




<https://doi.org/10.22363/2687-0088-36282>

EDN: XRMONI

Research article / Научная статья

Informality in academic English texts by Arabic and British scholars: A corpus study

Abuelgasim S. E. MOHAMMED   and Abdulaziz B. SANOSI 

*College of Science & Humanities, Prince Sattam Bin Abdulaziz University,
Al-Kharj, Saudi Arabia*
 ab.ibrahim@psau.edu.sa

Abstract

There has long been a widespread belief that academic writing has to maintain a high degree of formality and impersonality. Despite this, it has been observed that there is a general tendency to use informal style in academic writing by writers from different disciplines. This informality manifests itself in the use of various linguistic devices that were previously observed only in spoken discourse or in informal communication. The aim of the study is to identify common informal features in English academic writing used by Arabic and British scholars and compare the level of informality used in English academic writing by representatives of the two lingua-cultures. The study adopts the corpus linguistics method. The one-million-word Arab Scholar Written English Corpus (ASAWEC) was compiled and analysed for informality features. The results were then compared to the British Academic Written English (BAWE) corpus. We focused on the use of informal features, such as broad references, initial conjunctions, first-person singular, second person pronouns, final preposition, listing expressions, split infinitive, and contractions. The results revealed a significant difference in the use of informality features in favour of the native speakers. The findings showed that Arabic scholars tend to use broad references and initial conjunctions, however, they rarely use contractions and split infinitives. Contrary to this, British scholars used the whole spectrum of informality features. The findings can provide insightful implications for researchers, journal editors and peer reviewers to account for informality levels in academic writing in different linguacultures. They may also be used in teaching English academic writing.

Keywords: *academic English, academic writing, informality features, corpus linguistics, Arabic English, British English*

For citation:

Mohammed, Abuelgasim S. E. & Abdulaziz B. Sanosi. 2024. Informality in academic English texts by Arabic and British scholars: A corpus study. *Russian Journal of Linguistics* 28 (3). 633–654. <https://doi.org/10.22363/2687-0088-36282>


© Abuelgasim S.E. Mohammed & Abdulaziz B. Sanosi, 2024



This work is licensed under a Creative Commons Attribution 4.0 International License
<https://creativecommons.org/licenses/by-nc/4.0/legalcode>

Неформальность в академическом английском в текстах арабских и британских ученых: корпусное исследование

А.С.Э. МОХАММЕД  , А.Б. САНОСИ 

*Колледж естественных и гуманитарных наук,
Университет принца Саттама бин Абдулазиза, Аль-Хардж, Саудовская Аравия*
 ab.ibrahim@psau.edu.sa

Аннотация

Долгое время существовало распространенное мнение, что академическое письмо должно поддерживать высокий уровень формальности и безличности. Несмотря на это, было замечено, что существует универсальная тенденция к использованию неформального стиля в академическом письме авторами, представляющими различные научные дисциплины. Неформальность проявляется в использовании ряда языковых средств, которые ранее встречались только в разговорной речи или в неформальном общении. Цель данного исследования – выявить общие черты неформальности, используемые арабскими и британскими учеными в академическом английском, и сравнить уровень неформальности, допускаемый в английском академическом письме представителями двух лингвокультур. В исследовании использован метод корпусной лингвистики. Был составлен и проанализирован на предмет выявления неформальных черт корпус арабского письменного английского (ASAWEC) объемом в один миллион слов. Далее результаты сравнивались с корпусом британского академического письменного английского (BAWE). Мы исследовали использование таких неформальных черт, как широкая отсылка, начальные союзы, местоимения первого лица единственного числа, местоимения второго лица, конечный предлог, перечислительные выражения, расщепленный инфинитив и сокращения. Наши результаты показали значительную разницу в использовании черт неформальности в пользу носителей английского языка. Они показали, что арабские ученые склонны использовать широкую отсылку и начальные союзы, но они редко используют сокращения и расщепленные инфинитивы, в то время как британские ученые используют весь спектр показателей неформальности. Полученные результаты могут быть полезными для исследователей академического дискурса, редакторов журналов и рецензентов, которые должны учитывать уровень неформальности в академическом письме, допускаемый представителями различных лингвокультур, также они могут использоваться при обучении английскому академическому письму.

Ключевые слова: академический английский, академическое письмо, черты неформальности, корпусная лингвистика, арабский английский, британский английский

Для цитирования:

Mohammed A.S.E., Abdulaziz B.S. Informality in academic English texts by Arabic and British scholars: A corpus study. *Russian Journal of Linguistics*. 2024. Vol. 28. № 3. P. 633–654. <https://doi.org/10.22363/2687-0088-36282>

1. Introduction

Academic writing is traditionally described as formal and impersonal (Matsuda & Nuri 2020). Nonetheless, it has recently been observed that informality is gradually becoming a feature of academic writing along with formality (Boginskaya 2023, Hyland & Jiang 2019, Praminath et al. 2018, Adel 2008, Chang & Swales 1999). Several researchers (e.g., Yang & Pan 2023, Leedahm 2015, Mair

1998) have recorded the occurrence of informal features in academic writing in different disciplines. This tendency is described as a “part of a contemporary zeitgeist” by Hyland and Jiang (2019: 217).

Written English communication is considered a key to knowledge circulation. The English Language, therefore, has dominated the world of academia as scholars across the world widely adopt it in almost all disciplines (Mastuda & Nouri 2020, Chang & Swales 1999). In order to be recognised, academics, both native and non-native, are required to publish their research in the English language. As a result, the ability to write in an appropriate academic style has become a fundamental skill for writers in all fields. Scholars for whom English is a non-native language are aware of the significant influence that academic writing in English exerts on their domestic academic culture (Chang & Swales 1999) and they believe that following strict academic writing conventions can also mark them as members of this specific discourse community (Richter, Gaskaree & Mirazi 2022) and help them in getting published internationally (Fendri & Mounir 2022).

Previous studies (Hyland & Jiang 2017, 2019, Chang & Swales 1999) have examined the use of informal features used in selected disciplines, including applied linguistics. The aforementioned studies conducted a comparative analysis of informal features across disciplines of interest. Nevertheless, there has been a paucity of scholarship devoted to the examination of informal features in published academic articles on English language studies authored by non-native and native scholars. The researchers believe that this situation is unfortunate as it has been found that academic writing is more challenging to EFL writers than to native academics because the former’s “mental schemata are constrained by their limited mastery of the foreign language’s formal and pragmatic features” (Fendri & Mounir 2022: 58). As informality is more influential in creating engagement with the readers and hence fostering communication, it is therefore believed that using a balanced informality in academic writing promotes communication. Therefore, the objective of this study is to examine the characteristics of informal language employed by Arab scholars in the context of published academic writing in English. The level of this employment is determined by recognising potential differences in this use from that of native speakers’ use, and identifying the most and least frequent informal features as they occur in Arab Scholars Academic Written English Corpus (ASAWEC) and British Academic Written English (BAWE). It is anticipated that the outcome of this undertaking will furnish valuable insights into the rhetorical approaches adopted by Arab scholars in their respective fields, thereby offering a foundation for the advancement of academic writing among Arab scholars.

Accordingly, the current study is sought to answer the following questions:

1. What are the informality features used by Arab scholars in their English publications?
2. What are the most and least common informal features in ASAWEC and BAWE?

of the interactional impact of their words on their readers. Furthermore, it embraces the utilisation of specific linguistic features that are not only identifiable, but are also countable (Hyland & Jiang 2019).

Informality is an attempt to facilitate two elements. The first one is a specific bond between the author and the readers who share a similar context. This could facilitate communication, ensure understanding and make the information more accessible. The second is to create a common ground of familiarity (Hyland & Jiang 2019). In this respect, familiarity forms common ground for both writers and readers. It is worth noting that appropriate criteria of conciseness need to be guaranteed, implying that researchers avoid an inclination toward informality.

Hyland (2005) defines informality as the use of language associated with everyday communication. It shows how authors establish reciprocal relationships with readers. It appears difficult to isolate it from formality necessary for academic writing. Heyllighen and Dewaele (1999) distinguish between formal and informal styles. The former is associated with detachment, precision and rigour. The latter is characterised by flexibility, directness, implicitness, involvement, and information. In the same vein, Hyland and Jiang (2017) suggest that formality, on the one hand, subsidises dependence on the context and vague phrases in the text. In addition, it avoids ambiguity and misunderstanding. On the other hand, informality refutes formality in writing to achieve a friendlier and more acceptable person, to create a shared point of view and to ensure interaction with readers. Therefore, informality facilitates interaction through writers’ use of some linguistic devices. However, the utilisation of these features should be approached with care, as informality is typically associated with the writer’s subjectivity and personality, which “may clash with the objectivity of academic writing” (Fendri & Mounir 2022: 53).

2.2. Approaches to identifying informality features

Features of informal language may be identified by the occurrence of specific linguistic items, such as first-person pronouns (I and We) and unattended anaphoric pronouns (this, these) among others (Hyland & Jiang 2019). There are three main approaches to detecting them: Multidimensional corpus analysis, Ambiguity avoidance and Style guides (see Figure 1).

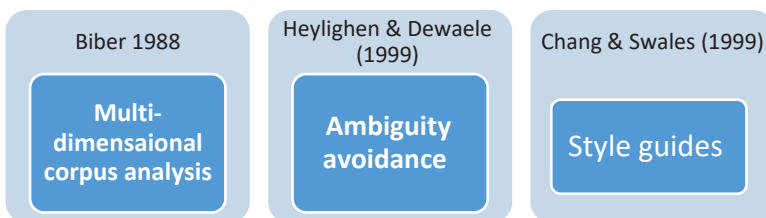


Figure 1. Informality features identification approaches

3. Is there any statistical difference in the use of informal features between Arab scholars and native speakers' academic written English?

It is considered important to answer these questions not only because they address the research gap in Arab academic writing, but also because they facilitate for a deeper understanding of academic writing among Arab scholars, which is of significance for several reasons. Firstly, there is a high increase in publications by Arab scholars, largely driven by the necessity for promotion and academic recognition. This has led to a culture of publish-or-perish, whereby scholars are compelled to publish their work in order to advance their careers. This increase in publications has highlighted the necessity for research-based insights to inform the development of training programs designed to enhance scientific publishing. Secondly, these scholars in question are primarily university professors who are also involved in teaching research-related activities, particularly in the case of the present study. It can be reasonably assumed that an improvement in the writing skills and knowledge of these scholar will have a positive impact on their students. Finally, an analysis of the style employed by Arab scholars in comparison to that of native speakers can yield insights into the influence of cultural differences in writing styles. Thus, the present study concentrates on the employment of informality features by Arab scholars and their comparison with those of British scholars'. The study adopts Chang and Swale's framework of informality features. In accordance with the corpus linguistics method, the analysis is conducted on two corpora: EFL corpus and a native corpus. The following sections present, a review of the relevant literature, the methodology, the results, a discussion of these results, and a conclusion.

2. Literature review

This section provides a review of the literature on informality features in academic writing and the approaches to identifying them. It also summarises some previous research on the use of informality elements in English language studies research.

2.1. Informality in academic writing

Informality in writing has been the focus of researchers for a considerably long time. Hyland and Jiang (2019), citing Gilbert and Mulkay (1984), describe informality as full of debates, hypothetical perceptions, cognitive devotion, and societal nepotism. Informality is believed to be a departure from a conventional objective academic posture. Informality, then, is the expression of private meaning, which implies a reciprocal relationship with the audience, a willingness to discuss claims and a favourable attitude towards subjectivity. In this respect, the use of informality features resembles the employment of *reader-oriented hedges*. Hyland (1996, cited in Boginskaya 2022) believes that these hedges not only facilitate the connection between the writer and the reader but also ensure that writers are aware

The first approach to identifying informality features is Biber's (1988) multidimensional corpus analysis. This model employs both quantitative and qualitative comparisons of various language registers (spoken and written). This distinction between oral and literate discourse helps clarify the concept of formality in communication, particularly in informal interactions. This method analysed the co-occurrence of 67 linguistic features, resulting in seven functional dimensions that predict register differences. The first dimension is *Interactivity/informative expressions*. Positive values signal high interactivity, while negative values propose strong information transmission. The second dimension is *narrative/non-narrative concern*. This is identified by positive values reflecting narrative texts (e.g., novels), whereas negative values indicate non-narrative discourse. The third one is *clear and context-dependent reference*. In this dimension, positive values show low dependence on context, while negative values indicate high dependence. The fourth dimension involves *explicit persuasive representation* in which positive values reveal a high degree of persuasion, while negative values signify low persuasion. The fifth dimension is *information abstract/non-abstract style*. In this dimension, positive values feature an abstract, formal style, while negative values indicate a non-abstract style. The next dimension is *fineness of immediate information organization* which distinguishes between real-time and non-real-time information transmission. The last dimension is *academic model expression*. Each dimension is composed of language features with complementary positive and negative loads, serving opposite communicative functions. This approach, known as multi-feature/multi-dimensional analysis (MFA/MDA), is widely used in corpus research to study register variation.

The second approach to informal feature detection was proposed by Heyligeln and Dewaele (1999) and it involves identifying features which contribute to text conciseness and context independence. by means of two methods. The first is listing vocabulary that has the specifying function and its reference is determined by its context in one group (formal). The second is itemizing features that presume an understanding of context in another group (informal). Heyligeln and Dewaele (1999) developed a formula to assess informality in an academic text. According to this formula, nouns, adjectives, prepositions, and articles of greater frequencies are features of formal text. In contrast, the prevalence of verbs, pronouns, and interjections contributes to making a text informal.

The third approach is based on a review of style manuals or guides (Hyland & Jiang 2019, Bennett 2009, Chang & Swales 1999). Bennett (2009) assessed style guides and identified that they contained constant reference to objectivity and formality that is attained through the comprehensive use of 'Latinated vocabulary' and detached structures. After surveying 40 style guides to formality, Chang and Swales (1999) noted persistent mentioning of first-person pronouns, wh-questions, listing expressions, and contractions as features of informality. Hyland and Jiang's (2019) survey of some sites of university language centres found corresponding suggestions to refrain from using these features.

Chang and Swales (1999) identify ten features of informality that are widely adopted in research. Subsequently, Hyland and Jiang (2017, 2019) modified this list by adding second-person singular instead of sentence fragments, which according to them, almost rarely appear in academic writing due to thorough review processes in journals. The final list includes: the first- person pronouns, anaphoric pronouns - broad references, split infinitives, starting a sentence with a conjunction or a conjunctive adverb, listing expressions (*and so on, etc., and so forth*), prepositions at the end of sentences, run-on sentences, sentence fragments, contractions, direct questions, and exclamations.

2.3. Previous studies

Several studies, such as Yang and Pan (2023), Hyland and Jiang (2017, 2019), Chang and Swales (1999), have been devoted to identifying informality features in academic writing. Most of these studies compared the use of informality features across disciplines. Yang and Pan (2023) found general compatibility between published advice and actual practice concerning the use of informal elements. They also noted the frequent use of informal linguistic features despite the style manuals recommendation against their use. Additionally, the study found variation in the use of informality features between linguistics and physics implying that academic writing is not consistent and may vary based on the discipline. Few other studies examined the use of informality features in studies focusing on English language studies. Following is a summary of some of these studies.

Tocalo et al. (2022) intended to analyse informality in Filipino ESL scholars' academic writing. The study adopted articles amounting to 1,000,000 words as its materials. These articles were published in 7 eminent local and international journals during 2010–2020. The study reported an increase in the use of informality features during that period. The most common informal features were unattended anaphoric pronouns and sentence-initial conjunctions/conjunctive adverbs. In addition, the study demonstrated that the least utilised features were the first-person pronouns and sentence-final prepositions.

Ebrahimi and Fakheri (2019) scrutinised informality occurrence in 50 applied linguistics published articles in two local Iranian journals published in 2014–2015. The study concluded that the most frequent informal features were unattended anaphoric pronouns and sentence-initial conjunctions, while the least occurring ones were exclamations and contractions.

Alipour and Nooreddinmoosa (2018) investigated the employment of informal features in applied linguistics research produced by Iranians and native speakers. The study examined 200 published research articles in 6 journals (3 Iranians and 3 British) published during 2012–2017. The Iranian corpus consisted of 692.046 words, while the British contained 910.740 words. The study found that informality occurred more frequently in British scholars' writings than in the Iranians'. The study identified no significant difference between the Iranian and British corpora in the utilisation of informality. Additionally, the most frequent informal features

appeared to be initial conjunctions, while the least common one were exclamation marks in both corpora.

Praminatih et al. (2018) analysed undergraduates' theses abstracts written by Indonesian EFL students to examine the use of informal language in 1992, 1996, 2000, 2004, 2008, 2012 and 2016. A total of 114 abstracts were subjected to analysis. The study concluded that some informal features, including sentence-initial conjunctions/conjunctive adverbs, sentence-final prepositions, contractions, and direct questions existed in the 114 abstracts. Moreover, the research revealed a decrease in the use of informality over time.

Sholihah (2018) attempted to discover the elements of informality in general and the most common ones in thesis proposals by EFL students. The data were elicited from 15 thesis proposals of English Teacher Education Department in Uin Sunan Ampel Surabaya. The study reported the use of nine informality features in the proposals. These features were unattended anaphoric pronouns, first-person pronouns, construction, initial conjunctions and conjunctive adverbs, sentence fragments, listing expressions, direct questions, adverbs in initial or final position, and second-person pronouns. In addition, the most common features of informality were initial conjunctions and conjunctive adverbs, first-person pronouns and unattended anaphoric pronouns.

Considering these studies, it can be suggested that informality features are commonly used in published English language studies. Only one study (Alipour & Nooreddinmoosa, 2018) compared the use of informality features by native and non-native speakers and found that the British corpus employed more informality features than the Iranian one. Unattended anaphoric pronouns emerged as the most common feature in two studies (Tocalo et al. 2022, Ebrahimi & Fakheri 2019).

It is also worth noting the limitations of these studies. For example, Tocalo et al. (2022) investigated a corpus of approximately 1.000.000 words drawn from articles published in 7 Filipino and international journals. The researchers neither disclosed further details about their materials nor did they provide more information about the journals. In addition, the number of the examined words is relatively small. Ebrahimi and Fakheri (2019) investigated 50 applied linguistic published research articles in two Iranian journals during 2014–2015. The number of articles was quite undersized. Alipour and Nooreddinmoosa (2018) investigated a British corpus of 910.740 words and an Iranian corpus of 692.046 words, which is considerably less than the British one by more than 300.000 words. Praminatih et al. (2018) confined their focus on undergraduate theses abstracts in 2016 that may not provide sufficient context for investigating informality features due to abstract restricted format. Thus, the constraints in all these studies may cause some shortcomings in their results.

3. Materials and methods

The present study adopted the corpus linguistic method as proposed by Biber et al. (2007) according to which the most appropriate way to analyse discourse

organisation is using a bottom-up approach. Both quantitative and qualitative techniques were implemented in the study as well as a comparative analysis with the previously used frameworks

3.1. Corpora

Two corpora were used in the study: Arab Scholar Academic Written English Corpus (ASAWEC) and an extracted sub-corpus from the British Academic Written English (BAWE) corpus.

3.1.1. ASAWEC

ASAWEC was compiled from research articles published by Arab scholars in renowned English language journals. The sources of the corpus files are detailed in Table 1 below.

Table 1. Details of ASAWEC sources

Journal	Article	Date of Pub.
Arab Journal of Applied Linguistics	11	2020–2022
Arab World English Journal	36	2021–2022
International Journal of Arabic-English Studies	35	2019–2022
Jordan Journal of Modern Languages and Literatures	36	2019–2022
Journal of Research in Language & Translation	8	2020–2022
Saudi Journal of Language Studies	19	2021–2022
Umm Al-Qura University Journal of Language and Literature	8	2019–2022
Total	153	

The chosen research articles focused mainly on applied linguistics, translation, and literature. The number of files selected from each journal was based on different criteria, such as frequency of publication and the online availability of the published issues.

3.1.2. BAWE

For reference purposes, the researchers used a mini corpus extracted from BAWE¹. Overall, 177 files were extracted from the corpus based on specific criteria to meet compatibility requirements with the main corpus. These criteria were believed to raise the compatibility levels between the two corpora. Nevertheless, minor variations between the two corpora existed, as evident in Table 2.

¹ BAWE was developed at the Universities of Warwick, Reading and Oxford Brookes, under the directorship of Hilary Nesi and Sheena Gardner (formerly of the Centre for Applied Linguistics [previously called CELTE], Warwick), Paul Thompson (Department of Applied Linguistics, Reading) and Paul Wickens (Westminster Institute of Education, Oxford Brookes), with funding from the ESRC (RES-000-23-0800). Source: The University of Warwick. https://warwick.ac.uk/fac/soc/al/research/collections/bawe/how_to_cite_bawe/, accessed 15th January. 2023.

Table 2. Characteristics of ASAWEC and BAWE corpora used in the study

Corpus	Files	Average File length	Tokens	Types	TTR*
ASAWEC	153	6563	1004155	45974	0.046
BAWE	177	5705	1004081	44968	0.045

*Note. TTR= Type-Token ratio

The major variation between the two corpora is represented by the difference in the numbers of the incorporated files. This variation is due to the length of BAWE files, which are shorter than those from ASAWEC. Also, a lower TTR in the BAWE suggests a higher lexical richness of the ASAWEC, possibly due to the writers’ use of advanced topics and levels of education.

3.2. Informality features framework

The present study adopted Chang and Swales’ (1999) framework, which was adapted from Hyland and Jiang (2017, 2019) for informality features. This framework has been used in numerous studies and has proven accurate and structured to determine linguistic devices that imply informality in academic writing. Chang and Swales (1999) built their list of informality features based on the frequency of mentioning of the features in manuals and handbooks on academic writing. They provide a list of ten most referred features along with the number of references that mention them. For the present study, the researchers used an adapted version of the eight elements, as shown in Figure 2. Two elements - direct questions and exclamations - were excluded from the list because they were referred to by style manuals and handbooks only twice (Chang & Swales 1999: 148). In addition, one informality feature - direct questions - was challenging to investigate as most of the texts include direct questions that are required by the genre itself, i.e. research questions, surveys and interviews. Consequently, it was difficult to identify whether these questions were reflective of the writing style of the writer or not.

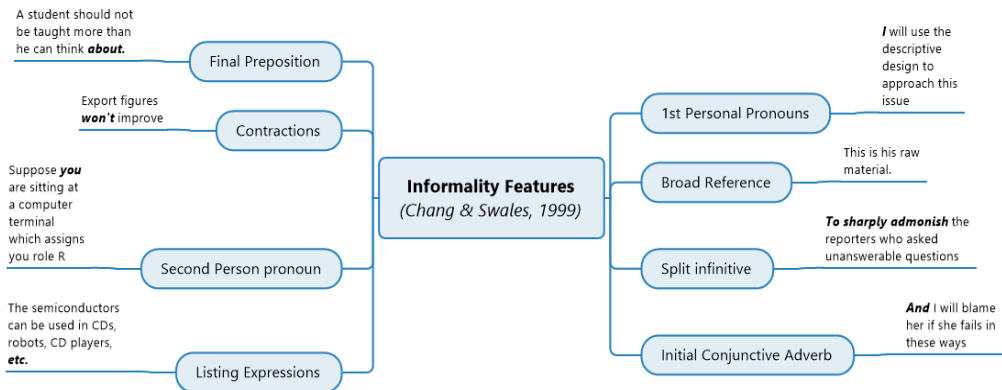


Figure 2. The adapted framework of informality features

Note. Adapted from: Chang & Swales (1999) and Hyland & Jiang (2017, 2019). Informal elements in English academic writing: Threats or opportunities for advanced non-native speakers? (p. 148).

We made this selection because, according to Hyland and Jiang (2019), the list is consistent with what is regarded as informal language. Additionally, the use of informal features entails communication and personal interaction with readers. Another reason is that the list caters to beginner researchers' anxiety about employing these informal features. Moreover, some of these items represent the conflict between linguistic forms and present writing applications. Finally, the list of informal items is widely adopted in informality research (e.g., Kuhi et al. 2020, Lee et al. 2019).

3.3. Procedures

The main corpus of the study, ASAWEC, was built following the steps and techniques of corpus building suggested by pioneer corpus linguists (Atkins et al. 1992, McEnery & Hardie 2012, Poole 2018, Tognini-Bonelli 2001). Initially, the journals were selected to meet the criterion of high academic standards. The selected journals are published by renowned Arab universities, and the scholars who published there are Arabs. Further validation of the identity of the authors was nevertheless conducted using the authors' full names, email addresses, and personal and academic websites. The articles which were authored or co-authored by non-Arab researchers were excluded. In this stage, a total of 185 PDF files were downloaded.

In the second stage, all the files were converted to *text* format and cleaned. The researchers used AntConcFileConverter² and EmEditor Professional³ software at this stage. Using the Regular Expression feature, all the noise data in the files were deleted. Elements such as Arabic characters, formatting and mathematical symbols, clichés, journal titles and bibliographical information were deleted. Furthermore, the references and appendices section of all the articles were removed. Ultimately, 153 files were selected and indexed in the corpus folder. The corpus files were then tagged using the POS tagging feature and re-uploaded to #LancsBox⁴ software for the analysis stage.

Regarding the BAWE sub-corpus, the researchers used the documentation file of the corpus and then sorted the files according to specific criteria. Each time a criterion was applied, the number of the selected file was reduced. Finally, 177 files were chosen from the filtered corpus to match the number of ASAWEC tokens. Figure 3 explains the process of criterion application and file selection. The process started with identifying the target language, L1 English language. Then, we focused on the disciplines covered by the collected articles. Three disciplines which were arts, humanities, and sociology were included. We then had to choose the genres to

² Anthony, L. 2017. AntFileConverter (version 1.2.1) [Computer Software]. Waseda University. <http://www.antlab.sci.waseda.ac.jp/>

³ Emurasoft, Inc. 2019. EmEditor Professional (Version 19.3.2) [Computer Software]. Filepuma. https://www.filepuma.com/download/emeditor_professional_64bit_19.3.2-23779/

⁴ Brezina, V., Weill-Tessier, P., & McEnery, A. 2020. #LancsBox v. 6.0 [software]. Available at: <http://corpora.lancs.ac.uk/lancsbox>.

be studied, which included research reports, essays, proposals, and case-studies. The students who produced these genres belonged to two grades, distinction and merit, and they were in levels three and four in undergraduates, but some postgraduates were also included. These filters produced 177 files to be included in our analysis. It is worth noting that these filters originated from the Excel file downloaded from BAWE.

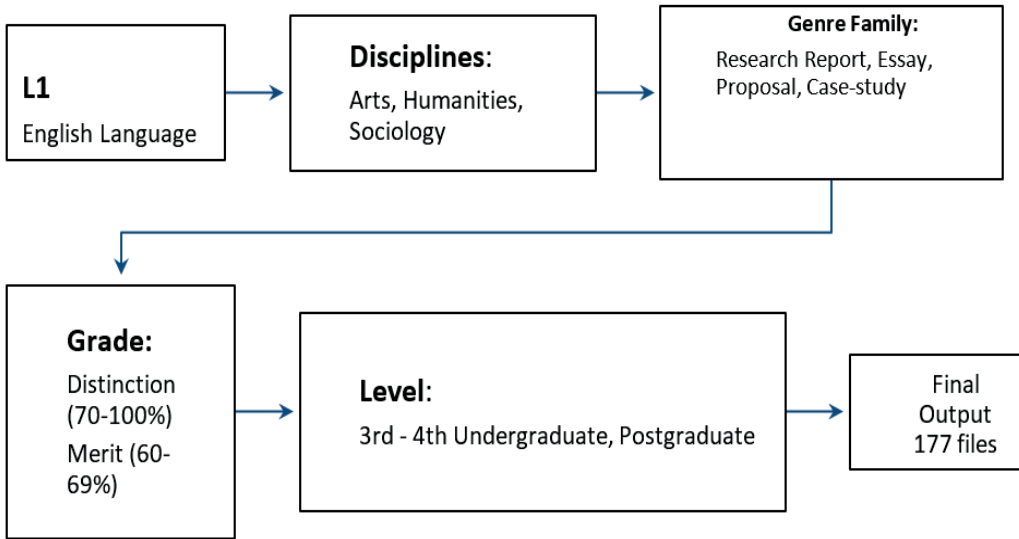


Figure 3. The process of selecting BAWE files

The two corpora were then uploaded to the corpus linguistic software #LancsBox for quantitative analysis.

3.4. Data analysis

The two corpora were analyzed using the Key Word in Context (KWIC) tool of the #LancsBox software. Each item of the eight informal academic writing features was explored using simple search, smart search, or RegEX search. For example, since complex structures such as split infinitives and contractions are predefined in the software, the researchers searched for them using the predefined search terms SPLIT INFINITIVE and CONTRACTION in the two corpora and generated the statistical significance of the differences from the software directly. See Figure 4 for an example of a smart search.

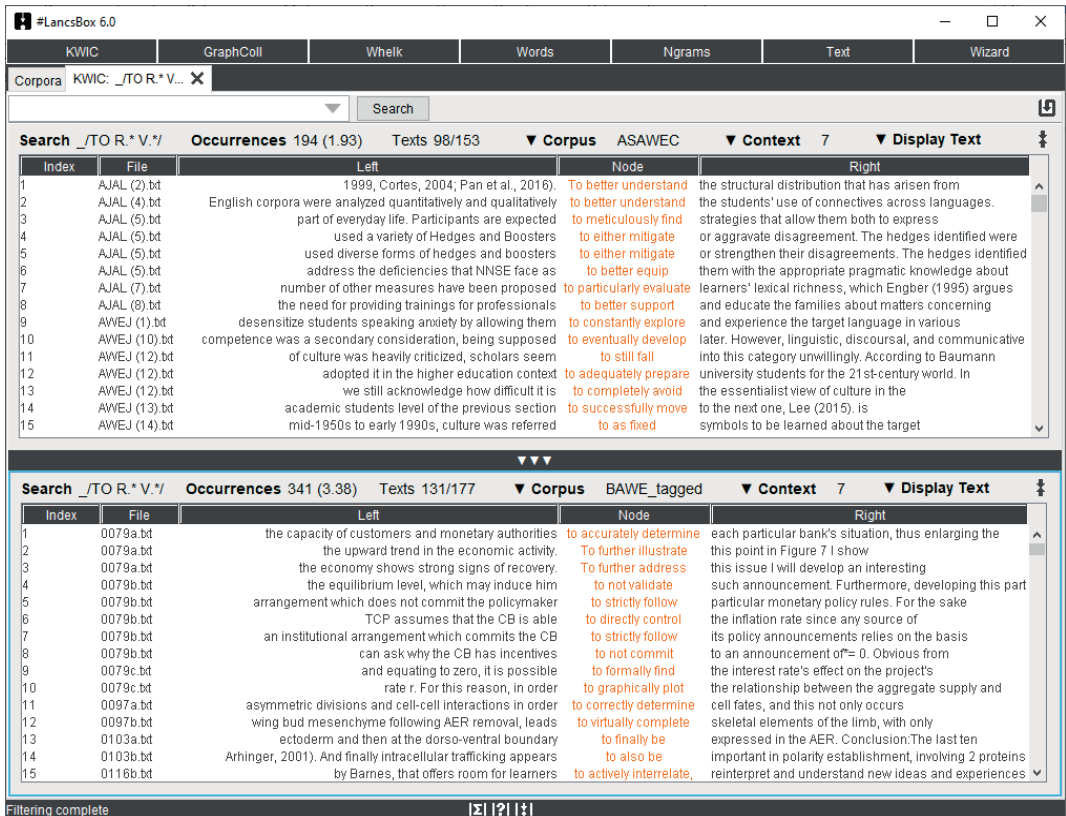


Figure 4. Smart search for the split infinitive

For search terms that require determining position, extra features like letter case and punctuation were used which we applied through the RegEX feature. For example, the search term */However/* is used to determine the occurrence of the conjunctive adverb *however* in the initial position, while the search term */in\.* /p was used to determine the occurrence of the preposition *in* at the end of a sentence. For other items that include single or compound lexical units, a simple search was applied.

After generating the frequency of occurrence of each item in the two corpora, the Welch Two Samples T-test was conducted to specify the statistical significance of the difference between the two groups. The results of the quantitative analysis are presented in the forthcoming section.

4. Results

The study results are presented in three parts. First, the Arab scholars' use of informality features is presented. Second, a comparison between Arab scholars and native English academic writers is reported. Finally, the statistical significance of the difference between the two groups of academic writers is shown.

4.1. Arab scholars' use of informality features

Using quantitative analysis of ASAWEC we found that Arab scholars used the informality features identified by Chang and Swales (1999) as follows.

Table 3. Informality features used by Arab scholars (sorted by relative frequency per 10k)

Feature	Relative frequency*	Range * %
Broad References	28.59	100
Initial conjunctions	25.31	38.3
First-person pronouns	12.85	73
Second-person pronouns	6.35	64.05
Final prepositions	2.4	7.68
Listing expressions	1.97	21.79
Split infinitives	1.94	64.05
Contractions	1.87	26.14

*Note. Frequency per 10,000 tokens.

* Range: Occurrence of an informality feature across analysed files

The most frequently used informality features by Arab scholars are broad references, initial conjunctive adverbs and conjunctions, and first-person pronouns, respectively. Table 4 displays a few examples of these features from ASAWEC.

Table 4. Examples of Arab scholars' use of informality features (top used features)

Feature	Example	Source (File)
Broad references	"... stance towards content and towards readers (hearers). This has resulted in various classifications of organisational.."	AJAL (11)
Initial conjunction	"Arabic counterparts of their mother tongue words. And that they often literally translate from Arabic"	AWEJ (7)
Personal Pronouns	" We extracted all the features and examined all the occurrences..."	IJAES (27)

It is also noted that broad reference, the most common informality feature in ASAWEC, is present in all the sample files. On the other hand, contractions, split infinitives, and listing expressions were the least used features in the corpus and unsystematically distributed across the corpus files.

4.2. Arab scholars' versus native speakers' use of informality features

When comparing the native language speakers' corpus BAWE with ASAWEC, we found that all the informality features were used in BAWE but not in ASAWEC. However, it is interesting that the top three informality features, i.e., broad reference, initial conjunctions, and personal pronouns, are identical across the two corpora. A slight difference is spotted in that initial conjunctions were the most used feature by the native speakers rather than broad reference. Figure 5 illustrates these differences.

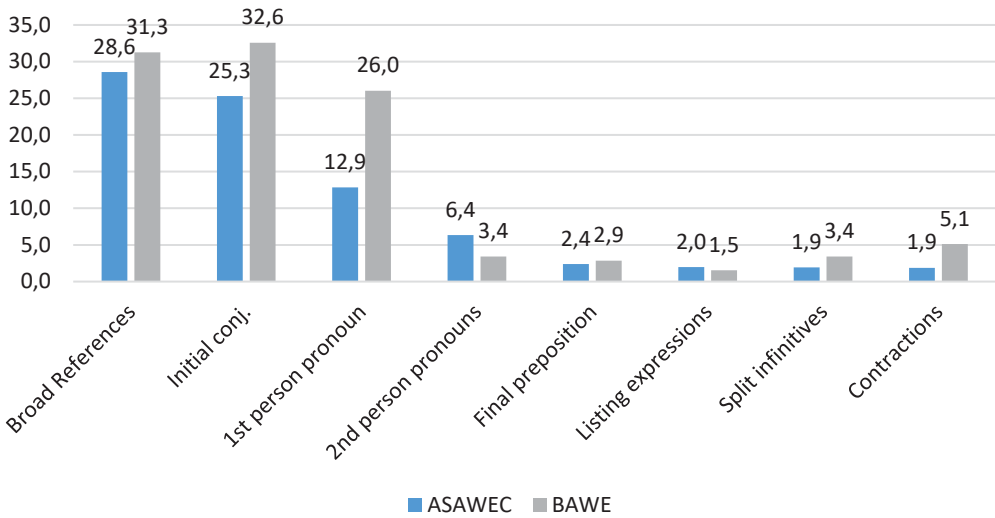


Figure 5. Informality features used in ASAWEC and BAWE (according to relative frequency per 10k)

The results also showed a minor difference in the least used features. Listing expressions and split infinitives are scarcely used by native speakers with relative frequencies of 1.5 and 3.4, respectively, while a considerable use of contractions amounted to 5.1 is spotted. Table 5 reports some examples of the use of these features in BAWE.

Table 5. Examples of native speaker use of informality features (least used features)

Feature	Example	Source (File)
Listing expressions	<i>“acting according to his desires or beliefs etc. “</i>	0407
Split infinitive	<i>“ for learners to actively construct meaning from the text”</i>	0116b
Contractions	<i>“.. but if the flow of information hasn’t improved then the objective hasn’t been met”</i>	0193d

A detailed comparison of the use of informality features in terms of relative frequency and range is presented in the Appendix.

4.3. Significance of the difference

Table 6 reports the significance of the differences in the use of informality features by Arabs and native academic writers.

The differences between Arabs’ and natives’ use of informality features were statistically significant in most of the studied features, initial references $P = 0.006$, first- person singular $P = 0.014$, second-person pronouns $P = 0.003$, split infinitive $P < 0.001$, and contractions $P < 0.001$. However, it seems that both groups use broad references, final prepositions, and listing expressions at approximately similar levels.

Table 6. Statistical significance of the differences

Feature	T-value	P-value	Significance
Broad References	303.92	0.34	No
Initial conjunction	297.26	0.006	Yes
First-person pronoun	299.86	0.014	Yes
Second-person pronouns	208.72	0.003	Yes
Final preposition	308.22	0.571	No
Listing expressions	258.76	0.147	No
Split infinitives	290.35	< 0.001	Yes
Contractions	316.54	< 0.001	Yes

5. Discussion

The analysis of the data revealed that Arab scholars used several informal features. Primarily, broad references or unattended references amounted to 28.59% of the total use of informal features. Broad references, such as *this*, *these*, *that* and *those*, are deemed informal as they regularly occur in various conversational situations. Strauss (1993) and Swales (2005) have shown that broad references appeared in 40% of the spoken language and 46% of dissertation discussions, respectively. In our study, broad reference was found in all the files implying that this feature is highly preferred by Arab scholars. Hyland and Jiang (2019) suggested that research style manuals recommended avoiding the use of broad references. However, Swales and Feak (2012) explain that the employment of broad references, such as ‘this’, serves sweeping from one sentence to another in texts. Thus, it appears that Arab scholars tend to facilitate the flow of information in their texts by applying this feature. The occurrence of broad references in ASAWEC accords with Swales and Feak’s (2012) proposition that use of broad references in academic writing indicates the writers’ position as experienced, skilful, and impressive. The result is compatible with the results by Tocalo et al. (2022), Sholihah (2018) and Ebrahimi & Fakheri (2019) who found that the corpora they studied commonly used unattended references.

Initial conjunctive adverbs and conjunctions are seen as informal since they represent features of spoken discourse, revealing incomplete sentences and unintentional speech (Hyland & Jiang 2019). This informality aspect is responsible for 25.31% of the total informal features in ASAWEC. This result implies enhanced rhetorical perception of Arab scholars related to cases of stylistic assortment accuracy. Moreover, the finding demonstrates that employment of initial conjunctive adverbs and conjunctions acts as an achievement of writing techniques illustrating a move in academic writing. Our finding accords with that of Tocalo et al. (2022), Ebrahimi et al. (2019), Alipour & Nooreddinmoos (2018) and Sholihah (2018). They independently reported that initial conjunctions and conjunctive adverbs were frequently used in their corpora.

The third most common informal feature in ASAWEC is personal pronouns amounting to 12.85% of the total features as presented in Table 3. This result is enhanced by the high dispersion of use, that is 73%, which means that personal

pronouns appear in 112 out of 153 files. Hyland & Jiang (2019) state that first-person pronouns outline informality as their use denotes a more common tendency in all written genres. In academic writing, non-native English speakers may use personal pronouns for reasons related to their language background. According to Hyland and Jiang (2019), employment of first-person pronouns to recognise authors does not imply personal power. This implies that personal pronouns' use in this context is not about confirming authority or dominance, but rather about acknowledging the contributions of others. This perspective may help to understand and interpret the use of personal pronouns in academic writing by non-native English speakers. Correspondingly, the results reveal Arab scholars' self-awareness as linguists. Hyland (2001) emphasises that the employment of *I* builds a voice and influential personality in academic research. Arab scholars' use of the first-person pronouns, therefore, signals their inclination toward showing their personality in their texts. This finding contradicts other research results. Tocalo et al. (2022), Pramintih et al. (2018) did not find the use of the first-person pronouns; rather, Tocalo et al. (2022) found that the first-person pronouns were the least used informality features.

Regarding the dispersion of informality features across the ASAWEC, it is noted that the frequency of applying informality features by Arab scholars is not always related to the distribution of that use. For example, while split infinitive is one of the least frequent features in the corpus, with only two hits per 10000 tokens, they are nevertheless the third used by the authors, i.e., they appeared in 98 out of 153 files which represent 64 % of the whole corpus. This indicates that while a considerable number of the authors used such a feature, the authors rarely repeat a feature once or twice in the same article.

The current study identified a difference in the number of informality features used by Arab scholars and native speakers. Overall, informality items appear more in BAWEC corpus than in ASAWEC. It is noted that the most common features in both corpora are the same. Nevertheless, the two corpora are slightly different in the least used features. This result indicates that the ASAWEC corpus underuses informality features that may be attributed to several reasons. The first one is the difficulty created by the gradual use of informal features in academic writing for students and inexperienced writers, especially those who write in English as a foreign language (Hyland & Jiang 2019). Chang and Swales (1999) propose that when EFL students learn the regulations of formal academic writing, this acts as a burden by itself which was further made complex by allowing the mixing formal and informal language. The second reason of ASAWEC's underuse of informality features is the influence of the authors' field and the perception of readers. This is because applied linguistics quality of academic writing tends to be less informal due to the attention they pay to their readers' expectations (Hyland & Jiang 2019). The third reason is that journal reviewers require less use of informal features. Whenever they find some of the informal aspects, they demand them to be reviewed. Another reason related to some of the informality features, such as sentence-initial conjunctions and conjunctive adverbs, and broad references is that

they are closely related to lexical bundles, and they are nearly idioms. Adel and Emran (2012) argued that these lexical bundles are liable to occur in native speakers' discourse more frequently than in non-natives'. Relating to the context of this study, Sanosi (2022) also found that native speaker writers outscored Arab scholars in using lexical bundles. The fourth reason refers to the non-native speakers' following of rules laid out by style manuals and author guides. They are more worried about the accurate formation of the language in their texts (Alipour & Nooreddinmoosa 2018). Although this worry is beneficial in that it assists in producing accurate discourse, it could be a drawback in the sense that it denotes that Arab scholars do not cope with the new inclinations of research. Updated EPA courses play a role in the diminishing use of informality features. Most EAP courses follow conventional instructional materials and methods. They need to cope with the latest trends in the field (Alipour & Nooreddinmoosa 2018). This result is compatible with that by Alipour and Nooreddinmoosa (2018). They reported that informality features appeared more frequently in the British scholars' corpus than in the non-native one.

It is worth mentioning, however, that the difference in employing informality features between the two corpora does not apply all features. Although significant difference occurs in 63 % of the studied articles, the two groups seem to use broad references, final prepositions, and listing expressions to approximately parallel extents, i.e., with no statistically significant difference. Considering that the investigated articles of ASAWEC have already been reviewed for writing mechanics, among other requirements for publication, it is envisaged that significant use of informal features might be adjusted by reviewers prior to publication. This observation should be acknowledged as a potential limitation of this study. Nevertheless, the data is still believed to be reflective of the reality of the use of informal features by Arab scholars since these review processes normally focus on specific features and cannot entirely change the authors' style. Based on this observation, future research should study the writing production of the scholar *as it is* or before the submission. Although this method is less convenient and may be difficult to apply on a wide range, it is more representative of scholars' employment of informality features in their academic writing.

It is mentioned that journal reviewers may not be in favour of use of informal features while the Chief Editors might recommend to their reviewers the use of informality aspects. Notwithstanding, we need to consider what was noted by Chang and Swales (1999) that foreign writers of English find it difficult to adopt informality in their writing. This may hinder the general movement of applying informality features in academic writing which is noted to be gradually dominating academic writing. This gradual increase in informality features' use is motivated by changes in the ways authors seek to assert unity with their audience. Another reason is the increase of impersonality representation as a rhetorical option in academic writing (Hyland & Jiang 2019). This is generally linked to changes in rhetorical practices that support authors' attempts to make their ideas clear to the audience. Thus, the use of informality features does not only imply that authors are adopting

the latest developments in academic writing, but it also indicates that they are adopting the changes in rhetorical practice. This is supported by Boginskaya's (2023) suggestion that academic language tends to be more impersonal because it is an outcome of "the social interaction between the writer and the reader" (p. 143).

6. Conclusions

Due to the increase in informality features use in scholarly writing, they have become an issue of interest to researchers (Boginskaya 2023). This study investigated informal aspects of academic writing as used by Arabic and British scholars. The results demonstrated a significant difference in informality aspects used in ASAWEC and BAWE. Results have shown that ASAWEC used most frequently broad references, initial conjunctive adverbs and conjunctions, and the first-person pronouns. Additionally, the most common informality features in both corpora are similar: broad references, initial conjunctive adverbs and conjunctions, and the first-person pronouns. Again, the two corpora are identical in the least regular informal features which were final prepositions, listing expressions, and split infinitives.

These results call for academic writing course revision in Arab universities to adapt to the latest trends in academic writing, which is the adoption of informality features in academic writing that has started to replace formality (Hyland & Jiang 2019, Praminath et al. 2018, Adel 2008).

This study is not without limitations that entail some future research recommendations. First, it studied a variety of areas in the English language, such as applied linguistics and literature. Further research is needed to investigate informality features in only one area in the English language. Second, the study selected published articles in local Arab journals. Including research published by Arab scholars in journals outside the Arab region may provide more insights into the topic. Third, the present study did not cater for style manuals and guides in Arab universities. Further research is strongly recommended to investigate these manuals to see if they reflect with the recent developments in the field of academic writing and whether they have an impact on scholars' writing.

Acknowledgements

This study is supported by funding from Prince Sattam bin Abdulaziz University (project number PSAU/2023/R/1444).

References

- Adel, Annelie. 2008. Metadiscourse across three varieties of English: American, British, and advanced learner English. In Connor Ulla, Ed Nagelhout & William Rozycki (eds), *Contrastive Rhetoric: Reaching to intercultural rhetoric*. Amsterdam: John Benjamins.
- Adel, Annelie & Britt Emran. 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>

- Alipour, Mohammad & Musa Nooreddinmoosa. 2018. Informality in applied linguistics research articles: Comparing native and non-native writings. *Eurasian Journal of Applied Linguistics* 4 (2). 349–373. <http://dx.doi.org/10.32601/ejal.464196>
- Atkins, Sue, Clear Jeremy & Ostler Nicholas. 1992. Corpus design criteria. *Literary and Linguistic Computing* 7 (1). 1–16. <https://doi.org/10.1093/lc/7.1.1>
- Bennett, Karen. 2009. English academic style manuals: A survey. *Journal of English for Academic Purposes* 8 (1). 43–54. <https://doi.org/10.1016/j.jeap.2008.12.003>
- Biber, Douglas. 1988. *Variation Across Speech and Writing*. New York: Cambridge University Press. <https://doi.org/10.1017/CBO9780511621024>
- Boginskaya, Olga. 2023. Interactional metadiscourse markers in English research article abstracts written by non-native authors: A corpus-based contrastive study. *Ikala* 28 (1). 139–154. <https://doi.org/10.17533/udea.ikala.v28n1a08>
- Boginskaya, Olga. 2022. Functional categories of hedges: A diachronic study of Russian research article abstracts. *Russian Journal of Linguistics* 26 (3). 645–667. <https://doi.org/10.22363/2687-0088-30017>
- Chang, Yu -Ying & John Swales. 1999. Informal elements in English academic writing: Threats or opportunities for advanced non-native speakers? In Christofer Candlin & Ken Hyland (eds.), *Writing: Texts, processes and practices*. London: Longman.
- Ebrahimi, Seyed Foad & Seyed Aqil. 2019. Features of informality in applied linguistics research articles published in Iranian local journals. *Romanian Journal of English Studies* 16. 135–143. <https://doi.org/10.1515/rjes-2019-0017>
- Fendri, Emna & Mounir Triki. 2022. The negotiation of authorial persona in dissertations literature review and discussion sections. *Russian Journal of Linguistics* 26 (1). 51–73. <https://doi.org/10.22363/2687-0088-27620>
- Gilbert, G. Nigle & Mulkay Micheal. 1984. *Opening Pandora's Box: A Sociological Analysis of Scientific Discourse*. Cambridge: Cambridge University Press.
- Heylighen, Francis & Jean-Marc Dewaele. 1999. *Formality of Language: Definition, Measurement, and Behavioural Determinants*. Internal Report, Center 'Leo Apostel', Free University of Brussels.
- Hyland, Ken. 2001. Humble servants of the discipline? Self-mention in research articles. *English for Specific Purposes* 20 (3). 207–226. [https://doi.org/10.1016/S0889-4906\(00\)00012-0](https://doi.org/10.1016/S0889-4906(00)00012-0)
- Hyland, Ken. 2005a. *Metadiscourse*. Amsterdam: Continuum.
- Hyland, Ken. 2005b. Stance and engagement: A model of interaction in academic discourse. *Discourse Studies* 7 (2). 173–192. <https://doi.org/10.1177/1461445605050365>
- Hyland, Ken & Feng Jiang. 2017. Is academic writing becoming more informal? *English for Specific Purposes* 45. 40–51. <https://doi.org/10.1016/j.esp.2016.09.001>
- Hyland, Ken & Jiang Feng. 2019. *Academic Discourse and Global Publishing*. London: Routledge.
- Lee, Joseph B. & James D. Maxwell. 2019. Breaking the rules? A corpus-based comparison of informal features in L1 and L2 undergraduate student writing. *System* 80. 143–1153. <https://doi.org/10.1016/j.system.2018.11.010>
- Kuhi, Davud, Sharghinezhad Behuaz & Rezaei Shirin. 2020. Informality in academic discourse: A cross-cultural and cross-disciplinary investigation. *The Asian ESP Journal* 16 (3). 180–1208.
- Leedham, Maria. 2015. *Chinese Students' Writing in English: Implications from a Corpus-driven Study*. London: Routledge.
- Mair, Christian. 1990. Corpora and the study of the major varieties of English: Issues and results. In Hans Lindquist, Staffan Klintbrog, Magnus Levin & Maria Estling (eds), *The major varieties of English: Papers from MAVEN 97, Växjö*. Växjö: Edman and Westerlunds Tryckeri.

- Matsuda, Paul K. & Nouri Melika. 2020. Informal writing and language learning. In Mark Dressman & William R. Randall (eds), *The handbook of informal language learning*. London: John Wiley & Sons Ltd.
- McEnery, Tony & Andrew Hardie. 2012. *Corpus Linguistics: Method, Theory and Practice*. Cambridge: Cambridge University Press.
- Poole, Robert. 2018. *A Guide to Using Corpora for English Language Learners*. Edinburgh: Edinburgh University Press.
- Pramatih, Gusti A., Deny A. Kwary & Viqi Ardaniah. 2018. Is EFL students' academic writing becoming more. *Journal of World Languages* 5 (2). 101–112. <https://doi.org/10.1080/21698252.2019.1570664>
- Richter, Kenneth G., Behruz L. 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>
- Sanosi, Abdulaziz B. 2022. The use and development of lexical bundles in Arab EFL writing: A corpus-driven study. *Journal of Language and Education* 8 (2). 108–123. <https://doi.org/10.17323/jle.2022.10826>
- Seoane, Elena & Lucia Loureiro-Porto. 2005. On the colloquialization of scientific British and American English. *ESP Across Cultures* 2. 106–118.
- Sholihah, Zaidatus. 2022. *The use of informal language in academic writing by the students of English Teacher Education*. Department at Uin Sunan Ampel Surabaya. Surabaya, Sunan Ampel State Islamic University dissertation.
- Strauss, Sausan. 1993. Why 'this' and 'that' are not complete without 'it'. In Katharine Beals, Gina Cooke, David Kathman, Sotaro Kita, Karl-Eric McCullough & David Testen (eds), *CLS 29: Proceedings of the 29th regional meeting of the Chicago Linguistic Society*. Chicago: Chicago Linguistic Society.
- Swales, John M. 2005. Attended and unattended 'this' in academic writing: A long and unfinished story. *ESP Malaysia* 11. 1–15.
- Swales, John & Christine B. Feak. 2012. *Academic Writing for Graduate Students: A Course for Non-native Speakers* (3 Ed.). Michigan: University of Michigan Press.
- Tocalo, Abdul Wahid, Racman Shainah P., Guiamelon Kyla, Nicole Alibai & Mama Bainrin M. 2022. Is Filipino ESL academic writing becoming more informal? *PASAA Journal* 64. 99–125.
- Tognini-Bonelli, Elena. 2001. *Corpus Linguistics at Work*. Amsterdam: John Benjamin's.
- Yang, Yiyang & Fan Pan. 2023. Informal features in English academic writing: Mismatch between prescriptive advice and actual practice. *Southern African Linguistics and Applied Language Studies* 41 (2). 102–119. <https://doi.org/10.2989/16073614.2022.2088579>

Appendix

Informality features in ASAWEC and BAWE

Feature	ASAWEC		BAWE	
	Rel. Freq*	Range %	Frequency	Range %
Broad References	28.59	153	31.27	176
Initial conjunctions	25.31	38.30	32.58	35.34
First-person pronouns	16.02	73.20	26.03	59.60
Second-person pronouns	6.35	64.05	3.41	36.36
Final prepositions	2.40	7.68	2.86	7.17
Listing expressions	1.97	21.79	1.54	11.93
Split infinitives	1.94	64.05	3.43	75
Contractions	1.87	26.14	5.13	58.52

* Relative frequency per 10,000 words.

Article history:

Received: 09 October 2023

Accepted: 25 August 2024

Bionotes:

Abuelgasim S.E. MOHAMMED is an Assistant Professor of Applied Linguistics at Prince Sattam bin Abdulaziz University in Saudi Arabia with 30 years teaching experience. His research interests include Applied Linguistics, English for Specific Purposes, English for Academic Purposes, materials design and corpus linguistics.

e-mail ab.ibrahim@psau.edu.sa

<https://orcid.org/0000-0001-9791-0905>

Abdulaziz B. SANOSI is a lecturer at the Department of English Language at Prince Sattam bin Abdulaziz University in Saudi Arabia. He teaches applied and theoretical linguistics courses. He is also a certified eLearning trainer and frequently provides training on eLearning applications and tools. His research interests include Corpus Linguistics, Discourse Analysis and Computer Assisted Language Learning.

e-mail a.assanosi@psau.edu.sa

<https://orcid.org/0000-0003-3447-2818>

Сведения об авторах

Абуэльгасим С.Э. МОХАММЕД – доцент кафедры прикладной лингвистики в Университете принца Саттама бин Абдулазиза, Саудовская Аравия. Имеет 30-летний опыт преподавания. В сферу его научных интересов входят прикладная лингвистика, английский для специальных целей, дизайн материалов и корпусная лингвистика.

e-mail: ab.ibrahim@psau.edu.sa

<https://orcid.org/0000-0001-9791-0905>

Абдулазиз В. САНОСИ – преподаватель кафедры английского языка в Университете принца Саттама бин Абдулазиза, Саудовская Аравия. Преподает прикладную и теоретическую лингвистику, проводит тренинги по приложениям и инструментам электронного обучения. Его научные интересы включают корпусную лингвистику, анализ дискурса и компьютерное обучение иностранным языкам.

e-mail: a.assanosi@psau.edu.sa

<https://orcid.org/0000-0003-3447-2818>