



ARAŞTIRMA MAKALESİ | RESEARCH ARTICLE

LEXICAL BUNDLES ACROSS DISCIPLINES: THE CASE OF RESEARCH  
ARTICLES IN THE SOCIAL SCIENCES

İlyas YAKUT

Dr. Öğr. Üyesi, Kahramanmaraş İstiklal Üniv,  
İnsan ve Toplum Bilimleri Fakültesi,

yakutilyas@gmail.com

0000-0001-9680-4561

Fatma YUVAYAPAN

Doç. Dr., Kahramanmaraş İstiklal Üniv,  
İnsan ve Toplum Bilimleri Fakültesi,

fyuvayapan@gmail.com

0000-0002-7924-0933

**Atıf/Citation:** Yakut, İ. & Yuvayapan, F. (2022). Lexical bundles across disciplines: The case of research articles in the social sciences. *İnönü Üniversitesi Uluslararası Sosyal Bilimler Dergisi*, 11(2), 272-291.

<https://doi.org/10.54282/inijoss.1143817>

**Abstract**

In order for an academic text to be considered appropriate in the community, it needs to exhibit disciplinary and cultural-based linguistic conventions. With the advances in corpus linguistics, scholars have been able to reveal the employment of these conventions in academic genres, one of which is lexical bundles. Simply defined as recurrent word combinations, lexical bundles reflect prominent functions in academic genres, as they deal with discourse organization, writer-reader negotiation, and stance construction, all of which achieve academic persuasion. Although the previous research has established the importance of lexical bundles, there is much less information about the disciplinary variations in the use of lexical bundles in academic genres. Adopting an automated frequency-driven approach, this research attempted to identify lexical bundles in research articles in the social sciences. Based on the investigation of 4-word lexical bundles in a corpus of research articles written between 2010 and 2019 in applied linguistics, marketing, and political sciences, we observed an impact of disciplinary variation on the overall lexical bundle usages. Concerning the structural and functional distributions of word strings, we observed differences across the disciplines, indicating that the academic communities might have a decisive role in text construction, yielding divergence across the disciplines in the social sciences. Despite differences, we also observed some similarities regarding the structural and functional sequences of bundles across the disciplines, indicating that the social sciences, an umbrella field in academia, has its own merits resulting in convergence across different disciplines.

**Keywords:** Corpus research, Social sciences, Academic writing, Research article, Lexical bundle

Received / Geliş Tarihi: 14.07.2022

Accepted / Kabul Tarihi: 11.10.2022

## FARKLI DİSİPLİNLERDE SÖZCÜK ÖBEKLERİ: SOSYAL BİLİMLERDEKİ ARAŞTIRMA MAKALELERİ ÖRNEĞİ

### Öz

Akademik bir metnin uygun görülebilmesi için disiplin ve kültürel temelli dilsel teamüllere uygun olması gerekmektedir. Derlem dilbilimindeki gelişmelerle birlikte, bilim insanları, akademik türlerdeki teamülleri – bunlardan birisi sözcük öbeği olan – belirlemeye başlamışlardır. En basit şekliyle tekrarlayan kelime kombinasyonları olarak tanımlanan sözcük öbekleri, akademik türlerde tamamı akademik ikna sürecini inşa eden diskur organizasyonunda, yazar-okuyucu müzakerelerinde ve yazar duruşunu yansıtmada önemli işlevler üstlenirler. Önceki araştırmalar sözcük öbeklerinin önemini ortaya koymuş olsa da, akademik türlerde kullanımındaki disiplin farklılıkları hakkında çok daha az bilgi vardır. Otomatik frekans-odaklı yaklaşımın benimsendiği bu araştırmada, sosyal bilimlerdeki araştırma makalelerindeki sözcük öbekleri belirlenmeye çalışılmıştır. Uygulamalı dilbilim, pazarlama ve siyaset biliminde 2010 ve 2019 yılları arasında yazılmış araştırma makalelerinden oluşan bir derlemde, dört kelimelik sözcük öbeklerinin araştırılmasına dayanan bu çalışmada, sözcük öbeği kullanımları üzerinde disiplin farklılıklarının bir etkisi olduğunu gözlemledik. Kelime dizilerinin yapısal ve işlevsel dağılımları ile ilgili olarak, disiplinler arasında farklılıklar gözlemledik. Bu durum akademik toplulukların metin inşasında belirleyici bir role sahip olabileceğini ve sosyal bilimlerdeki disiplinler arasında farklılıklara yol açabileceğini gösterdi. Gözlemlenen farklılıklarla beraber, disiplinler arasında sözcük öbeklerinin yapısal ve işlevsel dizileri ile ilgili bazı benzerlikler de gözlemledik. Bu durum ise sosyal bilimlerin akademide bir kapsayıcı alan olarak farklı disiplinler arasında sözcük öbeği kullanımında benzeşmelerin oluşmasında etkili olabilecek normlara sahip olduğunu göstermektedir.

**Anahtar Kelimeler:** Derlem araştırması, Sosyal bilimler, Akademik yazım, Araştırma makalesi, Sözcük öbeği

### INTRODUCTION

Being a member of an academic community in today's academic world, where English is the lingua franca, means being a persuasive academic voice to be heard in that community. Academic writing is not simply an act of reporting the interpretation of data. There may be more than one objective interpretation of the data, and writers must draw on these various explanations using disciplinary linguistic resources, which grounds the essence of academic persuasion. As Hyland (2005a) explains, readers may not accept these interpretations, so writers need to anticipate the possible rejections of their interpretations and claims. To achieve this, they must be familiar with the persuasive conventions of their disciplines, framing interpretations and claims in ways that their readers may find convincing.

What constitutes academic persuasion in academic genres is rooted in disciplinary and culturally constructed values reflected through linguistic devices. Utilizing these devices, academic writers move from a position of a reporter, from summarizing the findings and literature to a position where they stamp their authorial self and interact with their readers to create convincing credibility in their academic communities. For Dontcheva-Navratilova (2018), academic persuasion is intentional and interactive, including the strategic employment of linguistic resources to convince readers, which requires the presuppositions of possible objections by the readers. Mur-Duenas (2018) draws our attention to the interpersonal aspects of academic writing. Academic writers use rhetorical resources strategically to meet the expectations of their academic communities. To be successful and credible writers, they must adapt to their disciplines' prevailing linguistic conventions.

However, disciplinary and culturally constructed academic linguistic conventions and practices are problematic notions for academic writers. As claimed by Zamel (1998), academic discourse possesses its distinguishing features “because it appears to require a kind of language with its own vocabulary, norms, sets of conventions, and modes of inquiry, academic discourse has come to characterize a separate culture [...]” (p.187). Hence, the deployment of linguistic devices to maintain persuasion in academic writing has been the focus of English for Academic Purposes (EAP) research. Among various studies aiming to explore text-construction processes in academic writing, e.g. stance devices (Biber, 2004, 2006; Kockelman, 2004), metadiscourse (Hyland & Tse, 2004; Adel, 2006), and anticipatory it (Rodman, 1991; Hewings & Hewings, 2002), lexical bundle usage has gained popularity in recent research as such strings fulfill various functions in the text.

## 1. LITERATURE REVIEW

‘Lexical phrases’, ‘lexical bundles’, ‘formulas’, ‘routines’, ‘fixed expressions’, and ‘pre-fabricated patterns’ (or ‘prefabs’) are umbrella terms used to refer to a wide range of multi-word sequences (Biber & Barbieri, 2007). In the present study, we choose to call them lexical bundles (LBs, hereafter). At the heart of the concept is the view that these fixed expressions are “important building blocks of discourse in spoken and written registers” (Biber & Barbieri, 2007, p. 263). Hyland & Jiang (2018) concentrate on LBs regarding pragmatic values in academic discourse. These recurrent expressions assist readers with particular propositional content so they can comprehend the texts in a short time. They also possess disciplinary-based norms that maintain readers’ engagement and mitigation of the author stance.

LBs are defined as “recurrent expressions, regardless of their idiomaticity, and regardless of their structural status” (Biber et al., 1999, p. 990). In this definition, idiomaticity and structural status must be explained to comprehend the concept better. LBs do not convey idiomatic meaning. Many long idioms like “kick the bucket” (meaning to die) cannot be labeled as bundles. Idioms and bundles are frequently used in fiction rather than actual face-to-face conversation. LBs are not considered to be complete structural units. Instead, they link two structural units. Although they begin at clause, the last words of them are actually the first components of a second structural unit.

Hyland (2008a) defines them as structural units of words to establish coherence in a text. LBs are sequences of words that are not combined by chance. On the contrary, they are distinctive features in a register that writers and readers must recognize as a key to engaging in a particular academic community. Naturalness reflects fluency in the use of LBs in academic communities, and the absence of such clusters may indicate the lack of competence of a novice in the community. In this sense, sensitivity to expert academics’ choices of LBs is necessary for gaining control of academic communities. Similarly, Cortes (2004) emphasizes the importance of these fixed expressions in fluent linguistic production in a specific register. Expert writers lean more on the use of fixed expressions; hence learning conventions of register use includes learning the employment of certain fixed expressions.

Cortes (2004) explains that identifying LBs builds upon two main approaches. In the first approach, groups of expressions that can be considered familiar by native speakers of the language are selected prior to the study (Nattinger & DeCarrico, 1992, Csomay & Cortes, 2010), while the second one utilizes a search tool to identify occurrences of LBs at a different length, at different cut-off frequency points. For Biber (2006), the frequency cut-off criteria used to identify LBs can be made clear depending on the number of words in the data. A multi-word sequence must occur in at least five different texts to be recognized as an LB. Hyland (2008a) states that the minimal cut-off set for recurrent sets of words to be labeled as LBs is at least ten times per million words. The use of 4-word bundles is more common in research (over ten times more frequent) since they present a wider range of structures (Cortes, 2004).

In their comprehensive study of English grammar, Biber et al. (1999) examined the most frequent "lexical bundles" in academic prose and conversation regarding structural and grammatical categorization. They observed that most of the bundles in conversation were clausal of the type (pronoun) + verb + (complement). In contrast, 60% of the bundles are phrasal, parts of noun phrases, or prepositional phrases in academic prose. In 2004, Cortes proposed a structural categorization of LBs: noun phrase with 'of' phrase fragment, noun-phrase with post-nominal clause fragment, prepositional phrase with embedded 'of' phrase, other prepositional fragments, verb (be) + complement (noun phrase), other expressions. In the same year, Biber et al. (2004) suggested a more detailed structural categorization of LBs: LBs that incorporate verb phrase fragments, LBs that incorporate dependent clause fragments, LBs that incorporate noun phrases, and prepositional phrase fragments.

Following Cortes (2004), who grouped the functions of LBs into two categories: referential bundles and text organizers, Biber et al. (2004) proposed three categories: stance expressions, discourse organizers, and referential expressions. Stance bundles indicate writers' attitudes or evaluations of propositional certainty, and discourse organizers establish the relationship between previous and upcoming content. In contrast, referential bundles directly refer to physical or abstract entities or textual contexts. Drawing on the analysis of research articles and MA and Ph.D. theses, Hyland (2008a) developed a functional category of LBs: research, text, and 'participant-oriented'. 'Research-oriented' bundles "help writers to structure their activities and experiences of the real world" (p. 49). 'Text-oriented' clusters are associated with the organization of the text, and 'participant-oriented' bundles concentrate on writer-reader interaction in texts.

The first serious discussions and analyses of LBs emerged at the beginning of the 20th century. Cortes (2004) examined LBs in terms of their structures and functions in published and student-disciplinary writing in history and biology. After identifying the most frequent bundles in journals in the two disciplines, the study attempted to determine whether these bundles were used in students' writings. The results showed that students did not employ the bundles preferred by professional academic writers. In the same year, Biber et al. (2004) analyzed the deployment of LBs in two registers: textbooks and classroom teaching and observed the important role of LBs in the construction of stance, discourse organization, and referential status. In a corpus of written and spoken university discourse, Biber & Barbieri (2007) found that the use of LBs is grounded in writers' or speakers' communicative purposes. Stance bundles constituted a large proportion of

LBs in spoken university registers, while referential functions were more frequent in written registers.

In the following years, Cortes conducted many studies on LBs. In 2006, they focused on teaching LBs in history classes and found an increased awareness of students toward these fixed expressions. In their comparative analysis of LBs in history research articles in English and Spanish in 2008, they reported common employment of these structures in both academic communities. In 2013, they examined the relationship between LBs and the moves in research article introductions and observed typical LBs in specific moves. Shin et al. (2018) examined the definite article use in LBs in L2 writing. The distribution of articles was limited by the adjoining noun within the noun phrase.

Several systematic studies have also been undertaken in the Turkish academic community. Nesi & Basturkmen (2006) examined LBs in a corpus of 160 university lectures. The corpus included lectures from the Corpus of British Academic Spoken English (BASE) and the Michigan Corpus of Academic Spoken Corpus (MICASE). LBs are a means of discourse signaling in lectures. Muşlu (2018) provided an in-depth analysis of LBs in argumentative essays and reported a heavy reliance on LBs by non-native writers of English. Still, native English writers did not frequently prefer using such bundles in their essays. Similarly, Karabacak & Qin (2013) found less frequent employment of LBs by American students compared to Turkish and Chinese students. In another study, Güngör & Uysal (2016) examined the deployment of LBs in research articles and mentioned that Turkish academics used more LBs than L1 English academics.

As highlighted above, researchers have shown an increasing interest in the use of lexical bundles from different perspectives, forms, structures, and functions (Cortes, 2004; Biber et al., 2004; Hyland, 2008a); diachronic analysis (Hyland & Jiang, 2018); and variations in cultural communities and disciplines (Hyland, 2008b; Lu & Deng, 2019; Muşlu, 2018; Karabacak & Qin, 2013; Dontcheva-Navrotilova, 2012). Several attempts have also been made to identify genre variations in argumentative essays (Karabacak & Qin, 2013); BA theses (Dontcheva-Navratilova, 2012); acknowledgment in research articles (Demirel & Hesamoddin, 2013); literature reviews (Wright, 2019); and Ph.D. theses in L1 and L2 English context (Yakut, Yuvayapan, & Bada, 2021), research articles (Cortes, 2004; Güngör & Uysal, 2016), and research articles and postgraduate genres (Hyland, 2008a).

Nevertheless, these studies point to variations in different academic genres and L1 and L2 English communities; how far they vary among disciplines remains unclear (Hyland, 2008b). It is now well established that academic writing is a socially-constructed process in which writers follow the linguistic norms of their disciplines. Thus, the awareness of these norms is a prerequisite to writing in academic disciplines. As one of the most predominant academic genres, research articles are supposed to be based on cultural and disciplinary linguistic norms, for writers tend to present their arguments incorporated into the disciplinary consensus to maintain persuasion. Hence, a systematic understanding of how these norms contribute to the construction of persuasion in academic genres is essential.

As one of the most common academic genres, research articles also require an awareness of these disciplinary-based norms, one of which is lexical bundles. Drawing on the previous literature, the present study aims to compare the deployment of lexical bundles in research articles in three disciplines in the social sciences (applied linguistics, marketing, and political sciences) to reveal disciplinary variations in the use of lexical bundles. Since English is the medium of instruction in the global academic world, this study explores the employment of these structures in English academic discourse. It focuses on 4-word lexical bundles and addresses the following research questions.

1. Are there any statistical differences in the LB usage across applied linguistics, marketing, and political sciences?
2. What are the structures and functions of LBs emerging in the research articles in the disciplines of applied linguistics, marketing, and political sciences?

## 2. METHODS

The study adopts an automated frequency-driven approach to explore and compare the use of 4-word lexical bundles in the English language with respect to their forms and functions. The data used in this study was originally compiled by Yakut, Genç, and Bada (2021) for corpus-based research related to pronoun usage. In this study, we used a part of the corpus to analyze the sub-corpora of applied linguistics, marketing, and political sciences for LB usage across the three disciplines. The corpus for this study consisted of 60 research articles gathered from three different academic journals, all indexed in SSCI according to Web of Science's 2017 impact factor. In each sub-corpus, there were 20 articles written between 2010 and 2019. The Applied Linguistics Corpus (ALC) consisted of 210.159 words, the Marketing Corpus (MC) included 241.315 words, and the Political Sciences Corpus (PSC) was composed of 224.635 words totaling 676.109 words.

In this study, we identified and explored the use, forms, and functions of 4-word LBs emerging in research articles written in the English language as Hyland (2008b) suggests that 4-word LBs are more prevalent in English compared to 5-word LBs and their forms and functions are more apparent than 3-word bundles. In addition, 4-word LBs in English seem to offer more obvious phrasal forms and functions, constituting a more meaningful unit that can be explored in a corpus.

As all the 4-word LBs do not necessarily represent the general inclination of the academic writers regarding LB usage in a specific discipline, we narrowed our LB search according to the criteria of frequency and range to exclude LBs formed by individual authors (Biber & Barbieri, 2007). As there is not a pre-defined criterion regarding the frequency and range thresholds in LB identification, we took a conservative approach following Hyland's suggestions (Hyland, 2008b, 2012). Thus, we decided to include LBs that emerge at least 20 times in every one million words and LBs that were used in 20% of the texts. Setting a frequency and range cut-off helped us to focus on recurring word strings forming a typical usage in a specific field (Pan et al., 2016).

Using AntConc version 4.1.0 (Anthony, 2022), we identified 4-word LBs used in each sub-corpus by setting the frequency and range cut-off criteria mentioned above. The LBs that are part

of a longer word string, e.g., a 5-word LB, were removed from the list to refrain from inflating the quantitative results. After getting the LB lists ready, we divided the bundles depending on their structural forms using the taxonomy suggested by Yakut, Yuvayapan, and Bada (2021), who revised and added new categories to the taxonomy developed by Biber et al. (1999) and updated by Hyland & Jiang (2018). Taking part of speech tagging as the basis of structural tagging, we divided LBs into four main categories, i.e., noun phrase (NP), prepositional phrase (PP), verb phrase (VP), and clause-related (CR), each of which includes several sub-categories (for more discussion see Yakut, Yuvayapan, & Bada, 2021).

In the second part of the annotation, we identified the functions of LBs using the functional categories suggested by Hyland (2008a, 2008b). Taking the possibility of LB's having multiple functions into account, we determined the primary function of each LB through an integrated reading of the data, including both vertical and horizontal reading. While vertical reading provided us to decide on the frequencies of occurrences of LBs, horizontal reading helped us find out the primary function of the LBs. The two researchers conducted structural tagging and functional annotation processes to avoid possible rater bias.

Upon completing the data tagging and annotation, the results in terms of overall, structural, and functional distributions of LB usages were presented in tabular forms using descriptive statistics. We also normalized the raw frequencies per 1000 words. As this research is a corpus-based study, we compared LB usage across the disciplines using the Log-Likelihood statistics to reveal whether disciplinary variation impacts LB usage. To better understand LB usage in the research articles, we tried to strengthen our claims using examples extracted from the corpus.

### 3. FINDINGS AND DISCUSSION

The main interest of this study was to explore and compare research articles in applied linguistics, marketing, and political sciences with regard to 4-word LB usages. Table 1 illustrates the overall distribution of 4-word LBs across the three corpora.

**Table 1.** The overall distribution of 4-word LBs in the three disciplines

	<b>ALC</b>	<b>MC</b>	<b>PSC</b>
<b>Corpus size in words</b>	210159	241315	224635
<b>LB type</b>	90	73	115
<b>N/1000</b>	0,43	0,30	0,51
<b>LB token</b>	754	685	1136
<b>N/1000</b>	3,59	2,84	5,1

As shown in Table 1, 90 LB types in ALC, 73 in MC, and 115 in PSC emerged. The normalized frequency of the emergences per 1000 words indicated that the authors in the political sciences used more LB types than the authors of applied linguistics, which is in line with Hyland (2008b). In MC, LB types were not found frequently as in the other two disciplines. According to the LL test results, the difference between PSC and ALC was not statistically significant (LL=+1,62), yet a statistically significant difference between PSC and MC was observed (LL=+12,71). In addition,

the statistical comparison of ALC and MC yielded a significant difference between the two groups despite observing a relatively small LL ratio (LL=+4,90). Hence, the results with respect to the comparison of LB types in the three corpora revealed that there might be a relationship between LB type emergence and the disciplines we explored in the current study.

The comparisons of LB tokens across the three corpora showed that 754 LB tokens emerged in ALC, 685 in MC, and 1136 in PSC. Similar to the LB type density, we found that the authors of political sciences employed more LB tokens in their research articles compared to the authors of applied linguistics and marketing. The LL test results also supported the descriptive statistics as we observed statistical differences across the disciplines. To illustrate, the LL difference between PSC and ALC was +54,39; it was +147,51 between PS and MC and +19,72 between ALC and MC. The results concerning LB token comparisons showed that academic writers of English in each discipline benefited from 4-word LBs while constructing their research articles with different rates, which suggests that LB usage is an indispensable part of text construction in academic writing. However, the discipline directly impacts the LB usage rates in terms of type and token.

Table 2 below illustrates the structural categorization of 4-word LBs in each discipline regarding their type and token dispersions.

**Table 2.** The structural distribution of 4-word LBs in the three disciplines

Major St.	Sub-St.	Type						Token					
		ALC		MC		PSC		ALC		MC		PSC	
		N	%	N	%	N	%	N	%	N	%	N	%
Noun-phrase	with embedded of-phrase	24	26,7	18	24,7	29	25,2	179	23,7	192	28,0	248	21,8
	other noun phrase	4	4,4	12	16,4	9	7,8	27	3,6	89	13,0	83	7,3
	<b>Sub-total</b>	<b>28</b>	<b>31,1</b>	<b>30</b>	<b>41,1</b>	<b>38</b>	<b>33,0</b>	<b>206</b>	<b>27,3</b>	<b>281</b>	<b>41,0</b>	<b>331</b>	<b>29,1</b>
Prepositional-phrase	with embedded of-phrase	22	24,4	22	30,1	37	32,2	236	31,3	232	33,9	429	37,8
	other prepositional phrase	15	16,7	7	9,6	17	14,8	149	19,8	41	6,0	209	18,4
	comparative expressions	2	2,2	1	1,4	1	0,9	22	2,9	13	1,9	10	0,9
	<b>Sub-total</b>	<b>39</b>	<b>43,3</b>	<b>30</b>	<b>41,1</b>	<b>55</b>	<b>47,8</b>	<b>407</b>	<b>54,0</b>	<b>286</b>	<b>41,8</b>	<b>648</b>	<b>57,0</b>
Verb-phrase	Copula be + NP/ADJP	2	2,2	4	5,5	5	4,3	10	1,3	58	8,5	36	3,2
	imperative + VP	1	1,1					6	0,8				
	Verb + to clause fragment	5	5,6	2	2,7	2	1,7	27	3,6	12	1,8	14	1,2
	with passive verb	3	3,3	2	2,7	4	3,5	20	2,7	12	1,8	24	2,1
	<b>Sub-total</b>	<b>11</b>	<b>12,2</b>	<b>8</b>	<b>11,0</b>	<b>11</b>	<b>9,6</b>	<b>63</b>	<b>8,4</b>	<b>82</b>	<b>12,0</b>	<b>74</b>	<b>6,5</b>
Clause-related	abstract subject			1	1,4					9	1,3		
	Anticipatory it	8	8,9	2	2,7	4	3,5	58	7,7	12	1,8	34	3,0
	as + fragments	1	1,1	1	1,4	1	0,9	5	0,7	8	1,2	8	0,7
	human subject					1	0,9					10	0,9
	if + fragments					1	0,9					5	0,4
	that + fragments	3	3,3			1	0,9	15	2,0			9	0,8
	there+fragments					2	1,7					12	1,1
	wh + fragments			1	1,4	1	0,9			7	1,0	5	0,4
<b>Sub-total</b>	<b>12</b>	<b>13,3</b>	<b>5</b>	<b>6,8</b>	<b>11</b>	<b>9,6</b>	<b>78</b>	<b>10,3</b>	<b>36</b>	<b>5,3</b>	<b>83</b>	<b>7,3</b>	

According to the descriptive statistics, prepositional phrase (PP) utilization in LB construction regarding the type distribution constituted 43,3% of all 4-word LBs in ALC, PP-based LBs were followed by noun phrase (NP) (31,1%), clause-related (CR) (13,3%) and verb phrase (VP) (12,2%) related bundles, respectively. This result is in agreement with those obtained by Hyland (2008b), who observed the frequent use of PP-based LBs in the social sciences, which can be tied to the discursive nature of these disciplines. Similarly, Biber et al. (2004) reported the frequent deployment of PP-based LBs in academic prose. In the social sciences, there may be more than one explanation of a result, so writers tend to identify these explanations and their relationships in the text.

In MC, however, NP and PP-based LBs constituted 82,2% of LB occurrences, each representing different forms in the data accounting for 41,1% of the type-based distribution of LBs. NP and PP-based LBs were followed by VP (11,0%) and CR (6,8%) successively. The academic authors of PS favored PP in their formulaic expressions the most, as they constituted 47,8% of all LB types, and NP (33,00%) followed PP. VP and CR were typified by the same amount of types (each of which was represented by 11 types), and each structure constituted 9,6% of LB types. As seen in the table, NP and PP-related LB type ratios constituted 74,4% of LBs in ALC, 82,2% in MC, and 80,8% which revealed the great reliance on NP and PP-related formulaic language in the disciplines explored in the study.

Regarding the token-related distributions of LB structures, which reveals how often the formulaic expressions were used in text construction, 54,0% of the LBs, constituting more than half of the total emergences, in ALC were PP-related LBs. PP-related bundles were followed by NP (27,3%), CR (10,3%) and VP (8,4%) bundles in the data. LB token distribution in PSC indicated similarities with the occurrences in ALC. PP-related LBs in PSC constituted 57,0% of all LBs, while NP-related LBs formed 29,1% of all occurrences, and these two LB categories were followed by CR (7,3%) and VP (6,5%). In MC, however, PP-related LBs constituted 41,8% of the total LB emergences, followed by NP-related LBs with 41,0%. VP and CR-related LBs emerged quite infrequently in MC compared to PP and NP-related LBs, constituting 17,3% of all LB occurrences. According to the LB token distributions across the disciplines, we can see that LBs in ALC and PSC are in the same order, while the preferences in MC with regard to VP and CR-related LB sequences were in different order compared to ALC and PSC distributions. In addition, the distinctive use of PP-related NPs in ALC and PSC is observable, while both NP and PP-related LBs were used at the same rates in MC. While CR was the third most common LB type in ALC and PSC, VP was the third most frequented LB type in MM.

Concerning the sub-categories of the structures, bundles with PP structure were mainly construed through 'with embedded of-phrase' in all the disciplines in line with Hyland (2008b), who explains that this structure reflects logical relations between the propositions. Within PP structure, 'with embedded of-phrase' were represented by 22 (56,41%) different types in ALC, 22 (73,33%) in MC, and 37 (67,27%) in PSC, suggesting that in marketing, the writers gave more prominence to create convincing links between the propositional content. In addition to type-related ratios, token distributions of PP sub-structures in each discipline supported the prominent 'with embedded of-phrase' usage as they constituted 57,99% in ALC, 81,12% in MC, and 66,20%

in PSC of all PP bundles in academic research articles. The excerpts below exemplify the most frequently employed PP-related LB ‘with embedded of-phrase’ in each discipline showing how the writers specify the context.

- (1) **At the end of** the course, students completed a feedback questionnaire, which asked them to assess the utility of individual AntConc tools. (ALC-2018)
- (2) **In the context of** the hospitality sector, Davis and Stone (1985) divide the service encounter into two elements: direct and indirect services. (MC-2010)
- (3) In fact, the United States Supreme Court struck down the Civil Rights Act of 1875 that prohibited certain non-state actors from discriminating **on the basis of** race in Civil Rights Cases (1883). (PSC-2014)

‘Other prepositional phrase’ structure was the second most common PP type in all academic disciplines, which is also supported by the token ratios. ‘Comparative expressions’ were the least frequented sub-category of PP-related bundles in the three disciplines. Excerpts between (4) and (6) illustrate ‘other prepositional phrase’ usage in each discipline, while (7) shows how the academic writers employed ‘comparative expressions’ in their research articles.

- (4) More specifically, the metaphors in the election campaign spots attest to a common metaphorical conceptualization based on the JOURNEY schema, which is motivated by basic metaphors like ACTION IS MOTION and GOALS ARE DESTINATIONS. **At the same time**, this schema allows for conceptual elements particularly relevant to political discourse [...]. (ALC-2016)
- (5) Positioning consumers as vulnerable has the potential to restrict agency through ignoring poorer consumers’ lived experiences, which may include happiness and fulfillment. **On the other hand**, some vulnerable consumers do not fit society’s views of what it means to be vulnerable [...]. (MC-2014)
- (6) Kant’s principle of positive law is characterized by its relative autonomy **with respect to the** moral law (PSC-2011)
- (7) To contribute to this emerging literature base, we use the TRI and consumer engagement frameworks **as well as the** concepts of trust and risk, which we review next. (MC-2019)

The second most common LB structure in each discipline was NPs. NPs constructed ‘with embedded of-phrase’ were the most prominent sub-structure in each discipline. With respect to their type distributions, we observed 24 (85,71%) individual types of NPs ‘with embedded of-phrase’ in ALC, 18 (60,00%) in MC, and 29 (76,32%) in PSC. According to descriptive results, authors of political sciences benefited from more varied NP types ‘with embedded of-phrase’ compared to the authors of applied linguistics and marketing. Regarding token, we observed that this sub-category emerged 179 times in ALC (86,89%), 192 in MC (68,33%), and 248 in PSC (74,92%).

In (8) and (10), the authors used the LBs with the function of describing the content at hand, while the author in (9) used the bundle to explain the quantity. The LBs in (8) and (9) emerged in all the disciplines we explored. Yet, *'the state of nature'*, which refers to the actual or hypothetical state of citizens without political association, emerged only in PSC.

- (8) This may be attributed to ***the content of the*** items that were designed to evaluate learners' metacognitive control over their learning-to-write process rather than a specific writing task or a genre. (ALC-2017)
- (9) Furthermore, we investigate ***the extent to which*** certain work-context factors moderate the values-to-CO relationship. (MC-2013)
- (10) While unfamiliar as an interpretation of the relationship between ruler and ruled in Hobbes's theory, the concept of trust has figured in recent years in game-theoretic analyses of ***the state of nature***. (PSC-2013)

Concerning the sub-categories of VP-related LBs, we observed fewer types and tokens in each discipline. In (11), we observed 'verb + to clause fragment' usage employed by an author of applied linguistics to specify why ambiguous verbal utterances were used. In (12), *are more likely to* is an example of 'copula be + NP/ADJP', and it was employed to illustrate the author's stance, while *can be seen as* – a 4-word LB with the function of engagement – in excerpt (13) is an example of 'VP with a passive verb'.

- (11) One could expect that participants who used ambiguous verbal utterances ***to refer to the*** PTS might have tried to resolve this ambiguity in the gestural mode. (ALC-2017)
- (12) [...], whereas consumers with lower levels of involvement ***are more likely to*** prefer its present-based benefits. (MC-2013)
- (13) NK's approach ***can be seen as*** building capabilities in order to claim basic human rights. (PSC-2018)

As for the clause-related LBs, we observed infrequent usage of them in the data. In (14), *as one of the* which is a type of 'as + fragment' was used with the function of 'quantification'. In the excerpt (15), *I would like to*, the LB, which was employed only by political sciences authors, was used as a text-structuring device. As 'human subject' LBs were employed only by political sciences authors, we might assume that such usages can be peculiar to specific disciplines in academic writing. *As if it were* in (16) exemplifies 'if + fragments', *there would be no* in (17) is an example of 'there + fragments', *this is consistent with* in (18) is an example of 'abstract subject' and *that there is a* in (21) is a sample of 'that + fragments'. While the LBs emerging in (16), (17), (18), and (21) are formed using different structures, they all indicate the author's stance. In (19), *when it comes to* is a 'wh + fragment' with the function of framing the text, while *it might be argued* in (20), which is an engagement marker, illustrates 'anticipatory it' usage in academic writing.

- (14) At the judicial stage, courts may serve ***as one of the*** institutional devices to protect the higher law of the constitution against encroachments by the ordinary law of legislation. (PSC-2014)

- (15) Instead, *I would like to* briefly recall some theses of Schmitt’s Political Theology [...]. (PSC-2012)
- (16) The problem with these liberal responses to BLM is that they respond to the demand of the social movement *as if it were* a matter of liberal principles of equality. (PSC-2016)
- (17) Without insurance, *there would be no* air traffic, freighters would not sail, skyscrapers would not be built, the production of electricity would be only very small scale, and surgeons would not operate. (2017-PSC)
- (18) *This is consistent with* the iterative nature of service and our argument that resource integration results in resource modification that influences the value potential of an actor’s resources (see below). (MC-2012)
- (19) Appearance is a critical aspect of a campaign, and voters often make their decisions based on which candidate they like (De Landtsheer et al., 2008) particularly *when it comes to* female politicians (Carlin & Winfrey, 2009). (MC-2015)
- (20) *It might be argued* that none of the cases discussed in this section challenges the study of Information Structure. (ALC-2018)
- (21) They show *that there is a* wide range of dispersion in each group, which may be explained by the fact that there are different thesis formats, depending on the complex nature of the writer’s research topics and the varied objects that have been studied. (ALC-2014)

Table 3 displays the functional distribution of 4-word LBs by showing the type, the token, and their percentages in each discipline.

**Table 3.** The functional distribution of 4-word LBs in the three disciplines

		Type						Token					
		ALC		MC		PSC		ALC		MC		PSC	
		N	%	N	%	N	%	N	%	N	%	N	%
Research-oriented	Description	7	7,78	10	13,70	34	29,57	67	8,89	103	15,04	318	27,99
	Location	5	5,56	2	2,74	5	4,35	68	9,02	16	2,34	34	2,99
	Procedure	12	13,33	8	10,96	3	2,61	91	12,07	50	7,30	28	2,46
	Quantification	5	5,56	6	8,22	3	2,61	33	4,38	51	7,45	29	2,55
	Topic	2	2,22	2	2,74		0,00	12	1,59	46	6,72		0,00
	<b>Sub-total</b>	<b>31</b>	<b>34,44</b>	<b>28</b>	<b>38,36</b>	<b>45</b>	<b>39,13</b>	<b>271</b>	<b>35,94</b>	<b>266</b>	<b>38,83</b>	<b>409</b>	<b>36,00</b>
Text-oriented	Framing	23	25,56	19	26,03	32	27,83	202	26,79	210	30,66	391	34,42
	Resultative	5	5,56	5	6,85	1	0,87	40	5,31	37	5,40	12	1,06
	Structuring	4	4,44	3	4,11	7	6,09	37	4,91	22	3,21	56	4,93
	Transition	8	8,89	6	8,22	8	6,96	84	11,14	48	7,01	104	9,15
	<b>Sub-total</b>	<b>40</b>	<b>44,44</b>	<b>33</b>	<b>45,21</b>	<b>48</b>	<b>41,74</b>	<b>363</b>	<b>48,14</b>	<b>317</b>	<b>46,28</b>	<b>563</b>	<b>49,56</b>
Participant-oriented	Engagement	6	6,67	1	1,37	2	1,74	41	5,44	5	0,73	11	0,97
	Stance	13	14,44	11	15,07	20	17,39	79	10,48	97	14,16	153	13,47
	<b>Sub-total</b>	<b>19</b>	<b>21,11</b>	<b>12</b>	<b>16,44</b>	<b>22</b>	<b>19,13</b>	<b>120</b>	<b>15,92</b>	<b>102</b>	<b>14,89</b>	<b>164</b>	<b>14,44</b>

Consistent with Hyland (2008b, 2012), the majority of the LBs with ‘text-oriented’ functions, dealing with discourse organization, emerged in each data. As the results suggest, 363 LBs were employed with ‘text-oriented’ features in ALC (48,14%), 317 in MC (46,28%), and 563 in PSC (49,56%), indicating that almost half of the LBs emerging in each data were used with ‘text-oriented’ features. ‘Research-oriented’ function, which can be realized by LBs that are used by academic writers to organize connections between the other texts and author’s research-related activities and experiences, was the second most prominent feature in each discipline as LBs with that function emerged 271 times in ALC (35,94%), 266 times in MC (38,83%), and 409 times in PS (36,00%) revealing that at least one-third of LBs emerged with the ‘research-oriented’ features in each discipline. LBs with ‘participant-oriented’ features, whose main concern is to show the author’s beliefs or assessments in the text or to engage readers in the text, were the least common group in each data as they constituted only up to 16,00% of the functions of the LBs observed in the three corpora. We found 120 LBs in ALC (15,92%) with ‘participant-oriented’ features, 102 in MC (14,89%), and 164 in PSC (14,44%).

As seen in the table, ‘text-oriented’ness was represented by four specific functions. Among the four sub-functions of ‘text-oriented’ LBs, the ‘framing’ function was the most prominent one in all disciplines despite observing fluctuating numbers of occurrences across the groups. LBs with ‘framing’ function emerged 202 times in ALC, 210 in MC, and 391 in PSC, showing that 55,65% of LBs with ‘text-oriented’ features were used for framing the text in ALC, 66,25% in MC, and 69,45% in PSC. Hyland and Zou (2020, p. 32) identify the functions of frame markers: “signal text boundaries, mark elements of text structure, label stages, announce discourse goals, sequence material or shift arguments.” Hence, successful coherent texts depend on the use of a number of related text structuring signals in writing to bring readers into the texts. In (22), *in the case of* is used to limit the scope of the main argument by connecting it to a thesis which can be accepted as a narrower variable within the main argument.

- (22) The performance standards are usually listed as assessment criteria or, ***in the case of*** a thesis, usually classified as guidelines for examiners. (ALC-2011)

‘Transition’ was the second most frequently preferred ‘text-oriented’ sub-function in each discipline. The authors of applied linguistics used 84, marketing authors employed 48, and political sciences authors used 104 LBs for transition purposes. According to the descriptive results, 23,14% of the ‘text-oriented’ LBs in ALC, 15,14% in MC, and 18,47 in PSC were used to signal transitions in the text. Apparently, the clear indication of logical links of the propositional content enables the writers to create textual cohesion. The frequent employment of transitions in the Social Sciences may be explained by the tendency of “the careful crafting of a coherent and persuasive discourse” (Hyland, 2004, p.147). The use of *in addition to the* in (23) is an example of LB with ‘transition’ function as the bundle at hand is used to add extra information within the text.

- (23) ***In addition to the*** ability to group heterogeneous phenomena, another aspect of presenting the world to the reader that merits attention is the way in which the time implied by catastrophes is rendered something that is manageable. (PSC-2017)

While ‘resultative’ function was in the third place in ALC and MC among ‘text-oriented’ LBs, ‘structuring’ function was the third most prevalent category in PSC. A prominent function of ‘resultative’ is to encode causal relations in claims and conclusions in the context, whereas ‘structuring’ helps writers announce the goals of the discourse. In (24), the author used *as a result of* to achieve coherence in the text by indicating that the sentence following the bundle will be the result of *the outburst* that the author mentioned in the previous sentences in their article. In (25), *in the next section* is a structuring bundle that draws readers’ attention to a specific part of the text.

(24) ***As a result of*** this outburst, Penn is confined to the bail-dock. (PSC-2013)

(25) As we will see ***in the next section***, features of discourse semantics become relevant in such cases, and can tilt the readings in different directions. (ALC-2019)

Concerning the sub-categories of ‘research-oriented’ function, the authors of applied linguistics employed LBs to explain a ‘procedure’ 91 times, accounting for 33,58% of ‘research-oriented’ LB usage. The procedure function was followed by ‘location’ and ‘description’ functions as they emerged 68 and 67 times in ALC, accounting for 25,09% and 24,72% of ‘research-oriented’ LBs. LBs with ‘quantification’ and ‘topic’ functions were used rather infrequently in ALC, as they emerged 33 and 12 times, constituting 12,18% and 4,43% of ‘research-oriented’ LB usage in the data. It seems that the writers in the discipline of applied linguistics tended to convey information about the processes and procedures to lead readers to particular evaluations.

The authors of marketing preferred ‘description’ function 103 times, accounting for 38,72% of ‘research-oriented’ LBs. ‘Quantification’ was in second place in MC as it appeared 51 times, totaling approximately 19,00% of ‘research-oriented’ LBs. Procedure and ‘topic’ functions were the third and fourth most common ‘research-oriented’ functions, respectively. The procedure function emerged 50 times in the data while the ‘topic’ function occurred 46 times, constituting 18,80% and 17,29% of ‘research-oriented’ LBs successively. ‘Location’ function was the least preferred function as that function was used 16 times by marketing discipline authors covering only six percent of the category. The ‘description’ function shows how the writers specify the aspects of the research, which also conveys the grounded basis of the research.

The authors of political sciences, however, did not resort to all sub-functions of the ‘research-oriented’ category. Out of 409 ‘research-oriented’ LBs, 318 were used with the function of ‘description’ accounting for 77,75% of all ‘research-oriented’ LB usages in PSC. The function of ‘location’ (8,31%) was used 34 times, and it was followed by ‘quantification’ (7,09%) and ‘procedure’ (6,85%) as they were used 29 and 28 times, respectively. As we observed an overwhelming LB usage with the ‘description’ function, the other functions within the ‘research-oriented’ category did not occupy a prominent place in the distribution of ‘research-oriented’ LBs in PSC. Contrary to the results observed in ALC and MC, we did not find any LBs with ‘topic’ function in PSC. According to the results observed within the sub-functions of ‘research-oriented’ LBs, we can mention a heterogeneous distribution of LBs in the political sciences. In contrast, the distributions in ALC and MC were relatively balanced.

In (26), the bundle at hand is used with the ‘description’ function as the statement following *in the name of* describes who is represented in the articulation of the partisans. In (27), *at the time of* is a ‘location’ marker as it indicates a specific time in the research.

- (26) Partisans articulate such commitments ***in the name of*** the whole rather than the part, even if they are aware that the views they put forward are subject to reasonable disagreement (PSC-2017)
- (27) We performed a content analysis on all briefs available ***at the time of*** the study, that is, 54 briefs from Studyka.com (552 pages) and 12 briefs from eYeKa.com (94 pages). (MC-2016)

The bundle – *in the course of* in (28) - has the ‘procedure’ feature as the following expression after the bundle explains a specific process in the text. In (29) the bundle has ‘topic’ function because the LB gives information about the field of research. ‘Quantification’ function is exemplified in (30) since the author gives information about the degree of the variety of the items following the LB.

- (28) According to Skinner, in the 1640s British republicans embraced Republican political theory, brought from the ancient to the modern world by Machiavelli ***in the course of*** his interpretation of Livy’s History of Rome. (PS-2010)
- (29) Their study discovered a consistent increase in stance features ***in academic writing in*** general with the total use of stance items increasing by approximately 50% from 1965 to 2015. (ALC-2019)
- (30) The U&G perspective has been applied to ***a wide range of*** media and communication technologies, such as video cassette recorders, cable television, [...] and mobile Internet. (MC-2017)

Bundles with the ‘participant-oriented’ function were represented by ‘engagement’ and ‘stance’. The majority of LBs with the ‘participant-oriented’ function carried ‘stance’ features in all the disciplines we explored, while LBs with ‘engagement’ features emerged rather infrequently. LBs with ‘stance’ features emerged 79 times, accounting for 65,83% of all ‘participant-oriented’ LBs in applied linguistics, whereas ‘engagement’ features were observed 41 times (34,17%). While the distribution of ‘engagement’ and ‘stance’ features were relatively balanced in the applied linguistics research articles, the usage rates of ‘stance’ features in marketing and political sciences were overwhelmingly dominant compared to the usage rates of ‘engagement’. LBs with ‘stance’ function emerged 97 times in MC and 153 times in PSC, each occupying more than 90% of all ‘participant-oriented’ LB usages. LBs with ‘engagement’ features were used only five times in MC and 11 times in PSC, accounting for 4,90% and 6,71% of all ‘participant-oriented’ LB usages in the two disciplines, respectively.

In (31), we see two strategies that provide a significant means of hedging. With the use of *could* and *argue*, the writer aimed to present an objective claim, although it was their personal evaluation. In doing so, they took a tentative stance on their claim. Example (32) indicates another

strategy of stance, explicit attitudinal stance. Here, the writer made a personal evaluation of a particular result. The use of *think* in the passive conveys an implicit strategy of pulling readers into the text and disguising the writer's presence.

- (31) While each of these four strengths interact with each other, ***it could be argued*** that engagement and agility are closely related to strategic communication decisions. (MC-2017)
- (32) However, ***it is important to*** note that some of the examples do, in fact, contain more relevant linguistic features. (ALC-2014)
- (33) But delegating monetary policy to unelected officials can also ***be thought of as*** a means of signaling a credible commitment to the price stability objective. (PSC-2019)

Clearly, 'participant-oriented' LBs were less common in the three disciplines. The writers in these disciplines seemed to be more concerned with the textual elements of the research. However, as Hyland (2005b) states, academic writing is "a persuasive endeavour involving interaction between writers and readers" (p. 173). Recall that the corpus of the present study was compiled from research articles published by noble journals indexed in Social Sciences Citation Index. Therefore, we might suppose that these writers are experts in their fields and conclude that they need to mitigate a credible representation of themselves with the use of stance LBs and claim solidarity with their readers with the employment of engagement LBs. They could only build a convincing argument by controlling the level of personality and engagement. However, they preferred to turn their attention to the textual issues of their research articles.

## CONCLUSION

Academic writing is considered one of the most prominent communication means between researchers and readers with its disciplinary and cultural-based language conventions. Even though writing does seem to be more mechanical compared to speaking, we need to highlight that academic writing does not simply cover the presentation of propositional content, which focuses on the grammatical side of the interaction; it also conveys functional and pragmatic domains that are related to interpersonal features of communication between the writer and the reader. For this reason, the language used in academic writing has been one of the primary research areas in linguistics.

Keeping the textual and functional properties of writing in mind, we explored the use, the structures, and the functions of 4-word LBs in research articles within three scientific disciplines of the social sciences: applied linguistics, marketing, and political sciences. Using descriptive and inferential statistics, we figured out whether there was a relationship between LB usage and scientific discipline within the social sciences.

The overall results revealed that LB usages were more prominent in PSC compared to ALC and MC and LB usage rates in ALC were more common in ALC compared to the occurrences in MC, indicating the possible effects of disciplinary variations on overall LB usages. Depending on the results, we can postulate that the authors of political sciences and applied linguistics chose to

shape their texts using more formulaic expressions compared to the authors of marketing. According to Hyland (2004), disciplinary conventions impact how authors construct their texts and use of language while arguing their assumptions and engaging the readers in the text. Although the English language, which has been a sine qua non in academic writing for publicizing research results in academia, was used in the articles we explored in this study, we clearly observed the paramount importance of discipline-specific conventions, yielding divergence in the use of LBs, in the process of construing research articles published in high-ranked academic journals.

Concerning the structural analysis of LBs emerging in the data, we observed that PP was the most frequented LB structure in the three disciplines. Having said that, PP-related LBs constituted more than 50% of LB usage in ALC and PC, while that structure constituted some 42% of all LB usage rates in MC. According to Hyland (2008b), PP-related bundles are employed quite frequently in the social sciences to show "logical relations between propositional elements" (p. 10). Interestingly, NP-related LBs in ALC and PSC constituted less than 30% of the structural distributions of LBs, whereas that category emerged in 41% of LB usages in MC. In addition, the usage rates of clause-related LBs were more common in ALC and PSC compared to VP-related bundles, while an opposite result was observed in MC. Despite observing differences in the sequences of LB structures across the disciplines, we found that PP and NP-related bundles constituted almost 80% of LB structures in the three disciplines. As highlighted by Hyland (2008b), the common usage of PP and NP-related bundles reveal the prominence of the social sciences knowledge on "the discursive exploration of possibilities and limiting conditions, identifying and elaborating relationships in argument" (p. 11).

Regarding the functional distributions of LBs in each data, we observed heavy reliance on 'text-oriented' and 'research-oriented' bundles in the three disciplines. This result suggests that despite publishing in noble journals in each discipline, the authors in the social sciences gave more importance to the organization of the text by disguising themselves and prioritizing the research itself. Due to focusing more on text organization, the authors used fewer 'participant-oriented' bundles in their texts, directing us to deduce that the expert authors in the social sciences do not want to establish solidarity between their readers. In addition, the results suggest that the authors prioritize their research rather than spotlighting their thoughts and claims explicitly related to the research.

The high usage rates of 'text-oriented' bundles might be related to the dynamics of the social sciences as the arguments in the social sciences are constructed through more discursive and evaluative language, resulting in more interpretation-oriented persuasion. Hyland (2004) contends that the authors in the social sciences "expect readers may think to head off objections or counterclaims and to gain a more sympathetic hearing for their own views" (p. 138). For Hyland (2008b), the assumptions in the social sciences are generally based on real-world data, yet "knowledge is typically constructed as plausible reasoning rather than as nature speaking directly through experimental findings" (p. 138). Hence, observing more 'text-oriented' and 'research-oriented' LBs is an anticipated result. However, we believe that authors of master's and doctorate theses, who are accepted as novice academic writers, might refrain from explicitly stating their

claims using stance and engagement markers as their readers, e.g., supervisors, are highly qualified scholars in their fields. Yet, the authors of high-ranked academic journals could use a more explicit language, i.e., signaling stance and engaging the readers in the text, while discussing their claims and constructing their texts published in noble journals to persuade their readers about the solidity of their claims. However, we are aware of the fact that the knowledge in the social sciences is constructed as plausible reasoning.

To put it in a nutshell, the results of this study suggest that academic communities within each discipline impact overall LB usage, their structures, and functions, which results in divergences across the disciplines. In addition to differences, we also observed similarities across the disciplines which makes us consider that the social sciences as the main field of these disciplines affects the construction of the texts yielding convergence at some points, e. g. frequent usage of PP and NP-related bundles, and prioritizing ‘text-oriented’ and ‘research-oriented’ bundle usage, in the three disciplines we explored. However, there is one important limitation to this study. The data for this study were compiled from a limited number of journals, all indexed in SSCI according to Web of Science’s 2017 impact factor. Thus, the results might not be generalized to the LB usage in research articles appearing in different journals with different publishers and different indexing services as other journal publishers’ approaches to academic text construction might differ.

As for the implications, one of the major responsibilities of EAP teachers is to help students comprehend and enhance the correct use of LBs in academic genres. The most prominent part of this process is to develop an awareness of LBs. Recognition of LBs could be done based on specialized corpora of academic genres through concordance programs. Students may search for the certain LBs used in each academic genre from different disciplines in L1 and L2 English contexts. In this way, they improve their knowledge of the use of LBs in academic genres and explore cultural and disciplinary variations. Later, they may be encouraged to employ these fixed expressions in their own academic texts.

**Conflict of Interest Statement:** The authors declared no potential conflict of interest regarding the research, authorship, and publication of this article.

**Support Financing Information:** The authors have received no financial support for the research, authorship, and publication of this article.

**Ethics Committee Decision:** For this research, ethical permission dated-numbered was not required.

**Author Contribution Rate:** The contribution rates of all authors are equal.

## REFERENCES

- Ädel, A. (2006). *Metadiscourse in L1 and L2 English*. John Benjamins Publishing. <https://doi.org/10.1075/scl.24>
- Anthony, L. (2022). *AntConc* (Version 4.1.0) [Computer Software]. Tokyo, Japan: Waseda University. Available from <https://www.laurenceanthony.net/software>
- Biber, D. (2004). Historical patterns for the grammatical marking of stance: A cross-register comparison. *Journal of historical pragmatics*, 5(1), 107-136. <https://doi.org/10.1075/jhp.5.1.06bib>
- Biber, D. (2006). Stance in spoken and written university registers. *Journal of English for Academic Purposes*, 5(2), 97-116. <https://doi.org/10.1016/j.jeap.2006.05.001>

- 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., & 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
- 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. (2006). Teaching lexical bundles in the disciplines: An example from a writing intensive history class. *Linguistics and Education*, 17(4), 391-406. <https://doi.org/10.1016/j.linged.2007.02.001>
- Cortes, V. (2008). A comparative analysis of lexical bundles in academic history writing in English and Spanish. *Corpora*, 3(1), 43-57. <https://doi.org/10.3366/E1749503208000063>
- 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>
- Csomay, E., & Cortes, V. (2010). Lexical bundle distribution in university classroom talk. In S. Gries, S. Wulff, & M. Davies (Eds.), *Corpus-linguistic applications* (pp. 153-168). Brill.
- Demirel, E. T., & Hesamoddin, S. A. (2013). Lexical bundles in research article acknowledgments: A corpus comparison. *Hacettepe University Journal of Education*, 28(2), 457-468.
- Dontcheva-Navratilova, O. (2012). Lexical bundles in academic texts by non-native speakers. *Brno Studies in English*, 38(2), 37-58. <https://doi.org/10.5817/BSE2012-2-3>
- Dontcheva-Navrátilová, O. (2018). Persuasion in academic discourse. In J. Pelclova and W. Lu (Eds.), *Persuasion in public discourse: Cognitive and functional perspectives* (pp. 227-258). John Benjamins.
- Güngör, F., & Uysal, H. H. (2016). A comparative analysis of lexical bundles used by native and non-native Scholars. *English Language Teaching*, 9(6), 176-188. <https://doi.org/10.5539/elt.v9n6p176>
- Hewings, M., & Hewings, A. (2002). "It is interesting to note that...": a comparative study of anticipatory 'it' in student and published writing. *English for Specific Purposes*, 21(4), 367-383. [https://doi.org/10.1016/S0889-4906\(01\)00016-3](https://doi.org/10.1016/S0889-4906(01)00016-3)
- Hyland, K. (2004). Disciplinary interactions: Metadiscourse in L2 postgraduate writing. *Journal of Second Language Writing*, 13(2), 133-151. <https://doi.org/10.1016/j.jslw.2004.02.001>
- Hyland, K. (2005a). A convincing argument: corpus analysis and academic persuasion. In Connor, U. & Upton, T. (Eds.), *Discourse in the Professions: Perspectives from Corpus Linguistics* (pp. 87-114). Benjamins.
- Hyland, K. (2005b). Stance and engagement: A model of interaction in academic discourse. *Discourse studies*, 7(2), 173-192. <https://doi.org/10.1177/1461445605050365>
- 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). Bundles in academic discourse. *Annual review of applied linguistics*, 32, 150-169. <https://doi.org/10.1017/S0267190512000037>
- Hyland, K., & Jiang, F. K. (2018). Academic lexical bundles: How are they changing? *International Journal of Corpus Linguistics*, 23(4), 383-407. <https://doi.org/10.1075/ijcl.17080.hyl>
- Hyland, K., & Tse, P. (2004). Metadiscourse in academic writing: A reappraisal. *Applied Linguistics*, 25(2), 156-177. <https://doi.org/10.1093/applin/25.2.156>
- Hyland, K., & Zou, H. J. (2020). In the frame: Signalling structure in academic articles and blogs. *Journal of Pragmatics*, 165, 31-44. <https://doi.org/10.1016/j.pragma.2020.05.002>
- Karabacak, E., & Qin, J. (2013). Comparison of lexical bundles used by Turkish, Chinese, and American university students. *Procedia-Social and Behavioral Sciences*, 70, 622-628. <https://doi.org/10.1016/j.sbspro.2013.01.101>
- Kockelman, P. (2004). Stance and subjectivity. *Journal of Linguistic Anthropology*, 14(2), 127-150. <https://doi.org/10.1525/jlin.2004.14.2.127>

- 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>
- Mur-Duenas, P. (2018). Promotional strategies in academic writing. In J. Pelclova and W. Lu (Eds.), *Persuasion in public discourse: Cognitive and functional perspectives* (pp. 259-277). John Benjamins.
- Muşlu, M. (2018). Use of stance lexical bundles by Turkish and Japanese EFL learners and native English speakers in academic writing. *Gaziantep University Journal of Social Sciences*, 17(4), 1319-1337. <https://doi.org/10.21547/jss.444386>
- Nattinger, J. R., & DeCarrico, J. S. (1992). *Lexical phrases and language teaching*. Oxford University Press.
- Nesi, H., & Basturkmen, H. (2006). Lexical bundles and discourse signalling in academic lectures. *International Journal of Corpus Linguistics*, 11(3), 283-304. <https://doi.org/10.1075/bct.17.03nes>
- Rodman, L. (1991). Anticipatory it in scientific discourse. *Journal of Technical Writing and Communication*, 21(1), 17-27. <https://doi.org/10.2190/PWJ6-AG95-MCQD-RG1W>
- Shin, Y. K., Cortes, V., & Yoo, I. W. (2018). Using lexical bundles as a tool to analyze definite article use in L2 academic writing: An exploratory study. *Journal of Second Language Writing*, 39, 29-41. <https://doi.org/10.1016/j.jslw.2017.09.004>
- Wright, H. R. (2019). Lexical bundles in stand-alone literature reviews: Sections, frequencies, and functions. *English for Specific Purposes*, 54, 1-14. <https://doi.org/10.1016/j.esp.2018.09.001>
- Yakut, I., Genç, B., & Bada, E. (2021). Epicene pronoun usage in the social sciences: The case of research articles. *Journal of English for Academic Purposes*, 52, 101005, 1-10. <https://doi.org/10.1016/j.jeap.2021.101005>
- Yakut, I., Yuvayapan, F., & Bada, E. (2021). Lexical bundles in L1 and L2 English doctoral dissertations. *Journal of Teaching English for Specific and Academic Purposes*, 9(3), 475-493. <https://doi.org/10.22190/JTESAP2103475Y>
- Zamel, V (1998). Questioning Academic Discourse. In V. Zamel and R. Spack (Eds.), *Negotiating academic literacies: Teaching and learning across languages and cultures* (pp. 187-197). Mahwah, NJ: Erlbaum.