Research article / Научная статья

Academic English melting pot: Reconsidering the use of lexical bundles in academic writing

Elena S. GRITSENKO[1] ✉ and Olivier Mozard T. KAMOU[2]

[1]Moscow State Institute of International Relations MGIMO University, Moscow, Russia
[1,2]HSE University, Nizhny Novgorod, Russia
✉elena.s.gritsenko@gmail.com

Abstract
Numerous research studies addressing the differences in the use of lexical bundles in academic English by L1 and L2 writers interpret these differences as a deficiency or deviation that L2 writers need to eliminate. In this paper, we argue that this "deviant" use is not essentially the product of insufficient knowledge of English and/or Anglophone norms of academic writing but rather a transfer of the academic conventions of non-native speakers, rooted in their local culture. To confirm this hypothesis, we reviewed some previous studies and analyzed the use of lexical bundles in dissertations and research papers written in English by graduate and post-graduate students from Russia and Cameroon. The Russian corpus (38 texts of 576,186 words) was compiled from publicly available papers and dissertations written by bachelor's and master's students at the Higher School of Economics; the Cameroonian corpus (21 papers of 680,146 words) was compiled from papers contributed by students and teachers of the University of Yaoundé I. Using content analysis, corpus analysis, and the comparative method, we found that the most significant differences in the use of lexical bundles were connected with the peculiarities of the Russian and Cameroonian academic writing styles and cultural norms. Our study, therefore, reinforces the need to consider a more inclusive and culturally sensitive approach to the use of lexical bundles by L2 academic writers and take into account their diverse linguistic and cultural backgrounds. It will expand our knowledge of the linguistic features of different varieties of English and provide a deeper understanding of academic traditions in different languages and cultures.
Keywords: *lexical bundles, academic English, world Englishes, cross-cultural variation, translanguaging, corpus analysis*

Плавильный котел академического английского: использование лексических связок в академическом письме

Е.С. ГРИЦЕНКО[1] ✉, О.М.Т. КАМУ[2]

¹Московский государственный институт международных отношений (университет) Министерства иностранных дел Российской Федерации, Москва, Россия
1,2Национальный исследовательский университет «Высшая школа экономики», Нижний Новгород, Россия
✉elena.s.gritsenko@gmail.com

Аннотация

Большинство исследователей, анализирующих использование лексических связок в академическом письме носителями и неносителями английского языка, интерпретируют выявленные отличия как отклонения, связанные с интерференцией и (или) недостаточным знанием норм академического письма. Цель данной статьи – показать, что по крайней мере часть выявляемых отличий обусловлена не недостаточными знаниями, а влиянием коммуникативных норм родной для неносителей английского языка культуры. В статье проанализирован ряд предыдущих исследований и проведен анализ использования лексических связок в диссертационных и выпускных квалификационных работах на английском языке студентов и аспирантов из России и Камеруна. Корпус российских примеров был отобран из находящихся в открытом доступе работ магистрантов Высшей школы экономики; корпус камерунских примеров – из работ, предоставленных студентами и преподавателями университета Yaoundé I. Всего было проанализировано 59 текстов: 38 работ российских студентов (576 186 слов) и 21 работа студентов и аспирантов из Камеруна (680 146 слов). Использовались методы контент-анализа, корпусного анализа и сравнительно-сопоставительный метод. Выявлено, что наиболее значимые отличия в использовании лексических связок обусловлены особенностями академического стиля и коммуникативной культуры авторов текстов. С опорой на полученные данные обосновывается целесообразность трактовки культурно обусловленных отличий в использовании лексических связок не как недостатка, а как проявления транслингвальных компетенций пищущих. Такой подход позволит расширить представления о лингвистических особенностях различных вариантов английского языка и обеспечит более глубокое понимание академических традиций в различных языках и культурах.

Ключевые слова: *лексические связки, академический английский, контактная вариантология английского языка, кросс-культурное варьирование, транслингвизм, корпусный анализ*

1. Introduction

Writing for academic purposes requires not only mastering disciplinary knowledge and following established academic writing conventions but also making some special efforts and commitments. One of these commitments is to engage in the global dissemination of knowledge, which is now an essential part of academic success for both young scientists and experienced researchers. English is the main language through which academic knowledge is communicated to a wider audience (Flowerdew 2014). It has become the so-called lingua franca of academia, prompting researchers whose L1 is not English to strive for a somewhat "native-like" competence in English or even "worry about the correctness of their language" (Hyland & Jiang 2022: 554).

The key role of English in the global transmission of knowledge contributed to the emergence of numerous studies focusing on what constitutes acceptable and "proper" L2 academic English. These studies have mostly been dominated by Western-centric views, with a distinct emphasis on native English-speaking writers and their norms of academic English. Due to this Western-centric perspective, such studies may have failed to fully capture the diversity and complexity of L2 English writing in various cultural and linguistic contexts and, therefore, imposed a form of hegemony of L1 academic writing conventions on non-native speakers of English. As argued by McKinley (2022), the main reason for the dominance of this Western-led perspective is the reluctance to take risks and pursue diverse aspects of L2 writing and instead rely on well-founded theories and research questions, which somehow limit the research potential in L2 writing.

An aspect of academic English that has received considerable research attention and for which L2 students and researchers have been mostly criticized is the use of recurrent multi-word units commonly referred to as lexical bundles. Lexical bundles are sequences of three or more words or "extended collocations which appear more repeatedly than expected by chance" (Hyland & Jiang 2018: 383) and include such structures as *it was found that*, *in the case of*, *may be due to, as can be seen, with respect to the, it is possible that, on the other hand, etc.* These multi-word expressions are important for academic writing because students and scholars are expected to be precise and concise in communicating their ideas (Salazar 2010), and lexical bundles serve as "the building blocks of discourse" and valuable indicators of fluent linguistic production (Biber 2009, Hyland 2008a).

Undoubtedly, as previous studies have shown, lexical bundles lie at the heart of efficient academic exchanges, and adequate use of such structures is a sign of good mastery of academic English. However, most of these studies tend to portray L2 academic writers as deficient or deviant in their use of lexical bundles (e.g.,

Chen & Baker 2010), whereas they may simply comply with deeply rooted norms of their own languages and cultures, which certainly differ from those of native English speakers. If we agree that English is, for the time being, the main language of academic research, it seems appropriate to consider a more inclusive and culturally sensitive approach to the use of lexical bundles by L2 academic writers and take into account their diverse linguistic and cultural backgrounds.

In this paper, we present an argument that L2 academic writers' "deficient" use of lexical bundles does not always result from their inadequate mastery of English but may reflect the communicative conventions of local culture, which differ from Anglo-American norms. This is not to say that we must overlook good and intelligible English, but rather be more tolerant towards L2 writers' linguistic and cultural identities transferred to their academic writing in English. Since writers with different language backgrounds all bring something different to the act of academic writing, we propose to revise the existing approach to lexical bundles research by making it less prescriptive. We suggest shifting the focus from how L2 writers should write to how they actually write and what it tells us about different academic styles and cultures.

2. Overview of previous research

As important indicators of fluent linguistic production (Hyland 2008a), lexical bundles appear to significantly shape the meaning and coherence of academic texts (Richter et al. 2022), guiding writers and readers through the content and linking ideas (Hyland & Jiang 2022). Lexical bundles are identified using corpus analysis software that retrieves them in accordance with predetermined frequency and distribution criteria (Biber 2006), which may vary between ten and twenty times per million words (Biber et al. 1999, Conrad 2004) or 10% of the texts (Hyland 2008) and occur in at least three to five different texts to guard against idiosyncratic uses (Biber & Barbieri 2007). Conventionally, lexical bundles in academic discourse have been classified into structural and functional types, as proposed by Biber et al. (1999) and Hyland (2008a), respectively (see Tables 1 and 2).

Table 1. **Structural classification of lexical bundles (Biber et al. 1999: 1014–1024)**

Structure	Examples
Noun phrase + of	*the end of the, the nature of the, the beginning of the,*
Other noun phrases	*the fact that the, one of the most, the extent to which*
Prepositional phrase + of	*at the end of, as a result of, on the basis of, in the context*
Other prepositional phrases	*on the other hand, at the same time, in the present study*
the Passive + prep phrase fragment	*is shown in figure, is based on the, is defined as the*
Anticipatory it + verb/adj	*it is important to, it is possible that, it was found that*
Be + noun/adjectival phrase	*is the same as, is a matter of, is due to the, be the result of*
Others	*as shown in figure, is likely to be, as well as the*

Table 2. **Functional Classification of lexical bundles (Hyland 2008a: 13–14)**

Research-oriented:
Help writers structure their activities and experiences of the real world
Location: *at the beginning of, in the present study* **Procedure:** *the use of the, the role of the, the purpose of the, the operation of the* **Quantification:** *the magnitude of the, the wide range of, one of the most* **Description:** *the structure of the, the size of the, the surface of the* **Topic:** *the currency board system*
Text-oriented:
Concerned with the organization of the text and its meaning as a message or argument
Transition signals: *on the other hand, in addition to the, in contrast to the* **Resultative signals**: *as a result of, it was found that, these results suggest that* **Structuring signals**: *in the present study, in the next section, as shown in figure* **Framing signals**: *in the case of, with respect to the, on the basis of, in the presence of*
Participant-oriented:
Focus on the writer or reader of the text
Stance features: *are likely to be, may be due to, it is possible that* **Engagement features:** *it should be noted that, as can be seen*

Several studies employed a lexical bundle approach to explore variation across academic disciplines and genres. Biber et al. (2004), for example, compared lexical bundles in four academic genres, including conversations, university textbooks, university classroom teachings, and published research articles, and revealed systematic differences. Hyland (2008a) conducted research on lexical bundles in articles, doctoral dissertations, and master's theses, in four disciplines: biology, electrical engineering, applied linguistics, and business studies. In his 3.5-million-word corpus, he found that lexical bundles not only play an essential role in academic discourse but also vary considerably across disciplines in terms of frequency and preferred uses.

Cortes (2004) examined lexical bundles in a corpus of over two million words from two academic disciplines. She studied published articles in history and biology to identify the most frequent four-word bundles, which she referred to as "target bundles," and examined how these bundles are used by history and biology students. The study revealed that students rarely used target bundles in academic writing, and the bundles they used did not coincide with those used by professionals. The author concluded that for effective use of lexical bundles, students should immerse themselves in reading academic texts. Richter et al. (2022) explored the use of lexical bundles in the discussion sections of quantitative, qualitative, and mixed-methods research articles published in ten highly rated international journals in the field of applied linguistics. Their findings indicate that "different methodological paradigms are characterized by different functional uses of lexical bundles," which "constrain writers' language preferences" (p. 625).

Many researchers examined lexical bundles from a language development perspective, focusing on L2 academic writers or comparing L1 and L2 writers and learners (Pan et al. 2015). For example, Chen and Baker (2010) compared the use of lexical bundles by Chinese EFL learners and undergraduate and professional

native speakers of English in order to identify possible issues in second language learning. Their analysis showed that professional native speakers' writing exhibited a wider range of bundles, whereas Chinese EFL students' texts had a smaller range. In addition, lexical bundles that had high frequency in professional native speakers' texts, such as *in the context of*, were underused by Chinese professionals and undergraduates. Chen and Baker also found that Chinese students tend to overuse certain bundles (for example, *all over the world*), which are rarely used by native speakers. Adel and Erman (2012) compared lexical bundles in native and non-native writing in a series of essays by Swedish and English undergraduate students of linguistics. Their study reported that English students used a larger number of lexical bundles, which were also more diverse than those used by Swedish students. At the functional level, the authors noted that native speakers rarely used discourse-organizing or text-oriented bundles and employed more stance- or participant-oriented bundles than non-native students did. Salazar (2010) studied lexical bundles with verbs in a corpus of medical research articles. The corpus contained Philippine journal articles and articles from the British Medical Journal. Quantitative analysis showed that Philippine researchers used fewer verbal bundles. The study also uncovered other structural and functional differences between British and Filipino writers.

A large part of the literature focusing on the differences in the use of lexical bundles by L1 and L2 writers tends to overlook the reasons behind those differences. Studies that address this aspect and link the use of lexical bundles to the linguistic and cultural background of L2 writers still describe them as faulty. For instance, Wei and Lei (2011) found that Chinese students (advanced EFL learners) tended to overuse lexical bundles containing passive structures and underuse participant-oriented bundles. The researchers linked it to students' preference for impersonality in academic writing. This preference can be explained by the influence of classical Chinese (Wenyan 文言) on the rhetorical conventions of Chinese academic discourse. Although the authors indicated that the differences may be rooted in the norms of the Chinese language and culture, they interpreted them as the result of students' insufficient exposure to readings and conscious learning of target bundles. In other words, they viewed the differences as a deficiency that needed to be corrected.

A similar interpretation was offered by Alamari (2020), who used a "move-bundle approach" to compare the use of lexical bundles in research articles published in Saudi and international journals and explained the identified differences by the "cultural traits concerning the communication of knowledge" (p. 14). For instance, he argued that while "indicating the gap" in the introduction, the Saudi authors tend to "avoid direct criticism of the work of others" since they believe that criticism is "inappropriate or less acceptable" and "may engender negative attitudes from other researchers" (ibid.). Alternatively, Saudi authors would state that the subject is understudied or "provide other justifications to convince the readers" (p. 15). This "cultural trait" may explain (a) why participant-oriented bundles are much less frequent in the Introduction sections of Saudi papers

(8% of the Saudi corpus compared to 23% in the "international" corpus) and (b) why research-oriented bundles are used twice as much by Saudi authors compared to authors in international journals (56% vs. 27%). Alamari interpreted these dissimilarities as insufficient awareness or exposure of Saudi writers to international academic (i.e., Anglo-American) conventions that they should follow.

One of the few papers that directly links the use of lexical bundles to literacy traditions and intellectual style is the article by Olga Dontcheva-Navratilova (2012). The purpose of her study was to compare how Czech and German students use lexical bundles to indicate authorial presence in English academic writing and reveal to what extent they have adapted their writing style to Anglo-American academic discourse conventions. The author provides a comprehensive description of cross-cultural differences between Anglo-American and Central European academic styles (p. 9–11) and points out that "unlike in the Anglo-American tradition, the focus of Central European academic writing is on conceptual and terminological clarity rather than persuasion and discourse organization, and, thus, authorial presence tends to be backgrounded" (p. 10). Yet, her conclusion suggests that an approximation of the variety and frequency of interpersonal bundles to Anglo-American standards is necessary for non-native writers to be socialized into the global academic discourse community.

These interpretations reflect the widespread ideology of English as a monocentric language with the "native speaker" as a point of reference. Even though this approach is still influential in lexical bundles research, there appears to be a gradual shift from the idea that L2 writers' use of lexical bundles is deficient. A new approach is emerging in the World Englishes paradigm and is supported by the ideology of English as a polycentric language, which needs to be more inclusive. In line with this new approach, Hyland and Jiang (2022) have put forward the idea of academic discourse as a melting pot of Englishes, where "different varieties are constantly in contact so that the ever-increasing participation of EAL[1] authors in global publishing will, very likely, slowly enlarge the variety of bundles we see in professional texts" (p. 569).

Our goal in this paper is, therefore, to justify a more inclusive approach to lexical bundles research that will take into account the status of English as a polycentric language and the growing interest in the linguistic features of its varieties. To support our argument, we will address the use of lexical bundles in English by Russian and Cameroonian students – two settings of non-native academic writing that, to the best of our knowledge, have not been the subject of investigation. Russia has a strong academic writing tradition, whose discourse patterns are likely to be transposed into the English texts. In Cameroon, English is used as the second and official language, one of the two languages of education[2], and its use in academia is likely to be influenced by the local communication norms and habits. Therefore, both settings seem promising for exploring cross-cultural variation in lexical bundles use, which so far remains underexplored.

---

[1] EAL – English as additional language
[2] The other language is French. In Cameroon, local languages are not used in education.

3. Data and methods

Two corpora form the basis of our analysis: the Russian corpus (RusC), which includes 19 bachelor and 19 master's theses written in English, totaling 576,186 words, and the Cameroonian corpus (CamC), which includes 8 postgraduate and 13 master's dissertations[3] written in English, totaling 680,146 words (see Table 3). These academic genres were selected because they represent the main research genres in the academic context (Hyland 2008a). The Russian corpus was compiled from the website of the Higher School of Economics, where full texts are published with students' consent and are publicly available. As no such platform was available in Cameroon, for the Cameroonian corpus, we collected the texts individually from students at the University of Yaoundé I and from some lecturers who generously provided them to be used in this study. All title pages, tables of content, direct quotations, bibliographical data, and appendices were manually excluded from the texts.

Table 3. **Word and text distribution in the Russian and the Cameroonian corpora**

	Genre	No. of texts	No. of words	Total
RusC	Bachelor theses	19	291,948	576,186
	Master's thesis	19	284,948	
CamC	Post-graduate dissertations	8	159,223	680,146
	Master's dissertations	13	537,133	
Total		59	1,256,332	1,256,332

As stated earlier, there are two criteria for lexical bundle identification: the frequency of occurrence criterion and the dispersion criterion. Frequency of occurrence is the first and most important aspect of lexical bundle identification; it is also largely based on the type of corpora. As explained by Biber et al. (1999), lexical bundles are more frequent in spoken discourse than in written discourse. Consequently, the minimum frequency of occurrence is 40 times per million words in spoken discourse, whereas in written discourse, the minimum frequency varies from ten (Biber et al. 1999) to twenty (Hyland 2008a, Cortes 2004), or twenty-five times (Chen & Baker 2010) per million words. Since there seems to be no agreement on the frequency of occurrence, we decided to select lexical bundles that occurred at least 20 times per million words, following Hyland (2008a) and Cortes (2004).

The second criterion was dispersion, that is, the number of texts in which these four-word sequences had to occur to be regarded as a lexical bundle. Setting the dispersion criterion was, as Biber et al. (1999) had noted, to exclude lexical bundles that may be characteristic of individual writing styles. We decided to follow Cortes'

---

[3] In Cameroon, a dissertation is typically written as part of the requirements for a master's degree, whereas a thesis is reserved for PhD candidates. A postgraduate dissertation, however, differs slightly. It is usually completed by students pursuing additional training after earning a bachelor's degree. If they seek a master's degree, they must then write a master's dissertation.

(2004) conservative stand and included in our corpus only those lexical bundles that appeared in at least five different texts or more.

Antconc 4.0.10, a freely available corpus analysis software, was used to automatically extract all four-word bundles using the N-gram function. However, not all four-word sequences were incorporated into the list of lexical bundles; we excluded topic- and context-dependent ones and combined overlapping bundles (Navarro & Martinez 2019, Chen & Baker 2010). Since our corpus contained texts from different disciplines, it seemed necessary to exclude discipline-specific bundles that could impede the comparability of our results. Therefore, such bundles as, for instance, *of the Russian Federation* (RusC) and *the teaching and learning* (CamC) were manually excluded. Moreover, section titles such as *Review of related literature* (CamC) and overlapping bundles were also excluded. The exclusion of overlapping bundles involved joining such bundles, as *it can be concluded* and *can be concluded that*, where a sequence *can be concluded that* was actually a part of a longer bundle (*it can be concluded that*), and when checked in the concordance line, was preceded by *it* and thus inflated the results. Therefore, such bundles were combined into *it can be concluded + (that)*. The final list of bundles included all four-word sequences that appeared at least 20 times in five different texts. These bundles were then classified and analyzed according to structural and functional criteria.

4. Results

In the Russian and the Cameroonian corpora, we identified 63 and 74 (types) four-world bundles, respectively, after manual exclusion of content- and topic-dependent bundles, of which 2361 (RusC) and 2717 (CamC) were unique occurrences (tokens). However, since the two corpora were slightly different in size, we applied a normalized frequency of 100,000 times per million words to the tokens in order to obtain a comparable size. We found some noticeable differences in the kinds of lexical bundles used by Russian and Cameroonian students, even though there are some shared bundles in both sub-corpora (see Table 4).

It is noteworthy that the two most conspicuous differences in the use of lexical bundles by Russian and Cameroonian students are linked to the specifics of the local language and culture. As for the structural features, research findings indicate that Russian students used more lexical bundles with *of*-fragments than Cameroonian students did (46.03% *vs*. 33,78%). All in all, 27 such bundles were identified in RusC: 11 noun phrase + *of*-phrase fragment (e.g., *the analysis of the, a wide range of, the majority of the, the development of the, the meaning of the, the nature of the, a part of the, the results of the, (at) + the end of the, the role of the, the purpose of this*) and 18 prepositional phrases with *of* (*in the context of, with the help of, for the development of, as a part of, in the form of , in the development of , in the process of , in the course of , at the beginning of + (the) , for the sake of , in the field of , in the case of , for the purpose of , as one of the, by the end of , on the basis of + (the) , as a result of, from the perspective of*). Apparently, the higher

share of such bundles in the Russian corpus reflects a characteristic feature of Russian academic discourse, where strings of nouns in the genitive case are very frequent (Vladimirova 2010: 23). *Of*-phrases are the most common way of representing such strings in English (Dobrynina 2019).

Table 4. **The list of 40 most common lexical bundles in the Russsian and the Cameroonian corpora**

RusC	Freq.	CamC	Freq.
At the same time + (the)	*145*	**As a result of +(the)**	106
One of the most	*93*	**Is one of the**	93
As well as the	*89*	**The fact that the**	80
It is important to + (note)	*83*	**As well as the**	**79**
Is one of the	*76*	**On the other hand**	76
In the context of + (the)	*62*	**In the course of + (the)**	**73**
(due) + to the fact that	*61*	**It is important to**	63
The analysis of the	*57*	At the end of + (the)	57
It is possible to	*55*	At the level of	52
On the basis of + (the)	*54*	In the same light	50
The results of the	*54*	Is made up of	50
The fact that the	*53*	**To the fact that**	50
On the other hand	*52*	In the sense that	47
In the process of	*50*	**One of the most**	47
As one of the	*42*	**To be able to**	47
For the development of	*42*	The use of the	46
It is necessary to	*40*	In a bid to	45
In the case of	*38*	Is the fact that	43
One of the main	*38*	**At the same time**	42
(at)+the end of the	*38*	The rest of the	42
In the form of	*36*	Is based on the	40
It can be concluded + (that)	*36*	Of the fact that	38
By the fact that	*35*	The nature of the	37
A part of the	*34*	The role of the	36
It is crucial to	*34*	For the purpose of	35
With the help of	*34*	The extent to which	35
As a result of	*33*	The fact that it	35
To be able to	*31*	Through the use of	35
We can see that	*31*	To find out if	35
In the course of	*30*	**In the process of**	34
In the field of	*30*	It should be noted (that)	34
It was decided to	*30*	The fact that they	34
As a part of	*29*	**In the case of**	33
Despite the fact that	*29*	In relation to the	32
In other words**,** the	*29*	On the fact that	32
It impossible to	*29*	That there is a	32
When it comes to	*29*	To the extent that	32
In the context of **+** (the)	*28*	Should be able to	31
The majority of the	*27*	**In the form of**	29
From the perspective of	*26*	In the table below	29

Another notable difference is a higher percentage of *Anticipatory it + verb/adjective* bundles in RusC than in CamC (11, 11% vs. 5,4%) and their higher frequency: six out of seven bundles of this type are among the 40 most frequent RusC bundles (their total frequency is 277), whereas in CamC there are only four bundles of this type, and only two of them are among the 40 most frequent ones, with a total frequency of 97 (see Table 4). The low share of *Anticipatory it + verb/adjective* bundles in CamC may be connected with their functionality: they convey stance and engagement and, as such, are avoided by Cameroonian students for cultural reasons, which we will explain below.

Since in Cameroon local languages are not used in academia and all education is conducted either in English or in French, the structural analysis of lexical bundles in CamC has not revealed any specific local preferences. Yet, Cameroonian conventions of academic English (both oral and written) are influenced by the norms of local communication culture, and our study has revealed some traces of this influence. Thus, although both Russian and Cameroonian students show a heavy reliance on research-oriented bundles, which account for over 42% of the bundles in the CamC and about 34% in the RusC (see Figure 1), the share of text-oriented bundles is relatively higher in the CamC than in the RusC (36,98% and 25,39%, respectively). Unlike Russian students, Cameroonian students use fewer participant-oriented bundles, which convey the writer's attitude, judgment, belief or evaluation (*e.g., it should be noted + (that), it is clear that, it is important to, should be able to, etc.*) and more text-oriented bundles, which are concerned with text organization and its meaning as a message or argument *(e.g., on the one hand, in the sense that, at the same time, in relation to the, as far as the , in the same light, of the opinion that, when it comes to, in line with the, is based on the, in the face of , has to do with , in the domain of).* Apparently, for Cameroonian students, it is more important to focus on the actual text, facts, or results of the study than to engage with the readers and highlight the authorial presence by conveying their personal opinion, evaluation, or judgement. This can be explained by the influence of Cameroonian culture, where the norms of social interaction are always governed by deference, which constrains direct criticism or public questioning of authority, manifesting a culture-specific understanding of (im)politeness (Larina & Ponton 2022). For Cameroonian students, therefore, these cultural values mean that even in academic research, where critical thinking is usually encouraged, criticizing more advanced researchers may not only result in stained relations but also in negative evaluation of peers. As a result, Cameroonian students tend to avoid challenging prevailing views and minimize the authorial presence in the text, focusing on how to convey their findings without critical appraisal of previous studies. It differs from Russia, as in Russian academic culture, despite the tendency to diminish the authorial presence in the text (Boginskaya 2022), the critical analysis of previous research is acceptable, hence a wider range and higher percentage of participant-oriented bundles (33,33% in RusC *vs*. 19,17% in CamC).
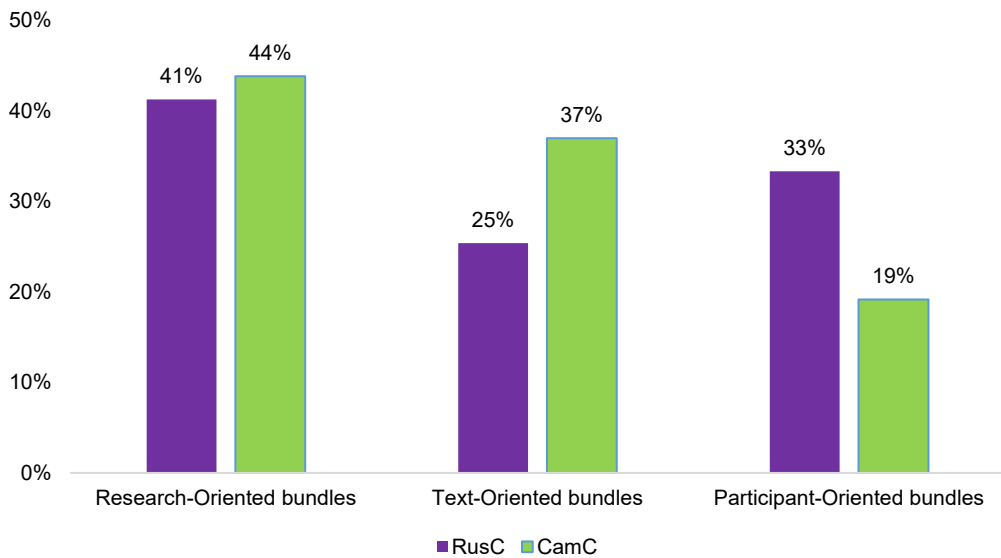
Figure 1. Functional Distribution of lexical bundles in Russian and Cameroonian corpora

5. Discussion

As indicated above, the use of lexical bundles by L2 English writers portrays some distinct traces of the local linguistic and cultural influence that significantly differ from Anglo-American usages. This influence is somewhat similar to interference from the language of the original in case of translations, "which — by their very nature — are produced under different constraints than native texts" (Grabowski, 2018: 405). The findings of our study presented in the previous section provided instances of such influence and showed that non-native learners and writers tend to accommodate their local communication norms to the use of lexical bundles. The limited use of participant-oriented bundles by Cameroonian students, that we mentioned above, aligns with broader aspects of academic writing in Cameroon, where there is a noticeable absence of engagement and interactional features. As noted by Nkemleke (2014), who analyzed the structure and citation practice in articles published by Cameroonian authors in local journals, more than 80% of research articles' introductions tend to avoid critical analysis of previous research (p. 197) and the most frequent verbs in citations are "points" and "reports", which "do not indicate writer's evaluation of the content of the proposition of knowledge claim" (p. 192).

Since there was no previous research on the use of lexical bundles in Russian and Cameroonian academic writing, to support our observations and enhance the understanding of "how and why language users make the choices they do when they speak/write" (Hyland, 2011: 106), we turned to the ethnographic study of Chinese scholars who interviewed Chinese postgraduate and graduate students "to probe possible reasons for their use of the typical sentence initial bundles identified in the self-built Chinese Masters and PhD thesis corpora" (Li et al. 2019: 37). It is

noteworthy that, according to this study, Chinese students also seem to be "highly conscious of their identity as student researchers and apprentice writers. They appeared more comfortable expressing their attitudes towards their own research rather than influencing the evaluations of their readers." (p. 46). For instance, the preference for the verb "to find" (instead of "to note") in the sentence "*It is interesting to find that there is wide networking coordination within the DMO in Zhu Jiayu and other stakeholders*" was explained by cultural appropriateness and respect for the readers':

(1) *I do not want to replace 'find' with 'note' and I am not comfortable to use 'note'. If my reader agrees with me and finds this interesting, it is interesting to them; otherwise, it is not. I am not willing to forcefully involve my readers and require them to pay attention to this point. Instead, I want to tell them that this is my finding. I think my readers should have their freedom. If they think this is an interesting finding, they will note this point. If they do not think so, then they do not share the same opinion with me* (p. 46).

The reluctance to use the word "interesting" was interpreted as striving for objectivity and the unwillingness to impose the author's opinion on the reader:

(2) *I rarely use 'interesting' in my writing. This word expresses my own feeling. Academic writing should be neutral…* (p. 47).

Similarly, avoiding the pronoun "I" was explained by the desire not to "flaunt one's self":

(3) *Strictly speaking, I should use 'the researcher' here. Academic writing should be objective and scientific, so I try not to use first person pronouns* (Ibid.).

Notwithstanding the clear ethno-cultural and ethno-linguistic roots of the above differences, the authors of the article linked them, among other things, to "a lack of rhetorical confidence" and "misunderstanding of the rhetorical conventions" of the foreign language in which the respondents were writing (i.e., English). The authors provided recommendations "for raising students' awareness of the common sentence starters in postgraduate academic writing" (p. 37), which included: "(1) raising students' awareness of the discrepancies between L1 and L2 bundles; (2) emphasizing bundle noticing in academic reading and writing; (3) increasing students' confidence as writers; (4) familiarizing Chinese students with rhetorical conventions of academic English" (ibid.).

Regardless of the pedagogical value of such recommendations, this approach raises a number of questions: Do L2 students and researchers need to have an Anglophone mentality to be effective academic writers? Is it necessary for them to fully detach themselves from their linguistic and cultural background? How can we find the golden mean between developing foreign language competence and retaining learners' cultural identity?

Some answers to these questions are provided in the works by Nabuyuki Hino (2018, 2021), who criticizes the traditional approach, which implies that non-native speakers should conform to native speaker norms. According to Hino, "with the dominance of native speakerism, Japanese learners of English have been taught to think and behave like Americans, where the criterion for good English has also been intelligibility to native speakers" (Hino 2021: 536). Comparing the native speaker rules to "linguistic mind-control," he claims that "the imposition of native speaker models deprives non-native speakers of the freedom of representing their cultural values" (531), which results in a "lack of diversity other than Anglo-American norms" (p. 532). He believes that users of English "must be liberated from native speaker norms in order to be allowed to fully represent their original identities" (p. 536). Referring to the seminal work of Smith and Nelson (1985), he suggests redesigning models for speaking and writing "to enhance global intelligibility, comprehensibility, and interpretability, reaching beyond the Inner Circle listeners and readers" (ibid.). As a theoretical foundation for seeking solutions to the problems of native-speakerism, Hino proposes three major paradigms, namely, WE (World Englishes), ELF (English as a Lingua Franca), and EIL (English as an International Language), and presents two methods of teaching English as an International Language: IPTEIL (Integrated Practice in Teaching English as an International Language) and CELFIL (Content and English as a Lingua Franca Integrated Learning), where "the former exposes learners to the linguacultural diversity of WE and the latter engages students in the interactional dynamism of ELF" (p. 540).

The approach discussed above aligns with Levisen's critique of the imposition of Anglo-centric linguistic frameworks on a global scale. Levisen argues that while his critique is not anti-English or anti-Anglo, it opposes the "eticization" of Anglo emics and the dominance of Anglo-English as a global metalanguage. He stresses that such Anglocentrism can lead to a conceptual monopoly and, in some cases, conceptual colonialism (Levisen 2024: 216). This reinforces our argument that the study of lexical bundles must move beyond the Anglocentric approach and instead reflect the diverse linguistic realities of L2 writers.

Another methodological alternative to the traditional perspective on lexical bundles research is the approach based on the theory of translingualism. According to this theory, "for multilinguals, languages are not discreet and separated but form an integrated system" or "a repertoire that is accessed for their communicative purposes" (Canagarajah, 2011: 1). Similarly, the linguistic competence of multilinguals "doesn't consist of separate competencies for each language, but a multicompetence that functions symbiotically for the different languages in one's repertoire" (ibid.). Therefore, cultural predilections for particular types of bundles can be viewed not as "a lack of familiarity with the common ways published writers create cohesive texts" (Hyland & Jiang 2022: 559) but as a manifestation of the translingual competences of EAL writers.

Both approaches emphasize the need to accommodate the diverse linguistic backgrounds of non-native speakers and foster authentic self-expression. The relevance of these approaches is supported by the studies that have found considerable variations in the use of lexical bundles not only among native and non-native speakers but also among writers with different first languages (Hyland & Jiang 2022) and different groups of native speakers, for instance, scholars in the humanities and natural sciences (Cortes 2004, Hyland 2008a) or reviewers in the fields of AI, engineering, and business (Kashiha 2023). These variations demonstrate that students and researchers follow the conventions of their particular academic disciplines and communities of practice. It justifies the culture-specific use of lexical bundles by L2 writers, provided that they do not violate the intelligibility (to both native and non-native speakers), style/register appropriateness, and other basic rules of academic writing. The proposed approaches will expand the boundaries of lexical bundles research by including cross-cultural variation in the scope of study and providing diversity beyond Anglo-American norms.

6. Conclusion

The main purpose of this investigation was to argue for a more inclusive approach to lexical bundles research, making it less prescriptive and more descriptive. Though our own investigation of lexical bundle use was limited in scope, our findings have shown that, in agreement with the results reported by previous studies, the linguistic and cultural backgrounds of writers with different L1s certainly contribute to the differences in the formulaic patterns of academic writing. One of the most prominent differences is a conscious effort by non-native authors to stress their research, its method, and procedures, focusing on the topic of the study rather than its presentation and engagement with readers.

Our study also confirmed that the L1 background of the writers influences their choice of bundles, including preferences for different structural and functional patterns. Although for Cameroonian students, whose texts we studied, English was the only language used in the academic setting, their choice of lexical bundles, especially participant-oriented bundles, differs considerably from that of native speakers of English. Moreover, due to cultural constraints, Cameroonian students used fewer participant-oriented bundles than even Russian students did. As for the Russian students, their use of lexical bundles reflects a characteristic feature of the Russian academic style, i.e., frequent use of strings of nouns in the genitive case. In other words, it is the native language and local culture that give a person cognitive potential with culturally appropriate connotations and models of argumentation necessary for accurate representation of complex meanings.

With this in mind, we argue for a shift of focus in the research on lexical bundles in academic English to make it more sensitive to the linguistic and cultural identities of L2 writers. We draw attention to the fact that some researchers (e.g., Hyland & Jiang 2022) have already started pointing out: what previous studies

usually viewed as a deficiency or deviation can be a manifestation of cultural peculiarities or local traditions of academic discourse. We believe that looking at lexical bundles in L2 writing from a "post-native-speakerist" standpoint is a promising venue for research: it will expand our knowledge of the linguistic features of different varieties of English and provide a more nuanced understanding of the existing variation in academic discourse conventions across languages and cultures. Further studies on lexical bundles use may concentrate on how academic English gets appropriated and adapted by L2 writers and how cross-cultural variation affects international academic norms.

References

Ädel, Annelie & Britt Erman. 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

Alamri, Basim. 2020. A comparative study of Saudi and international journals of applied linguistics: The move–bundle connection approach. *Journal of Language and Education* 6 (2). 9–30. https://doi.org/10.17323/jle.2020.10531

Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad & Edward Finegan. 1999. *Longman Grammar of Spoken and Written English*. London: Longman.

Biber, Douglas, Susan Conrad & Viviana Cortes. 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, Douglas & Federica Barbieri. 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, Douglas. 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

Boginskaya, Olga A. 2022. Functional categories of hedges: A diachronic study of Russian research articles abstracts. *Russian Journal of Linguistics* 26 (3). 645–667. https://doi.org/10.22363/2687-0088-30017

Canagarajah, Suresh. 2011. Translanguaging in the classroom: Emerging issues for research and pedagogy. *Applied Linguistics Review* 2. 1–28. https://doi.org/10.1515/9783110239331.1

Chen, Yu-Hua & Paul Baker. 2010. Lexical bundles in L1 and L2 Academic Writing. *Language Learning & Technology* 14 (2). 30–49.

Cortes, Viviana. 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

Dobrynina, Oksana L. 2019. Academic Writing for publication purposes: The infelicities of style. *Higher Education in Russia* 28 (10). 38–49. (In Russ., abstract in Eng.) https://doi.org/10.31992/0869-3617-2019-28-10-38-49

Dontcheva-Navratilova, Olga. 2012. Lexical bundles indicating authorial presence: A cross-cultural analysis of novice Czech and German writers' academic discourse. *Discourse and Interaction* 6 (1). 7–24. https://doi.org/10.5817/DI2013-1-7

Flowerdew, John. 2014. *Academic Discourse.* London: Routledge.

Grabowski, Łukasz. 2018. Stance bundles in English-to-Polish translation: A corpus-informed study. *Russian Journal of Linguistics* 22 (2). 404–422. https://doi.org/10.22363/2312-9182-2018-22-2-404–422

Hino, Nobuyuki. 2018. Pedagogy for the post-native-speakerist teacher of English. In Stephanie A. Houghton & Kayoko Hashimoto (eds.), *Towards Post-Native-Speakerism: Dynamics and shifts,* 217–233. Singapore. Springer. https://doi.org/10.1007/978-981-10-7162-1_12

Hino, Nobuyuki. 2021. Language education from a post-native-speakerism perspective: The case of English as an international language. *Russian Journal of Linguistics* 25 (2). 528–545. https://doi.org/10.22363/2687-0088-2021-25-2-528-545

Hyland, Ken. 2008a. As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes* 27. 4–21. https://doi.org/10.1016/j.esp.2007.06.001

Hyland, Ken. 2008b. Academic clusters: Text patterning in published and postgraduate writing. *International Journal of Applied Linguistics* 18 (1). 41–62. http://dx.doi.org/10.1111/j.1473-4192.2008.00178.x

Hyland, Ken. 2011. Academic discourse. In Ken Hyland & Brian Partridge (eds.), *Continuum companion to Discourse Analysis,* 171–184. London: Continuum.

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

Kashiha, Hadi. 2023. Beyond words in evaluation: Formulaic language in critical reviews of research articles across disciplines. *Russian Journal of Linguistics* 27 (2). 251–275. https://doi.org/10.22363/2687-0088-34320

Larina, Tatiana & Douglas Mark Ponton. 2022. I wanted to honour your journal, and you spat in my face: Emotive (im)politeness and face in the English and Russian blind peer review. *Journal of Politeness Research* 18 (1). 201–226. https://doi.org/10.1515/pr-2019-0035

Levisen, Carsten. 2024. *Postcolonial Semantics: Meaning and Metalanguage in a Multipolar World*. Berlin, Boston: De Gruyter. https://doi.org/10.1515/9783111337432

Li, Liang, Franken Margaret & Wu Shaoqun. 2019. Chinese postgraduates' explanation of the sources of sentences initial bundles in their thesis writing. *RELC Journal* 50 (1). 37–52. https://doi.org/10.1177/0033688217750641

McKinley, Jim. 2022. An argument for globalized L2 writing methodological innovation. *Journal of Second Language Writing* 58. 100945. https://doi.org/10.1016/j.jslw.2022.100945 https://www.sciencedirect.com/science/article/pii/S1060374322000807?via%3Dihub (accessed June 8, 2024).

Navarro, Gil Noelia & Elena C. Martínez. 2019. Lexical bundles in learner and expert academic writing. *Bellaterra Journal of Teaching & Learning Language & Literature* 12 (1). 65–90. https://doi.org/10.5565/rev/jtl3.794

Nkemleke, Daniel. 2014. Researching the research article in Cameroon English. *Brno Studies in English* 40 (1). 175–203. https://doi.org/10.5817/BSE2014-1-9

Richter, Kenneth G., Behruz Lotfi Gaskaree & Milad Mirzai. 2022. A functional analysis of lexical bundles in the discussion sections of applied linguistics research articles: A cross-paradigm study. *Russian Journal of Linguistics* 26 (3). 625–644. https://doi.org/10.22363/2687-0088-27752

Salazar, Dominica. 2010. Lexical bundles in Philippine and British scientific English. *Philippine Journal of Linguistics* 41. 94–109.

Smith, Larry E. & Cecil L. Nelson. 1985. International intelligibility of English: Directions and resources. *World Englishes* 4 (3). 333–342.

Vladimirova, Tatiana L. 2010. *Language and Style of Academic Texts.* [Yazyk i stil' nauchnogo teksta]. Tomsk: Izd-vo Tomskogo politekhnicheskogo universiteta. (In Russ.).

Wei, Yaoyu & Lei Lei. 2011. Lexical bundles in the Academic Writing of Advanced Chinese EFL learners. *RELC Journal* 42 (2). 155–166. https://doi.org/10.1177/0033688211407295

Bionotes

Elena S. GRITSENKO is Dr. Habil. in Linguistics, Professor of the English Department at MGIMO University. Her research interests include sociolinguistics of globalization, discourse analysis, language and gender. She authored and co-authored monographs, book chapters and articles in peer-reviewed journals.
e-mail: elena.s.gritsenko@gmail.com
https://orcid.org/0000-0001-8669-360X

Olivier Mozard T. KAMOU is a PhD student at HSE University in Nizhny Novgorod. His research focuses on a corpus-based approach to indexicality and (im)politeness in online communities. His other research interests include corpus linguistics, computational pragmatics and conventionality.
e-mail: otiotsopkamu@hse.ru, kamouolivier@gmail.com
https://orchid.org/0000-0001-5291-3037

Сведения об авторах:

Елена Сергеевна ГРИЦЕНКО – доктор филологических наук, профессор, профессор кафедры английского языка № 3, МГИМО МИД России. Сфера научных интересов – социолингвистика, контактная вариантология английского языка, язык и культура, язык и гендер. Является автором и соавтором публикаций на русском и английском языках, включая монографии, главы в книгах и статьи в международных рецензируемых научных журналах.
e-mail: elena.s.gritsenko@gmail.com
https://orcid.org/0000-0001-8669-360X

Оливье Мозард Т. КАМУ – аспирант департамента иностранных языков Национального исследовательского университета «Высшая школа экономики». Его диссертационное исследование посвящено корпусному анализу индексальности и вежливости в онлайн-сообществах. Другие исследовательские интересы включают корпусную лингвистику, вычислительную прагматику и конвенциональность.
e-mail: otiotsopkamu@hse.ru, kamouolivier@gmail.com
https://orchid.org/0000-0001-5291-3037