions on Engineering and Technology Education, 15*(1), 19–22. https://doi.org/10.22158/elsr.v4n2p1

Hyland, K. (2004). *Disciplinary discourses: Social interactions in academic writing* (2nd ed.). University of Michigan Press. https://doi.org/10.2307/3587657

Hyland, K. (2005a). Stance and engagement: A model of interaction in academic discourse. *Discourse Studies, 7*(2), 173–192. https://doi.org/10.1177/1461445605050365

Hyland, K. (2005b). *Metadiscourse: Exploring interaction in writing*. Continuum. https://doi.org/10.5040/9781350063617.0007

Hyland, K. (2008a). 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. (2008b). 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. (2012). English for academic purposes and discourse analysis. In J. P. Gee & M. Handford (Eds.), *The Routledge handbook of discourse analysis* (pp. 412–424). Routledge. https://doi.org/10.4324/9780203809068.ch29

Hyland, K. (2016). Academic publishing and the myth of linguistic injustice. *Journal of Second Language Writing, 31*, 58–69. https://doi.org/10.1016/j.jslw.2016.01.005

Hyland, K., & Jiang, F. (2022). Bundles in advanced EAL authors' articles: How do they compare with World Englishes practices? *World Englishes, 41*(4), 554–570. https://doi.org/10.1111/weng.12605

Işık-Taş, E. E. (2018). Authorial identity in Turkish-language and English-language research articles in sociology: The role of publication context in academic writers' discourse choices. *English for Specific Purposes, 49*, 26–38. https://doi.org/10.1016/j.esp.2017.10.003

Kachru, B. B. (1985). Standards, codification and sociolinguistic realism: The English language in the Outer Circle. In R. Quirk & H. G. Widdowson (Eds.), *English in the world: Teaching and learning the language and literatures* (pp. 11–30). Cambridge University Press.

Kaya, F., & Yağız, O. (2020). Move analysis of research article abstracts in the field of ELT: A comparative study. *Journal of Language and Linguistic Studies, 16*(1), 390–404. https://doi.org/10.17263/jlls.712854

Kim, S., & Kessler, M. (2022). Examining L2 English university students' uses of lexical bundles and their relationship to writing quality. *Assessing Writing, 51*, Article 100589. https://doi.org/10.1016/j.asw.2021.100589

Koester, A. (2010). Building small specialised corpora. In A. O'Keeffe & M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 66–79). Routledge. https://doi.org/10.4324/9780203856949-6

Lan, G., Zhang, Q., Lucas, K., Sun, Y., & Gao, J. (2022). A corpus-based investigation on noun-phrase complexity in L1 and L2 English writing. *English for Specific Purposes, 67*, 4–17. https://doi.org/10.1016/j.esp.2022.02.002

Li, J., & Schmitt, N. (2009). The acquisition of lexical phrases in academic writing: A longitudinal case study. *Journal of Second Language Writing, 18*(2), 85–102. https://doi.org/10.1016/j.jslw.2009.02.001

Liu, J., & Lu, Y. (2019). A corpus-based comparative study on lexical bundles in native and Chinese scholars' English abstracts: Taking linguistics and chemistry as an

example. *Chinese Journal of Applied Linguistics, 42*(4), 488–502. https://doi.org/10.1515/cjal-2019-0029

Lorés, R. (2004). On research-article abstracts: From rhetorical structure to thematic organisation. *English for Specific Purposes, 23*, 280–302. https://doi.org/10.1016/j.esp.2003.06.001

Lu, X., & Deng, J. (2019). With the rapid development: A contrastive analysis of lexical bundles in dissertation abstracts by Chinese and L1-English doctoral students. *Journal of English for Academic Purposes, 39*, 21–36. https://doi.org/10.1016/j.jeap.2019.03.008

Lyu, M., & Gee, R. W. (2020). Lexical bundles in thesis abstracts by L1-Chinese learners of English and U.S. students. *English Language Teaching, 13*(1), 141–155. https://doi.org/10.5539/elt.v13n1p141

Martín, P., & Burgess, S. (2023). "Our study offers insight into …": Rhetorical promotion in English and Spanish conference abstracts. *International Journal of Applied Linguistics.* Advance online publication. https://doi.org/10.1111/ijal.12483

Martin-Martin, P. (2002). A genre-based investigation of abstract writing in English and Spanish. *Revista Canaria de Estudios Ingleses, 44*, 47–64.

Öztürk, Y. (2014). *Lexical bundle use of Turkish and native English writers: A corpus-based study* (Unpublished master's thesis). Anadolu University, Eskişehir, Türkiye.

Pan, F., Reppen, R., & Biber, D. (2016). Comparing patterns of L1 versus L2 English academic professionals: Lexical bundles in telecommunications research journals. *Journal of English for Academic Purposes, 21*, 60–71. https://doi.org/10.1016/j.jeap.2015.11.003

Paquot, M., & Granger, S. (2012). Formulaic language in learner corpora. *Annual Review of Applied Linguistics, 32*, 130–149. https://doi.org/10.1017/S0267190512000098

Pérez-Llantada, C. (2014). Formulaic language in L1 and L2 expert academic writing: Convergent and divergent usage. *Journal of English for Academic Purposes, 14*, 84–94. https://doi.org/10.1016/j.jeap.2014.01.002

Phillipson, R. (2008). Lingua franca or lingua frankensteinia? English in European integration and globalisation. *World Englishes, 27*(2), 250–267. https://doi.org/10.1111/j.1467-971x.2008.00555.x

Pho, P. D. (2008). Research-article abstracts in applied linguistics and educational technology: A study of rhetorical structure and authorial stance. *Discourse Studies, 10*(2), 231–250. https://doi.org/10.1177/1461445607087010

Ren, H., & Li, Y. (2011). A comparative study on the rhetorical moves of abstracts in published research articles and master's foreign-language theses. *English Language Teaching, 4*(1), 162–167. https://doi.org/10.5539/elt.v4n1p162

Salager-Meyer, F. (1997). Scientific multilingualism and "lesser" languages. *Interciencia, 22*, 197–201.

Salazar, D. (2014). *Lexical bundles in native and non-native scientific writing: Applying a corpus-based study to language teaching.* John Benjamins. https://doi.org/10.1075/scl.65

Samraj, B. (2005). An exploration of a genre set: Research-article abstracts and introductions in two disciplines. *English for Specific Purposes, 24*(2), 141–156. https://doi.org/10.1016/j.esp.2002.10.001

Scott, M., & Tribble, C. (2006). *Textual patterns: Key words and corpus analysis in language education*. John Benjamins. https://doi.org/10.1075/scl.22

Sidek, H. M., Mat Saad, N. S., Baharun, H., & Idris, M. M. (2016). An analysis of rhetorical moves in abstracts for conference proceedings. *International E-Journal of Advances in Social Sciences, 2*(4), 24–32. https://doi.org/10.18769/ijasos.80136

Soler, J. (2021). Linguistic injustice in academic publishing in English: Limitations and ways forward in the debate. *Journal of English for Research Publication Purposes, 2*(2), 160–171. https://doi.org/10.1075/jerpp.21002.sol

Sun, S. A., & Crosthwaite, P. (2022). "The findings might not be generalisable": Investigating negation in the limitations sections of PhD theses across disciplines. *Journal of English for Academic Purposes, 59*, Article 101155. https://doi.org/10.1016/j.jeap.2022.101155

Swales, J. M. (1997). English as Tyrannosaurus Rex? *World Englishes, 16*(3), 373–382. https://doi.org/10.1111/1467-971x.00071

Swales, J. M. (2004). *Research genres: Explorations and applications*. Cambridge University Press. https://doi.org/10.1017/CBO9781139524827

Swales, J. M., & Feak, C. B. (2012). *Academic writing for graduate students* (3rd ed.). University of Michigan Press. https://doi.org/10.3998/mpub.2173936

Tardy, C. (2004). The role of English in scientific communication: Lingua franca or Tyrannosaurus Rex? *Journal of English for Academic Purposes, 3*(3), 247–269. https://doi.org/10.1016/j.jeap.2003.10.001

Tseng, F. P. (2011). Analysis of move structure and verb tense of research-article abstracts in applied-linguistics journals. *International Journal of English Linguistics, 1*(2), 27–39. https://doi.org/10.5539/ijel.v1n2p27

Tüfekçioğlu, B., & Albayrak, F. (2022). Statistical analysis from a pedagogical perspective of high-frequency words in Turkish academic vocabulary. *International Journal of Eurasian Education and Culture, 7*(18), 1757–1793. https://doi.org/10.35826/ijoecc.596

Uysal, H. H. (2014). Turkish academic culture in transition: Centre-based state policies and semiperipheral practices of research, publishing and promotion. In K. Bennett (Ed.), *The semiperiphery of academic writing: Discourses, communities and practices* (pp. 165–188). Palgrave Macmillan. https://doi.org/10.1057/9781137351197_10

Van Bonn, S., & Swales, J. M. (2007). English and French journal abstracts in the language sciences: Three exploratory studies. *Journal of English for Academic Purposes, 6*(2), 93–108. https://doi.org/10.1016/j.jeap.2007.04.001 (ScienceDirect)

Wei, Y., & Lei, L. (2011). Lexical bundles in the academic writing of advanced Chinese EFL learners. *RELC Journal, 42*(2), 155–166. https://doi.org/10.1177/0033688211407295

Wood, A. (2001). International scientific English: The language of research scientists around the world. In J. Flowerdew & M. Peacock (Eds.), *Research perspectives on English for academic purposes* (pp. 81–83). Cambridge University Press. https://doi.org/10.1017/cbo9781139524766.008