ISSN 2411-7390

jle

JOURNAL OF LANGUAGE & EDUCATION

Volume 10 Issue 3, 2024

HIGHER SCHOOL OF ECONOMICS
NATIONAL RESEARCH UNIVERSITY

EDITOR-IN-CHIEF

EDITORIAL BOARD

EDITORIAL TEAM

CONTENT

EDITORIAL

RESEARCH PAPERS

REVIEW PAPERS

BOOK REVIEWS

https://doi.org/10.17323/jle.2024.23769

Data Commentary in Research Publications: A Scoping Review

Elena Tikhonova [1, 2], **Lilia Raitskaya** [3]

[1] HSE University, Moscow, Russian Federation

[2] RUDN University, Moscow, Russian Federation

[3] MGIMO University, Moscow, Russian Federation

ABSTRACT

Introduction: The modality of scholarly written communication offers verbal and visual modes that are supposedly to be interrelated. The verbal component is thoroughly studied whereas the visual, including tables, charts, graphs and others are mainly described from a technical perspective. The editorial review aims to synthesize research on data commentary as a comparatively new rhetorical genre.

Method: The review adhered to the PRISMA extension for scoping reviews, and the PPC framework. The eligibility criteria include problem, concept, context, language, types of sources, databases (Scopus, SpringerLink). The time range of the search was subject to availability of indexed publications.

Results: The searches and consequent screening by titles, abstracts, keywords and full texts identified 19 relevant publications. The books on academic writing focused on the three themes: general guidelines on data commentary, functions, and rhetorical structure of such commentaries, figure legends and notes to tables. The functional analysis of the publications singled out a prevailing framework, including rhetorical, ontological, epistemological, and argumentative functions. The rhetorical steps in data commentaries follow patterns suitable for the type of a visual. The review outlines generic steps described and proved by the research publications included in the review.

Conclusion: Though the research field is scattered, and no definite trends were specified, the potential of the field is rising as the implications of such studies are significant. A constant trend towards more visualization of the new knowledge requires more research on the interrelations between the verbal and the visual, with a special accent on data commentary.

KEYWORDS

data commentary, rhetorical steps in data commentary, tables, figures, rhetorical functions

INTRODUCTION

Any act of communication may involve various means to transmit the message (Johns, 1998). In written communication, besides the text as its basis, visual objects considerably add to the modality of communication (Hemais, 2014). The visual components in research publications entailing figures, tables, and other visual materials (Miller, 1998) communicate condensed information to the readers (Parija & Kate, 2017). Royce (2002) points out that the verbal and the visual "create meanings in a complimentary mode". A synergistic interaction of verbal and visual is emphasized (Gross & Harman, 2014).

New knowledge production is overwhelmingly disseminated via research publications in peer-reviewed journals1. The way new knowledge is presented is essential to be correctly understood and perceived as verbal and visual modes create complementary meanings (Hemais, 2014). Being an integral part of research tradition, visual thinking is converted into visual language accompanying verbal mode to communicate reasoning and

Citation: Tikhonova, E., & Raitskaya, L. (2024). Data commentary in research publications: A systematic scoping review. *Journal of Language and Education, 10*(3), 5-24. https://doi.org/10.17323/jle.2024.23769

Correspondence:
Elena Tikhonova,
etihonova@hse.ru

Received: September 8, 2024

Accepted: September 16, 2024

Published: September 30, 2024

[1] PublishingState.com. Academic Journals: The Pillar of Academia. https://publishingstate.com/academic-journals-pillar-of-academia/2023/

research results to the reader (Cocchiarella, 2015, p. VI). For years, researchers have occasionally turned to relationships between verbal and visual components in the text, including Kress & van Leeuwen (2006) who based their research on the functional grammar theory (Halliday, 1985), visual persuasion and presentation of data (Miller, 1998; Miller, 2004), distribution of visual representations across scientific genres (Tang, 2023).

Though visual presentation of data in research publications entails heterogenous forms (Ariga & Tashiro, 2022), we are to focus essentially on tables and charts of all types as research publications tend to be saturated with tables, figures, and other illustrations of the kind that occupy "from one third to one half of the page" of an average research article (Miller, 1998)[2]. Our special interest lies in what is termed as "data commentary" (Swales & Feak, 2012). The genre, or the type of the text, where such a commentary is placed determines the specificity of data commentary, its length and rhetorical steps it follows (Swales & Feak, 2012). Whereas verbal rhetoric of scholarly publications has been thoroughly studied, starting with the profound work by John Swales (1990), research on data commentary is quite scattered and seems occasional. It is considered as "a relatively new genre" (Parviz & Lan, 2023).

Data commentary (DC) in research or review articles includes paragraphs in the scholarly text preceding or following a visual object as well as notes and legends relating to tables, figures, and other visuals (Swales & Feak, 2012). The significance of further research on DC is rooted in its implications for Academic Writing as a discipline and journal guidelines. Not each academic writing course focuses on rhetoric of data commentary as most instructors prefer to dwell upon the verbal components of the text as DC created in a multimodal context that is quite a challenge not only for student writers but for experienced researchers and instructors (Parviz & Lan, 2023). The reasoning behind selective attention to visual rhetoric is also linked to various limitations within academic writing courses (with a course time limit as the frontrunner), and many other challenges a novice researcher faces to be primarily met.

The emerging field of research on DC requires reviewing for setting the scope of the research area and specifying gaps in the knowledge. To this end, no review has been found. This editorial review aims to synthesize research on data commentary in scholarly publications in peer-reviewed journals and academic writing books and identify the scope and cohesion of the field.

To attain the aim, we are to answer the following review questions:

RQ#1: What are the prevailing themes in the research field?

RQ#2: What guidelines do academic writing books offer regarding commentary on visuals?

RQ#3: What functions do data commentaries perform in research publications?

RQ#4: What rhetoric steps do researchers outline regarding data commentary in various contexts?

METHOD

Protocol

While starting the present scoping review, we meticulously developed a research protocol. The authors hereby certify that this review report constitutes a faithful, precise, and transparent description of the conducted review. No deviations from the protocol were registered. Any departures from the original study design have been duly elucidated. This scoping review stick to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) extension for Scoping Reviews (Tricco et al., 2018) and the framework proposed by Arksey and O'Malley (2005).

Search Eligibility Criteria

In the review, the problem (population), concept, and context (PCC) framework was applied to devise an effective search strategy where each criterion was justified (Table 1).

Search Strategies

The search to attain the aim and to reply to the review questions was conducted in a two-stage mode. At the first stage, the Scopus and SpringerLink databases were thoroughly searched to identify relevant publications. The search was conducted using a range of the keywords: "data commentary", "commentary on data", and "visual commentary" as of September 2, 2024. All other potential keywords were applied in pre-protocol searches but failed to bring any results.

At the second stage, the search in the refence lists was done in the publications selected from the Scopus and SpringerLink databases after screening of the titles, abstracts and full texts. The full-text publications eligible for the review were identified after screening. In addition, after all relevant books on academic writing were selected on the SpringerLink, the authors sifted their reference lists to find more relevant publications. Those with full texts were included in the review.

[2] This claim is true regarding some sections of research articles, including the results and discussion sections, and occasionally the introduction and method section subject to disciplines.

Table 1

Eligibility Criteria

Criterion	Inclusion	Exclusion	Rationale
Problem	Data commentary	All publications that do not deal with data commentary	The review focuses on data or visual commentary. The problem is defined as the interpenetration of the visual and the verbal in a scholarly text
Concept	Rhetorical concept applicable to data commentary (or commentary on visuals)	Other concepts	The aim of the review is to identify the scope and recent trends of rhetorical and other relevant research on data commentary
Context	Relevant research articles on rhetoric of data commentary and academic writing books	Other concepts	The review dwells upon the rhetoric of data commentary
Language	English	Other languages	The object of all research in focus is scholarly publications in English. The language choice is also identified by its status as a lingua franca of international science.
Time period	All publications available in the database	N/A	The pilot pre-protocol searches found that the relevant publications were scattered across a long period of time, starting from the late1990s
Types of sources	In the Scopus database: full texts of articles, reviews, conference papers, books, and book chapters; In the SpringerLink: books	Unavailable sources, unavailable full texts	This review aims to get a comprehensive understanding of the field
Geographical location	Any location	None	Getting international perspective
Database	Scopus SpringerLink	Other bases than Scopus and SpringerLink	The Scopus and SpringerLink databases were selected as two of the biggest covering publications related to visual rhetoric
Areas of Research	All	N/A	As the review focuses on the rhetorical concept, publications rarely go beyond communication and linguistics, studying rhetorical specificity in research across all sciences ultimately may be classified otherwise

Study Selection

Both authors identified research publications and books subject to the eligibility criteria enumerated in Table 1. After the Scopus and SpringerLink filters (language and types of sources) had been applied, each reviewer independently screened the titles, and then the abstracts of the identified documents. The publications were subsequently tagged by each reviewer with “to include” or “to exclude” marks. When occasional disagreements arose, they were settled by mutual consent. No disputed issue required lateral expertise.

The full texts were found via the publishers or at request applying to the authors of the publications. Each full text was profoundly read and independently analysed by each reviewer. Relevant publications were identified. The relevant publications found in the reference list of the selected studies were also included subject to full text.

Data Extraction

With title and review questions determined under the PCC framework, pre-protocol pilot searches made us identify the basic structure of the extracted data we would require:

1. Data from the selected books on academic writing related to commentary on data and visuals, its rhetoric functions and steps;
2. Data from the reviewed publications regarding commentary on tables, figures and other data presented in a visual mode;
3. Data from the articles and books under review containing any reference to rhetorical steps in commentaries on data.

All raw data were double-checked by the authors.

RESULTS

Search and Selection Results

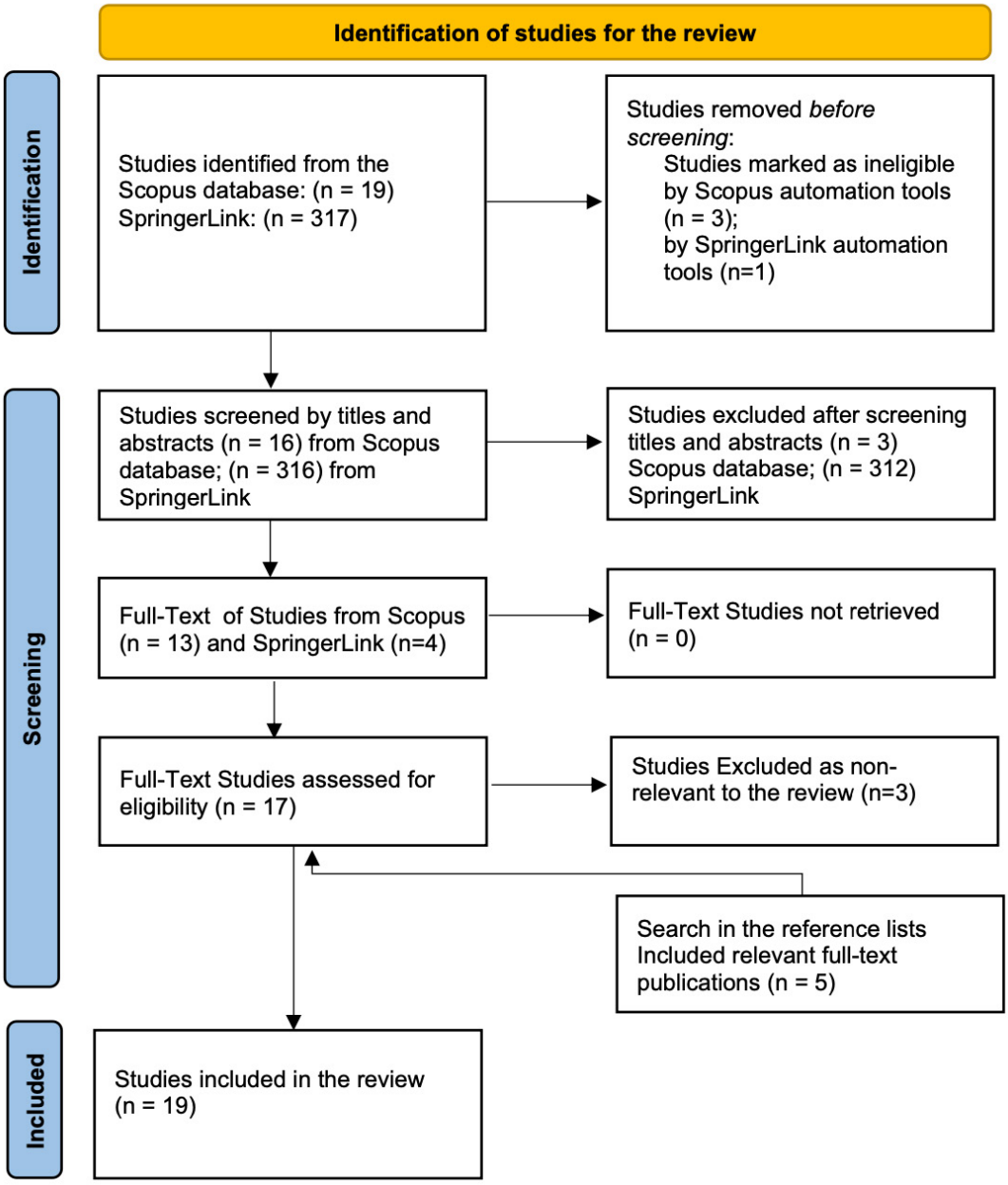
The search results were fixed as of September 2, 2024. A total of 336 documents were initially found, including 19 records in the Scopus database, and 317 records in the SpringerLink. After applied filters (language; in the SpringerLink – language and type of publication), the total decreased from 336 to 332 studies that were eligible for title and abstract screening. After the title and abstract screening, 315 documents were deemed irrelevant and excluded, including 312 books from the SpringerLink and three articles from the Scopus database. The remaining seventeen publications had full texts. After full-text publications had been screened, three article were not included in the final analysis. A thorough search in the reference lists of the selected full-text books and articles brought another three full-text books and two full-text articles. The total of 19 publications was finally selected for further analysis. The PRISMA flow-chart (Figure 1) depicts the whole identification and screening procedure.

Characteristics of the Research Field

The publications indexed in the Scopus database included ten research articles and one book. The timeline for the selected records began with 1998 and ended in 2024, though

Figure 1

Selection of Publications for the Review



the distribution of publications was uneven: 1998, 2012, 2014, 2019, and 2021 with one record each, 2018, 2023 and 2024 with two records. Five publications appeared in the *English for Specific Purposes Journal*; two were published in *ESP Today. Journal of Contemporary Ethnography, Journal of Language and Education, Research in Science & Technological Education,* and *Journal of English for Academic Purposes* brought out one article each. The sampling entailed one book (Swales & Feak, 2012). The eleven publications were authored by 24 researchers, 2.2 authors per record. The authors had twelve affiliations, with Jilin University as the frontrunner (two authors). Most authors were from China (4), Canada (2), and the USA (2). Brazil, Hong Kong, Sweden and Iran accounted for one author each. The affiliation of one author was unidentified. Social Sciences and Arts & Humanities accounted for 52.4 % and 47.6 % of the publications respectively.

The 316 books out of the 317 books were selected from the SpingerLink database as they were published in English. Screening eliminated 312 books as they had no mentioning of data commentary or a description of tables, chart or other visuals.

Both the pre-protocol searches and selection of studies of the review proved that the Scopus and SpringerLink databases had few indexed publications on the rhetorical aspects of commentaries on tables and figures (otherwise termed as "data commentaries"), an effort to single out thematic clusters remained a supplementary task subject to any emerging trends that would be detected during the review process. The total number of the articles and books under review finally worked out at 19 (Table 2). Thus, the ultimate number of the publications was too limited to specify any trends. The field did not show any growth patterns and may be considered next to non-existent at present.

Table 2

Themes of the Publications under Review

Publication	Type of Publication Indexed in Database (if any)	Theme
1. Dougherty & Ilyankou, 2021	Book	Academic writing Data storytelling
2. Du et al., 2021	Article Scopus	DC Figure legends
3. Eriksson & Nordrum, 2018	Article Scopus	Data commentary in master's thesis projects
4. Graves, 2014	Article	Academic writing Visual functions
5. Hemais, 2014	Article Scopus	Visuals in marketing articles
6. Jagadeesh et al., 2023	Book SpringerBook	Academic writing General guidelines on tables and charts DC
7. Johns, 1998	Article Scopus	The visual and the verbal in macroeconomics
8. Lövei, 2021	Book	Academic writing DC Notes to tables Figure legends
9. Lui et al., 2023	Article Scopus	Figure legends (rhetorical moves and phrase frames)
10. Miller, 2004	Book	Academic writing Notes to tables
11. Moghaddasi et al., 2019	Article Scopus	Visual moves in mathematics research articles
12. Parija & Kate, 2017	Book SpringerLink	Academic writing Tables and graphs
13. Parviz & Lan, 2023	Article Scopus	Commentaries on visuals (rhetorical moves and phrasal complexity)
14. Swales & Feak, 2012	Book Scopus	Academic Writing Data commentary (structure) General guidelines on DC

Publication	Type of Publication Indexed in Database (if any)	Theme
15. Van den Scott, 2018	Article Scopus	Visual methods in ethnography
16. Wallwork & Southern, 2020	Book SpringerBook	Academic writing Results. Tables
17. Wallwork, 2023	Book SpringerBook	Academic writing Commentary on tables and figures. Legends and captions
18. Wu et al., 2024	Article Scopus	Commentaries on visuals (rhetorical moves and phrase frames)
19. Zhang et al., 2024	Article Scopus	Data commentary in discourse of economics

Table 2 offers a list of the publications under review and a description of their themes. While sifting the publications (research articles, conference papers, book chapters, and books) found in the Scopus and SpringerLink databases both during the pre-protocol pilot and review searches, we eliminated records on general issues of academic writing and rhetoric moves and steps (genres) of scholarly publications as well as quite many publications relating to technical aspects of data presentation in science. Those publications did not dwell upon data commentary or related issues.

Instructions on Data Commentary in Research Publications in Academic Writing Books

To study today's approaches introduced in Academic Writing as a discipline, we extracted and analysed the raw data from the four academic writing books published by Springer and another four books brought out by other publishers. The choice of Springer books was determined by the authors' personal experiences in teaching Academic English and Academic Writing at several universities. Those books are not merely textbooks but may be classified as comprehensive manuals for researchers. Those four books were the result of our final selection after we had screened 316 books on academic writing and failed to find any mentioning of rhetorical structure or functions of data commentary in 312 of them. Most of the books had no sections on data presentation. The books under review were analysed to extract the data (Table 3).

All academic writing books under review contained instructions on data commentary. The form and scope of the instructions followed various patterns and had individual focuses. According to the raw data (Table 3), the purposes of data commentary that are covered by the books included in the review may be boiled down to highlighting the results, supporting an argument, assessing theory or data reliability, comparing data, evaluating data, discussing data implications, and making recommendations (Swales & Feak, 2012). Most books offered a DC structure (Swales & Feak, 2012; Jagadeesh et al., 2023; Wallwork, 2023; Miller, 2004). The books specified the Results Section as the most appropriate for DC (Wallwork, 2023; Wallwork & Southern, 2020). Two of the books indicated that a narrative should interpret pictures about figures (Dougherty & Ilyankou, 2021; Miller, 2004).

Less attention was paid to legends and table notes. Wallwork (2023) states that legends, being short, are to be self-explanatory. No double presentation and comprehensive information in legends are prescribed by Lövei (2021). Miller (2004) concentrates on a distribution of information among the title, column and row labels, and notes to tables. It is the only book in our review where we found instructions relating to description of variables by types (specifying direction and magnitude of association, considering statistical significance, considering types of variables, units, and distribution, using quantitative comparisons, and organizing the text to coordinate with a table or chart. Thus, DC was extended to cover any text dealing with data that included DC.

The textual analysis shows that the publications under review fall under the following themes:

general guidelines on DC (Jagadeesh et al., 2023; Wallwork, 2020; Swales & Feak, 2012; Miller, 2004); rhetorical structure of DC (Swales & Feak, 2012; Graves, 2014; Eriksson & Nordrum, 2018); and figure legends and notes to tables (Lövei, 2021; Miller, 2004; Parija & Kate, 2017; Wallwork, 2023; Wallwork & Southern, 2020).

The review found that the extent of the instructions relating to DC was uneven, both in length and depth. The objectives of the books may justify the authors' choices. Moreover, at present, there is no definite and widely accepted standard for chapters on data commentary as compared with other more traditional themes in Academic Writing courses. Further developments in the research field might lead to a more conventional pattern in the instructional literature on DCs.

Table 3

Instructions on Data Commentary in Research Publications in Academic Writing Books

Publication	Extracted Data
Swales & Feak, 2012	Data Commentary In many disciplines, the data is displayed in a table, graph, figure, or some other kind of non-verbal illustration. The data may come from a source, or it may be the outcome of your own work - that is, your results. ... Like many other aspects of academic writing, data commentaries are exercises in positioning yourself. There are, as a result, both dangers and opportunities. One danger is to simply repeat in words what the data has expressed in non-verbal form - in other words, to offer description rather than actual commentary or interpretation. An opposite danger is to read too much into the data and draw conclusions that are not well supported. The art of the commentary is for you to find the right strength of claim in discussing the data and then to order your statements in some appropriate way (perhaps in order of interest or relevance)...
Swales & Feak, 2012	It is not easy to predict precisely what you might need to do in a data commentary, but some of the more common purposes are to • highlight the results of research • use the data to support a point or make an argument in your paper • assess theory, common beliefs, or general practice in light of the given data • compare and evaluate different data sets • assess the reliability of the data in terms of the methodology that produced it • discuss the implications of the data • make recommendations
Swales & Feak, 2012	Structure of Data Commentary Data commentaries usually have these elements in the following order. 1. location elements and/or summary statements 2. highlighting statements 3. discussions of implications, problems, exceptions, recommendations, or other interesting aspects of the data
Swales & Feak, 2012	Location Elements and Summaries Many data commentary sections begin with a sentence containing a location element and a brief summary of what can be found in a visual display of information... ... location statements direct readers to view important information in a table, chart, graph, or other figure. Even though research indicates that readers often look at the visual information before reading, location statements are expected. They are considered to be a form of metadiscourse-sentences or phrases that help readers make their way through a text by revealing such things as organization, referring readers to relevant parts of a text, or establishing logical connections.
Parija & Kate, 2017	• Tables, illustrations, and graphs represent data in a format that is easy to understand and grasp at a glance. • They are a substitute for, and not an addition to, voluminous descriptions in the body of the article. • Tables are best when there is more text to display and the data is qualitative. They organize data into understandable classifications.
Wallwork & Southern, 2020	Results: Do not write long descriptions of your results if these could easily be put in a table. And do not repeat information that is clearly shown in a table, instead interpret it. If you can put your results in a table, then use a table. Treat the table and the text as two distinct elements: the table provides the information; the text interprets it.
Wallwork & Southern, 2020	Tables: In captions, and when referring to figures and tables, use the least words possible. Begin the sentence with Figure 3 / Table 5 shows / reports / highlights / reveals etc. Remember that if the first word of your sentence is Table 1, Figure 7 etc., the words table and figure need to be written in full. When associated with a number, table, figure etc. require an initial capital letter (Table 2, Figure 3).
Jagadeesh et al., 2023	Substantial duplication of information in text, figures and tables should be avoided. If a flow chart is used, only the key points could be highlighted in the text while referring to the figure...

Publication	Extracted Data
Jagadeesh et al., 2023	Text–table dichotomy It is important to ensure that text and tables are complementary to each other and not merely repetitive. Describing all parameters that are depicted in the table is not required. Only some salient features and concise description in the text is sufficient to inform the reader as to what is described in the tables…
Jagadeesh et al., 2023	Do not insert any table without the corresponding reference in the text. As with tables, a figure should also be self-explanatory with an informative but precise heading. Other components of a figure include legends, data labels, axis titles, etc. It is noteworthy that text, tables and figures serve different purposes in presenting information, however, repetition of data should be avoided. All figures should be cited in the text and numbered in the order of citation/appearance in the manuscript.
Wallwork, 2023	How should I comment on my tables and figures? When writing *Results sections* you should use the tables and figures to illustrate points in the text, rather than making them the subject of your text.
Wallwork, 2023	A typical mistake when writing the main text is to repeat information from the table… When commenting on a table, your job is to: • interpret / discuss the results • bring to the reader's attention anything that is particularly meaningful or significant • add further details that help to explain the results or which enable them to be compared with previous results…
Wallwork, 2023	What about legends and captions? A typical mistake is to repeat word for word the caption / legend to your figures and tables within the main text… They should be as short as possible and be sufficiently detailed to enable your readers to understand the figure or table without having to read your text. It is vital that you pay attention to legends as some readers may only look at your figures and tables, without even reading the paper itself!
Wallwork, 2023	Every figure and table included in the paper MUST be referred to from the text. Use sentences that draw the reader's attention to the relationship or trend you wish to highlight, referring to the appropriate figure or table only parenthetically…
Wallwork, 2023	Avoid sentences that give no information other than directing the reader to the figure or table. Like the title of the paper itself, each legend should convey as much information as possible about what the table or figure tells the reader: • what results are being shown in the graph(s) including the summary statistics plotted • the organism studied in the experiment (if applicable), context for the results: the treatment applied or the relationship displayed, etc. • location (ONLY if a field experiment), • specific explanatory information needed to interpret the results shown (in tables, this is frequently done as footnotes) • culture parameters or conditions if applicable (temperature, media, etc) as applicable, and, • sample sizes and statistical test summaries as they apply.
Dougherty & Ilyankou, 2021	The goal of data visualization is not simply to make pictures about numbers, but also to craft a truthful narrative that convinces readers how and why your interpretation matters.
Lövei, 2021	There are general design rules for figures that are worth mentioning here. The first is the prohibition of double data presentation. A set of data can be presented in only one way – either in text, on a figure, or in a table. Single values and trends can be mentioned and discussed in the text, but larger parts of the whole dataset cannot be presented in more than one way. A second rule is that figures, together with their captions, have to be self-explanatory: the reader should understand what is pictured on the figure, without reference to the text or to other figures. Note that the interpretation of the figure does not need to be given here – that goes into the text …
Lövei, 2021	Tables, just as figures, must also be self-explanatory: collectively, the title, table headings, and footnotes must allow the reader to understand the content of the table, without reference to the text.

Publication	Extracted Data
Miller, 2004	Notes to Tables Put information that does not fit easily in the title, row, or column labels in notes to the table. Spell out abbreviations, give brief definitions, and provide citations for data sources or other background information. To keep tables concise and tidy, limit notes to a simple sentence or two, referring to longer descriptions in the text or appendixes if more detail is needed. If a table requires more than one note, label them with different symbols or letters, rather than numbers, which could be confused with exponents, then list the notes in that order at the bottom of the table following the conventions for your intended publisher... If you are using secondary data, provide a source note to each table, citing the name and date of the data set or a reference to a publication that describes it. If all tables in your article, report, or presentation use data from the same source, you might not need to cite it for every table.
Miller, 2004	As you write about the patterns shown in your tables, proceed systematically, comparing numbers either across the columns or down the rows of your table. To describe both types of patterns, create separate paragraphs for the "down the rows" and "across the columns" comparisons...
Miller, 2004	Writing about numbers often involves portraying the distribution of a variable or describing the association between two or more variables. These tasks require several of the principles and tools introduced in the preceding chapters: specifying direction and magnitude of association (chapter 2), considering statistical significance (chapter 3), considering types of variables, units, and distribution (chapter 4), using quantitative comparisons (chapter 5), and organizing the text to coordinate with a table or chart (chapters 6 and 7).
Miller, 2004	Systematically introduce and explain the numeric evidence in your exhibits — tables, charts, maps, or other diagrams — building a logical sequence of analyses.
Miller, 2004	To describe a table or chart that encompasses more than one type of pattern, organize your narrative into paragraphs, each of which deals with one topic or set of closely related topics. For instance, a de- scription of a chart portraying trends in unemployment over two decades for each of several occupations might be organized into two paragraphs, the first describing trends over time and whether they are consistent for all the occupation categories, the second comparing levels of unemployment across occupational categories at one point in time and whether that pattern is consistent across time.
Miller, 2004	Start each paragraph with a sentence that introduces the topic of that paragraph and generalizes the patterns. Then present numeric evidence for those conclusions. A handful of numbers can be presented in a sentence or two. For more complex patterns, report the numbers in a chart or table, then describe the patterns using the "generalization, example, exception" (GEE) approach. Refer to each table or chart by name as you describe the patterns and report numbers presented therein.

Functions of Commentaries on Data in the Reviewed Research Publications

Many authors (Mishra, 2004; Liu et al, 2023; Moghaddasi et al., 2019; Morell, 2015; Graves, 2014) in the reviewed publications stick to frameworks based on the following major functions: rhetorical, ontological, epistemological, and argumentative functions (Appendix 2). O'Toole (1996) and Miller (1998) also included a compositional function in their analyses. Moghaddasi et al., 2019 cited a publication by Gross and Harmon (2014) who offered iconic, symbolic, indexical functions as a framework. Explaining was considered as a rhetorical function in Moghaddasi et al. (2019). In rhetorical studies, researchers tend to label functions with phrases. In the reviewed studies, they were essentially expressed by infinitives. There was some overlapping in functions. Sometimes, the researchers assigned various meanings to the functions. The prevailing functions described by infinitives entail:

- To announce results (Wu et al., 2024)
- To clarify information (Clymo, 2014; Franzblau and Chung, 2012; Saver, 2006; Zhang et al., 2024; Miller, 1998; Moghaddasi et al., 2019)
- To communicate simplified information (Graves, 2014)
- To compare findings with others (Wu et al., 2024)
- To conceptualize intractable phenomena (Goodwin, 2001; Moghaddasi et al., 2019)
- To connect the study (Wu et al., 2024)
- To consolidate information (Prus, 1987; van den Scott, 2018; Wu et al., 2024)
- To describe experiments (Wu et al., 2024)
- To discuss (Wu et al., 2024)
- To discuss the implications (Swales & Feak, 2012; Zhang et al., 2024)
- To enhance interpretability (Clymo, 2014; Franzblau and Chung, 2012; Saver, 2006; Zhang et al., 2024)
- To establish presumptions (Wu et al., 2024)
- To highlight the more significant information (Swales & Feak, 2012; Zhang et al., 2024)
- To interpret results (Wu et al., 2024)
- To introduce mathematical concepts (O'Halloran, 2010; Moghaddasi et al., 2019)
- To locate data (Swales & Feak, 2012; Zhang et al., 2024)
- To persuade the reader of the validity of the argument (Miller, 1998; Hemais, 2014)

- To popularize a complex reasoning (Du et al., 2021)
- To present results (Wu et al., 2024)
- To prove argument (Miller, 1998; Moghaddasi et al., 2019; Wu et al., 2024)
- To provide an overview between mathematical participants (O'Halloran, 2010; Moghaddasi et al., 2019); background (Wu et al., 2024); evidence (Lui et al., 2023)
- To save space (Clymo, 2014; Franzblau and Chung, 2012; Saver, 2006; Zhang et al., 2024)
- To summarize information (Swales & Feak, 2012; Clymo, 2014; Franzblau and Chung, 2012; Saver, 2006; Zhang et al., 2024; Wu et al., 2024)
- To support the argument (Graves, 2014; Miller, 1998; Hemais, 2014; O'Toole, 1996; Gross & Harmon, 2014; Moghaddasi et al., 2019)

The functions lay the foundation for the moves and steps described in the reviewed publications. The functions expressed by infinitives easily fit into the functions listed above (argumentative, ontological, epistemological, compositional, iconic, symbolic, indexical functions), with nearly all simultaneously being rhetorical functions.

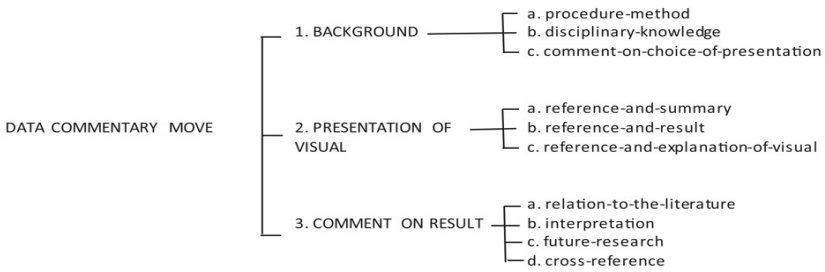
Rhetorical Steps in Data Commentary

The rhetorical steps to follow in DC depend on its type. Judging by the reviewed publications on data commentary at large (Swales & Feak, 2012; Du et al., 2021; Lui et al, 2023; Cargill & O'Connor, 2013; Parviz & Lan, 2023; Wu et al., 2024), the rhetorical steps were approached differently (Table 4). Swales and Feak (2012) outline steps within a DC that precedes or follows a table or a chart, limiting those steps to location elements, highlighting statements and discussion of data aspects. In the same vein, Parviz and Lan (2023) outline the rhetorical steps that are patterned in the data commentaries of students, elaborating some of the steps by Swales and Feak (2012). Location elements were merged with the presention of visual information, discussion of data

Table 4

Rhetorical Steps Typical of Commentary on Data

Publication	Extracted Data	References
Swales & Feak, 2012	Structure of Data Commentary Data commentaries usually have these elements in the following order. 1. location elements and/or summary statements 2. highlighting statements 3. discussions of implications, problems, exceptions, recommendations, or other interesting aspects of the data	
Du et al., 2021	...in science writing figure legends have a general form with five parts which usually occur in sequence: (1) A title which summarizes what the figure is about; (2) Details of results or models shown in the figure or supplementary to the figure; (3) Additional explanation of the components of the figure, methods used, or essential details of the figure's contribution to the results story; (4) Description of the units or statistical notation included; (5) Explanation of any other symbols or notation used.	Cargill & O'Connor, 2013
Liu et al., 2023	...in science writing, figure legends have a general form with five parts, which usually occur in sequence (p. 31): (1) A title that summarizes what the figure is about. (2) Details of results or models shown in the figure or supplementary to the figure. (3) Additional explanation of the figure's components, methods used, or essential details of the figure's contribution to the results. (4) Description of the units or statistical notation included. (5) Explanation of any other symbols or notation used.	Cargill & O'Connor, 2013
Liu et al., 2023	Moves 1. Title 2. Account of experimental details 3. Definition of graphic items 4. Reporting of statistical information 5. Reference of sources of data 6. Result statement 7. Interpretation of results	

Publication	Extracted Data	References
Liu et al., 2023	Nature advises authors that "each figure legend should begin with a brief title for the whole figure and continue with a short description of each panel and the symbols used."[1] Similarly, PLOS ONE requires legends to "describe the key messages of a figure: provide a description of the figure that will allow readers to understand it without referring to the text" and "define all non-standard symbols and abbreviations." [2]	Nature PLOS ONE
Parviz & Lan, 2023	Rhetorical Functions Found in Data Commentary Move 1: Presenting Visual Information Step 1: Providing an explanatory note to set the scene Step 2: Indicating the location of the data Move 2: Highlighting Visual Information; Comparing and Contrasting Key Points Step 1: Describing the facts (with/without providing statistical evidence) Move 3: Commenting on Visual Information Step 1: Personal asides Move 4: Concluding Visual Information	
Wu et al., 2024	Swales & Feak (2012) … suggested that there are three key elements of "data commentary" (as they name it), including summary statements with reference to the visual, highlights of specific information in the visual, and discussions on meanings and implications of the highlighted information.	Swales and Feak, 2012
Wu et al., 2024	• Functional framework for CoVs[3] • Introduction • Providing background • Stating the presumptions • Connecting the current study Data • Describing/ rationalizing experiments • Presenting results • Interpreting results • Discussion • Summarizing the present study • Comparing findings with other studies • Explaining or consolidating findings	
Eriksson & Nordrum, 2018	…the three main moves of the model are *background*, *presentation of visual* and *comment on result*. These moves then contain several submoves or steps (Swales and Feak 2012, 331). It is not necessary to use all moves and submoves in single data commentary. For example, many data commentaries do not include a background move, and if they do, only one of the sub-moves *procedure-method*, *disciplinary-knowledge* or *comment-on-choice-of-presentation* is usually present…	Swales & Feak, 2018
Eriksson & Nordrum, 2018		

[1] https://www.nature.com/nmat/for-authors/preparing-your-submission.

[2] https://journals.plos.org/plosone/s/figures.

[3] Commentary on visuals

Publication	Extracted Data	References
Eriksson & Nordrum, 2018	Figure 1. A moves model of data commentary on result-reporting visuals in chemical engineering. The moves should be read from left to right. an example of a data commentary from a master's thesis in chemical engineering could be a background move realised by submove 1a) reminding the reader of how the data presented in the visual was obtained, followed by a presentation-of-visual move realised by submove 3a) providing a reference to the figure and the main result (e.g. *Figure 1 shows that there is an increase in* (...)), and last a comment-on-result move realised by submove 3d) giving a cross-reference to where in the master's thesis the result in the visual is discussed (e.g. *This increase will be further discussed in Section 4.2*).	

was classified as commenting on and concluding visual information (Parviz & Lan, 2023).

Analysing data legends as a separate genre of data commentary, Du et al. (2021) and Liu et al. (2023) refer to Cargill and O'Connor (2013), presenting detailed steps, including a title, information relating to results, components of a figure, scales and units, other notation or symbols. The steps in the legend that is treated as a DC are subject to a strict order due to limitations (space and no wordiness). Lui et al. (2023) offer their rhetoric framework of graphic legends based on a corpus of articles in four science disciplines, adding more steps such as statistical information (where the research design and statistical methods need explaining) and reference to sources of data (necessary for locating data).

All rhetoric steps extracted from the reviewed publications were summed up in Table 5 as generic moves in data commentary. Each step or rhetorical function found in the publications was analysed, merged into a generic group and included into the table. Moves – Title & Introduction, Data, and Discussion – were specified via generic steps typical of research publications that the authors of the reviewed articles and books had outlined.

We maintain that moves and steps depend more on the type of data commentary. The second essential feature is the type of the publication they are used in. For instance, titles are common for all data commentary, with legends limited essentially to charts, graphs and other visuals of the kind. Tables and graphs in the introduction of a research publication serve more as an argument or illustration. Thus, they may lack in many steps that are considered generic for those if such visuals are placed in the Results section.

DISCUSSION

The findings of this scoping review reveal considerable variability in how data commentary (DC) is approached within academic writing literature, as illustrated by the reviewed publications (Table 2). This variability is evident not only in the depth and length of coverage but also in the diversity of focus, rhetorical functions, and instructional approaches to DC. The review indicates that despite data commentary being an integral part of academic writing, it remains largely underrepresented in conventional academic writing textbooks. Our analysis of 316 academic writing books found only a small fraction (four texts, Table 3) containing sections explicitly addressing DC as a distinct component of scholarly discourse. This finding aligns with the current lack of standardization around DC content in academic writing, suggesting that while many academic disciplines recognize the importance of DC, a codified instructional framework has yet to emerge.

The range of functions attributed to data commentary in the reviewed sources highlights the complexity of its role in academic writing. Major DC functions identified include highlighting results, supporting arguments, assessing data reliability, evaluating implications, and synthesizing comparative data. This diversity of functions is mirrored by the differences in DC rhetorical structures across sources, with frameworks ranging from general guidelines to highly specific steps in the Results section. Publications such as those by Swales & Feak (2012) and Wallwork (2023) consistently emphasize the need for DC to present data without redundancy, thus distinguishing between tables and textual interpretation. However, these sources vary in their recommendations for specific rhetorical steps within DC, such as the inclusion of location statements or the use of metadiscourse, revealing a need for flexibility in DC structuring based on disciplinary and methodological contexts.

In addition, the analysis reveals that many academic writing books emphasize the role of legends and notes, particularly in guiding readers through complex visual data. For instance, Lövei (2021) and Miller (2004) stress that captions and notes should independently convey sufficient information for readers to understand a table or figure without referring to the main text. However, only a few sources, such as Miller (2004), go further by detailing the specific informational elements required in captions, such as abbreviations and quantitative comparisons, thus providing a more granular approach to DC. Such guidance is valuable for researchers who need to convey complex statistical or experimental data succinctly, suggesting that textbooks could benefit from more comprehensive, standardized coverage of DC-specific conventions for captions, legends, and textual commentary.

Table 5

Generic Moves in Data Commentary

Moves & Steps	Swales & Feak, 2012	Du et al., 2021	Cargill & O'Connor, 2013	Liu et al., 2023	Parviz & Lan, 2023	Wu et al., 2024	Eriksson & Nordrum, 2018
Title & Introduction							
A title (which summarizes what the figure is about)		+	+	+			
Location elements and/or summary statements	+			+	+	+	
Providing background, describing experiment					+	+	+
Highlighting statements	+				+	+	
Data							
Presentation of the visual							+
Definition of graphic items		+	+	+			
Description of the units or statistical notation included		+	+				
Reporting of statistical information			+	+			
Reference of sources of data				+			
Commenting on visual information					+	+	+
Discussion							
Connecting to the current study						+	
Discussions of aspects of the data	+						
Concluding visual information					+		
Result statement		+	+	+		+	
Interpretation of results				+	+	+	
Comparing and explaining findings with others					+	+	

The identified rhetorical steps in DC vary widely based on the type of data commentary and its position within a publication. Generic moves, such as those defined by Swales & Feak (2012), begin with location elements, proceed with highlight statements, and culminate in a discussion of implications or exceptions. In contrast, publications focused on figure legends and tables, such as those by Liu et al. (2023) and Cargill & O'Connor (2013), outline a different set of steps tailored to the structure and space constraints of figure captions. These sources emphasize brevity and precision, underscoring the role of captions as both self-explanatory and supportive of the main text. As illustrated in Table 5, generic moves for DC generally align with either a Results-focused or legend-focused structure, depending on the data role in the publication. This distinction underscores the importance of adapting DC strategies to the publication type and the nature of the data being discussed.

CONCLUSION

In achieving the aim of the review of research on data commentary, we found that there was a wide gap in the field of scientific communication and research on rhetorical interrelations between the verbal and the visual in the data presentation. The scattered and occasional distribution of sparse publications communicates an outline of a potentially wider field of research, including verbal and visual meanings and interrelations, rhetoric functions and moves in data commentary and figure legends. The books on academic writing quite rarely focus on DC.

The two major aspects of research on data commentary – DC rhetorical and other functions and moves in DC - show that researchers are unanimous in their approaches. The prevailing functions in the classifications entail argumentative, rhetorical, ontological, and epistemological functions. Though they are elaborated to include dozens of sub-functions expressed in the infinitive phrases. The moves are considered given the location of a visual and its function. Generic moves were exposed for data commentary and commentary and notes in legends.

The limitation of the review relates to some research that are not indexed in the international databases. Further studies of data commentary on discipline-based corpora of diverse genres of research publications might benefit the research field. More elaborated taxonomy and DC moves are top on the agenda. Such studies are in need for academic writing courses as visualization of data are on the rise, with visuals being increasingly employed to communicate scientific information and especially research results in the visual mode. These directions of research are essential for the field development as their implications range from academic communication, academic writing, writing for publication to improved quality of research reporting and better practices of scholar journal publishing.

DECLARATION OF COMPETITING INTEREST

None declared.

AUTHOR CONTRIBUTIONS

Elena Tikhonova: conceptualization; data curation; formal analysis; investigation; methodology; resources; software; validation; visualization; writing – original draft; writing – review & editing.

Lilia Raitskaya: conceptualization; data curation; formal analysis; investigation; methodology; resources; software; validation, visualization; writing – original draft; writing – review & editing.

REFERENCES

Ariga, K., & Tashiro, M. (2022). Change in the graphics of journal articles in the life sciences field: analysis of figures and tables in the journal "*Cell*". *History and Philosophy of the Life Sciences, 44*, 33. https://doi.org/10.1007/s40656-022-00516-9

Arksey, H., & O'Malley, L. (2005). Scoping studies: towards a methodological framework. *International Journal of Social Research Methodology, 8*(1), 19–32. https://doi.org/10.1080/ 1364557032000119616

Cargill, M., & O'Connor, P. (2013). *Writing scientific research articles: Strategy and steps.* Wiley-Blackwell.

Clymo, R. S. (2014). *Reporting research: A biologist's guide to articles, talks, and posters.* Cambridge University Press.

Cocchiarella, L. (Ed.). (2015). *The visual language of technique* (vol. 2: Heritage and expectations in research). Springer International Publishing Switzerland.

Dougherty, J. & Ilyankou, I. (2021). *Hands-on data visualization. Interactive storytelling from spreadsheets to code*. O'Reilly.

Du, Z., Jiang, F., & Liu, L. (2021). Profiling figure legends in scientific research articles: A corpus-driven approach. *Journal of English for Academic Purposes*, *54*, 101054. https://doi.org/10.1016/j.jeap.2021.101054

Eriksson, A., & Nordrum, L. (2018). Unpacking challenges of data commentary writing in master's thesis projects: An insider perspective from chemical engineering. *Research in Science & Technological Education*, *36*(4), 499-520. https://doi.org/10.1080/02635143.2018.1460339

Franzblau, L. E., & Chung, K. C. (2012). Graphs, tables, and figures in scientific publications: The good, the bad, and how not to be the latter. *The Journal of Hand Surgery*, *37*(3), 591-596. https://doi.org/10.1016/j.jhsa.2011.12.041

Goodwin, C. (2001). Practices of seeing visual analysis: An ethnomethodological approach. In T. Van Leeuwen & C. Jewitt (Eds.), *Handbook of visual analysis* (pp. 157-182). Sage.

Graves, H. (2014). The rhetoric of (interdisciplinary) science: Visuals and the construction of facts in nanotechnology. *Poroi, 10*(2), 1-19. https://doi.org/10.13008/2151-2957.1207

Gross, A. G., & Harmon, J. E. (2014). *Science from sight to insight: How scientists illustrate meaning*. University of Chicago Press.

Halliday, M. A. K. (1985). *An introduction to functional grammar*. Edward Arnold.

Hemais, B.J.W. (2014). Word and image in academic writing: A study of verbal and visual meanings in marketing articles. *Journal of English for Specific Purposes at Tertiary Level, 2*(2), 113-133.

Jagadeesh, G., Balakumar, P., & Senatore, F. (Eds.). (2023). *The quintessence of basic and clinical research and scientific publishing*. Springer Nature. https://doi.org/10.1007/978-981-99-1284-1

Johns, A.M. (1998). The visual and the verbal: A case study in macroeconomics. *English for Specific Purposes*, *17*(2), 183-197.

Kress, G., & van Leeuwen, T. (2006). *Reading images: The grammar of visual design* (2nd ed.). Routledge.

Lövei, G.L. (2021). *Writing and publishing scientific papers a primer for the non-English speaker*. Open Book Publishers. https://doi.org/10.11647/OBP.0235

Lui, L., Jiang, F., & Du, Z. (2023). Figure legends of scientific research articles: Rhetorical moves and phrase frames. *English for Specific Purposes*, *70,* 86-100. https://doi.org/10.1016/j.esp.2022.11.005

Miller, J.E. (2004). *The Chicago guide to writing about numbers. The effective presentation of quantitive information*. The University of Chicago Press.

Miller, T. (1998). Visual persuasion: A comparison of visuals in academic texts and the popular press. *English for Specific Purposes*, *17*(1), 29-46.

Mishra, P. (2004). The Role of abstraction in scientific illustration: Implications for pedagogy. In C. Handa (Ed.), *Visual rhetoric in a digital world: A sourcebook* (pp. 177 – 194). Bedford/St. Martins.

Moghaddasi, S., Graves, H.A.B., Graves, R. (2019). "See Figure 1": Visual moves in discrete mathematics research articles. *English for Specific Purpose*s, *56*, 50-67. https://doi.org/10.1016/j.esp.2019.08.001

Morell, T. (2015). International conference paper presentations: A multimodal analysis to determine effectiveness. *English for Specific Purposes*, *37*, 137-150. http://dx.doi.org/10.1016/j.esp.2014.10.002

O'Halloran, K. L. (2010). The semantic hyperspace: Accumulating mathematical knowledge across semiotic resources and modalities. In F. Christie & K. Maton (Eds.), *Disciplinarity: Functional linguistic and sociological perspectives* (pp. 217-236). Continuum.

O'Toole, M. (1996). A systemic-functional semiotics of art. In P. Fries & M. Gregory (Eds.), *Discourse in society: Systemic functional perspectives*. Ablex Publishing Corporation.

Parija, S.C., & Kate, V. (Eds.) (2017). *Writing and publishing a scientific research paper.* Springer Nature. http://doi.org/10.1007/978-981-10-4720-6

Parviz M., & Lan G. (2023). A corpus-based investigation of phrasal complexity features and rhetorical functions in data commentary. *Journal of Language and Education, 9*(3), 90-109. https://doi.org/10.17323/jle.2023.16044

Prus, R. (1987). Generic Social Processes: Maximizing Conceptual Development in Ethnographic Research. *Journal of Contemporary Ethnography, 16*(3), 250–93. https://doi.org/10.1177/0891241687163002

Royce, T. (2002). Multimodality in the TESOL classroom: Exploring visual-verbal synergy. *TESOL Quar*terly, *36*(2), 191-204. http://doi.org/ 10.2307/3588330

Saver, C. (2006). Tables and figures: adding vitality to your article. *AORN Journal*, *84*(6), 945-950. https://doi.org/10.1016/S0001-2092(06)63991-4

Swales, J. M. (1990). *Genre analysis: English in academic and research settings*. Cambridge University Press.

Swales, J.M., & Feak, C.B. (2012). *Academic writing for graduate students: Essential tasks and skills* (3rd ed.). University of Michigan Press.

Tang, K.-S. (2023). Distribution of visual representations across scientific genres in secondary science textbooks: analysing multimodal genre pattern of verbal-visual texts. *Research in Science Education*, *53*(2), 357-375. https://doi.org/10.1007/s11165-022-10058-6

Tricco, A.C., Lillie, E., Zarin, W., O'Brien, K.K., Colquhoun, H., Levac, D., Moher, D., Peters, M.D.J., Horseley, T., Weeks, L., Hempel, S., & Akl, E.A. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. *Annals of Internal Medicine*, *169*(7), 467–73. https://doi.org/10. 7326/M18-0850

van den Scott, L.-J.K. (2018). Visual methods in ethnography. *Journal of Contemporary Ethnography*, *47*, 6, 719-728. https://doi.org/10.1177/0891241618806972

Wallwork, A. (2023). *English for writing research papers*. Springer Nature. https://doi.org/10.1007/978-3-031-31072-0

Wallwork, A., & Sothern, A. (2020*), 100 tips to avoid mistakes in academic writing and presenting.* Springer Nature. https://doi.org/10.1007/978-3-030-44214-9

Wu, J., Zhao, C.G., Lu, X., & Jin, T. (2024). A rhetorical function and phraseological analysis of commentaries on visuals. *English for Specific Purposes*, *73*, 33-45. https://doi.org/10.1016/j.esp.2023.09.001

Zhang, L., Jang, R., & Zhang, J. (2024). 'Table 1 shows that': A local grammar of graphic data commentary in discourse of Economics. *English for Specific Purposes*, *74*, 68-82. https://doi.org/10.1016/j.esp.2024.01.001

APPENDIX 1

Publications Included in the Review

Dougherty, J. & Ilyankou, I. (2021). *Hands-on data visualization. Interactive storytelling from spreadsheets to code*. O'Reilly.

Du, Z., Jiang, F., & Liu, L. (2021). Profiling figure legends in scientific research articles: A Corpus-driven approach. *English for Specific Purposes, 54*, 101054. https://doi.org/10.1016/j.jeap.2021.101054

Graves, H. (2014). The Rhetoric of (Interdisciplinary) Science: visuals and the construction of facts in nanotechnology. *Poroi, 10*(2), 1-19. https://doi.org/10.13008/2151-2957.1207

Eriksson, A., & Nordrum, L. (2018). Unpacking challenges of data commentary writing in master's thesis projects: An insider perspective from chemical engineering. *Research in Science & Technological Education, 36*(4), 499-520. https://doi.org/10.1080/02635143.2018.1460339

Hemais, B.J.W. (2014). Word and image in academic writing: A study of verbal and visual meanings in marketing articles. *Journal of English for Specific Purposes at Tertiary Level, 2*(2), 113-133.

Jagadeesh, G., Balakumar, P., & Senatore, F. (Eds.). (2023). *The quintessence of basic and clinical research and scientific publishing*. Springer Nature. https://doi.org/10.1007/978-981-99-1284-1

Johns, A.M. (1998). The visual and the verbal: A case study in macroeconomics. *English for Specific Purposes, 17*(2), 183-197.

Lövei, G.L. (2021). Writing and publishing scientific papers a primer for the non-English speaker. Open Book Publishers. https://doi.org/10.11647/OBP.0235

Lui, L., Jiang, F., & Du, Z. (2023). Figure legends of scientific research articles: Rhetorical moves and phrase frames. *English for Specific Purposes, 70*, 86-100. https://doi.org/10.1016/j.esp.2022.11.005

Miller, J.E. (2004). *The Chicago guide to writing about numbers. The effective presentation of quantitive information.* The University of Chicago Press.

Moghaddasi, S., Graves, H.A.B., Graves, R. (2019). "See Figure 1": Visual moves in discrete mathematics research articles. *English for Specific Purposes, 56*, 50-67. https://doi.org/10.1016/j.esp.2019.08.001

Parija, S.C., & Kate, V. (2017). *Writing and publishing a scientific research paper.* Springer Nature. http://doi.org/10.1007/978-981-10-4720-6

Parviz M., & Lan G. (2023). A Corpus-based investigation of phrasal complexity features and rhetorical functions in data commentary. *Journal of Language and Education, 9*(3), 90-109. https://doi.org/10.17323/jle.2023.16044

Swales, J.M., & Feak, C.B. (2012). *Academic writing for graduate students: Essential tasks and skills* (3rd ed.). University of Michigan Press.

Van den Scott, L.-J.K. (2018). Visual methods in ethnography. *Journal of Contemporary Ethnography, 47*(6), 719-728. https://doi.org/10.1177/0891241618806972

Wallwork, A. (2023). *English for writing research papers*. Springer Nature. https://doi.org/10.1007/978-3-031-31072-0

Wallwork, A., & Sothern, A. (2020). *100 tips to avoid mistakes in academic writing and presenting.* Springer Nature. https://doi.org/10.1007/978-3-030-44214-9

Wu, J., Zhao, C.G., Lu, X., & Jin, T. (2024). A rhetorical function and phraseological analysis of commentaries on visuals. *English for Specific Purposes, 73*, 33-45. https://doi.org/10.1016/j.esp.2023.09.001

Zhang, L., Jang, R., & Zhang, J. (2024). 'Table 1 shows that': A local grammar of graphic data commentary in discourse of Economics. *English for Specific Purposes, 74*, 68-82. https://doi.org/10.1016/j.esp.2024.01.001

APPENDIX 2

Functions of Data Commentaries in the Reviewed Research Publications

Publication	Extracted Data	References	Function
Du et al., 2021	Almost a half of scientific data are represented visually (Hyland, 2006, p. 53), so visual representations are not mere add-ons or ways to popularise a complex reasoning but are an essential part of academic discourse …	Hyland, 2006	To popularize a complex reasoning
Graves, 2014	…the authors use visual representations of their data to accomplish a complex range of activities from informational to rhetorical to ontological to epistemological. "rhetorical" as the role visuals play as evidence supporting argumentative claims or persuading readers of the validity of the interpretation offered for the data. By "ontological," … the visual's role in instantiating (serving as visual "proof") that the scientific phenomenon exists/is real. … "epistemological" to refer to the visual's role as the argument transforms it into a knowledge claim that reciprocally strengthens the argument.	Mishra, 2004	Rhetorical function Ontological function Epistemological function
Graves, 2014	The illustrations are artists' rendering of the concepts meant to communicate simplified information, while the table and charts present complex, highly mediated information that require significant viewer background knowledge and engagement to communicate meaning.		To communicate simplified information
Graves, 2014	…the visuals function as critical pieces of evidence that support the argument …		To support the argument
Hemais, 2014	In articles in academic journals, visuals such as graphs and diagrams provide significant support for the arguments in the text… Visuals are an aid in persuading the reader of the validity of the authors' arguments, since "the last line of defense and foundation of the research argument is the findings themselves, almost always presented in the form of visual display" (Miller, 1998: 30).	Miller, 1998	To support the argument To persuade the reader of the validity of the argument
Liu et al., 2023	…visuals in scientific discourse are not merely an alternative form of data presentation but also serve as "evidence providers" (Morell, 2015, p.138) that help to testify to scientific hypotheses or validate academic claims.	Morell, 2015	To provide evidence
Liu et al., 2023	In mathematics, for example, figures are employed to realize ontological, argumentative, and epistemological functions (Moghaddasi et al., 2019).	Moghaddasi et al., 2019	Argumentative function Ontological function Epistemological function

Publication	Extracted Data	References	Function
Miller, 1998	The last line of defense and foundation of the research *argument* is the findings themselves, almost always presented in the form of visual display… Photographs, graphs, and tables also give the illusion of direct access to the data, which makes this portion of the argument particularly convincing… …visuals in academic articles provide data to convince the reader of the validity of the findings and allow the readers to see how the data were obtained and to interpret the data themselves… The most important use of the visuals in the academic texts is to *support the argument*. The figures and tables invite the readers to see for themselves as if the data rather than the scientist are carrying the argument… The corresponding visual realization for 'theme' is what O'Toole (1996) calls the *compositional function*. This function involves framing, horizontals and verticals, proportion, line, geometric forms, and color cohesion pattern (highlighted by the iconic overlay) among the squares themselves.	O'Toole (1996)	To support the argument Compositional function
Moghaddasi et al., 2019	…visuals … perform three functions: ontological, argumentative, and epistemological… Morell (2015) assigns three functions to non-verbal (that is, visual) material (NVM)[3] in her analysis of conference presentations: illustrative, decorative, and expository. Illustrative NVM contains a verbal component, such as illustrating a process using a flow chart. Decorative NVM creates backgrounds and usually appears in social sciences. Expository[2] NVM fills evidence-providing roles and appears more commonly in sciences and engineering. Morrell's categories identify dominant functions, yet she notes the functions are mixed.	Morell, 2015	Argumentative function Ontological function Epistemological function
Moghaddasi et al., 2019	Regarding visuals in articles in biology, Miller (1998) concludes that visuals in RAs[4] both 'prove' and 'clarify'	Miller, 1998	To prove the argument To clarify
Moghaddasi et al., 2019	O'Halloran (2010) argues that each semiotic resource fulfils particular functions: images 'provide an intuitive overview of the relations between mathematical participants,' language and images 'introduce and conceptualize mathematical concepts and problems' (p. 4), and symbolism 'formalise[s] those relations and solve[s] the problem' (p. 5).	O'Halloran, 2010	To provide an overview between mathematical participants To introduce mathematical concepts
Moghaddasi et al., 2019	Gross and Harmon (2014) also propose that throughout various argument stages visuals fulfil different semiotic meanings: 1) iconic (i.e., they represent the world); 2) symbolic (i.e., they stand for aspects of the world); and 3) indexical (i.e., they show causal relationships in the world). Gross and Harmon argue that viewers interpret this meaning by placing the visuals in the context of argumentative structures… they assign a subordinate position to visuals, stating that they cannot be arguments - a view not shared in all scientific disciplines.	Gross & Harmon, 2014	Iconic function Symbolic function Indexical function To support the argument
Moghaddasi et al., 2019	Mathematization, the primary purpose of some visuals, is defined by Goodwin (2001) as those contextually-driven practices aimed at transforming intractable phenomena into mathematically tractable visuals such as graphs and diagrams.	Goodwin, 2001	To conceptualize intractable phenomena

[3] NVM is "non-verbal material"

[4] RAs stand for "research articles"

Publication	Extracted Data	References	Function
Moghaddasi et al., 2019	Graves (2014) argues that visual data in nanotechnology accomplish complex activities from informational to rhetorical to ontological to epistemological: visual data can constitute the evidence that develops and supports the claims, but it can also be the fact itself or simultaneously form the foundation and the structure for new knowledge.	Graves, 2014	Argumentative function Ontological function Epistemological Function Informational function
van den Scott, 2018	Visuals are data. They are also tools to gather and record data. The visual can help us to understand generic social processes (Prus, 1987). Visuals can add a layer of depth to our analysis, can offer patterns for analysis…	Prus, 1987	To consolidate the information
Wu et al., 2024	As for relevant verbal accounts for the visuals, four main functions were identified, including establishing presumptions, announcing results, proof, and discussion.		To establish presumptions To announce results To prove To discuss
Wu et al., 2024	Functional framework for CoVs[5] • Introduction • Providing background • Stating the presumptions • Connecting the current study • Data • Describing/ rationalizing experiments • Presenting results • Interpreting results • Discussion • Summarizing the present study • Comparing findings with other studies • Explaining or consolidating findings		To provide background To state the presumptions To connect the study To describe experiments To present results To interpret results To summarize the study To compare findings with others To consolidate findings
Zhang et al., 2024	By using graphics, scientific writers can 1) reduce reading time by summarizing key information, 2) supplement the main text to clarify complex information, 3) add visual effect to text to enhance interpretability of knowledge, and 4) reduce word counts to save space (Clymo, 2014; Franzblau and Chung, 2012; Saver, 2006).	Clymo, 2014; Franzblau & Chung, 2012; Saver, 2006	To summarize information To clarify information To enhance interpretability To save space
Zhang et al., 2024	Swales and Feak note that academic writing involves the task in which writers need to discuss data typically displayed in tables, charts or figures. They termed this task 'data commentary' which consists of three sub-tasks to complete: 1) pointing out the location of data and summarizing the graphic content (e.g., Table 5 shows the most common modes of infection for U.S. business.), 2) highlighting the information in graphics from the more significant to the less significant (e.g., As can be seen, in the majority of cases. However, it is alarming to note that…), and 3) discussing implications, problems, etc…(Swales and Feak, 1994, p. 80).	Swales and Feak, 2012	To locate data To summarize information To highlight the more significant information To discuss implications

[5] Commentary on visuals.

https://doi.org/10.17323/jle.2024.21618

Variation in Academic Writing: A Corpus-Based Investigation on the Use of Syntactic Features by Advanced L2 Academic Writers

Muhammad Ahmad [1], Muhammad Asim Mahmood [1], Ali Raza Siddique [1], Muhammad Imran [2,3], Norah Almusharraf [4]

[1] Government College University, Faisalabad, Pakistan

[2] Prince Sultan University, Saudi Arabia

[3] Khazar University, Azerbaijan

[4] Prince Sultan University, Saudi Arabia

ABSTRACT

Background: Writing means communication through words whereas academic writing means making careful use of words to communicate ideas to a range of readers and audiences. Therefore, academic writing reflects specificities related to audience, context/discipline and purpose of the use. These specificities result in ample differences in terms of language use.

Purpose: This study investigates disciplinary variation in the use of different syntactic (i.e., clausal, intermediate and phrasal) features in academic writing produced by the Pakistani advanced writers of English as an L2 specializing in different disciplines of arts and humanities, life sciences, physical sciences and social sciences.

Method: For the said purpose, the corpus has been developed from dissertation texts produced by the Pakistani doctoral candidates from 16 academic disciplines of four disciplinary divisions. The analysis has been performed using AntConc Software after tagging with Multidimensional Analysis, and TagAnt Taggers.

Results: The results reveal mixed findings. On the one hand, the results show variation in the use of syntactic features that is observed to be marked by the difference in the frequency of the different types of the said features across disciplines. On the other hand, the results show a similarity in the use of syntactic features that has been evidenced by the finding that the most and least frequently used features are identical across disciplines.

Conclusion: These results suggest both heterogeneity and homogeneity in the use of syntactic features by the Pakistani advanced L2 academic writers. The results of this study have implications for educators, policy makers, and syllabus designers to ensure discipline-specific instruction, and incorporation of the discipline-specific syntactic features into the academic curricula for supporting academic writing development skills in the students particularly at the advanced level of education.

KEYWORDS

academic disciplines, academic writing, advanced L2 academic writers, clausal features, disciplinary variations, intermediate features, phrasal features

Citation: Ahmad, M., Mahmood, M. A., Siddique, A. R., Muhammad, I., & Norah, A. (2024). Variation in academic writing: A corpus-based investigation on the use of syntactic features by advanced L2 academic writers. *Journal of Language and Education, 10*(3), 25-39. https://doi.org/10.17323/jle.2024.21618

Correspondence:
Muhammad Ahmad
E-mail: ahmad453@yandex.com

Received: May 15, 2024
Accepted: September 16, 2024
Published: September 30, 2024

INTRODUCTION

Syntactic complexity refers to "the range and degree of sophistication of syntactic structures" (Lu, 2014, p. 130), and equates with linguistic complexity and syntactic maturity (see Ortega, 2003). However, on a general level, syntactic complexity refers to the degree of elaboration, sophistication, and variation of the syntactic structures used in the discourse (Dong et al., 2023; Lu, 2017; Norris & Ortega, 2009; Ortega, 2015; Yin et al., 2021). It (syntactic complexity) has been the favorite subject of experts (Biber et al., 2011; Biber & Gray, 2016; Gray, 2015; Lan & Sun, 2018; Wang & Lowie, 2021) in the research on academic writing for

the last decade. Consequently, linguistic complexity studies, also called syntactic complexity studies (see Xue & Ge, 2021) yielded diverse and insightful understandings of the linguistic variation in academic writing using different variables like academic genre, academic discipline, functional and rhetorical move structures, language background, and writing proficiency (Dong et al., 2023).

For example, Biber et al. (2016) compared the use of syntactic features, that is, clausal, intermediate, and phrasal features in the less- and more-proficient writers of English as a first language (L1) and reported the frequent use of phrasal and clausal features by the more- and less-proficient writers respectively. Ansarifar et al. (2018) studied phrase structures in the academic writing produced by the Iranian master and doctoral-level writers of English as a second language (L2) by comparing the results to those of expert writers. Their findings showed significant and no variation in the use of phrase structures by the Master's and doctoral level writers, respectively, from the expert writers. A recent research (Yin et al., 2021) investigated variation in the academic writing produced by emerging and expert writers. Their analysis revealed significant variation in the use of 14 syntactic measures by emerging and expert writers.

Similarly, other variables were explored in different studies with interesting findings like academic genre (see Nasseri, 2021; Yoon & Polio, 2017), functional and rhetorical move structures (see Khamaiseh, 2023; Lu et al., 2020; Saricaoglu et al., 2021), language background (see Ahmad et al., 2023a; Lan et al., 2022; Lu & Ai, 2015), and academic discipline (see Biber & Gray, 2016; Gray, 2015; Staples et al., 2016). Among these variables, disciplinary variation in the use of syntactic features has been treated as a "notable strand of inquiry" (Dong et al., 2023), and the findings of some studies (Biber & Gray, 2016; Casal et al., 2021; Gray, 2015; Staples et al., 2016) revealed syntactic features displaying disciplinary variation in different texts of the academic discourse.

This study explores the under-representation of Pakistani L2-related dissertations in academic writing research, particularly in the context of syntactic variation in Pakistan. While there is ample research on academic writing in general, including research articles, essays, and textbooks, there is a lack of studies focusing specifically on dissertations. Similarly, while there is extensive research on academic writing in Pakistan, there is a notable gap in research on dissertations, which are crucial for studying disciplinary variation in L1 and L2 academic writing at advanced education levels (see Biber & Gray, 2016; Casal et al., 2021; Gray, 2015; Staples et al., 2016). In this context, the aim of this research is to address this gap and present compelling findings to encourage further exploration of dissertations across disciplines (Table 1).

Secondly, this research employs the methodology gap. The reason for this choice is that the past corpus-based research (conducted in Pakistan and abroad) approached it mostly through multidimensional perspectives (see literature review section) utilizing subscription-based tools for corpus and data analyses. This research proposes that the syntactic features can be studied with the help of online available open-access tools like TagAnt, MAT Taggers, and the AntConc software, and expects to introduce future researchers on academic writing towards the use of cost-effective tools for analyses. For this purpose, this research employs a list of formulaic patterns (see corpus analysis in the methodology section).

Lastly, many studies (e.g., Ansarifar et al., 2018; Lan & Sun, 2018) explored disciplinary variation in L1 and L2 academic writings. The exploration of L2 academic writing, specifically in Pakistan, has been overlooked. This pioneering research endeavors to bridge this gap and sheds light on this crucial area.

LITERATURE REVIEW

Previous Research on Disciplinary Variation

Syntactic complexity in academic writing, which is caused by phrasal structures, varies "based on the discipline to which it belongs" (Elliott, 2019, p. 10). This variation has been the subject of a good number of previous studies. For example, Gray (2015) investigated more than 70 linguistic features in the academic writing corpus developed from 270 articles from six disciplines: applied linguistics, biology, history, philosophy, physics, and political science. Results obtained through the comprehensive analyses based on grammatical/lexical survey, structural complexity exploration, and the Multidimensional Analysis (MDA) showed variations in the use of linguistic features, including phrasal structures across disciplines. This study contributed to the methodological considerations for future corpus-based research on academic writing across disciplines by going beyond traditional methods of analysis and considering varied realizations of academic discourse both across and within disciplines. Biber and Gray (2016) investigated core grammatical and structural features in academic writing from a wide array of disciplines related to humanities, popular science, social science, and specialist science. Results obtained through the quantitative analyses revealed an increase in the diversification and specialization within the disciplines.

Staples et al. (2016) investigated syntactic complexity features in the academic texts produced by the university-level writers of English as an L1. They conducted analyses across academic levels, genres, and disciplines. Results relevant to this present study revealed the frequent use of phrasal features across disciplines (arts and humanities, life sciences, and physical sciences) with the exception of noun+ of phrases that were frequent in the social sciences. Furthermore, 'premodifying nouns' were found in frequent use in

life sciences and physical sciences, whereas nouns, nominalizations, and attributive adjectives were frequently found in social sciences. Similarly, prepositional phrases and of genitives were frequent in arts and humanities. These findings aligned with, Biber and Gray's (2016), and Gray's (2015) findings, which demonstrated that academic writing heavily relies on phrasal complexity features. The extent of this reliance varied across disciplines: science disciplines relied on phrasal complexity features to the greatest extent, followed by the social sciences and then the arts and humanities. Jalilifar et al. (2017) explored nominalization structures in the academic texts from hard (physics) and soft science (applied linguistics) disciplines. Their results revealed marked variations in the use of the said structures. For example, academic texts from hard sciences contained nominals with both pre-and post-modifications, whereas the academic texts from soft sciences contained nominals with relative clauses as post-modifiers. Another study (Elliott, 2019) examined the use of noun phrase structures across 16 disciplines from four academic divisions. The findings demonstrated that advanced-level students' academic writing varied in how they used the said structures across disciplines. These findings were in line with the findings of the previous studies (Biber & Gray, 2016; Gray, 2015; Jalilifar et al., 2017; Staples et al., 2016).

Some of the most recent studies also investigated the same variables in academic writing. For example, Casal et al. (2021) investigated eight syntactic complexity measures in academic writing from three disciplines (i.e., applied linguistics, economics, and psychology) of social sciences. The findings revealed that academic writing from applied linguistics had the most complex structures, while academic writing from economics had the least complex structures. In detail, noun phrase per clause was frequent in applied linguistics, non-finite subordination and phrasal coordination were common in psychology, and finite clausal subordination was frequent in economics. The disciplinary variations are clearly evident in these results. Lu et al. (2021) investigated disciplinary variations in the relationship between the syntactic complexity structures and rhetorical move steps of the introduction sections in research articles. The corpus for this research comprised the texts of 400 research articles from core disciplines of two disciplinary divisions, that is, engineering (chemical engineering and electrical engineering) and social sciences (anthropology and sociology). The results revealed significant variations in terms of the syntactic complexity measures across disciplines. Another recent study (Ziaeian & Golparvar, 2022) used fine-grained clausal and phrasal indices to investigate syntactic complexity in the discussion sections of research articles from three academic disciplines (i.e., applied linguistics, chemistry, and economics). The results showed significant variations in the use of the said structures across disciplines. For example, clausal indices were frequently observed in applied linguistics and economics, whereas phrasal features were frequently observed in chemistry. Saricaoglu and Atak (2022) explored variation in terms of lexical and syntactic complexity markers in the academic writing produced by Turkish students. The results obtained through manual and automated analyses revealed ample variations in the use of complement clauses, passives, and the words placed before the main verbs. These findings helped them explain the relationship between L2 writing proficiency levels and linguistic features. A recent study (Tian & Zhang, 2023) investigated nominalizations in the academic writing produced by writers from linguistics, shipbuilding, and oceanography engineering disciplines. The results of this study also showed significant variations in the use of the nominalizations across disciplines. Another most recent study (Dong et al., 2023) investigated the disciplinary variations of the syntactic complexity structures in academic writing across 31 disciplines from four disciplinary divisions. The corpus for this research was obtained from a British Academic Written English (BAWE) source and analysed through an automatic process. The results revealed significant variations in coordination, length, sophistication, subordination, and sentence complexity across disciplines and disciplinary divisions. This study differed from the above-reviewed studies in the sense that it not only covered a broad array of academic disciplines but also discussed the results from a form-functional perspective. The studies (reviewed in this section) sufficiently confirm the existence of variation in the use of lexical and syntactic structures in the academic discourse across genres, educational levels, language backgrounds (L1 and L2), and disciplines.

Previous Research on Pakistani Academic Writing

A number of previous researchers investigated Pakistani academic writing and reported interesting findings. For example, Aziz et al. (2016) investigated linguistic variation in Pakistani academic writing across two disciplines (i.e., biological and health sciences and physical sciences). They prepared the corpus from doctoral dissertations and analyzed it through MDA. Their findings reported significant variations at the dimensions 1-3 and similarities at the dimensions 4-5. Similarly, Azher et al. (2019) investigated register variation in Pakistani academic writing across humanities, sciences, and social sciences disciplines employing MDA. The results revealed significant variations in Pakistani academic writing in terms of different dimensions that underlined discipline-and register-specific pedagogies with reference to Pakistani English. Another MDA-based study (Rashid & Mahmood, 2019) investigated linguistic variation in Pakistani academic writing, preparing a corpus from research articles across humanities, sciences, and social sciences disciplines. The analysis revealed interesting findings related to the variations across disciplines. For example, the academic writing from social sciences was observed to be more informational, impersonal, non-narrative, and non-personal compared to the academic writing from sciences and humanities disciplines.

In addition to exploring linguistic and register variation, structural variation has also been examined in Pakistani academic writing. For instance, Qasim et al. (2017) investigated structural variations specifically within this context. For this purpose, they developed the corpus from the texts of conclusion sections of Master's theses from humanities and social sciences, and science and technology disciplines. Their results highlighted variations in the structures of the conclusion sections written by Pakistani Master's level thesis writers. This study was presented as useful material for the students' familiarization with the structural features of the conclusion sections of Master's theses. Abid et al. (2022) conducted a cross-cultural MDA of the academic writing produced by Chinese and Pakistani academic writers and reported interesting results on cross-cultural variations in academic writing. This study is unique in that it reports the uniqueness of Pakistani academic writing compared to that of Chinese academic writing. Recently, Fatima et al. (2023, p. 50) investigated linguistic variations in the dissertation abstracts written by Pakistani doctoral-level academic writers across 16 disciplines employing MDA. The results showed distinct variations across disciplines and supported the idea of Pakistani English being "a separate linguistic entity with unique characteristics." Another recent study (Pervez et al.,2024) investigated linguistic variation in the discussion sections of Pakistani English research articles and reported interesting results related to the linguistic variations according to the different dimensions proposed in Biber (1988). Thus, the results of the MDA-based studies (Abid et al., 2022; Azher et al., 2019; Aziz et al., 2016; Fatima et al., 2023; Pervez et al., 2024; Qasim et al., 2017; Rashid & Mahmood, 2019) conducted in Pakistani context show that Pakistani academic writing depicts variations.

However, recent non-MDA-based studies reported different results. For example, Ahmad et al. (2022) conducted a corpus-based study to explore disciplinary variations in phrasal features in Pakistani academic writing produced by doctoral students from arts and humanities, and life sciences disciplines. Results revealed the frequent use of nouns in Pakistani academic writing across the said disciplines. These results concluded that Pakistani academic writing does not reflect disciplinary variation, which is the salient feature of academic writing. Another recent study (Ahmad et al., 2023b) investigated the salient features characterizing Pakistani academic writing across hard and soft sciences disciplines. The results revealed that Pakistani academic writing from social sciences relies on phrasal features more than academic writing from hard sciences. This study explored a wide array of linguistic features in Pakistani academic writing. However, its scope was limited to the four sub-disciplines, that is biology, physics (hard sciences), and history and linguistics (soft sciences). Considering this limitation, Ahmad et al. (2023c) conducted another corpus-based study on variation in Pakistani academic writing across four disciplinary divisions: arts and humanities, life sciences, physical sciences, and social sciences. The results revealed homogenous as well as heterogeneous use of the syntactic features. The heterogeneity was observed in relation to the frequency of different types of syntactic features, whereas the homogeneity was reported in relation to the highest and lowest used features; that is, the highest and the lowest used features were the same across the four disciplinary divisions. These results concluded that Pakistani academic writing does not reflect disciplinary variation. This practice was reported being contrary to the expert convention. Therefore, Pakistani academic writers were suggested to appropriately use the syntactic features in accordance with the expert convention in the relevant discipline.

Thus, the results of studies (Ahmad et al., 2022, 2023b, 2023c) differ from those of MDA-based studies (Azher et al., 2019; Aziz et al., 2016; Fatima et al., 2023; Pervez et al., 2024; Qasim et al., 2017; Rashid & Mahmood, 2019). This suggests the need for further confirmation of these differences in Pakistani academic writing by exploring variations across a wide array of academic disciplines (Table 1).

METHOD

Research Design

This is a corpus-based descriptive study which presents an investigation of the syntactic features to report variation in the L2 academic writing. The details of materials and methods employed in this study are described below.

Research Corpus

The corpus for this study was developed from dissertation texts written by Pakistani doctoral candidates across different academic disciplines. The list of disciplines employed in this study was chosen from Nesi and Gardner (2012). The choice for the list of disciplines proposed by Nesi and Gardner (2012) was made because it provided a wide range of academic disciplines representing four broad disciplinary divisions (see Table 1 for the details regarding disciplines, disciplinary divisions, and corpus distribution) compared to the other divisions (e.g., Becher' 1981; Biglan, 1973).

The corpus development process involved procedural steps. First of all, the dissertations were retrieved in portable document format (PDF) from the Pakistan Research Repository (PRR). PRR is an online database hosted by the Higher Education Commission of Pakistan. PRR contains dissertations written by the Master's and doctoral candidates at Pakistani universities and provides free online access to the researchers beyond borders. Secondly, the PDFs were converted into MS Word format using the freely online available software

iLovePDF[1]. Thirdly, the converted files were cleansed. In this process, the preliminary pages, headings, tables, figures, references, headers, footers, equations, and formulas were removed. Finally, the remaining texts were saved in Notepad files. Data compiled in the Notepad files formed the corpus that was ready to be processed by the corpus software for analysis purposes.

The dissertation was considered in this study for several reasons: it is an important genre of academic writing (see Hyland, 2004) at the graduation level due to demonstrating students' ability and expertise to contribute to their disciplines, and it has its own conventions, purposes and structures that distinguish it from other types of academic writing. In addition, the dissertation is vital for sharing discipline-specific knowledge of the writers (Housseine & Oifaa, 2020; Parry, 1998) particularly for characterizing the knowledge of the disciplinary community that is constructed using diverse linguistic features which are important to study in order to find how these features contribute to the disciplinary context in Pakistani academic writing (Azher et al., 2019).

Secondly, Pakistani academic writing was considered in this study for the reasons that: the study of different genres of Pakistani academic writing is essential to portray the comprehensive picture of the Pakistani academic writing (Rashid & Mahmood, 2019); Pakistani academic writing is an important form of Pakistani English (Fatima et al., 2023; Kachru et al., 2006; Mahboob, 2008; Mahmood, 2009;Pervez et al., 2024; Rahman, 1990; Talaat, 1993) which is a legitimate variety of World Englishes, and according to Azher et al. (2019), Pakistani academic writing is a form which invites the interest of the linguists as well as researchers for the further strengthening of the Pakistani English.

Syntactic Features

A list of syntactic features (Tables 2-5) was adapted from Staples et al. (2016). The reason for this choice was based on

Table 1

Corpus Distribution across Disciplinary Divisions and Respective Disciplines

Disciplinary Division	Discipline	Number of Texts	Number of Words
Arts and Humanities	Philosophy	10	48576
	English Literature	10	572299
	History	10	27753
	Linguistics	10	489827
Total		***40***	***1138455***
Social Sciences	Politics	10	344893
	Sociology	10	316658
	Law	10	498929
	Economics	10	87492
Total		***40***	***1247972***
Physical Sciences	Mathematics	10	48010
	Physics	10	109361
	Engineering	10	143281
	Computer Science	10	157668
Total		***40***	***458320***
Life Sciences	Psychology	10	103623
	Food Sciences	10	172120
	Biology	10	179080
	Agriculture	10	313620
Total		***40***	***768443***
Grand Total		***160***	***3613190***

Note. Ahmad (2022)

[1] iLovePDF. https://www.ilovepdf.com/word_to_pdf

the notion that the said features are the latest development in academic writing research and have been empirically tested for the study of syntactic complexity, disciplinary as well as generic variation, level of academic writing development, and so on. This study considered the said features to report disciplinary variation in the academic writing produced by advanced academic writers across disciplines (Table 1).

Corpus Analysis

Corpus analysis was also completed in a number of procedural steps. First of all, the corpus was tagged through MAT (Multidimensional Analysis Tagger) and TagAnt Taggers. Both of these taggers are available online for free access and are used for the tagging of a large number of corpora, whereas MAT facilitates analyses to discover variations in the corpora. In the second step, the tagged corpus was processed in the AntConc, another freely available software, for analysis. In this regard, different formulas (as used in Ahmad, 2022, p. 87-90) were applied. For example, attributive adjectives were searched through four formulas based on four descriptors: bracketed with the relevant formula, i.e., *_DT *_JJ *_NN (Determiner + Adjective + Noun); *_DT *_JJ *_JJ*_NN (Determiner + Adjective + Adjective + Noun); *_DT *_JJ *_NOMZ (Determiner +Adjective + Nominalization); and *_DT *_JJ *_NN *_NN (Determiner + Adjective + Noun+ Noun). This process provided the frequencies/examples of the said features. In the third step, the frequencies were separately extracted in MS Excel sheets for presentation as results (Tables 2-5) of the study.

RESULTS AND DISCUSSION

Syntactic Variation across Disciplines in Arts and Humanities

This study examined syntactic features in various disciplines (i.e., English literature, linguistics, history, and philosophy) of arts and humanities,focusing particularly onvariationinthe use of clausal, intermediate, and phrasal features. The results revealed notable differences in usingthe said features across the selected disciplines (Table 2). In English literature and linguistics, the frequency of clausal features was 2,340 and 2,091, respectively. Intermediate features were used 12,155 times in English literature and 11,854 times in linguistics, while phrasal features were observed 251,075 and 249,650 times, respectively. Conversely, history and philosophy showed considerably lower frequencies: history used 100 clausal features, 391 intermediate features, and 15,098 phrasal features; philosophy used 147 clausal features, 851 intermediate features, and 22,846 phrasal features.

These findings indicate the syntactic complexity of English literature and linguistics, reflecting their rhetorical and analytical requirements. The high frequency of clausal and phrasal features in these disciplines suggests a preference for complex as well as elaborated sentence structures, which is consistent with the requirements for nuanced argumentation and comprehensive analysis (Biber & Gray, 2016; Casal et al., 2021; Dong et al., 2023; Elliott, 2019; Gray, 2015; Jalilifar et al., 2017; Lu et al., 2021; Saricaoglu & Atak, 2022; Staples et al., 2016; Tian & Zhang, 2023; Ziaeian & Golparvar, 2022).

In contrast, the lower usage of these features in history and philosophy indicates a different approach to academic writing, potentially due to the narrative and abstract nature of these fields, which may not necessitate complex syntactic constructions. This observation aligns with previous research indicating discipline-specific variations in syntactic complexity (Elliott, 2019; Staples et al., 2016).

However, these results (Table 2) reveal a divergence from certain expert norms, particularly in the lower use of clausal features in history and philosophy compared to what is observed in the practices of experts in arts and humanities. This discrepancy may be attributed to the differences in educational practices, the influence of the writers' first language, or varying familiarity with international academic standards. Such variations challenge the results of recent studies (Ahmad et al., 2022, 2023b, 2023c), which reported less disciplinary variation in Pakistani academic writing. Thus, the results of this studyi ndicate that significant variation does exist in Pakistani academic writing, though it may not fully align with the expert trends.

Syntactic Variation across Disciplines in Social Sciences

This study examines syntactic variation in Pakistani advanced L2 academic writing within the social sciences, focusing on law, politics, sociology, and economics. The results reveal significant differences in the use of clausal, phrasal, and intermediate features across these disciplines, reflecting their unique rhetorical and communicative demands.

The results (Table 3) indicate notable variations in the frequency of syntactic features. Specifically, the frequency of clausal features was highest in law, with 1,852 instances, followed by politics, with 1,226; sociology, with 822; and economics, with 164. Phrasal features were also unevenly distributed, with law leading at 224,319 instances, followed by politics at 177,164, sociology at 159,891, and economics at 48,005. Intermediate features showed a similar pattern, with the law again at the lead with 10,987 instances, followed by politics with 8,508, sociology with 4,726, and economics with 1,750. Overall, Pakistani social sciences writers used 4,064 clausal, 609,379 phrasal, and 25,971 intermediate features. This distribution indicates a predominant reliance on phrasal features across all disciplines, with clausal features being

Table 2

Use of Syntactic Features across Disciplines in Arts and Humanities

FEATURES	FREQUENCIES IN DIFFERENT DISCIPLINES				TOTAL
Clausal Features	**Philosophy**	**English Literature**	**History**	**Linguistics**	
Finite adverbial clauses	22	588	6	402	1018
WH complement clauses	0	49	0	45	94
Verb + that-clauses	4	60	0	66	130
Clausal coordinating conjunctions	121	1643	94	1578	3436
Total	***147***	***2340***	***100***	***2091***	***4678***
Intermediate Features					
Adverbs	73	1042	54	1035	2204
Linking adverbials	251	3025	129	3287	6692
Extraposed Adjective + that clauses	1	32	0	26	59
Noun + that-clauses	30	607	10	526	1173
WH relative clauses	18	283	5	199	505
That relative clauses	119	2282	48	1966	4415
Verb + to-clauses	80	1223	10	1337	2650
Desire verb + to-clauses	0	0	0	0	0
Raising structures and extraposed adjective + to- clauses	0	4	1	3	8
Noun + to-clauses	190	2589	87	2419	5285
Verb + ing-clauses	9	210	3	146	368
Passive voice verbs	0	19	1	12	32
Passive nonfinite relative clauses	80	839	43	898	1860
Total	***851***	***12155***	***391***	***11854***	***25251***
Phrasal Features					
Nouns	13469	158977	9045	156660	338151
Attributive adjectives	2737	28769	1696	28058	61260
Premodifying nouns	2078	22726	1732	21860	48396
Nominalizations	2961	21434	1230	24158	49783
of genitives	895	9825	688	9811	21219
Prepositional phrases	706	9344	707	9103	19860
Total	***22846***	***251075***	***15098***	***249650***	***538669***

Note. Adapted from Ahmad (2022).

the least frequently used. The varying frequencies across disciplines highlight clear disciplinary variation in syntactic preferences.

The syntactic variation (observed in this study) corresponds with the expert research on academic writing, indicating that different disciplines have distinct syntactic conventions. For instance, the high frequency of clausal and phrasal features in law corresponds with the disciplinary requirement for complex argumentative structures. This finding is consistent with Biber and Gray (2016), Casal et al. (2021), Dong et al. (2023), Elliott (2019), Gray (2015), Jalilifar et al. (2017), Lu et al. (2021), Saricaoglu and Atak (2022), Staples et al. (2016), Szczygłowska (2022, 2023), Tian and Zhang (2023), Ziaeian and Golparvar (2022), who documented that disciplines with intricate argumentative demands, like law, employ more so-

phisticated syntactic constructions. Conversely, the lower use of these features in economics aligns with Elliott (2019) and Staples et al. (2016), who observed that disciplines focused on quantitative analysis, such as economics, prefer simpler and more straightforward syntactic structures to maintain clarity and precision.

As observed in this study, the frequent use of linking adverbials and to-clauses in law supports the notion that such features are critical for articulating complex legal arguments. This finding reflects the emphasis on clarity and logical coherence in legal writing, highlighted by Gray (2015) and Jalilifar et al. (2017). On the other hand, the less frequent use of these features in sociology may be attributed to its focus on thematic exploration and narrative style, where complex syntactic structures are less focused.

These results (Table 3) contrast with some recent non-MDA-based research (Ahmad et al., 2022, 2023b, 2023c) in Pakistan, suggesting moderate disciplinary variation in advanced L2 academic writing. The results of this present study reveal that significant variation does exist, potentially due to thein-depth analysis and the inclusion of a wider range of disciplines.

The results (Table 3) highlight the importance of recognizing and teaching discipline-specific syntactic conventions in Pakistani academic writing. Improved instruction that addresses the unique syntactic needs of each field could improve writing effectiveness and alignment with the expert standards. For example, law students might benefit from training in complex clausal structures and phrasal features, while economics students could focus on achieving clarity through simpler syntax.

Syntactic Variation across Disciplines in Physical Sciences

The analysis of syntactic features in the advanced L2 academic writing of Pakistani students within the physical sciences reveals notable variations across different disciplines. The frequency (Table 4) of clausal features varied: engineering exhibited the highest frequency (407 instances), followed by computer science (311), physics (289), and mathematics (191). In contrast, phrasal features were predominant, with computer science leading at 73,184 instances, followed closely by physics (70,365), engineering (65,588), and mathematics (57,899). The distribution of intermediate features varied as well, with computer science showing the highest frequency (2,828 instances), followed by physics (2,494), engineering (2,253), and mathematics (1,757).

Overall, Pakistani academic writers in the physical sciences used 1,198, 267,036, and 9,332 clausal, phrasal, and intermediate features (Table 4). This distribution indicates a clear preference for phrasal features, with clausal features being the least frequently used across all disciplines. The variation in the frequency of different syntactic features indicates the presence of disciplinary variation, reflecting the specific rhetorical and communicative needs of each discipline.

The syntactic variation observed in the physical sciences (Table 4) aligns with the expert research practices, confirming that disciplinary demands shape academic writing. The higher frequency of phrasal features across all disciplines supports findings from Biber and Gray (2016), Dong et al. (2023), Elliott (2019), Gray (2015), Jalilifar et al. (2017), Saricaoglu and Atak (2022), Staples et al. (2016), Tian and Zhang (2023), Ziaeian and Golparvar (2022), who observed that academic writing relies on dense noun phrases to convey complex information efficiently. The predominance of phrasal features in disciplines like computer science and physics reflects the practical nature of these disciplines, where precise and concise communication is essential.

The lower use of clausal features in mathematics, as observed in this study (Table 4), is consistent with Staples et al. (2016) and Dong et al. (2023), who noted that mathematical writing prioritizes brevity and clarity, avoiding complex clausal structures that might obscure the logical flow of arguments. Similarly, the varied use of intermediate features, such as linking adverbials and finite adverbial clauses, across disciplines illustrates the differing needs for explicit logical connections in disciplines like engineering and computer science. These structures help clarify the relationships between technical processes and outcomes.

These results (Table 4) are also consistent with Casal et al. (2021) and Lu et al. (2021), who emphasized that syntactic choices in academic writing are closely tied to the epistemological and communicative practices of each discipline. The variation in the use of clausal and phrasal features across physical science disciplines confirms that advanced L2 academic writing is discipline-specific and varies within broader disciplinary categories.

The results (Table 4) have significant implications for academic writing pedagogy in the physical sciences. The observed syntactic variation suggests that discipline-specific instruction could enhance writing effectiveness. For example, computer science students might benefit from focusing on mastering phrasal structures that support technical descriptions, while mathematics students could be guided toward using concise and clear syntactic forms that align with mathematical writing conventions.

Table 3

Use of Syntactic Features across Disciplines in Social Sciences

FEATURES	FREQUENCIES IN DIFFERENT DISCIPLINES				TOTAL
Clausal Features	**Politics**	**Sociology**	**Law**	**Economics**	
Finite adverbial clauses	120	90	248	13	471
WH complement clauses	21	9	26	8	64
Verb + that-clauses	54	51	94	15	214
Clausal coordinating conjunctions	1031	672	1484	128	3315
Total	***1226***	***822***	***1852***	***164***	***4064***
Intermediate Features					
Adverbs	715	531	867	110	2223
Linking adverbials	1903	1392	3060	709	7064
Extraposed Adjective + that clauses	6	8	36	2	52
Noun + that-clauses	208	103	386	25	722
WH relative clauses	87	75	206	15	383
That relative clauses	923	562	1461	171	3117
Verb + to-clauses	452	233	757	102	1544
Desire verb + to-clauses	0	0	0	0	0
Raising structures and extraposed adjective + to-clauses	1	0	1	0	2
Noun + to-clauses	2756	1067	2756	347	6926
Verb + ing-clauses	95	88	95	66	344
Passive voice verbs	41	18	41	3	103
Passive nonfinite relative clauses	1321	649	1321	200	3491
Total	***8508***	***4726***	***10987***	***1750***	***25971***
Phrasal Features					
Nouns	109199	98777	135561	28532	372069
Attributive adjectives	17303	16769	21112	5292	60476
Premodifying nouns	18705	16052	20098	5416	60271
Nominalizations	16884	15761	28966	4842	66453
of genitives	8882	6808	9437	2259	27386
Prepositional phrases	6191	5724	9145	1664	22724
Total	***177164***	***159891***	***224319***	***48005***	***609379***

Note. Adapted from Ahmad (2022).

Syntactic Variation across Disciplines in Life Sciences

This study examined syntactic variation in the academic writing of Pakistani PhD students in life science disciplines, including agriculture, biology, food sciences, and psychology. The results reveal notable differences in using clausal, phrasal, and intermediate features.

The frequency (Table 5) of clausal features varied among disciplines: agriculture (191 instances), biology (293), food sciences (168), and psychology (226). Phrasal features were predominant across all life sciences disciplines, with biology leading (81,215 instances), followed by food sciences (61,581), agriculture (57,899), and psychology (55,042). Intermediate features were observed in the following fre-

Table 4

Use of Syntactic Features across Disciplines in Physical Sciences

FEATURES	FREQUENCIES IN DIFFERENT DISCIPLINES				TOTAL
Clausal Features	**Mathematics**	**Computer Science**	**Physics**	**Engineering**	
Finite adverbial clauses	18	34	34	46	132
WH complement clauses	8	16	10	9	43
Verb + that-clauses	20	9	29	2	60
Clausal coordinating conjunctions	145	252	216	350	963
Total	***191***	***311***	***289***	***407***	***1198***
Intermediate Features					
Adverbs	120	176	177	277	750
Linking adverbials	589	871	1027	641	3128
Extraposed Adjective + that clauses	6	3	9	3	21
Noun + that-clauses	27	82	60	62	231
WH relative clauses	26	32	49	56	163
That relative clauses	219	393	234	175	1021
Verb + to-clauses	59	254	146	143	602
Desire verb + to-clauses	0	0	0	0	0
Raising structures and extraposed adjective + to-clauses	0	0	0	0	0
Noun + to-clauses	394	549	359	460	1762
Verb + ing-clauses	77	151	60	123	411
Passive voice verbs	4	2	2	7	15
Passive nonfinite relative clauses	236	315	371	306	1228
Total	***1757***	***2828***	***2494***	***2253***	***9332***
Phrasal Features					
Nouns	35315	44825	42277	39672	162089
Attributive adjectives	5791	7593	8696	7153	29233
Premodifying nouns	6717	7521	7763	7816	29817
Nominalizations	5681	8179	6587	6969	27416
of genitives	2629	2692	2576	1871	9768
Prepositional phrases	1766	2374	2466	2107	8713
Total	***57899***	***73184***	***70365***	***65588***	***267036***

Note. Adapted from Ahmad (2022).

quencies: biology (1,951), psychology (1,891), agriculture (1,757), and food sciences (1,479).

In total, the use of clausal, phrasal, and intermediate features across life sciences disciplines was 878, 255,737, and 7,078, respectively (Table 5). Consistent with the findings in the physical sciences, phrasal features were predominant, while clausal features were the least frequently used. This variation across disciplines indicates the presence of disciplinary differences in syntactic preferences.

The observed syntactic variation in life sciences aligns with the established research practices, reinforcing the idea that academic writing characterizes the specific needs of each discipline. The dominance of phrasal features in life sciences supports the findings of Biber and Gray (2016) and

Gray (2015), who emphasized that academic writing relies on dense noun phrases to convey detailed information efficiently. The preference for phrasal structures in disciplines such as biology and food sciences can be attributed to the need for precise and comprehensive communication.

The lower frequency of clausal features in disciplines like agriculture and food sciences is consistent with Dong et al. (2023) and Staples et al. (2016), who found that certain scientific disciplines prioritize clarity and conciseness, avoiding complex clausal structures that could obscure the presentation of information. The variation in intermediate features, such as linking adverbials, reflects the different needs for explicit logical connections in academic writing.

These findings are also consistent with Casal et al. (2021) and Lu et al. (2021), who highlighted that syntactic choices are closely linked to the epistemological and communicative practices of each discipline. The variation in clausal and phrasal features across life sciences confirms that advanced academic writing is not only discipline-specific but also varies within broader disciplinary categories.

These findings suggest that academic writing instruction in the physical and life sciences would benefit from a focus on discipline-specific syntactic conventions. A change in instruction could improve writing effectiveness by addressing the unique syntactic needs of each field. For example, students in biology might benefit from focused instruction on mastering phrasal structures to support detailed technical descriptions, while those in agriculture could be guided toward using concise syntactic forms that align with the conventions of their field.

The results presented in Tables 2, 3, 4, and 5 reveal variations in the use of clausal, phrasal, and intermediate features by the Pakistani advanced L2 academic writers from arts and humanities, social sciences, life sciences, and physical sciences. These results seem to conform to the notion that variation is the characteristic feature of academic writing

Table 5

Use of Syntactic Features across Disciplines in Life Sciences

FEATURES	FREQUENCIES IN DIFFERENT DISCIPLINES				TOTAL
Clausal Features	**Psychology**	**Food Sciences**	**Biology**	**Agriculture**	
Finite adverbial clauses	57	17	24	18	116
WH complement clauses	10	2	7	8	27
Verb + that-clauses	32	13	14	20	79
Clausal coordinating conjunctions	127	136	248	145	656
Total	***226***	***168***	***293***	***191***	***878***
Intermediate Features					
Adverbs	135	94	162	120	511
Linking adverbials	650	523	634	589	2396
Extraposed Adjective + that clauses	0	3	2	6	11
Noun + that-clauses	49	25	41	27	142
WH relative clauses	22	9	28	26	85
That relative clauses	304	145	258	219	926
Verb + to-clauses	200	47	93	59	399
Desire verb + to-clauses	0	0	0	0	0
Raising structures and extraposed adjective+ to-clauses	0	0	0	0	0
Noun + to-clauses	329	246	302	394	1271
Verb + ing-clauses	29	68	66	77	240
Passive voice verbs	2	5	1	4	12
Passive nonfinite relative clauses	171	314	364	236	1085
Total	***1891***	***1479***	***1951***	***1757***	***7078***

FEATURES	FREQUENCIES IN DIFFERENT DISCIPLINES				TOTAL
Clausal Features	**Psychology**	**Food Sciences**	**Biology**	**Agriculture**	
Phrasal Features					
Nouns	36388	38485	50941	35315	161129
Attributive adjectives	5927	6196	8560	5791	26474
Premodifying nouns	4201	7838	10365	6717	29121
Nominalizations	5036	4837	5931	5681	21485
of genitives	2114	2712	3277	2629	10732
Prepositional phrases	1376	1513	2141	1766	6796
Total	***55042***	***61581***	***81215***	***57899***	***255737***

Note. Adapted from Ahmad (2022).

at the advanced level of education. However, this variation is at the level of frequency only. When we see the overall use of the said features it becomes evident (see Tables 2-5) that phrasal features are in the highest use across the four academic disciplinary divisions. This shows that Pakistani academic writing is not influenced by disciplinary variation. Thus, these results corroborate with the results presented in Ahmad et al. (2022). Furthermore, these results corroborate with the results discussed in Ahmad et al. (2023c) on the account that these results show the heterogynous and homogenous use of the syntactic features. The heterogeneity is marked by the difference in the use of the syntactic features, whereas the homogeneity is evidenced by the similarity in using the highest and lowest used features.

Syntactic variation is a characteristic feature that is reflected in the academic writing produced by the advanced level academic writers across disciplines (Ahmad et al., 2022, 2023b, 2023c; Biber & Gray, 2016; Casal et al., 2021; Dong et al., 2023; Gray, 2015; Jalilifar et al., 2017; Lu et al., 2021; Saricaoglu & Atak, 2022; Staples et al., 2016; Szczygłowska, 2022, 2023; Tian & Zhang, 2023). In fact, "Writing is discipline-specific, and writing talent is a function of the relationship between the individual and the domain. "The term "domain" refers to the discipline or field of writing. Every domain involves a group of individuals sharing "the same domain knowledge" and ideas that "emanate from these individuals" (Olthouse, 2013, p. 260). These ideas are shared in academic writing using different linguistic and syntactic devices and the use of these devices varies from discipline to discipline. That is why syntactic variation is essential in academic writing produced by writers specializing in different disciplines. Therefore, this study suggests that Pakistani advanced-level academic writers follow this practice to produce expert-like academic writing. This can be achieved by mastering the use of linguistic devices (Ahmad et al., 2019).

CONCLUSION

This study investigated variation in the use of syntactic features, that is, clausal, intermediate, and phrasal features in the academic writing produced by the Pakistani advanced L2 academic writers across four disciplinary divisions. The results revealed mixed findings. On one hand, the results showed variation in the use of the said features. This variation was marked by the difference infrequencies of the different types of phrasal, clausal, and intermediate features. On the other hand, the results demonstrated a similarity in the use of the said features, with the highest and lowest frequently used features being the most similar across disciplines. For example, phrasal features were identified as the most prevalent across all disciplines, while clausal features were identified as the least frequent. These results indicated the absence of variation that is the salient feature of advanced academic writing.

These results highlight the need for educators to integrate discipline-specific instruction into their teaching practices of the syntactic features. Such as, educators should change their pedagogical approaches to address the syntactic demands of different disciplines. By focusing on the specific syntactic features prevalent in each field, educators can better support students in developing advanced academic writing skills that align with disciplinary expectations.

For policymakers and syllabus designers, these results highlight the importance of incorporating discipline-specific syntactic features into academic curricula. This approach is supported by the notion, which emphasize that curricula

should reflect the syntactic demands of various academic disciplines to better prepare students for their respective disciplines.

This study paves the way for further research into syntactic practices across disciplines. Future studies could explore additional aspects (e.g., genre, register, and level of education) of syntactic variation and its implications for academic writing. Thus, further research could provide deeper insights into how syntactic practices influence academic communication across different disciplines, contributing to more effective and impactful scholarly writing.

ACKNOWLEDGEMENT

The authors would like to thank Prince Sultan University, Saudi Arabia for its technical support.

DECLARATION OF COMPETITING INTEREST

None declared.

AUTHOR CONTRIBUTIONS:

Muhammad Ahmad: conceptualization; writing – original draft; formal analysis; writing – review and editing; visualization.

Muhammad Asim Mahmood: supervision; review and editing.

Ali Raza Siddique: methodology; writing– review and editing; visualization.

Muhammad Imran: resources-provision of study materials; review and editing.

Norah Almusharraf: writing– review and editing.

REFERENCES

Abid, A., Manzoor, H., & Siddique, A. R. (2022). Cross-cultural examination of argumentative English essays: A multidimensional analysis of Pakistani and Chinese learners. *Linguistic Forum, 4*(4), 40–48. https://doi.org/10.53057/linfo/2022.4.4.6

Ahmad, M., Mahmood, M. A., & Siddique, A. R. (2019). Organisational skills in academic writing: A study on coherence and cohesion in Pakistani research abstracts. *Languages*, *4*(4), 1–26. https://doi.org/10.3390/languages4040092

Ahmad, M. (2022). *Phrasal complexity in Pakistani academic writing: A corpus-based comparative study of doctoral dissertations across disciplines* [Unpublished doctoral Ddssertation]. Government College University.

Ahmad, M., Mahmood, M. A., & Siddique, A. R. (2022). Exploring disciplinary variation in Pakistani academic writing: A corpus-based research on doctoral dissertations. *Pakistan Languages and Humanities Review*, *6*(4), 51–60. https://doi.org/10.47205/plhr.2022(6-IV)06

Ahmad, M., Mahmood, M. A., & Siddique, A. R. (2023a). Determining the L2 academic writing development stage: A corpus-based research on doctoral dissertations. *International Review of Applied Linguistics in Language Teaching*. https://doi.org/10.1515/iral-2023-0028

Ahmad, M., Mahmood, M. A., & Siddique, A. R. (2023b). Features characterizing academic writing: A corpus-based research on dissertations from hard and soft science disciplines. *City University Research Journal of Literature and Linguistics*, *6*(1), 61–78.

Ahmad, M., Mahmood, M. A., & Siddique, A. R. (2023c). Variation in academic writing: A corpus-based research on syntactic features across four disciplinary divisions. *Novitas-ROYAL (Research on Youth and Language)*, *17*(2), 50–65. https://doi.org/10.5281/zenodo.10015816

Ansarifar, A., Shahriari, H., & Pishghadam, R. (2018). Phrasal complexity in academic writing: A comparison of abstracts written by graduate students and expert writers in applied linguistics. *Journal of English for Academic Purposes, 31*, 58–71. https://doi.org/10.1016/j.jeap.2017.12.008

Azher, M., Faiz, R., Izhar, A., Nisa, R., & Ali, S. (2019). Revealing disciplinary variation in Pakistani academic writing: A multidimensional analysis. *International Journal of English Linguistics, 9*(2), 258–272.https://doi.org/10.5539/ijel.v9n2p258.

Aziz, A., Pathan, H., & Ali, S. (2016). Linguistic variation across major disciplinary groups of Pakistani academic writing: Multidimensional analysis of doctoral theses. *ARIEL-An International Research Journal of English Language and Literature*, *27*, 27–60.

Becher, T. (1981). Towards a definition of disciplinary cultures. *Studies in Higher Education, 6*(2), 109–122. https://doi.org/10.1080/03075078112331379362

Biber, D. (1988). *Variation across speech and writing*. Cambridge University Press.

Biber, D., Gray, B., & Poonpon, K. (2011). Should we use characteristics of conversation to measure grammatical complexity in L2 writing development? *TESOL Quarterly, 45*(1), 5–35. https://doi.org/10.5054/tq.2011.244483

Biber, D., & Gray, B. (2016 Eds.). Academic writing: Challenging the stereotypes. In D. Biber & B. Gray (Eds.), *Grammatical complexity in academic English: Linguistic change in writing* (pp. 1–42). Cambridge University Press. https://doi.org/10.1017/CBO9780511920776

Biber, D., Gray, B., & Staples, S. (2016). Contrasting the grammatical complexities of conversation and academic writing: Implications for EAP writing development and teaching. *Language in Focus, 2*(1), 1–18. https://doi.org/10.1515/lifijsal-2016-0001

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

Casal, J. E., Lu, X., Qiu, X., Wang, Y., & Zhang, G. (2021). Syntactic complexity across academic research article part-genres: A cross-disciplinary perspective. *Journal of English for Academic Purposes, 52*, 1–12. https://doi.org/10.1016/j.jeap.2021.100996

Dong, J., Wang, H., & Buckingham, L. (2023). Mapping out the disciplinary variation of syntactic complexity in student academic writing. *System, 113*, 1–15. https://doi.org/10.1016/j.system.2022.102974

Elliott, T. (2019). *Variation in use of nouns as nominal premodifiers in advanced student writing across academic disciplines* [Unpublished doctoral dissertation]. Iowa State University.

Fatima, N., Siddique, A. R., & Ahmad, M. (2023). Linguistic variations in the abstracts of Pakistani dissertations: A multidimensional analysis across disciplines. *University of Chitral Journal of Linguistics and Literature*, *7*(1), 50–80. https://doi.org/10.33195/btm19817

Gray, B. (2015). *Linguistic variation in research articles: When discipline tells only part of the story*. John Benjamins Publishing Company. https://doi.org/10.1075/scl.71

Housseine, B., & Oifaa, T. (2020). The significance of English scientific writing proficiency for publishing purposes: The case of Moroccan EFL PhD students at the Euromed University of Fes. *Linguistic Forum,* 2(3), 13–19. https://doi.org/10.53057/linfo/2020.2.3.3

Hyland, K. (2004). *Disciplinary discourses: Social interactions in academic writing*. University of Michigan Press.

Hyland, K. (2016). Methods and methodologies in second language writing research. *System, 59*, 116–125. https://doi.org/10.1016/j.system.2016.05.002

Jalilifar, A., White, P., & Malekizadeh, N. (2017). Exploring nominalization in scientific textbooks: A cross-disciplinary study of hard and soft sciences. *International Journal of English Studies, 17*(2), 1–20.https://doi.org/10.6018/ijes/2017/2/272781

Kachru, B., Kachru, Y., & Nelson, C. L. (2006). *Handbook of world Englishes*. Blackwell.

Khamaiseh, M. (2023). Toward a model for analyzing the rhetorical move structure of the master thesis introductions in applied linguistics. *Linguistic Forum,* 5(4), 1–29. http://dx.doi.org/10.53057/linfo/2023.5.4.1

Lan, G., & Sun, Y. (2018). A corpus-based investigation of noun phrase complexity in the L2 writings of a first-year composition course. *Journal of English for Academic Purposes, 38*, 14–24.https://doi.org/10.1016/j.jeap.2018.12.001

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

Lu, X. (2014). *Computational methods for corpus annotation and analysis*. Springer.

Lu, X., & Ai, H. (2015). Syntactic complexity in college-level English writing: Differences among writers with diverse L1 backgrounds. *Journal of Second Language Writing, 29*, 16–27. https://doi.org/10.1016/j.jslw.2015.06.003

Lu, X. (2017). Automated measurement of syntactic complexity in corpus-based L2 writing research and implications for writing assessment. *Language Testing, 34*(4), 493–511. https://doi.org/10.1177/0265532217710675

Lu, X., Casal, J. E., & Liu, Y. (2020). The rhetorical functions of syntactically complex sentences in social science research article introductions. *Journal of English for Academic Purposes, 44*, 1–16. https://doi.org/10.1016/j.jeap.2019.100832

Lu, X., Casal, J. E., Liu, Y., Kisselev, O., & Yoon, J. (2021). The relationship between syntactic complexity and rhetorical move-steps in research article introductions: Variation among four social science and engineering disciplines. *Journal of English for Academic Purposes, 52*, 1–13. https://doi.org/10.1016/j.jeap.2021.101006

Mahboob, A. (2008). Pakistani English: Morphology and syntax. In R. Mesthrie, B. Kortmann & E. Schneider (Ed.), *4 Africa, South and Southeast Asia* (pp. 578–592). De Gruyter Mouton. https://doi.org/10.1515/9783110208429.2.578.

Mahmood, M. A. (2009). *A corpus-analysis of Pakistani English* [Unpublished doctoral dissertation]. Bahauddin Zakariya University.

Nasseri, M. (2021). Is postgraduate English academic writing more clausal or phrasal? Syntactic complexification at the crossroads of genre, proficiency, and statistical modelling. *Journal of English for Academic Purposes*, *49*, 100940. https://doi.org/10.1016/j.jeap.2020.100940

Nesi, H., & Gardner, S. (2012). *Genres across the disciplines: Student writing in higher education*. Cambridge University Press.

Norris, J. M., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics, 30*(4), 555–578. https://doi.org/10.1093/applin/amp044

Olthouse, J. M. (2013). MFA writers' relationships with writing. *Journal of Advanced Academics, 24*(4), 259–274. https://doi.org/10.1177/1932202X13507972

Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college level L2 writing. *Applied Linguistics, 24*(4), 492–518. https://doi.org/10.1093/applin/24.4.492

Ortega, L. (2015). Syntactic complexity in L2 writing: Progress and expansion. *Journal of Second Language Writing, 29*, 82–94. https://doi.org/10.1016/j.jslw.2015.06.008

Parry, S. (1998). Disciplinary discourse in doctoral education. *Higher Education, 36*, 273–299. https://doi.org/10.1023/A:1003216613001

Pervez, N., Siddique, A. R., & Ahmad, M. (2024). Analyzing linguistic variations in the discussion sections of Pakistani English research articles: A multidimensional study. *LinguisticaSilesiana, 45*(1), 103–134. https://doi.org/10.24425/linsi.2024.150392

Rahman, T. (1990). *Pakistani English: The linguistic description of a non-native variety of English*. National Institute of Pakistan Studies.

Qasim, S., Hussain, Z., & Mahmood, M. A. (2017). A critical analysis of structuralist variations in academic writing. *Journal of Research (Humanities)*, *53*(1), 217–240.

Rashid, A., & Mahmood, M. A. (2019). Linguistic variations across disciplines: A multidimensional analysis of Pakistani research articles. *Global Social Sciences Review*, *4*(1), 34–48. http://dx.doi.org/10.31703/gssr.2019(IV-I).04

Saricaoglu, A., Bilki, Z., & Plakans, L. (2021). Syntactic complexity in learner-generated research paper introductions: Rhetorical functions and level of move/step realization. *Journal of English for Academic Purposes, 53*, 1–11. https://doi.org/10.1016/j.jeap.2021.101037

Saricaoglu, A., & Atak, N. (2022). Syntactic complexity and lexical complexity in argumentative writing: Variation by proficiency. *Novitas-ROYAL (Research on Youth and Language)*, *16*(1), 56–73.

Staples, S., Egbert, J., Biber, D., & Gray, B. (2016). Academic writing development at the university level: Phrasal and clausal complexity across level of study, discipline, and genre. *Written Communication, 33*(2), 149–183. https://doi.org/10.1177/0741088316631527

Szczygłowska, T. (2022). Lexical verbs of epistemic modality in academic written English: Disciplinary variation. *Linguistica Silesiana*,*43*, 91–111. https://doi.org/10.24425/linsi.2022.141219

Szczygłowska, T. (2023). Adverbial cohesion in English-medium academic prose: Disciplinary and linguacultural considerations. *Linguistica Silesiana, 44*(2), 81–108. https://doi.org/10.24425/linsi.2023.146648

Talaat, M. (1993). Lexical variation in Pakistani English. In R. J. Baumgardner (Ed.), *The English language in Pakistan* (pp. 55–62). Oxford University Press.

Tian, M., & Zhang, Y. (2023). Exploring nominalization in academic writing: A comparative study of shipbuilding and oceanography engineering and linguistics. *Athens Journal of Philology, 10*(2), 101–122. https://doi.org/10.30958/ajp.10-2-2

Wang, M., & Lowie, W. (2021). Understanding advanced level academic writing on syntactic complexity. In *Proceedings of the 35th Pacific Asia Conference on Language, Information and Computation* (pp. 455–465). Association for Computational Linguistics.

Xue, Q., & Ge, T. (2021). A corpus-based study on phrasal complexity in computer science abstracts of novice and advanced writers. *Open Journal of Modern Linguistics, 11*, 808–822. https://doi.org/10.4236/ojml.2021.115062

Yin, S., Gao, Y., & Lu, X. (2021). Syntactic complexity of research article part-genres: Differences between emerging and expert international publication writers. *System, 97*, 1–14. https://doi.org/10.1016/j.system.2020.102427

Yoon, H. J., & Polio, C. (2017). The linguistic development of students of English as a second language in two written genres. *Tesol Quarterly, 51*(2), 1–27. https://doi.org/10.1002/tesq.296

Ziaeian, E., & Golparvar, S. E. (2022). Fine-grained measures of syntactic complexity in the discussion section of research articles: The effect of discipline and language background. *Journal of English for Academic Purposes, 57*, 101116. https://doi.org/10.1016/j.jeap.2022.101116

https://doi.org/10.17323/jle.2024.16080

Evaluative Stancetaking in English-Medium Academic Prose: A Study of Research Article Abstracts by Russian and Chinese L2 Writers

Olga Boginskaya

Irkutsk National Research Technical University, Russian Federation

ABSTRACT

Background: Globalization has created the academic community's need to learn English in order to publish internationally and caused intensive research into academic prose by non-native writers with the aim of revealing prevailing culture-and discipline-specific rhetoric structures and suggesting ways of improving academic writing skills.

Purpose: This contrastive study explored preferences in the employment of stance features in English-medium research article abstracts by second language writers from two different cultural backgrounds (Russia and China) assuming that variations in stancetaking are culturally shaped.

Method: Hyland's (2005b) taxonomy of stance resources was adopted for the current study as the most comprehensive one including a wide range of writer-oriented features. This taxonomy can help identify pragmatic functions of linguistic markers used for stancetaking in academic prose. The methods of quantitative and qualitative analysis were applied.

Results: A contrastive analysis of the findings showed that the Russian and Chinese academic communities manifest different stancetaking preferences. The quantitative analysis revealed that Chinese-authored RA abstracts contained considerably more stance features than those written by their Russian counterparts. Most quantitative differences between the application of stance features by Russian and Chinese authors were statistically significant. It was also revealed that while the Chinese academic writers seemed to be more careful in making claims, anticipating and acknowledging, the Russian scholars chose to create an impression of certainty and assurance, instilling confidence in their readers. The differences in the employment of stance features identified in the study are likely to reflect culture-specific writing peculiarities of the Chinese and Russian academic communities which favour slightly different discursive strategies.

Conclusion: The findings carry pedagogical implications for academic writing course designers and can enhance L2 writers' familiarity with the culture-specific academic writing conventions in the knowledge domain.

KEYWORDS

stance, research article abstract, academic discourse, cross-cultural variation

Citation: Boginskaya, O. (2024). Evaluative stancetaking in English-medium academic prose: A study of research article abstracts by Russian and Chinese L2 writers. *Journal of Language and Education, 10*(3), 40-52. https://doi.org/10.17323/jle.2024.16080

Correspondence:
Olga Boginskaya
E-mail: Katiefan100@gmail.com

Received: October 10, 2022
Accepted: May 16, 2024
Published: September 30, 2024

INTRODUCTION

Globalization has created the need to learn academic English in order to teach disciplines or to communicate research results on the global academic arena. Many scholars from non-Anglophone countries are required to publish their research papers for promotion as universities rely on Science Citation Index indicators for their ranking. Over the past 15 years there has been a dramatic increase in the number of published English-language research articles by second language (L2) academic writers. This movement has caused intensive research into L2 English academic texts with the aim of revealing prevailing culture-specific rhetorical structures.

The motivation behind the selection of RA abstracts by Russian and Chinese

scholars for a contrastive analysis was significant cultural differences on the one hand and similarities in the academic contexts in which the Russian and Chinese academic communities have been developing in the early twenty first century on the other one. In both cultures, English was not used as a language of science and education. However, due to the process of globalization of education and science, English has been gaining influence there which is confirmed by the expansion of English language education, the initiation of the policy of using English as a medium of instruction and a growing number of English-medium publications by Russian and Chinese scholars supported by government and university policies (Boginskaya, 2024; Korotkina, 2018; He, 2017; Lei & Jiang, 2019). To enhance their research productivities, both Russian and Chinese universities are putting pressure on their scholars to publish their papers in international journals. This study may therefore shed light on how L2 academic writers from two culturally different academic communities increasingly favoring English as a medium of instruction and publication present their research results, acknowledge findings of other scholars and construct a dialogue with readers.

In addition, whilst the use of rhetorical patterns in RA abstracts has received much attention, little empirical research appears to have compared stance features in academic texts by L2 writers from a cross-cultural perspective. The literature review revealed that English-medium academic texts produced by L2 writers have mostly been analysed in terms of their distinctions from L1 academic prose. Differences in stance features in English academic texts by L2 writers with different cultural backgrounds have not received their deserved attention. It is possible that this analysis can reveal stance features in Chinese- and Russian-authored academic prose which has never been explored to date in terms of differences. It is assumed that despite a relative uniformity of RA abstracts imposed by generic requirements, there is significant intercultural variation in the stancetaking preferences of Chinese and Russian L2 writers influenced by culture-specific writing conventions. The present study focuses therefore on Russian- and Chinese-authored RA abstracts, in particular on variation in the employment of stance features in this academic genre, and seeks to answer the following questions:

(1) Are there any cross-cultural differences between Russian- and Chinese-authored RA abstracts in terms of the categories of stance the authors opt for?

(2) Are there any cross-cultural differences between Russian- and Chinese-authored RA abstracts in terms of the frequency of occurrence of the stancetaking categories and their types?

The following section describes previous studies on cultural preferences in the use of stance features in academic discourse and proceeds to discuss the concept of stancetaking as a rhetorical strategy.

LITERATURE REVIEW

Previous Studies on Cultural Preferences in the Use of Stance Features in Academic Prose

Comparative studies on stance in L1 and L2 academic writing have revealed significant differences in various cultural groups, highlighting the complex interplay between language, culture, and academic discourse. These differences are not simply about individual variations but rather reflect broader cultural norms regarding the role of the writer, the relationship between writer and reader, and the nature of knowledge itself.

The literature review indicated a large number of studies conducted by Chinese scholars, contrasting rhetorical features in Chinese-authored academic texts and those written by L1 English writers. Xiong's (2007) study, for example, made a structure comparison of metadiscourse patterns in English- and Chinese-authored papers and identified significant differences in the metadiscourse choices made by L1 and L2 Chinese writers. Hu and Cao (2011) made an attempt to contrast hedges and boosters in RA abstracts published in Chinese and English journals and revealed the predominance of boosters in Chinese-authored abstracts and a larger number of hedges in academic prose by L1 writers. In contrast, Hu & Wang's (2014) research into metadiscourse in Chinese- and Anglophone-authored RAs revealed that the Chinese scholars are more indirect than their Anglophone counterparts. They attribute this rhetorical feature to a Chinese culture-specific emphasis on harmonious relationships and desire to avoid face-threatening acts. It is interesting that the authors of these two studies which yielded totally different results explained variations in the use of hedges and boosters by Chinese and English authors by the influences of culture-specific rhetorical practices. Hu and Cao (2011) claimed that Chinese rhetorical norms encourage the framing of ideas in non-polemical terms and rely more on experiential knowledge in giving less attention to possible counterarguments, while Hu & Wang (2014) argued that due to a tendency to maintain interpersonal harmony that is highly valued in China, Chinese authors tend to avoid using boosters and confronting another person directly. Xu and Nesi (2019), who explored differences in evaluative style in RA introductions written by Chinese and British scholars, revealed that Chinese writers make more categorical assertions using more boosters, while British ones tend to acknowledge alternative views and are more explicit about their own attitudes towards the research issue and previous studies giving preference to hedging devices.

In addition to hedges and boosters, self-mention markers were also a focus of attention in a number of studies. For example, Wu & Zhu's (2015) contrastive study (2015) based on a corpus of English-medium research articles by English and Chinese scholars yielded different results and showed that English writers are more ready to present their self by taking on the role of discourse constructor, arguer and evaluator, whereas their Chinese counterparts were inclined to show their collective self and take on the role of researcher. Xia (2018) investigated research articles by English and Chinese writers across the four disciplines and revealed no consistent frequency differences in the self-mention usage.

With this profusion of studies into rhetorical features of Chinese-authored research articles, English-language academic prose by Russian writers has been explored in a scarce number of studies (Belyakova, 2017; Boginskaya, 2022; Krapivkina, 2014; Pyankova, 1994; Zanina, 2016). Pyankova (1994) studied differences between English and Russian academic texts and found that Russian scholars underuse self-mention markers and overuse passive and impersonal structures. Zanina (2016) analysed English- and Russian-medium RAs and revealed differences between the two sub-corpora in terms of the frequencies of occurrence of hedges and their subtypes. Zanina found that the overall frequency of hedges is significantly higher in the English RAs, and there are differences between the shares of the types of hedges in two sub-corpora. Belyakova (2017) carried out a cross-cultural comparison of English-medium RA abstracts by L2 (Russian) and L1 academic writers in the field of geoscience to investigate their metadiscourse features and found that writers from Russian academia disguise themselves to a larger extent, which was evident from the lower frequency of self-mention markers.

Taking prior research altogether, it seems differences in stance features in English-medium academic texts by L2 writers with different cultural backgrounds have not received their deserved attention. I assume that despite a relative uniformity of RA abstracts imposed by generic requirements, there is significant intercultural variation in the metadiscourse preferences of writers influenced by the culture-specific academic writing conventions or having adopted the Anglophone academic writing style. To fill the gap in the research into L2 academic writing, the present study focuses on Russian- and Chinese-authored RA abstracts, in particular on variation in the employment of stance features in this academic genre.

Stancetaking as a Rhetorical Strategy in Academic Discourse

Previous research has confirmed that stance is a crucial feature of academic discourse, and writers make choices on using stancetaking rhetorical devices to interact with an audience in different academic genres. Crismore and Farnsworth (1990) claimed that the employment of stance features such as hedges and boosters increase the persuasiveness of academic texts. Abdi (2002) argued that stance features help establish credibility. The same conclusions were made by Hyland (1998) who revealed that stance serves the persuasive function. According to Çiftçi and Akbaş (2021), expressing stance in academic discourse is crucial for constructing authorial identities and negotiating with readers.

The term 'stance' was introduced by Biber and Finegan (1988) who defined it as the linguistic expression of commitment to the proposition. In their later study, they referred personal feelings, attitudes, value judgments, and assessments to this category (Biber et al., 1999, p. 966). Since then, stance has been interpreted from diverse perspectives. Gray and Biber (2012) described it as the tool used for encoding opinions and assessments. Hyland (2005b) described stance as a type of evaluation, conceptualizing it as an attitudinal dimension that includes features used by writers to present themselves and convey their judgements and opinions. The engagement, one more type of evaluation, was referred to a dimension where the writer acknowledges others, recognizes the presence of his/her readers, focuses their attention, includes them as discourse participants, and guides them to interpretations (Hyland, 2005b). Hyland distinguished between three components of stance: evidentiality, affect, and presence. Evidentiality, as Hyland (2005b) put it, refers to the writer's expressed commitment to the reliability of the proposition and its potential impact on the reader; affect involves a wide range of attitudes towards what is said; and presence concerns the extent to which the writer projects him/herself into the text. These three components are realized in the four stance features: (1) hedges used to withhold complete commitment to a proposition, (2) boosters that help express certainty in what is being said and mark involvement with the topic and solidarity with an audience, (3) attitude markers used to indicate the writer's affective attitude to propositions, and (4) self-mentions that manifest the explicitness of an authorial presence. Hyland's taxonomy of stance resources was adopted for the current study as the most comprehensive one including a wide range of writer-oriented features. In addition, this taxonomy can help identify pragmatic functions of linguistic markers used to construct stance in RA abstracts by culturally diverse authors. The types of boosting were identified based on Hyland and Zou's (2021) typology. For the purposes of this study, I slightly modified this model to reflect the rhetorical peculiarities of the genre under discussion. Specifically, I added one more type of boosting – solidarity markers (e.g., *well-known, widely-known, common, as we know*) which also seem to enhance the degree of commitment to the propositional content, as they create a sense of solidarity with readers, thus building credibility and imparting confidence. These markers were referred to boosting devices due to their strong ability to enhance persuasiveness of proposition and suppress alternative views which can exist beyond well-known facts. Hedging devices found in the corpus were analysed using the model by the same authors (Hyland &

Zou, 2021) who identified three ways of conveying respect for colleagues' views. The taxonomies adopted in the current study are presented in Table 1.

In an attempt to analyse attitudinal stance, the model proposed by Dueñas (2010) was used (Table 2). This model is a simplified version of Swales and Burke's (2003) taxonomy that includes seven categories of evaluative adjectives: acuity, aesthetic appeal, assessment, deviance, relevance, size, and strength.

METHOD

Corpus Design

The present study was conducted on a corpus of RA abstracts derived from six Scopus-indexed journals in the field of engineering: *Energies, Metal Powder Report, Renewable Energy, Symmetry, Materials Today, and npj 2D Materials and Applications*.

Having identified the target journals, 234 research article abstracts (117 Chinese-authored and 117 Russian-authored) were randomly selected to ensure a good degree of objectivity and comparability of texts. The RA abstracts were distributed between the two sub-corpora: sub-corpus 1 (SC1) includes Chinese-authored RA abstracts and sub-corpus 2 (SC2) includes Russian-authored RA abstracts. The number of words in each sub-corpus is 12,875 and 11,574, which makes 24,449 words altogether. The corpus is considered large enough to provide sufficient examples of the target rhetorical features. To eliminate the impact of a publication period, only RA abstracts from the most recent issues of each journal published between 2017 and 2022 were selected in order to exhibit the linguistic characteristics of present-day academic discourse. Only one RA abstract from every author was selected in order to avoid the influence of an individual writing style. The origin of the authors was judged by their names and affiliations. The first author of each article with a Russian name and an affiliation with a Russian university was taken to be a Russian author, and the first author of each article with a Chinese name and an affiliation with a Chinese university was taken to be a Chinese author. It is assumed that only the persons listed as authors are responsible for the language used in the RA abstracts. Copyeditors' contribution is disregarded since it is difficult to separate it from that of the authors. All the journals selected to build the corpus have a large readership and rather high prestige in their fields (Q1-Q2 in Scopus). They impose strict requirements on the quality of English used in research articles. This was the motive of selecting these journals for the current study.

The corpus was built to ensure comparability in terms of genre (RA abstracts), authors' origin (Russia and China), field (engineering sciences), and currency. This methodological equivalence, as Hu & Wang (2014, p. 18) put it, allows the researcher "to establish a common platform for making meaningful comparisons and drawing reliable and valid conclusions about cross-disciplinary and cross-linguistic differences/similarities".

Table 1

Types of Hedges and Boosters

Categories	Types	Function
Hedges	Plausibility hedges	Signal that a claim is based on assumptions
	Downtoners	Mitigate the intensity of a statement
	Rounders	Indicate an approximation
Boosters	Certainty markers	Indicate the writer's epistemic conviction
	Extremity markers	Emphasize the upper edge of a continuum
	Intensity markers	Amplify the emotive strength of a statement
	Solidarity markers	Signal well–known facts and values

Table 2

Types of Attitudinal Stance

Types	Function
Assessment markers	signal the writer's evaluation of the study (novelty, usefulness, validity)
Significance markers	signal relevance or importance of the study
Emotion markers	refer to the writer's affective position and generate the same sentiment in readers

Procedure

Since the study aimed to compare the use of stance in English-medium RA abstracts written by L2 English writers from two different academic cultures, the methods of quantitative and qualitative analysis were applied. The quantitative analysis assisted with WordSmith Tools 5 was conducted to reveal the frequency of stance features in RA abstracts selected to build the corpus. First, stance features were identified using this text analysis software, which provides details about the text and can ensure the accuracy of research results. Hyland's (2005b) framework of stance features (hedging, boosting, attitude, and self-mention) was adopted as the initial model for revealing stance features. Second, the markers found in the corpus were manually analyzed in context to determine their pragmatic functions and to ensure that they serve as stance features. Finally, the occurrences of stance markers were classified and combined. The inter-group (Russian versus Chinese authors) contrastive analysis was conducted to find potential similarities and differences between the groups. After classifying stance features and calculating their frequencies, Rayson's (2008) Log Likelihood Calculator[1] was used to find out whether the differences stance features are significant.

A careful analysis of the context was conducted to classify stance features by their categories and types within each category and interpret differences revealed. The identification of stance features seems to be a complicated procedure as it is necessary to decode context-dependent lexical items to interpret how the stancetaking strategy is realized. To ensure in-depth exploration into the use of stance, examples were taken from the corpus being studied and explanations were provided to describe the rhetorical functions of stance features found in the corpus.

The following are the steps of the analysis as it appears in the paper:

(1) Deriving RA abstracts from the electronic versions of the selected journals and exporting them to two Microsoft Word files by the names and affiliations of the authors.
(2) Compiling two sub-corpora containing Chinese-authored RA abstracts and Russian-authored RA abstracts.
(3) Examining each sub-corpus for presence of stance features.
(4) Labeling each instance as a stance feature based on contextual analysis in order to be certain about its function as stancetaking.
(5) Arranging the stance features into groups based on the above-mentioned taxonomies of stance and its types.
(6) Counting the raw numbers of stance features for each group in each sub-corpus.
(7) Normalizing the occurrences of the stancetaking categories and their types found in each sub-corpus to 1000 words.
(8) Summarizing the results in a table format.
(9) Interpreting the rhetorical functions of the stance features found in the two corpora.

RESULTS

Quantitative Analysis of Stance Features in RA Abstracts by Chinese and Russian Writers

The outcome of the quantitative analysis shows similarities and differences in the use of stance features by Russian and Chinese authors in terms of frequencies.

Overall, I found 718 stance features in SC1 and 550 ones in SC2. Hedges were the most frequent stance resources in the Chinese-authored texts. Attitude markers ranked second in SC1 and first in SC2. Boosters exhibited the greatest difference with nearly twice as many cases in the Russian sub-corpus. Self-mention markers ranked third in SC1 and were the least frequently used in SC2. The results suggest that researchers from both cultural backgrounds are conscious of the need to engage with the content and readers in a different way. The difference was significant for almost all stance features (log Likelihood = 41,87 for hedges, 33,95 – for boosters, 7,48 – for self-mention markers, $p<0.01$). The difference between the corpora was statistically insignificant only for the total number of attitude markers (log Likelihood = 0,01). The raw and normalized frequencies of stance features for SC1 and SC2 are shown in Table 3.

The types of stance features were also analysed within each stance category. The results are presented in Tables 4-7.

Table 4 manifests that hedging devices found in the two sub-corpora differ both in terms of frequencies and types. The difference between the total use of hedges was statistically significant (log Likelihood = 41,87, $p<0.01$). The table also shows that the general trends in the ranking of the types of hedging are similar: while plausibility hedges clearly prevail in both sub-corpora, rounders rank third. The difference between the two subcorpora in terms of the types of hedging was statistically significant for plausibility hedges (log Likelihood = 26,83, $p<0.01$) and downtoners (log Likelihood = 19,07, $p<0.01$). For rounders it was 0,22.

Similar to hedges, boosters differed both numerically with a statistically significant difference (log Likelihood = 33,95) and functionally. The difference was statistically significant for certainty markers (log Likelihood = 7,29, $p<0.01$) and intensity markers (log Likelihood = 32,03). For extremity and

[1] Rayson's Log Likelihood Calculator. https://ucrel-api.lancaster.ac.uk/

solidarity markers, no statistically significant difference was revealed between the two subcorpora (log Likelihood values were 0,78 and 0,91, respectively). Table 5 illustrates that all the four types of boosting are used more frequently by the Russian authors. The distribution of these types within each sub-corpus is also different. In the Russian sub-corpus, intensity boosters rank first followed by certainty, extremity and solidarity markers. In the Chinese sub-corpus, certainty markers prevail. Intensity markers rank second followed by extremity and solidarity markers. In both sub-corpora, the solidarity markers are infrequent.

The analysis revealed no statistically significant difference in the total application of attitude markers by the Chinese and Russian authors (log Likelihood = 0,01). However, the difference was statistically significant for the types of attitudes: for assessment markers, log Likelihood = 40,6 and for significance markers log Likelihood = 35,47, $p<0.01$.

Table 3

Stance Features in the Two Sub-Corpora (Raw and Normalized Frequencies).

Stance features	SC1	SC2
Hedges	355 (31.1)	179 (15.1)
Boosters	59 (5.2)	128 (10.8)
Attitude markers	209 (18.3)	189 (16)
Self-mention markers	95 (8.2)	54 (4.5)
Total	***718 (62.8)***	***550 (46.4)***

Table 4

Types of Hedging in the Two Sub-Corpora (Raw and Normalized Frequencies)

Hedges	SC1	SC2
Plausibility hedges	219 (19.2)	109 (9.2)
Downtoners	123 (10.8)	56 (4.7)
Rounders	13 (1.1)	14 (1.2)
Total	***355 (31.1)***	***179 (15.1)***

Table 5

Types of Boosting in the Two Sub-Corpora (Raw and Normalized Frequencies)

Boosters	SC1	SC2
Certainty markers	31 (2.7)	51 (4.3)
Extremity markers	10 (0.9)	13 (1.1)
Intensity markers	16 (1.4)	60 (5.1)
Solidarity markers	2 (0.2)	4 (0.3)
Total	***59 (5.2)***	***128 (10.8)***

Table 6

Types of Attitudes in the Two Sub-Corpora (Raw and Normalized fFequencies)

Attitude markers	SC1	SC2
Assessment markers	55 (4.8)	131 (11.1)
Significance markers	154 (13.5)	58 (5.9)
Total	***209 (18.3)***	***189 (16)***

The types of self-mention markers are presented in Table 7.

The results also revealed a statistically significant difference between the total use of self-mention markers (log Likelihood = 7,48, p<0.01) by Russian and Chinese authors

Quantitative results do not fully explicate the ways in which culturally diverse academic writers deploy the stance features. Thus, a detailed qualitative analysis of their functions is presented below.

Qualitative Analysis of Stance Features in RA Abstracts by Chinese and Russian Writers

Hedges

Hedges downplay a writer's commitment to a proposition, modifying its certainty, helping to acknowledge alternative viewpoints, and steering the reader to the conclusion or reasoning of the writer's choice. Here is an example of the plausibility hedge derived from the Chinese sub-corpus that indicates that the statement is based on an assumption rather than facts.

1. Our results **suggest** that population-specific assemblies are necessary for genetic and medical analysis. (SC1)

The humility-indicating hedge signals an awareness of alternative viewpoints and seeks to avoid potential criticism. In (2) and (3), the plausibility hedges also signal that the claims are based on author's assumptions rather than facts:

2. *Determining the maximally economically efficient HP capacity **may** be the key limiting factor for the potential range of solutions (SC2)*

3. *Nickel (Ni) is ubiquitous in the environment and evidence **has suggested** that Ni **can** cause ocular surface inflammation. (SC1)*

In interpreting their research results, the authors draw conclusions using speculative language to avoid commitment to their claims. This approach to reasoning is helpful in achieving this rhetorical purpose.

Discrepancies in the use of downtoners by the Russian and Chinese authors indicates that they tend to show some professional modesty and soften claims in a different way. In (4), the downtowner *often* protects the writer against inaccuracy of research results. In (5), *quite* as a downtoner mitigates the intensity of the statement and lessens the certainty of the authorial claim.

4. *However, its efficacy is **often** limited by the immunosuppressive tumor microenvironment (TME) in solid tumors. (SC1)*

5. *The stable operation is **quite** important to the safety of the engine. (SC2)*

In (6), the downtoners *usually* and *almost* might convey a certain qualification with regard to the degree of accuracy of the conclusions demonstrating that the statement might be inaccurate (Hyland, 1998).

6. *The main problem of fluid sampling is due to the fact that even a small pressure drawdown **usually** leads to the formation of a two-phase mixture in the bottom hole area, and it is **almost** impossible to take representative samples with downhole samplers or a formation tester (SC1).*

One more type of hedging distinguished by Hyland and Zou (2021) – rounders indicating an approximation – was surprisingly rather scarce in both sub-corpora. In hard sciences, which present a large amount of statistical data, rounders are considered to be more common than in soft sciences which deal with verbal rather than numeric data (Hyland & Zou, 2021). However, the corpus features the authors' tendency to present precise numerical data without approximating it. Here is an example of the rounder found in the Chinese-authored RA abstract:

7. In the middle of fatigue, the reinforcement material can reduce the deterioration value of the bridge deck by **approximately** 50%. (SC1)

By making the number a little fuzzy, the adverb employed as a rounder expresses approximation, thereby reducing accuracy of the claim.

Tale 7

Types of Self-Mention in the Two Sub-Corpora (Raw and Normalized Frequencies)

Self-mention markers	SC1	SC2
First-person plural pronoun (we)	66 (5.8)	35 (3)
Possessive adjectives (our)	29 (2.4)	19 (1.5)
Total	***95 (8.2)***	***54 (4.5)***

Boosters

In contrast to hedges, boosters function by "presenting the proposition with conviction while marking involvement, solidarity and engagement with readers" (Hyland, 2005a, p. 145). An analysis has revealed the higher normalized frequency of these devices in SC2, which indicates that Russian writers tend to occupy a stronger stance and are keener to express their convictions and highlight the significance of their work. Here are two examples from the corpus.

8. *It was **evident** from the study's findings that the pilot tunnels excavation and the arches installation accounted for 67% and 23.1% of the total surface settlement, respectively.* (SC1)

9. *Despite their priority, molecular and genetic aspects of diabetes pathogenesis are poorly understood; however, the involvement of oxidative stress in this process is **undoubted**.* (SC2)

The certainty markers used in the above examples help remove any doubts about the claims closing down potential opposition. They help authors build a strong voice and indicate the writer's epistemic conviction. In addition to conveying an authorial assertive stance, these devices play a crucial role in realizing the fact-based approach to reasoning.

10. *The calculations **demonstrate** that the developed algorithms have high speed and high performance in detecting deviations of the electrical power quality*. (SC2)

11. *This work **proves** that the adding of an integrated catalyst layer is a promising strategy to directly utilize methanol for Ni-YSZ anode-supported SOFCs.* (SC1)

In the above examples, the authors anticipate possible responses from the reader but choose to prevent them. The boosting verb *demonstrate* is used to indicate that the claims are based on accurate data rather than on authors' assumptions. The verb *prove* expresses conviction with which the authors communicate their research results obtained through the experiment.

12. *For the same design parameters, the creep damage was **evidently** greater than the fatigue damage*. (SC1)

The adverb is used here to signal accepted truth – that is, it downplays the author's involvement by implying that the claim is one that is already generally accepted in mechanics.

Intensity items, one more type of boosting, function by amplifying the emotive strength of a statement. In contrast to certainty markers, they add affective color to claims rather than concern epistemic assurance (Hyland & Zou, 2021).

13. *The structure–property relationships are **particularly** emphasized*. (SC1)

14. *When released to the environment, the rocket fuel unsymmetrical dimethylhydrazine (UDMH) undergoes oxidative transformations, resulting in the formation of an **extremely** large number of nitrogen-containing transformation products*. (SC2)

The writers consider the issues they are going to discuss fundamental and make attempts to encourage their audiences to perceive them in the same vein.

Extremity markers, that rank third in both sub-corpora, "emphasize the upper edge of a continuum" (Hyland & Zou, 2021, p. 8), as in here:

15. *Human error is **the most common** accident in industrial systems*. (SC1)

16. *Energetic materials constitute one of **the most important** subtypes of functional materials used for various applications*. (SC1)

By upgrading the propositions, the writers emphasize the frequency of human errors (15) and the importance of the subtypes of materials under study (16) without the need for elaboration.

Finally, solidarity markers, which were the least frequent boosting items in both sub-corpora, contribute to the persuasiveness of authorial claims through the appeal to shared knowledge. These devices are used to argue that knowledge claims are widely accepted or known in academia and demonstrate that the authors expect their readers to be familiar with certain facts and feel solidarity thus taking the audience's knowledge for granted. The following example illustrates the case.

17. *As it is **widely known**, along with the manufacturer and the consumer, the authorities that implement the state policy on ensuring the safety of products put into circulation in the country*. (SC2)

Attitude Markers

Assessment markers signal the writer's evaluation of the study emphasizing interesting, crucial or debatable findings. This rhetorical strategy helps promote and evaluate research. Here are two examples from the Chinese sub-corpus.

18. *This study provides not only the first systematic understanding about the physics of CE, but also demonstrates that the triboelectric nanogenerator (TENG) is an **effective** method for studying the nature of CE between any materials*. (SC1)

19. *Our work not only provides a **new effective** way to reprogram TME in vivo, but also shed light on the design of novel bioorthogonal nanozymes for cancer immunotherapy*. (SC1)

Assessing the efficiency and novelty of the methods are key features of research, particularly among engineering scholars whose studies are typically aimed at revealing innovative ways of solving practical problems.

Significance attitude markers are used to show the role of research results and present a valid argument, as in the examples below.

20. *Well test equipment setup becomes much more compact and less weight; the costs of drilling time are reduced, which is **viably important** for well testing on the Arctic conditions*. (SC2).

21. *Thus, the theoretical and practical **significance** of the study lies in revealing the features and problems of the Iraq oil industry infrastructure*. (SC2)

The significant type of attitudes markers is used here to evaluate the research results. The authors highlight the importance of their studies for the body of disciplinary knowledge.

Finally, it is not surprising that the emotional type of attitude, which refers to how the writer feels rather than evaluate the research, does not appear in the corpus. The RA abstract acts as a time-saving tool intended to represent the content as accurately and concisely as possible. Due to its length being limited to 200-250 words, the writer has no space to build an emotive stance. Therefore, the absence of emotion attitude markers in both sub-corpora might be due to the generic features of the texts under consideration rather than culturally shaped differences.

Self-Mention Markers

Self-mention sends an indication to the reader of the perspective from which the statement should be interpreted (Hyland, 2005a). Although it is often taught to avoid personal pronouns in RA articles, an analysis shows a regular use of self-mentions to emphasize the importance that should be given to authorial claims or choices. Graff and Birkenstein (2010), for example, advise to abandon the perceived prohibition about the use of personal pronouns, because they will not eliminate the subjective opinions and may hurt writers' abilities to distinguish their views from other people's perspectives. In the same vein, Hyland (2005a) argues that self-mention markers are important to emphasize the writer's contribution. In my present corpus, however, in contrast to some other studies (Fløttum, 2012; Hyland, 2001), self-mentions are not frequent stance features. The analysis revealed that engineering writers tend to avoid creating an authorial presence. Here are rare examples from the corpus that feature the use of self-mentions to present a discoursal self.

22. ***Our** conclusion is that electron transfer is the dominant mechanism for CE between solid–solid pairs*. (SC1)

23. *In this paper, **we** present the results of dispersion of thermodynamically immiscible polypropylene (PP)*. (SC2)

In (22), the possessive adjective *our* helps the authors to outline their conclusion, i.e. to express their knowledge claim about the topic. In (23), *we*-pronoun is used for creating an identity of the author as an architect of the text who shows how the text is organized, i.e. for effecting the rhetorical function of presenting the findings.

Interestingly, the first-person plural pronouns appeared only in the co-authored texts in both sub-corpora. In those 10 single-authored articles selected to build the corpus, there were no occurrences of these pronouns. This finding suggests that both Russian and Chinese authors tend to avoid expressing their collective identities, indicating they are members of a larger community, which is a typical function of *Pluralis Majestatis*. In the Russian academic community, for example, *we* often signals the author's desire to enhance the significance of the work presenting his/her claims as the opinion of a scientific school (Krapivkina, 2014).

What is more, no occurrences of first-person singular pronouns were found in the corpus, which might be explained by the fact that almost all the RAs selected to build the corpus are co-authored. Only nine RAs in SC2 and one RA in SC1 are single authored, but the pronouns *I* and *me* were not found in these texts. This finding contracts the results obtained by some other scholars. For example, Seone (2013) traced an increase in the use of first-person singular pronouns in hard science articles. The same trend towards informality was emphasized by Hyland and Jiang (2017). However, like the present study, Harwood's (2005) research revealed that the hard sciences show a preference for the first plural pronoun.

Summing up the similarities and differences in the use of self-mentions in the two sub-corpora (no occurrences of first-person singular pronouns and the heavier use of first-person plural pronouns by Chinese authors), it seems that Chinese authors are more aware of the role of these devices in persuading readers to accept their claims. The lower frequency of self-mentions in SC2 shows that Russian engineering scholars tend to be more objective in their academic writing and hide their personality behind impersonal constructions emphasizing research results rather than an authorial stance

DISCUSSION

Conducted from a cross-cultural perspective, this study aimed to explore culturally shaped variations in the employment of stance features in a corpus of English-medium RA abstracts written by Russian and Chinese engineering authors which previously did not attract much attention of linguists. The study was based on the assumption that variations in stancetaking are culturally shaped. This assumption relied on previous studies in the field of contrastive rhetoric (Alonso-Almeida, 2014; Belyakova, 2017; Hryniuk, 2018; Isık-Tas, 2017; Mikolaychik, 2019; Walková, 2018; Wu & Zhu, 2015, etc.), which emphasized that despite internationally recognized generic requirements, there is significant intercultural variation in the rhetorical preferences of non-native writers influenced by culture-specific writing conventions.

A comparison of the RA abstracts has showed that the Russian and Chinese engineering communities manifest different stancetaking preferences. The research questions guiding this study asked about cross-cultural differences between Russian- and Chinese-authored RA abstracts in terms of the categories of stance the authors opt for and their frequencies. The study revealed significant differences between the two sub-corpora, both quantitative and qualitative. While the Chinese authors seemed to be much more careful in making claims and presenting findings thus securing their academic credibility, the Russian ones preferred to suppress alternatives and leave little room to the reader's own interpretations thus creating an impression of certainty and assurance and instilling confidence in the academic audience. Regarding the attitude markers, the differences were also significant. The Russian writers used attitude markers to assess the efficiency and novelty of the methods developed while their Chinese counterparts exploit evaluative features to highlight the significance of their studies. The higher share of significance attitude markers found in the Chinese sub-corpus might be explained taking into account the very competitive nature of the Chinese academic community in which it seems necessary to promote one's research as regards its contribution to the body of disciplinary knowledge in order to be published. The use of self-mentions was also different in the two sub-corpora. They were more frequent in the Chinese sub-corpus, which signals that the Russian scholars opted for an impersonal style of writing favored by the Russian academic writing conventions.

The differences in the employment of stance features identified in the study allow me to suggest that the Chinese and Russian academic communities favour slightly different writing strategies. This conclusion is in line with the results obtained by other scholars, who emphasize the role of culture in academic writing (AlGhamdi & Suleiman Alyousef, 2022; Belyakova, 2017; Boginskaya, 2023; Isık-Tas, 2017; Mikolaychik, 2019; Shchemeleva, 2015; Walková, 2018).

The role played by the disciplinary writing culture in engineering RA abstracts published in English is outweighed by the culturally shaped conventions, and the disciplinary beliefs and understandings do not predominate over the cultural ones. This conclusion has been already made by Hyland (2013), according to whom writing is bound up with culture since it makes available certain taken-for-granted ways of organizing our understanding. The analysis revealed that in an effort to be more confident, impersonal and objective, engineering researchers from Russian academia avoid using hedging devices which can reduce assertiveness of the argument and precision of research results. The ability to present arguments as established facts and make assertive knowledge claims is encouraged. This finding is in line with the conclusion made by Prokhorov (2006) who described Russian academic discourse as uncompromising, categorical, and featuring assertions. For Russian writers, it is therefore a tricky facet to find a balance between the disciplinary objectivity and impersonality requirements and the need to demonstrate the writer's personality and mitigate claims favored by the international academic writing conventions.

Regarding the Chinese writers, despite the fact that Chinese academia valorizes objectivity and impersonality and tends to downplay "measured uncertainty, guarded commitment, tentativeness, subjectivity, and possibility of multiple interpretations" (Hu & Cao, 2011, p. 2805), Chinese authors seemed to be less overtly critical than their Russian counterparts, taking a more tentative approach which is in line with the Confucian beliefs and dogmas (Lee, 2015). In addition, in Chinese-authored academic prose an awareness of the Anglophone academic writing conventions is more evident than in the Russian-authored ones. It seems that Chinese engineering scholars have mastered the Anglophone standards of academic writing better than their Russian counterparts.

Therefore, the study fully confirmed the assumption made in this and previous studies and revealed that the Russian and Chinese academic communities manifest different stancetaking preferences.

It should be recognized that the research results presented here are limited due to a small number of RA abstracts collected to build the corpus. Due to this limitation, the research results can be interpreted only as trends in the discipline which may be confirmed or rejected by comparative research based on a larger corpus.

Implications

The findings suggest that it is necessary to teach stance building strategies to L2 writers and raise their awareness of cultural, disciplinary and generic differences in the use of interactional devices in academic prose. Stancetaking conventions are not always easily understood by L2 writers due

to a lack of explicit practice. These points make stancetaking deserve a prominent place in EAP courses.

The need for teaching stance features emphasized in the current study supports the findings of previous research which revealed that L2 academic writers find it difficult to give a credible representation of themselves through the use of stance resources (Hyland, 2005). As Vande Kopple (2021) suggests, meanings conveyed by stance can be nuanced and L2 writers must carefully examine linguistic elements, meanings, and probable effects of those meanings within a particular context. It is advisable to help students to understand both cultural, genre-specific and disciplinary stancetaking variations through a systemic instruction.

Firstly, when teaching stance, EAP teachers should use common stance features. Secondly, more examples of how to use stance devices in different academic genres should be introduced by EAP teachers. They should be taken from academic prose by L1 academic writers in the discipline. Thirdly, it seems that explicit teaching of stance in different academic genres can help raise awareness of their interactional aspect among students and increase their ability to interact with the targeted audience and make their claims more persuasive.

The following exercises can be used to develop stancetaking competence in novice L2 academic writers.

(1) Underline the stance feature used in the excerpt.

(2) Identify the purpose of using the hedges in the excerpt.

(3) Produce the more persuasive argument using the boosting device.

(4) Reduce the degree of commitment in the following statement using the hedging device.

(5) Rewrite the following sentence using appropriate stance feature.

In EAP classes with students majoring in different disciplinary fields, the findings of the present study may be highlighted through consciousness-raising classroom activities. There are examples of these activities in various disciplines that can be implemented. EAP teachers might encourage their students to compare the use of stance features in different disciplines and draw students' attention to differences between them. The teacher may ask students to read two or three academic texts from different disciplines and compare stance use. The students may be tasked to report their findings during classroom sessions.

EAP teachers may also guide students to write their academic texts with a greater sense of responsibility, for example, by using stance features intentionally in their writing. Explicit instruction of linguistic features, including how to use stance devices appropriately, by evaluating academic texts written by other students can also help improve the academic writing skills.

CONCLUSION

In a globalized world, nations with greater academic power such as the USA and the UK are located at the center of academic knowledge production controlling high-impact academic journals and prescribing communication rules. This is one of the reasons why most international journals require authors to submit only English-medium manuscripts thus ensuring an academic monopoly for Anglophone writers. In most non-Anglophone countries, including China and Russia, the universities have imposed policies to promote publications in international academic journals with the aim to increase the country's share of global research output. To be efficient, researchers should publish their findings in English, which requires high English language proficiency and knowledge of the academic writing conventions to conform to the expectations of global academia and successfully integrate into it.

The aim of the present study was to contribute to a better understanding of stance as a crucial feature academic writing through a contrastive analysis of L2-authored academic texts and to provide an answer to the question of how culture manifests itself in academic communication. The materials for the study were derived from six Scopus-indexed journals in the field of engineering.

The results confirmed the assumption about the reflection of cultural contexts in academic prose by L2 writers. A comparison of the RA abstracts has showed that the Chinese and Russian academic communities manifest different stancetaking preferences. Cultural values appear to be determinants of academic writers' rhetorical behaviour affecting the ways they express the commitment to their claims and interact with the reader.

The significance of the present research lies in showing how and to what extent Russian and Chinese L2 writers use stance features in their academic writing. It also demonstrates that L2 writers realize that in order to interact successfully with an audience and to promote their research results, they need to follow the international writing conventions. I hope that this study brings some pedagogical implications both for novice writers and EAP teachers. Apart from pointing to the crucial role of writer-reader interaction, it could help to raise novice writers' awareness of how stance features contribute to the pragmatic effect of academic prose. The study could also be useful to EAP teachers by providing them with some valuable insights into culture-specific L2 academic writing and indicating those areas which deserve more attention in EAP course.

Through a study of interactional preferences of writers from a larger number of disciplines, we will learn more about rhetorical practices and values which would help novice writers learn academic style features typically used in a disciplinary community to produce knowledge in an accepted way. This analysis was limited to written academic discourse. It will be of interest to see if disciplinary differences in stancetaking can also be observed in oral presentations of research results. This will be pedagogically useful for students as it will draw their attention to the stancetaking discrepancies between oral academic genres and make them sensitive to the nuances of oral academic discourse. Diachronic variation in the use of stance features in L2 academic prose could be also of interest.

DECLARATION OF COMPETITING INTEREST

None declared

REFERENCES

Abdi, R. (2002). Interpersonal metadiscourse: An indicator of interaction and identity. *Discourse Studies*, *4*(2),39–45. http://dx.doi.org/10.1177/14614456020040020101

Alghamdi, G.A., & Suleiman Alyousef, H. (2022). The construction of knowledge claims in three disciplines: An exploration of hedging and boosting strategies in research articles written in English by Arab and Anglophone writers. *Journal of Language and Education*, *8*(2), 32-48. https://doi.org/10.17323/jle.2022.12363

Alonso-Almeida, F. (2014). Evidential and epistemic devices in English and Spanish medical, computing and legal scientific abstracts: A contrastive study. In M. Bondi & R. Lorés Sanz (Eds.), *Abstracts in academic discourse: Variation and change* (pp. 21-42). Peter Lang.

Belyakova, M. (2017). English-Russian cross-linguistic comparison of research article abstracts in geoscience. *Estudios de Lingüística Universidad de Alicante, 31*, 27-45. https://doi.org/10.14198/ELUA2017.31.02

Biber, D., & Finegan, E. (1988). Adverbial stance types in English. *Discourse Processes*, *11*(1), 1–34. https://doi.org/10.1080/01638538809544689

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Pearson.

Boginskaya, O. (2022). Creating an authorial presence in English-medium research articles abstracts by academic writers from different cultural backgrounds. *International Journal of Language Studies, 16*(2), 49-70.

Boginskaya, O. (2023). Interactional metadiscourse markers in english research article abstracts written by non-native authors: A corpus-based contrastive study. *Ikala, Revista de Lenguaje y Cultura, 28*(1), 139-154. https://doi.org/10.17533/udea.ikala.v28n1a08

Boginskaya, O. A. (2024). A comparison of explicit and implicit approaches to EAP teaching to postgraduate students. *Higher Education in Russia, 33*(2), 148-161. https://doi.org/10.31992/0869-3617-2024-33-2-148-161

Çiftçi, H., & Akbaş, E. (2021). Stancetaking in spoken ELF discourse in academic settings: interpersonal functions of I don't know as a face-maintaining strategy. *Eurasian Journal of Applied Linguistics*, *7*(1), 484–502. https://doi.org/10.32601/ejal.911499

Crismore, A., & Farnthworth, R. (1990). Metadiscourse in popular and professional science discourse. In W. Nash (Ed.), *The writing scholar: Studies in academic discourse* (pp. 118-136). Sage.

Duenas, P. M. (2010). Attitude markers in business management research articles: A cross-cultural corpus-driven approach. *International Journal of Applied Linguistics, 20*(1), 50-72. https://doi.org/10.1111/j.1473-4192.2009.00228.x

Fløttum, K. (2012). Variation of stance and voice across cultures. In K. Hyland & C. Sancho Guinda (Eds.), *Stance and voice in written academic genres* (pp. 218-231). Palgrave Macmillan. https://doi.org/10.1057/9781137030825_14

Graff, G., & Birkenstein, C. (2010). *They say/ I say*. Norton.

Gray, B., & Biber, D. (2012). Current conceptions of stance. In K. Hyland & C. S. Guinda (Eds.), *Stance and voice in written academic genres* (pp. 15–33). Palgrave Macmillan.

He, D. (2017). The use of English in the professional world in China. *World Englishes, 36*(2), 571-590. https://doi.org/10.1111/weng.12284

Hryniuk, K. (2018). Expert-like use of hedges and boosters in research articles written by Polish and English native-speaker writers. *Research in Language, 16*(3), 263-280. https://doi.org/10.2478/rela-2018-0013

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

Hu, G & Wang, G. (2014). Disciplinary and ethnolinguistic influences on citation in research articles. *Journal of English for Academic Purposes, 14*(1), 14–28. https://doi.org/10.1016/j.jeap.2013.11.001

Hyland, K. (2001). Humble servants of the discipline? Self-mention in research articles. *English for Specific Purposes, 20*(3), 207-26. https://doi.org/10.1016/S0889-4906(00)00012-0

Hyland, K. (2005a). *Metadiscourse: Exploring interaction in writing*. Continuum.

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

Hyland, K. (2013). Writing in the university: Education, knowledge and reputation. *Language Teaching, 46,* 53–70. https://doi.org/10.1017/S0261444811000036

Hyland, K. & Jiang, F. K. (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, K., & Zou, H. (2021). "I believe the findings are fascinating": Stance in three-minute these'. *Journal of English for Academic Purposes, 50,* 100973. https://doi.org/10.1016/j.jeap.2021.100973

Isık-Tas, E.E. (2017). Authorial identity in Turkish language and English language research articles in Sociology: The role of publication context in academic writers' discourse choices. *Journal of English for Specific Purposes, 49,* 26-38. https://doi.org/10.1016/j.esp.2017.10.003

Korotkina, I.B. (2018). Academic writing in Russia: The urge for interdisciplinary studies. *Higher Education in Russia*, *10*(27), 64-74. https://doi.org/10.31992/0869-3617-2018-27-10-64-74

Krapivkina, O. (2014). Pronominal choice in academic discourse. *Middle-East Journal of Scientific Research, 20*(7), 833-843. https://doi.org/10.5829/idosi.mejsr.2014.20.07.13676

Lee, Z. (2015). Etiquette behavior as the realization of Confucian values. *Journal of Leningrad State University, 5*(3), 92-100.

Lei, J., & Jiang, T. (2019). Chinese university faculty's motivation and language choice for scholarly publishing. *Ibérica, 38*, 51–74.

Lorés Sanz, R. (2006). I will argue that: First person pronouns as metadiscoursal devices in research article abstracts in English and Spanish. *ESP across Cultures, 3*, 23-40.

Mikolaychik M.V. (2019). Lexical hedging in english abstracts of Russian economics research articles: A corpus-based study. *Science Journal of Volgograd State University. Linguistics, 19*(5), 38-47. https://doi.org/10.15688/jvolsu2.2020.5.4

Prokhorov, Iu. E., & Sternin, I. A. (2006). *Russkie: Kommunikativnoe povedee* [Russians: Communicative behavior]. Flinta.

Pyankova, T. (1994). *A practical guide for the translation of Russian scientific and technical literature into English*. Letopis.

Seone, E. (2013). On the conventionalisation and loss of pragmatic function of the passive in Late Modern English scientific discourse. *Journal of Historical Pragmatics, 14*(1), 70-99. https://doi.org/10.1075/jhp.14.1.03seo

Shchemeleva, I. (2015). The development of stance-taking strategies in L2 students' academic essays: The case of a content-based Russian-American teleconference course. *Journal of Language and Education, 1*(4), 45-53. https://doi.org/10.17323/2411-7390-2015-1-4-45-53

Walková, M. (2018). Author's self-representation in research articles by Anglophone and Slovak linguists. *Discourse and Interaction, 11*(1), 86-105. https://doi.org/10.5817/DI2018-1-86

Wu, G., & Zhu, Y. (2015). Self-mention and authorial identity construction in English and Chinese research articles: A contrastive study. *Linguistics and the Human Sciences,10*(2),133-158. https://doi.org/10.1558/lhs.v10i2.28557

Yakhontova, T. (1997). The signs of a new time: Academic writing in ESP curricula of Ukrainian universities. In Duszak, A. (Ed.), *Culture and styles of academic discourse* (p. 323-341). Mouton de Gruyter.

Xia, G. (2018). A cross-disciplinary corpus-based study on English and Chinese native speakers' use of first-person pronouns in academic English writing. *Text & Talk, 38*(1), 93-113. https://doi.org/10.1515/text-2017-0032

Xiong, D. A (2007). Comparison between English and Chinese metadiscourse*. Journal of Chongqing Jiaotong University*, *7*(6), 101-105.

Xu, X., Nesi, H. (2019). Evaluation in research article introductions: A comparison of the strategies used by Chinese and British authors. *Text & Talk, 39*(6), 797-818. https://doi.org/10.1515/text-2019-2046

Zanina, E. (2016). Strategic hedging: A comparative study of methods, results and discussion (and conclusion) sections of research articles in English and Russian. *Journal of Language and Education, 2*(2), 52-60. https://doi.org/10.17323/2411-7390-2016-2-2-52-60

https://doi.org/10.17323/jle.2024.21568

Mitigating Plagiarism in ESL Academic Writing: Evaluating the Efficacy of Educational Intervention

Nazirah Mahmud [1], **Mohd Haniff Mohd Tahir** [2]

[1] University of Szeged, Hungary
[2] Universiti Pendidikan Sultan Idris, Malaysia

ABSTRACT

Background: Plagiarism is a serious academic misconduct demanding mitigation to uphold the integrity of original work. Undeniably, with the advancement of technological age, plagiarising becomes easier and harder to detect. However, the reliance on technology significantly rises the likelihood of fostering academic dishonesty among the students. Therefore, it is imperative to disseminate explicit education on plagiarism reduction strategies, particularly on the importance of mastering correct referencing techniques.

Purpose: The present study applied a mixed-method approach to explore the effectiveness of an intervention called Educational Intervention (EI) on enhancing ESL students' academic referencing skills namely in in-text citation abilities, paraphrasing skills and writing reference list using APA style.

Method: The EI consisted of explicit instruction that involved structured teaching sessions integrated with two primary instructional strategies: lectures and practical exercises on referencing in academic writing. A total of 70 participants participated in the study. Data collection involved administering tests, assigning exercises and assignments, and conducting focus group interviews. The study utilised a quasi-experimental design to analyse the effectiveness of the intervention. Quantitative data were analysed using one-way ANOVA and paired sample t-test to assess the effectiveness of the EI. Qualitative data were examined through thematic analysis to explore participants' experiences and perceptions.

Results: Data analyses of the plagiarism test indicated that the experimental group showed significantly greater improvements in referencing skills compared to the control group, with statistical significance at $p < .05$ for all assessed skills. However, findings for the gathered exercises and assignments demonstrated improvements but not statistically significant in their referencing skills except for paraphrasing. Nonetheless, feedback from focus group interviews indicated positive responses toward EI and its role in enhancing referencing skills.

Conclusion: Overall, while EI proved effective in enhancing ESL undergraduates' referencing skills, there remains room for improvement to fully realize its potential. The study holds a global importance of imparting explicit education on referencing skills, offering educators worldwide a strategic measure aimed at reducing plagiarism act.

KEYWORDS

Plagiarism, Educational Intervention (EI), English as a Second Language (ESL), Academic Writing, Higher Education.

Citation: Mahmud N., & Mohd Tahir M.H. (2024). Mitigating Plagiarism in ESL Academic Writing: Evaluating the Efficacy of Educational Intervention. *Journal of Language and Education, 10*(3), 53-68. https://doi.org/10.17323/jle.2024.21568

Correspondence:
Nazirah Mahmud
E-mail: nazirah.binti.mahmud@edu.u-szeged.hu

Received: May 2, 2024
Accepted: September 16, 2024
Published: September 30, 2024

INTRODUCTION

Plagiarism is a serious academic misconduct that is becoming increasingly prevalent in the academic landscape (Macdonald & Carroll, 2006; Puga, 2014; Shahabuddin, 2009). Research across various continents including Pakistan (Ramzan et. al, 2012), South Africa (Singh et.al, 2020) and Australia (Smedley, Crawford & Cloete, 2015; 2019) has reported a rising trend of plagiarism among undergraduates in higher education institutions. In Malaysia, several studies have similarly reported that plagiarism is an alarming issue within universities, with a growing number of cases being recorded among undergraduates (Abusafia et

al., 2018; Al-Shaibani et al., 2016; Mustapha et al., 2017; Zejno, 2018). This widespread academic misconduct highlights that plagiarism is a global issue, affecting learners worldwide who engage in it extensively within their institutions.

Vuori et al. (2004) pointed out that plagiarism is not just about breaking rules, it is influenced by various cultural and educational factors. In Malaysia, students often struggle with plagiarism as their previous school system does not fully prepare them for the demands of university, where critical thinking and analytical skills are crucial (Yang & Lin, 2009). Unlike more homogenous educational systems, Malaysian universities are characterised by their multiethnic, multicultural and multilingual student populations, where learners navigate complex linguistic and cultural landscapes. This diversity adds another layer of difficulty as different students may have varying understandings of academic integrity and plagiarism. It highlights the importance of considering each students' background when addressing plagiarism (Sowden, 2005). Strict adherence to academic integrity in Malaysian universities can also be overwhelming. This challenge is greater for students learning in English, as it is their second language. According to Riasati and Rahimi (2013), many students find it challenging to write academically in English, especially when it comes to understanding and using complex academic texts. This struggle is compounded by their low proficiency in the target language, which can lead them to plagiarise to cope with academic pressures and deadlines (Mohd Habali & Fong, 2016; Al-Zubaidi & Richards, 2010; Mousavi & Kashefian-Naeeini, 2011).

The present study was motivated by the positive outcomes reported by Smedley et al. (2015) regarding the effectiveness of an intervention known as Educational Intervention (EI) in enhancing students' referencing skills. Building on these findings, this study aims to evaluate the effectiveness of Smedley et al.'s EI with a different population. The focus is on improving students' skills in writing in-text citations, paraphrasing and compiling reference lists in APA style These skills are critical for reducing instances of plagiarism among English as a Second Language (ESL) undergraduates. The research will address the following questions:

RQ#1: How far does the use of EI improve ESL undergraduates' in-text citation abilities?

RQ#2: To what extent does the use of EI improve ESL undergraduates' para-phrasing skills?

RQ#3: To what degree does the use of EI improve ESL undergraduates' abilities to write a correct reference list using APA style?

RQ#4: What are ESL undergraduates' perceptions of the use of EI in improving their abilities to write in-text citation, paraphrase and write a reference list using APA style?

LITERATURE REVIEW

Social Constructivism Theory

Learning is a process that involves facilitators who aid learners in gaining knowledge on various topics. Lev Vygotsky's (1934) theory of social constructivism posits that the construction of knowledge occurs within the learners' social context. Vygotsky argues that the process of knowing is significantly influenced by others that are mediated through community and culture. According to Au (1998), the roles of culture and context in society such as interactions with peers, teachers, family, and the broader environment contribute to the learning process. Guidance from a more knowledgeable individual fosters social interaction, underscoring the importance of social contexts in governing the learning process. Thus, acquiring knowledge in a specific area often requires expert guidance to ensure that learners achieve a full understanding of the subject. This concept is reflected in the Zone of Proximal Development (ZPD), a key element of Vygotsky's Social Constructivism Theory. The ZPD refers to "the difference between the child's actual level of development and the level of performance achieved in collaboration with adult" (p.209). The theory suggests that learners acquire knowledge more effective through joint efforts with an expert. This process helps them with the internalization of new concepts, methods and skills. Roosevelt (2008) argues that, from a Vygotskian perspective, the primary goal of learning is to use mediating activities or tasks with the assistance of a facilitator to create culturally meaningful learning experiences. Learners are expected to collaborate with more competent individuals or peers to complete assigned tasks. This approach suggests that optimal learning occurs when learners can complete similar tasks independently after receiving initial guidance. This process helps expand their ZPD for that activity.

In short, social constructivism posits that learning occurs when learners engage in interactions with peers or experts while addressing real-life tasks (Woo & Reeves, 2007). Social constructivism serves as the theoretical framework for this study. It explores the process of acquiring scientific concepts, specifically the development of referencing skills. This theory is also pertinent to the study, given the involvement of experts in facilitating the acquisition of the subject matter.

Using Educational Intervention (EI) to Reduce Plagiarism Act

Malaysian researchers have also addressed the issue of student awareness of plagiarism in higher learning institutions (Mohamed Mohan et al., 2020; Noorashikin Hussein et al., 2016; Zejno, 2018). However, limited efforts have been made to implement and evaluate intervention strategies aimed at reducing plagiarism among students. Khazriyati (2016) examined the effectiveness of a 'course talk' on plagiarism in

reducing plagiarism among Malaysian university students, finding it successful in raising awareness about the severity of plagiarism. Nevertheless, the study did not address issues related to students' ability to reference sources correctly.

Educating undergraduates about plagiarism and its implications is crucial for ensuring that they understand the seriousness of such act. Smedley et al. (2015) conducted a quantitative study with 150 nursing students. They received EI, which included lectures and exercises on academic writing, understanding plagiarism, and proper referencing. The results indicated that participants generally showed improvement in their knowledge and understanding of plagiarism, including referencing skills, after the intervention. Furthermore, the researchers noted improvements in paraphrasing and assignment referencing in the semester following the intervention. In a subsequent study in 2019, the researchers extended their evaluation of the EI and again found positive outcomes in students' understanding of referencing skills post-treatment.

A literature review by Fatemi and Saito (2020) on unintentional plagiarism revealed deficiencies in support programs, particularly regarding the time spent on contact and educating students on the importance of proper referencing techniques. Sowell (2018) noted that academic referencing proficiency can be time-consuming and complex, requiring students to learn and adhere to specific rules, especially when dealing with formatting. As undergraduates transition from secondary school to academic writing at the university level, they may find academic referencing unfamiliar and challenging. Akakura (2012) emphasised that these rules should be explicitly taught with a consistent focus during instruction. Providing students with thorough education on these rules is essential to ensure they receive adequate exposure (Bennett et al., 2011; Culwin, 2006; Schuetze, 2004).

Academic referencing requires mastery of three key skills: writing in-text citation, paraphrasing and compiling a reference list. These skills demand adherence to specific formats and rules to maintain consistency and comply with standard guidelines. Blum (2009) argued that learning proper citation is challenging and requires a deliberate and meticulous teaching approach. Additionally, incorporating practical exercises into instruction has been identified as an effective strategy for enhancing referencing skills. According to Landau et al. (2002), engaging students in paraphrasing exercises increased their awareness of plagiarism and guided them toward producing accurate paraphrases. Stander (2020) further supported this approach, highlighting the positive impact of exercises on students' paraphrasing techniques. Moreover, studies by Moniz et al.'s (2008) indicated that integrating lectures with practical exercises effectively enhances students' understanding of plagiarism and improves their referencing skills. Thus, the combination of lectures and exercises proves to be an effective approach for teaching correct reference writing.

Teaching and exercises alone do not provide students with a complete understanding of citation rules, but incorporating feedback is essential for evaluating their performance. Fazilatfar et al. (2018) found that while students improved in following standard citation rules, they still made errors in citation writing. The lack of feedback on their citations may contribute to these errors, supporting Pecocari's (2006) argument that insufficient feedback can lead students to incorrectly assume their citations are accurate. Wingate et al. (2011) and Hortsmanhof and Brownie (2013) also emphasise the importance of feedback in helping students achieve correct academic writing.

Furthermore, the internet has consistently been identified as a factor that increases the risk of plagiarism among students due to the easy access to information and websites enabled by modern technologies (Dias & Bastos, 2014; Liliana Cuervo-Sánchez & Etxague, 2023; Ramzan et al., 2012; Sprajc et al., 2017). For example, tools like QuillBot poses a potential risk which may expose students to excessive reliance and hinder them from fully mastering their writing skills. However, Kusuma (2020) offered a different perspective. He suggested that QuilBot helped teachers and students by providing paraphrasing support that might be difficult to do manually. This assistance can reduce plagiarism in writing. Technological advancements can help writers produce well-structured work that follows the necessary rules. However, over-reliance on these tools may eventually harm their ability to create original content.

Based on the issues and literature discussed, the present study aims to underscore the importance of providing explicit instruction on proper reference writing as a strategy to reduce plagiarism in students' academic work, particularly in this era of rapid technological advancement.

METHOD

Research Design

This study utilised a quasi-experimental design to collect data. According to Price et al. (2014), quasi- experimental studies resemble true experimental research but differ in that they do not involve random assignment. Participants were divided into two groups: an experimental group and control group. The experimental group received the EI treatment, while the control group was taught using the implicit instruction method currently employed in the class. Pretests and post-test were conducted to evaluate the effectiveness of the intervention.

Participants

This study was conducted at a private university in Malaysia, employing a purposive sampling technique to select participants. Purposive sampling allows researchers to target

individuals who possess specific traits relevant to the study (Turner, 2019). Participants were selected from a pool of 100 ESL undergraduates enrolled in an English Academic Writing (EAW) course during that semester. The participants had the flexibility to choose their group based on their semester timetable, resulting in the use of intact groups. Consequently, participants were assigned to groups according to the available EAW course sections. All participants were from different academic programs, and informed consent was obtained from them prior to the experiment. A total of 70 undergraduates participated in the study, with each group-experimental and control consisted of 35 participants.

Research Procedure

The experimental group received explicit instruction that involved structured teaching sessions integrated with two primary instructional strategies: lectures and practical exercises. These sessions were carefully designed to enhance the participants' ability to accurately perform in-text citation, paraphrasing, and compile reference lists according to APA guidelines. The instructional content was prepared using two key resources: "Writing Academic English" by Alice Oshima and Ann Hogue (2006) and the instructional guide "Avoiding Plagiarism, Self-Plagiarism, and Other Questionable Writing Practices: A Guide to Ethical Writing" by Miguel Roig (2015).

The explicit instruction was divided into two distinct stages. In the first stage, participants were provided with theoretical knowledge and practical techniques related to APA style through interactive lectures. These sessions were conducted by an experienced English language instructor who encouraged active participation and discussions among students, fostering a collaborative learning environment. Students were also guided to locate relevant information from various sources, including books, journal articles, and online materials such as newspaper articles and webpages. In the second stage, participants engaged in a series of practical exercises designed to test their ability to apply the knowledge gained during the lectures. These exercises focused on writing in-text citations, paraphrasing, and creating reference lists using APA style.

In contrast, the control group received implicit instruction. This approach did not involve direct teaching or detailed explanations of the targeted writing skills. Instead, participants were exposed to in-text citation, paraphrasing, and APA referencing through reading journal articles and academic papers. The instructor for the control group provided various reading materials but did not offer explicit guidance or engage interactively with the students. Participants were expected to observe and emulate the writing techniques and referencing formats presented in the readings independently.

While the experimental group's instruction was characterized by high levels of interaction between the instructor and students, as well as among the students. In contrast, the control group had minimal interaction. The instructor's role was limited to overseeing the reading sessions.

Table 1 below shows the comparison of instructional approaches utilised in this research.

Research Instrument

Three instruments were utilized in this study: the Avoiding Plagiarism Test (APT), participants' documents in the form of exercises and assignments, and a focus group interview. The APT, which consisted of 50 questions in multiple-choice and true-or-false formats, was adapted from an online test available on the library website of Royal Roads University in 2021. Adaptations were made to ensure that the questions were contextually relevant to the participants. However, the original test is no longer accessible, as the website has transitioned to a new format featuring pop quizzes as of 2024[1]. See Appendix A for a selection of sample questions for the APT test. The test measured participants' understanding and ability to write in-text citations, paraphrase, and create reference lists in APA style, aiming to reduce plagiarism in ESL undergraduates' academic writing. Before the main study, a pilot test was conducted with 30 participants to assess the reliability of the APT. To assess the reliability of the questionnaire used in this study, a test-retest reliability method was employed. Pearson Product- Moment correlation coefficient was utilised, and the result of the test demonstrated a relia-

Table 1

Summary Comparison of Instructional Approaches

Aspect	Experimental Group (Educational Intervention)	Control Group (Implicit Instruction)
Instructional Strategy	Explicit teaching via lectures and exercises	Passive exposure through reading materials
Content Focus	In-text citation, paraphrasing, APA reference lists	In-text citation, paraphrasing, APA reference lists
Interaction	High (instructor- student, student-student)	Low (minimal instructor -student interaction)
Instructor Role	Active (lecturing, guiding)	Passive (overseeing sessions)

[1] For further details see Royal Roads University, Quoting, summarizing, and paraphrasing (2024) at https://libguides.royalroads.ca/quoting.

bility coefficient of r = .607. This confirmed that the APT was reliable and consistent over two-time points.

The second instrument involved collecting written evidence in the form of exercises and assignments focused on the referencing skills addressed in the study. These collections were conducted three times which were during Week 4, Week 8 and Week 12 over a 14-week period, specifically with the experimental group, as the control group did not engage in exercises as part of their instruction.

Qualitative data was gathered through a semi-structured focus group interview with 18 participants from the experimental group. A total of 10 questions were designed based on Adams (2015) guidelines. These guidelines emphasized the formulation of open-ended questions aligned with the research themes and the incorporation of prompts to enhance participant engagement and facilitate deeper exploration of the topics. The questions focused on capturing participants' view on the intervention's effectiveness, exploring their experiences and challenges during the process and collecting suggestions for improvement. See Appendix B for a complete list of the interview questions. The interviews were recorded using an audio recorder and later transcribed for analysis.

Analysis

The participants' test scores from both the pre-test and post-test of the APT, along with their exercise and assignments scores were analysed using one-way ANOVA and paired sample t-test. These two statistical tests were used to evaluate intervention effectiveness by analysing differences within and between groups. Meanwhile, the transcript from the focus group interview was analysed using thematic analysis to generate the qualitative findings of the study. It began with thoroughly reading the interview transcripts to become familiar with the data and note initial impressions. Key data was then organized and coded. The codes were later grouped into themes, which were reviewed to ensure coherence and distinctiveness. Finally, themes were defined to capture their core meanings and findings were reported to support study triangulation.

RESULTS

The results are presented in alignment with the research questions guiding this study. This section provides an analysis of the inferential statistics obtained from the APT tests and excerpts from the interview transcript to support each of the formulated research questions.

The Effectiveness of EI on In-text Citation Abilities

The first research question investigates the effectiveness of EI on participants' ability to write correct in-text citation. Table 2 presents the results of a one-way ANOVA for both pre-test and post-test of the experimental and control groups. The one-way ANOVA results for the pre-test showed no significant difference between the experimental and control groups. The scores were not significant at $p > .05$ ($F(1, 68) = .148$, $p = .702$). This finding suggests that participants in both groups had similar levels of knowledge about writing in-text citations before the intervention. Meanwhile, the post-test results showed that the mean score for the experimental group was higher than that of the control group, with a mean difference of MD = 4.000. The table also indicates a significant effect of the EI on participants' APT score for in-text citation questions, with $p < .05$ ($F(1, 68) = 10.593$, $p = .002$). This result demonstrates that the treatment administered to the experimental group was more effective, as reflected in the higher and statistically significant APT mean score for the experimental group.

Table 3 illustrates the results of the paired sample t-test of for the APT scores of participants in the experimental group. The p-value for this pair was $p < .05$ ($t(34) = -3.845$, $p = .001$), indicating that the scores were statistically significant. Cohen's d was estimated at d = 1.221, which is considered a large effect according to Cohen's (1992) guidelines. This finding suggests that the treatment was effective in improving participants' ability to write in-text citation as evidenced by the significant improvement in their APT score post treatment.

Table 4 depicts the results of the repeated measures ANOVA for the exercises and assignments completed by partic-

Table 2

Result of One-Way ANOVA for Pre-Test and Post-Test (In-Text Citation)

APT	Group	N	M	SD		df1	df2	F	Sig.
Pre-test	Experimental Group	35	14.686	4.086	0.400	1	68	.148	.702
	Control Group	35	15.086	4.861					
Post-test	Experimental Group	35	19.086	5.453	4.000	1	68	10.593	.002
	Control Group	35	15.086	4.810					

ipants in the experimental group. The results indicate that the mean scores of participants' exercises improved over time, increasing from M = 6.600 in EA1 to M = 7.086 in EA2, and finally to M = 7.514 in EA3. This trend suggests an improvement in participants' ability to write in-text citations towards the end of the intervention. However, Cohen's d was estimated at d = 0.230, which is deemed a small effect. The mean scores for participants' exercises and assignments showed minimal improvement. The increase was statistically insignificant with a p-value of 0.326, which is greater than 0.05. Therefore, it can be concluded that the intervention had a statistically insignificant effect on the ESL undergraduates' ability to write in-text citation in their exercises and assignments (F (2, 68) = 1.139, p = .326).

The Effectiveness of EI on Paraphrasing Skills

The second research question examined the effectiveness of EI in improving participants' ability to write correct paraphrase. A one-way analysis of variance (ANOVA) was conducted on the pre-test and post-test scores for paraphrasing questions in the APT for both participants in the experimental and control groups. The results are presented in Table 5. The pre-test results indicated insignificant difference at p > .05 (F (1, 68) = 1.603, p =.210), suggesting that the pre-test scores between experimental and control groups were not significantly different. This finding implies that participants' paraphrasing skills were at a similar level in both groups before the study began. Further analysis of the post-test scores revealed that the experimental group (M=19.086) outperformed the control group (M = 13.086) with a mean difference of = 6.000, indicating that participants in the experimental group performed better on the APT compared to those in the control group. The difference in scores between the groups was also statistically significant at p < .05 (F (1, 68) = 21.975, p =.000). This data demonstrates that EI was more effective in improving participants' ability to write correct paraphrase in the APT than the implicit instruction used in the control group.

The APT score for the pre-test and post-test in the experimental group were further analysed using a paired sample t-test, as shown in Table 6. The analysis revealed that the improvement in scores was statistically significant at p <.05 (t (34) = -4.449, p =.000) with an estimation of Cohen's d at d = 1.215, indicating a large effect. Therefore, this result confirms that the implementation of EI was effective in enhancing participants' paraphrasing abilities, as evidenced by significant improvement in the participants' APT score post intervention.

Table 7 presents the analysis of the repeated measures ANOVA for the exercises and assignments on paraphrasing skills within the experimental group. The results showed a steady increase in the average mean score, from M =3.514 for EA1 to M = 5.486 for EA2, and finally to M = 6.300 for

Table 3

Result of Paired-Sample T-test for the Pre-Test and Post-Test of the Experimental Group (In-Text Citation)

	N		SD	SEM	t	df	Sig.	d.
Pre-test – Post-test	35	-4.400	6.770	1.144	-3.845	34	.001	1.221

Table 4

Result of Repeated Measures ANOVA for Exercises and Assignments of the Experimental Group (In- Text citation)

	N	M	SD	df1	df2	F	Sig.	d.
Exercise/ Assignment 1 (EA1)	35	6.600	2.603	2	68	1.139	.326	0.230
Exercise/ Assignment 2 (EA2)	35	7.086	2.454					
Exercise/ Assignment 3 (EA3)	35	7.514	2.884					

Table 5

Result of One-Way ANOVA for Pre-Test and Post-Test (Paraphrase)

APT	Group	N	M	SD		df1	df2	F	Sig.
Pre-test	Experimental Group	35	14.229	5.504	1.600	1	68	1.603	.210
	Control Group	35	12.629	5.059					
Post-test	Experimental Group	35	19.086	5.431	6.000	1	68	21.975	.000
	Control Group	35	15.086	4.810					

EA3. This trend indicates a gradual improvement in participants' ability to write paraphrases in their exercises and assignments throughout the intervention period. Cohen's d was also estimated at d = 1.485, signifying a large effect. The analysis further reported that the substantial improvement in scores for the exercises and assignments was statistically significant at p= .000. Therefore, these findings indicate that EI was effective in enhancing participants' ability to write correct paraphrases in their exercises and assignments (F (2, 68) = 48.561, p = .000).

The Effectiveness of EI on Writing Reference List using APA style

The third research question aimed to evaluate the effectiveness of Smedley et al.'s (2015) EI in improving ESL undergraduates' abilities to write a correct reference list using APA style. A one-way ANOVA was conducted on the pre-test and post-test scores of the APT for both the experimental and control groups, with the results presented in Table 8. The pre-test results showed a p-value of p > .05 (F (1, 68) = 3.786), p =.056), indicating no significant difference between the groups. This suggests that the participants in both groups had a similar level of knowledge regarding APA-style reference lists before the intervention. Subsequently, another one-way (ANOVA) was performed to assess the effectiveness of the intervention by comparing the post-test APT scores between the experimental and control groups. The data revealed that the mean score for the experimental group (M =19.943) was higher than that of control group (M =17.200), with a significant difference between the groups at p < .05 (F (1, 68) = 8.011, p =.006). This finding indicates that participants in the experimental group performed better in writing a reference list using APA style than those in the control group.

The post-test data of the experimental group were further analysed using a paired sample t-test to determine the effectiveness of the intervention specifically on writing a correct reference list. Table 9 presents the analysis, which shows a significant difference in scores at p <.05 (t (34) = -7.737, p =.000). Cohen's d was estimated at d = 1.577, indicating a large effect. Thus, the APT data suggest that the EI was ef-

Table 6

Result of Paired-Sample T-test for the Pre-Test and Post-Test of the Experimental Group (Paraphrase)

	N		SD	SEM	t	df	Sig.	d.
Pre-test – Post-test	35	-4.857	6.459	1.092	-4.449	34	.000	1.215

Table 7

Result of Repeated Measures ANOVA for Exercises and Assignments of the Experimental Group (Paraphrase)

	N	M	SD	df1	df2	F	Sig.	d.
Exercise/ Assignment 1 (EA1)	35	3.514	1.915	2	68	48.561	.000	1.485
Exercise/ Assignment 2 (EA2)	35	5.486	1.502					
Exercise/ Assignment 3 (EA3)	35	6.300	1.030					

Table 8

Result of One-Way ANOVA for Pre-Test and Post-Test (Reference List Using APA Style)

APT	Group	N	M	SD		df1	df2	F	Sig.
Pre-test	Experimental Group	35	14.743	3.665	1.657	1	68	3.786	.056
	Control Group	35	16.400	3.457					
Post-test	Experimental Group	35	19.943	3.741	2.743	1	68	8.011	.006
	Control Group	35	17.200	4.344					

Table 9

Result of Paired-Sample T-test for the Pre-Test and Post-Test of the Experimental Group (Reference List Using APA Style)

	N		SD	SEM	t	df	Sig.	d.
Pre-test – Post-test	35	-5.200	3.976	.672	-.737	34	.000	1.577

fective in improving participants' ability to write a reference list using APA style after the intervention.

Additionally, a repeated measures ANOVA was conducted on the exercises and assignments submitted by the experimental group to explore whether a significant change occurred in participants' ability to write a correct reference list using APA style. The results are displayed in Table 10. The mean scores of the participants increased from M = 5.857 in EA 1 to M = 5.971 in EA 2, and finally to M = 6.586 in EA 3, indicating a gradual improvement in their ability to write correct reference lists over the course of the intervention. However, the effect size, measured by Cohen's d was estimated at d = 0.197. This indicated a small effect size. Despite the observed improvement, the change was not statistically significant F (2, 68) = .729, p = .486. Therefore, it cannot be proven that the use of EI significantly improved ESL undergraduates' ability to write a reference list using APA style in their exercises and assignments.

ESL Undergraduates' Perceptions on the Effectiveness of EI on their referencing skills

The fourth research question was to identify the ESL undergraduates' perception on the use of EI in improving their referencing skills. Four key themes emerged from the transcriptions. These themes include the positive aspects of EI, the negative aspects of EI, the impact of EI on students' abilities to write in-text citations, paraphrasing and compiling reference lists according to APA style, and their overall opinions on the benefits of EI as an initiative to prevent plagiarism in academic writing.

Positive Aspects of EI

A common theme among the responses was that the intervention provided crucial initial exposure to the concept of plagiarism and strategies to avoid it. Participants indicated that this was their first experience learning about plagiarism prevention, with the EI intervention serving as an essential introduction to the topic (T1 and T2).

T1: "It is my first time learning on how to avoid plagiarism. So, the lectures and exercises help me a lot" (P1)

T2: "This is my first time knowing all the elements regarding plagiarism" (P6)

Participants also highlighted the practical benefits of the lectures and exercises. Some indicated that the resources particularly useful when completing tasks that required accurate referencing (T3, and T4). Meanwhile, the knowledge gained from EI was also noted to have broader applicability with participants mentioning that the skills learned could utilised in other subjects as well as future writing projects (T5 and T6).

T3: "For me, the EI helps me to provide exposure to the formats on how to write reference correctly" (P5)

T4: "I always referred to the lecture notes because it is simpler and easier to understand than using reference books which tend to be lengthy" (P8)

T5: "The information that I gained from EI can also be applied for other subjects" (P6)

T6: "Recently, I am able to apply the knowledge that I gained from EI on my finance report assignment" (P15)

Negative Aspects of EI

Nonetheless, EI has its own loopholes. Participants expressed a need for more discussion sessions to clarify their understanding and address mistakes (T7). Some also added that they need more examples or sample answers to enhance their comprehensions (T8).

T7: "I need more discussion sessions to avoid confusion on the mistakes that I made" (P4)

T8: "I think more examples and sample answers will help me to reach full understanding of the topic" (P10)

Other concerns included the lack of variety in the exercises and the need for more engaging visual aids. Some participants suggested incorporating a wider range of exercises (T9) and visual tools like mind maps to aid in learning (T10)

Table 10

Result of Repeated Measures ANOVA for Exercises and Assignments of the Experimental Group (Reference List Using APA Style)

	N	M	SD	df1	df2	F	Sig.	d.
Exercise/ Assignment 1 (EA1)	35	5.857	4.110	2	68	.729	.486	1.485
Exercise/ Assignment 2 (EA2)	35	5.971	2.203					
Exercise/ Assignment 3 (EA3)	35	6.586	2.328					

T9: "Variety of exercises maybe can be integrated so that it can provide more guidance on the types of questions that will be asked" (P12)

T10: "I think the EI can be improved by adding on some visual notes like the mind map. It will be easier for me to remember the format" (P18)

Impact of EI on Students' Abilities to Write In-Text Citations, Paraphrasing and Compiling Reference Lists according to APA Style

All participants agreed that the EI significantly enhanced their ability to accurately write in-text citation, paraphrasing and compiling reference lists in APA style.

Overall Opinions on the Benefits of EI as an Initiative to Prevent Plagiarism in Academic Writing

The interview concluded with participants sharing their overall assessment of the EI's importance in helping them avoid plagiarism in their academic writing. Participants unanimously agreed that the intervention was valuable as it provided them with essential knowledge to effectively prevent plagiarism in their work.

DISCUSSION

Based on the result, it was reported that participants who underwent EI demonstrated a statistically significant increase in mean scores on their post-test for in-text citation skills. The mean score for the experimental group was higher than the control group. This indicates that EI was more effective in improving the ability to write in-text citations compared to the implicit teaching methods used in the control group. Focused teaching by the instructor on how to correctly write in-text citations provided learners with essential knowledge and skills. This aligns with Teeter's (2015) assertion that the learning process involves active roles from both instructors and learners. Since writing in-text citations requires adherence to specific formats and rules, expert guidance is crucial for demonstrating the correct techniques. Interviews showed that this was the participants' first exposure to plagiarism. This highlights the need for structured education to effectively teach essential skills, such as writing in-text citations.

The findings of this study demonstrated that sufficient exposure to the skill through a series of lectures and exercises significantly improved the learners' ability to write in-text citations. This aligns with Blum's (2009) argument that mastering proper citation is challenging and requires "slow, careful teaching" (p.13). The guidance provided by knowledgeable instructors underscores the importance of social contexts in achieving the learning objectives, specifically the ability to write accurate in-text citations. This suggests that explicit teaching is more effective than implicit methods, as learners need specific guidance on adhering to academic referencing formats. However, the study also revealed contrasting result when analysing the exercises and assignments. Although the scores for in-text citation improved across the collections, the change was not statistically significant. This lack of significance was attributed to small differences in mean scores and persistent mistakes in applying standard citation rules. Participants often confused in-text citations within the text and outside the text, applying the same format incorrectly across their submissions. The most common error was the misplacement of parentheses in citations. Fazilatfar et al. (2018) identified this challenge as well, noting that students continued to make citation errors despite showing improvement. The lack of feedback was highlighted as a key factor contributing to these persistent errors. Pecocari (2006) argued that insufficient feedback could lead students to mistakenly believe their citations are correct. This study's focus on lectures and exercises without incorporating feedback sessions likely contributed to the ongoing uncertainty among participants. Without feedback, participants only attended lectures and completed exercises. This lack of feedback may have prevented them from fully internalizing correct citation practices. Engaging instructors to review and provide feedback on learners' work could help students recognize and correct their citation errors. This approach may reduce the likelihood of them repeating the same mistakes. This aligns with the arguments of Wingate et al. (2011) and Hortsmanhof and Brownie (2013), who emphasize the necessity of feedback in guiding students toward accurate academic writing.

The results indicated that the use of EI had significantly improved participants' paraphrasing skills as evidenced by the statistically significant increase in their post-test mean scores. The experimental group performed better than the control group. EI proved to be more effective in improving the ability to write correct paraphrases compared to the implicit teaching method used with the control group. Having a knowledgeable instructor demonstrate proper paraphrasing techniques helped students to understand and master the topic. This enabled them to paraphrase correctly. This scenario aligns with Vygotsky's concept of the Zone of Proximal Development (ZPD) (1978), where the instructor acts as a facilitator, guiding students as they acquire new knowledge. Introducing and educating students on essential topics, such as proper paraphrasing techniques, should be a priority. Interviews conducted during the study revealed that participants were initially unaware of these referencing skills. Fatemi and Saito (2020) highlighted that insufficient explicit instruction can lead to unintentional plagiarism due to a lack of awareness among learners. Therefore, the findings of this study underscore the importance of explicit teaching, as it was shown to be effective in improving participants' paraphrasing skills following the intervention. Sowell (2018) further emphasised that mastering proper academic

referencing can be challenging and time-consuming, requiring a collaborative effort between teachers and students for successful learning outcomes.

The results from the paraphrasing exercises and assignments indicate a statistically significant improvement in participants' performance, as shown by the increasing trend in mean scores over time. Participants demonstrated a better ability to write their own paraphrases without heavily relying on the original texts. However, inaccuracies in their work were mainly due to low English proficiency. This issue was beyond the scope of the study. Notably, most submitted paraphrases reflected the participants' efforts to accurately retain the original meaning without using the exact wording, and they also included proper source acknowledgement. Given that effective paraphrasing requires students to use their own words, incorporating only lectures proved insufficient. Consequently, additional exercises were implemented to help participants fully grasp the skill. This approach significantly improved their paraphrasing abilities. The outcome aligns with previous research by Landau et al. (2002). Their study found that engaging in paraphrasing exercises increased students' awareness of plagiarism and guided them in creating appropriate paraphrases. Similarly, a more recent study by Stander (2020) supported this, reporting that paraphrasing exercises enhanced students' abilities to accurately paraphrase. However, concerns were raised regarding the potential impact of internet accessibility on the study's findings. The availability of online paraphrasing tools such as QuillBot and Paraphrasing-tool may have influenced participants to produce plagiarism-free paraphrases through simple copy-pasting, rather than by developing their own writing skills. Kusuma (2020) noted that while QuillBot can assist teachers and students in generating paraphrases they might struggle to write manually, overreliance on such technology can lead to unoriginal work and hinder skill development. Since this study was conducted online, it was challenging to completely prevent access to these tools. To obtain more valid findings on the effectiveness of the intervention, future studies should be conducted in a controlled physical environment. This would minimise external variables like internet access.

This study also explored the skill of constructing a reference list using APA style. The findings revealed that participants in the experimental group achieved higher mean scores on questions related to APA reference list formatting in their post-test compared to their pre-test. Moreover, the experimental group outperformed the control group. This suggested that the EI method was more effective than the implicit teaching method used in the control group in improving the participants' ability to write a reference list using APA style. The success of the EI method can be attributed to its structured approach. Explicit instruction on APA format was delivered through a series of lectures and exercises. Akakura (2012) emphasises that teaching rules explicitly is essential and should be prioritised in instructional design. In this study, lectures that detailed APA formatting rules and guidelines were conducted to ensure that participants first gained a thorough understanding of the necessary conventions. This foundation is crucial as educating students on the rules and formats of APA style helps them acquire the functional knowledge required to apply these rules accurately in their writing (Bennett et al., 2011; Culwin, 2006; Schuetze, 2004). APA style varies depending on the type of source being cited whether it is a journal article, website, book, or other materials. The intervention's exercises allowed participants to practice different techniques, strengthening their ability to accurately construct reference lists. This improvement in understanding and applying APA style is consistent with the findings of Moniz et al. (2008). Their research showed that lectures and exercises were effective in enhancing students' knowledge of plagiarism and developing their referencing skills.

A different outcome was observed in the analysis of the exercises and assignments related to writing a reference list using APA style. Although the mean scores for these tasks showed a gradual increase over time, the improvement was not statistically significant. This lack of significance could be due to the minimal differences in mean scores from the first to the last collection, leading to an overall insignificant result. The primary issue seemed to be the participants' difficulty in adhering to the correct APA conventions, as inaccuracies in the presentation of the reference lists were still apparent in their submission. Similar findings were reported by Smedley et. al (2015) who noted that while students showed some improvement in writing references for their assignments, persistent issues with APA formatting remained evident. However, this study's intervention may have been insufficient in this regard, as participants expressed the need for more examples and discussion sessions in the future. Additionally, the lack of visual aids, such as mind maps, was identified as a contributing factor to the results. Participants suggested that incorporating more visual tools could help them better understand and memorise APA format guidelines.

It is also crucial to recognize that the use of technology, particularly the internet, may have undermined the participants' efforts to produce original work. The internet has long been associated with an increased risk of plagiarism, as the accessibility of information and websites has been greatly facilitated by new technologies (Dias & Bastos, 2014; Liliana Cuervo-Sánchez & Etxague, 2023; Ramzan et al., 2012; Sprajc et al., 2017) The availability of websites like "Citation Machine" and "Cite this for Me" which offer free citation generation, may have contributed to the minimal differences in participants' scores, leading to an insignificant improvement in their citation skills. While it might be assumed that these websites would assist participants in correctly formatting citations and reference lists according to APA style, the findings suggest otherwise. Participants continued to make errors in their references for exercises and assignments, despite having access to these online tools during the study,

which was conducted using online platforms. This unrestricted access likely limited the researcher's ability to control the use of the internet during the intervention. One possible explanation for these findings is the participants' lack of understanding regarding the sources they needed to cite. The requirement to cite various types of materials such as books, journal articles, magazines and web-based contents which each with different formatting rules, may have led to confusion. This confusion could have resulted in incorrect information being input into the citation websites that lead to errors in the generated reference lists.

The interviews with the participants revealed that the EI significantly improved their referencing skills. They described the intervention as beneficial, noting that it enhanced their ability to accurately reference sources and provided valuable guidance for completing their tasks. Despite the benefits of the EI intervention, participants identified areas for improvement, particularly the need for more discussion sessions and feedback. They emphasised that additional examples and guidance would enhance their understanding of the topics. This aligns with Fazilatfar et al. (2018), who found that insufficient feedback contributed to incorrect citations in student work, underscoring the necessity of follow-up guidance.

CONCLUSION

This study concludes that EI is effective in enhancing ESL undergraduates' understanding of plagiarism, particularly in developing their referencing skills, such as in-text citation, paraphrasing, and writing reference lists using APA style. However, the credibility of the intervention could be improved to achieve optimal effectiveness, given the mixed findings reported. Enhancements, such as the integration of feedback sessions as suggested by the participants could lead to more significant improvements in the undergraduates' referencing skills. This study underscores the global importance of providing explicit education on referencing skills, offering educators a strategic approach to reducing plagiarism, particularly in ESL contexts.

It is also crucial to consider that ESL learners might have encountered challenges during the EI process due to their limited English proficiency. This limitation could have made it difficult for them to fully master the required skills. To fully understand the English text was one thing, but to critically use it as a source in their writing adds another level of difficulty. This struggle was evident in the participants' paraphrasing tasks. Many had trouble constructing grammatically correct paraphrases, likely due to limited grammar and vocabulary. However, this issue was not thoroughly discussed in the study, as its scope was limited to addressing referencing skills.

Future research could incorporate improved measures that address both the lack of knowledge and skills related to plagiarism, and the limited English proficiency of ESL learners. This dual focus could help to reduce plagiarism more effectively. Besides, the present study was also limited to a period of 14 weeks which was relatively enough to bring out findings. However, a longitudinal study would be better in establishing the effectiveness of the intervention. Future studies could opt for a longitudinal approach to see if extended instruction leads to greater improvement in undergraduates' referencing skills.

DECLARATION OF COMPETITING INTEREST

None declared

AUTHOR CONTRIBUTIONS:

Nazirah Mahmud: conceptualisation; formal analysis; investigation; methodology; project administration; resources, validation; visualization; writing - original draft; writing – review & editing.

Mohd Haniff Mohd Tahir: conceptualisation; formal analysis; supervision; validation; visualization; writing - reviewing & editing.

REFERENCES

Abusafia, A. H., Nurhanis S. R., Dariah M. Y., & Mohd Z. M. N.. (2018). Snapshot of academic dishonesty among Malaysian nursing students: A single university experience. *Journal of Taibah University Medical Sciences, 13*(4), 370-376. https://doi.org/10.1016/j.jtumed.2018.04.003

Akakura, M. (2012). Evaluating effectiveness of explicit instruction on implicit and explicit L2 knowledge. *Language Teaching Research, 16*(1), 9-37. https://doi.org/10.1177/1362168811423339

Al-Shaibani, G.K.S., Fauzilah Md Husain, & Mahfoodh, O. (2016). A qualitative investigation into the understanding of plagiarism in a Malaysian Research University. *Journal of Applied Linguistic and Language Research, 3*(7), 337-352.

Al-Zubaidi, K., & Richards, C. (2010). Arab postgraduate students in Malaysia: Identifying and overcoming the cultural and language barriers. *Arab World English Journal, 1*(1), 107–129. https://dx.doi.org/10.24093/awej/vol1no1.5

Au, K.H. (1998). Social Constructivism and the school literacy learning of students of diverse backgrounds. *Journal of Literacy Research, 30*(2), 297-319. https://doi.org/10.1080/10862969809548000

Bennett, K. K., Behrendt, L. S., & Boothby, J. L. (2011). Instructor perceptions of plagiarism: Are we finding common ground? *Teaching of Psychology*, *38*, 29-35. https://doi.org/10.1177/0098628310390851

Blum, S. (2009). *My word!: Plagiarism and college culture*. Cornell University Press.

Cuervo-Sánchez, S. L., & Etxague, I. (2023). *The four P's on the Internet: Pornography, plagiarism, piracy and permission* [Las cuatro P en Internet: Pornografía, plagio, piratería y permisos]. *Comunicar, 76*, 85-96. https://doi.org/10.3916/C76-2023-07

Culwin, F. (2006). An active introduction to academic misconduct and the measured demographics of misconduct. *Assessment & Evaluation in Higher Education, 31*(2), 167-182. https://doi.org/10.1080/02602930500262478

DeGeeter M., Harris K., Kehr, H. Ford, C., Lane, D. C., Nuzum, D.S., & Compton, C. (2014). Pharmacy students' ability to identify plagiarism after an educational intervention. *American Journal of Pharmaceutical Education, 78*(2), 1-6. https://doi.org/10.5688/ajpe78233

Dias, P. C., & Bastos, A. S. (2014). Plagiarism phenomenon in European countries: Results from GENUIS project. *Procedia- Social and Behavioral Sciences, 116*, 2526-2531. https://doi.org/10.1016/j.sbspro.2014.01.605

Fatemi, G., & Saito, E. (2020). Unintentional plagiarism and academic integrity: The challenges and needs of postgraduate international students in Australia. *Journal of Further and Higher Education*, *44*(10), 1305-1319. https://doi.org/10.1080/0309877X.2019.1683521

Fazilatfar A.M., Elhambakhsh, S.E., & Allami H. (2018). An investigation of the effects of citation instruction to avoid plagiarism in EFL academic writing assignments. *Sage Open, 8*(2), 1-13. https://doi.org/10.1177/2158244018769958

Horstmanshof, L., & Brownie, S. (2013). A scaffolded approach to discussion board use for formative assessment of academic writing skills. *Assessment & Evaluation in Higher Education*, *38*(1), 61-73. https://doi.org/10.1080/02602938.2011.604121

Khazriyati S. (2016). Creating Awareness of Plagiarism among Postgraduates in a Postgraduate course through a talk. *Jurnal Pendidikan Malaysia, 41*(1), 47-51.

Kusuma, I. P. I. (2020). *Mengajar Bahasa Inggris dengan teknologi: Teori dasar dan ide pengajaran* [Teaching English with technology: Basic theories and teaching ideas]. Deepublish.

Landau, J. D., Druen, P., & Arcuri, J. A. (2002) Methods for helping students avoid plagiarism. *Teaching of Pscychology, 29*(2), 112-115. https://doi.org/10.1207/S15328023TOP2902_06

Macdonald, R., & Carroll, J. (2006). Plagiarism – A complex issue requiring a holistic institutional approach. *Assessment and Evaluation in Higher Education, 31*(2), 233–245. http://dx.doi.org/10.1080/02602930500262536

Moniz, R., Fine. J. & Bliss, L. (2008). The effectiveness of direct- instruction and student-centered teaching methods on students' functional understanding of plagiarism. *College & Undergraduate Libraries, 15*(3), 255-279. https://doi.org/10.1080/10691310802258174

Mousavi, H. S., & Kashefian-Naeeini, S. (2011). Academic writing problems of Iranian post- graduate students at National University of Malaysia (UKM). *European Journal of Social Sciences, 23*(4), 593-603.

Mustapha, R., Hussin, Z., Siraj, S. & Darusalam, G. (2017). Academic dishonesty among higher education students: The Malaysian evidence (2014 – 2016). *KATHA - The Official Journal of the Centre for Civilisational Dialogue, 13*(1), 73–93. https://doi.org/10.22452/KATHA.vol13no1.4

Mohamed Mohan, N. M. M., Muslim, N. & Jahari, N. A. (2020). Plagiarism: Exploring students' awareness level in higher education setting. *Selangor Business Review, 5*(2), 43- 55.

Mohd Habali, A.H., Fong, L.L. (2016). Plagiarism in Academic Writing Among TESL Postgraduate Students: A Case Study. In: Fook, C., Sidhu, G., Narasuman, S., Fong, L., Abdul Rahman, S. (Eds.), *7th International Conference on University Learning and Teaching (InCULT 2014) Proceedings.* (pp. 729-740). Springer. https://doi.org/10.1007/978-981-287-664-5_57

Parfitt, M. (2012). *Writing in response*. Bedford.

Pecorari, D. (2006). Visible and occluded citation features in post-graduate second-language learning. *English for Specific Purposes, 25,* 4-29. https://doi.org/10.1016/j.esp.2005.04.004

Power, L.G. (2009). University student's perceptions of plagiarism. *The Journal of Higher Education, 80,* 643-662. http://dx.doi.org/10.1353/jhe.0.0073

Price, P., Jhangiani, R., & A.Chiang, I. (2014). *Research methods in Psychology* (2nd ed.). BC Open Textbook Project.

Puga, J. L. (2014). Analyzing and reducing plagiarism at university. *European Journal of Education and Psychology, 7*(2), 131-140. https://doi.org/10.1989/ejep.v7i2.1

Ramzan, M., Munir, M.A., Siddique, N., & Asif, M. (2012). Awareness about plagiarism amongst university students in Pakistan. *Higher Education, 64*(1), 73-84. https://doi.org/10.1007/s10734-011-9481-4

Riasati, M. J., & Rahimi, F. (2013). Why do Iranian postgraduate students plagiarize? A qualitative investigation. *Middle East Journal of Scientific Research, 14*(3), 309–317. https://doi.org/10.5829/idosi.mejsr.2013.14.3.522

Roosevelt, F.D. (2008). Zone of proximal development. *Encyclopedia of Educational Psychology.* SAGE publication. https://doi.org/10.4135/9781412963848

Schuetze, P. (2004). Evaluation of a brief homework assignment designed to reduce citation problems. *Teaching of Psychology*, *31*(4), 257-259. https://doi.org/10.1207/s15328023top3104_6

Shahabuddin, S. (2009). Plagiarism in academia. *International Journal of Teaching and Learning in Higher Education, 21*(3), 354-359.

Singh, D., Steenkamp, M., Harmse, T., & Botha, J.C. (2020). Engaging the students voices to improve referencing skills and practices in higher education: A South African case study. *South African Journal of Higher Education, 34*(5), 122-135. https://doi.org/10.20853/34-5-4199.

Smedley, A., Crawford, T. & Cloete, L. (2019). An evaluation of an extended intervention to reduce plagiarism in Bachelor of nursing students. *Nursing Education Perspectives, 41*(2), 106-108. https://doi.org/10.1097/01.NEP.0000000000000492

Smedley, A., Crawford, T., & Cloete, L. (2015). An intervention aimed at reducing plagiarism in undergraduate nursing students. *Nurse Education in Practice*, *15*, 168-173. https://doi.org/10.1016/j.nepr.2014.12.003

Sowell, J. (2018), Beyond the plagiarism checker: Helping Non-Native English Speakers (NNESs) avoid Plagiarism. *English Teaching Forum, 56*(2), 2-15.

Sprajc, P., Urh, M., Jerebic, J., Trivan, D. & Jereb, E. (2017). Reasons for plagiarism in higher education. *Organizacija, 50*(1), 33-45. https://doi.org/10.1515/orga-2017-0002

Stander, M. (2020). Strategies to help university students avoid plagiarism: A focus on translation as an intervention strategy. *Journal of Further and Higher Education*, *44*(2), 156-169. https://doi.org/10.1080/0309877X.2018.1526260

Teeter, J. (2015). Deconstructing attitudes towards plagiarism of japanese undergraduate in EFL academic writing classes. *English Language Teaching, 8*(1), 95-109. http://dx.doi.org/10.5539/elt.v8n1p95

Turner, D. P. (2019). Sampling methods in research design. *The Journal of Head and Face Pain*, *60*(1), 8-12. https://doi.org/10.1111/head.13707

Vygotsky, L. S. (1934). Thinking and speech. In R. W. Rieber & A. S. Carton (Eds.)., *The collected works of L. S. Vygotsky (*vol. 1: Problems of General Psychology, pp. 39-288). Plenum Press.

Wingate, U., Andon, N., & Cogo, A. (2011). Embedding academic writing instruction into subject teaching: A case study. *Active Learning in Higher Education*, *12*(1), 69-81. https://doi.org/10.1177/1469787410387814

Woo, Y., & Reeves, T. C. (2007). Meaningful interaction in web-based learning: A social constructivist interpretation. *The Internet and Higher Education, 10*(1), 15–25. https://doi.org/10.1016/j.iheduc.2006.10.005

Zejno, B. (2018). Plagiarism in academic writing among students of higher learning institutions in Malaysia: An Islamic perspective. *Journal of Education and Social Sciences, 9*(3), 1-14.

APPENDIX A

Avoiding Plagiarism Test (APT)

The test utilized both true/false and multiple-choice formats and comprised four sections:

1. Section A: Background knowledge of plagiarism
2. Section B: In-text citations
3. Section C: Paraphrasing
4. Section D: Reference List

Below are selective questions for each section. Full access to the complete set of questions will be available upon request.

Section A: Background knowledge of plagiarism

1. Plagiarism can be defined as
   A. Paraphrasing another's idea with explicit attribution to the author
   B. Acknowledging the sources where your ideas build upon by using their exact words.
   C. Representing another person's work as your own.
   D. None of the above.

2. What **BEST** describes the rightness or wrongness of plagiarism?
   A. It is always wrong because it is considered as an act of theft and fraud.
   B. It is not a right or wrong kind of act.
   C. In certain situations, plagiarism is acceptable.
   D. It is not wrong to commit plagiarism.

Section B: In-text citations

3. In- text citation is **NOT COMPULSORY** in the writer's body of text even though the writer uses outside resources to support his/her arguments in the paper (True/ False)

4. When do you use "et al." in your work?
   A. When the work has more than two authors and being mentioned for the second time in your body of text.
   B. When the work has more than two authors and is being mentioned for the first time in your body of text.
   C. When you want to decrease the number of words of your body of text
   D. All of the above.

Section C: Paraphrasing

5. A paraphrase has quotation marks around it. (True/ False)

6. Which of the following should you **NOT** do when paraphrasing?
   A. Restate your information and ideas accurately.
   B. Use your own language and style.
   C. Reference the source.
   D. Change just one or two words in a sentence.

Section D: Reference List

7. Reference list is important as it allows readers to have access to the same sources for themselves. (True/ False)

Royal Roads University. (n.d.). Create a preliminary document plan. http://library.royalroads.ca/writing-centre/writing-essay-start-here/create-preliminary-document-plan

8. Determine what is the type of resource for the reference list provided above.
   A. Wiki post
   B. Blog post
   C. Technical or research report
   D. Webpage

APPENDIX B

Focus Group Interview Questions

1. Tell me about your experience as a participant in the Educational Intervention (EI).
2. What are the positive aspects that you like about the intervention?
3. What are the negative aspects that you do not like about the intervention?
4. Did the EI help you to improve your ability to write in-text citation correctly? If yes, why. If no, why?
5. Did the EI help you to improve your ability to write paraphrase accurately? If yes, why. If no, why?
6. Did the EI help you to improve your ability to write a reference list using APA style correctly? If yes, why. If no, why?
7. In your judgement, do you think that the EI benefits you?
8. Why do you think that way? (prompt)
9. Do you think that the EI is a good initiative to help you to avoid plagiarism act?
10. Why do you think that way? (prompt)

https://doi.org/10.17323/jle.2024.16290

To Type or To Write: The Effect of Writing Modes and Time Constraints on Students' Writing Quality

I Gusti Ngurah Agung Wijaya Mahardika [1], **IGA Lokita Purnamika Utami** [2]

[1] Universitas Hindu Negeri I Gusti Bagus Sugriwa Denpasar, Bali, Indonesia

[2] Universitas Pendidikan Ganesha, Bali, Indonesia

ABSTRACT

Background: The incorporation of technology into English writing class has prompted the use of computer typing as an alternative writing mode to handwriting. However, previous studies investigating the effect of writing modes on writing quality have delivered conflicting results.

Purpose: Considering the increasing prevalence of computers in English writing classes and the time limitation commonly employed by teachers in writing courses, the present study attempts to identify the interaction effect of writing modes and time constraints on EFL students' writing quality.

Method: This study adopted a counterbalanced quantitative design; data were collected from 30 EFL undergraduate students majoring in English Education, using an adapted version of Jacobs' ESL Composition Profile. All participants were subjected to four writing conditions, based on the combination of typing using a computer or writing by hand and a 30-minute or 60-minute time allotment.

Results: The findings of the present study suggest that there is no interaction effect between writing modes and time constraints on writing quality $F(1, 116) = .086$, $p = .770$, and despite the higher scores obtained by the essays typed with computer, writing modes have no significant effect on writing quality $F(1, 116) = .820$, $p = .367$. The results also suggest that time constraints significantly affect writing quality $F(1, 116) = 14.308$, $p = <.001$.

Conclusion: Due to the absence of writing modes effect on writing quality, English teachers are recommended to permit both writing modes in writing essays and to provide more time for students to write, especially in a creative writing environment as opposed to an examination environment.

KEYWORDS

writing modes, time constraints, writing quality, computer typing, handwriting

INTRODUCTION

The incorporation of technology into writing class has elevated the usage of technology as one of the factors that influence students' writing quality, including the effect of computer typing as an alternative writing mode to handwriting. The relationship between writing modes, i.e., computer typing and handwriting, on writing has been investigated by several studies. However, these studies' results are contradictory (Aberšek et al., 2018; Chen et al., 2011; Kimmons et al., 2017; Lee, 2004; Zhu et al., 2016).

Present English learning is taking place in a somewhat different setting than in the past which brought these two writing modes into a very unique rivalry. On the one hand, EFL students claim that writing by hand helps improve their concentration, especially in the drafting stage of the writing (Lund, 2016). Studies in general also suggest that students achieved better information retention, perform better in answering conceptual questions (Mueller & Oppenheimer, 2014); and able to recall more information (Aragón-Mendizábal et al., 2016; Mangen et al., 2015). On the other hand, govern-

Citation: Mahardika, I.G.N.A.W., & Utami I.L.P. (2024). To type or to write: The effect of writing modes and time constraints on students' writing quality. *Journal of Language and Education, 10*(3), 69-78. https://doi.org/10.17323/jle.2024.16290

Correspondence:
I Gusti Ngurah Agung Wijaya Mahardika
E-mail: wijayamahardika@gmail.com

Received: November 11, 2022
Accepted: May 16, 2024
Published: September 30, 2024

ments as well as learning institutions started conducting paperless classrooms, in which all the books are digitalized and students use computers for their academic writing and even in answering exams (Meishar-Tal & Shonfeld, 2018; Siddiqui & Muntjir, 2017). This made an investigation into writing modes vital because a better understanding of the effect of writing modes will eventually contribute valuable empirical findings to the EFL writing body of knowledge.

Besides writing modes, other variables have also been investigated in relation to writing quality, such as writing apprehension (Daly, 1978), pre-writing planning (Amiryousefi, 2017), and time limitation (Caudery, 1990; Ghanbari et al., 2015; Kenworthy, 2006; Zare et al., 2016). In relation to writing modes, time limitation is also very important, because time limitation may influence students' writing quality. Despite the logic that more times means more chance for the students to perform, studies found that it is not always the case. A study on the effect of time pressure and peer feedback on Iranian EFL students taking Oxford Proficiency Test found that peer feedback affected students' writing while time pressure did not have any effect (Ghanbari et al., 2015), similar results were also reported by other researchers (Caudery, 1990; Knoch & Elder, 2010). These results contradict those from Zare, et al. who found that time pressure influences Iranian EFL students' writing coherence and cohesion. The study found that students exposed to time pressure, i.e., limited time to write, performed worse and applied more irrelevant sentences in their writing compared to those given normal time to write (Zare et al., 2016). This result is similar to the findings of Kenworthy, who studied Hong Kong university students and found that students with limited time produced more grammatical mistakes in their writing compared to students with normal time (Kenworthy, 2006). These results highlight the importance of time in writing, for adequate time appears to help students in producing better quality writings.

Considering the contradictory results of previous studies on the effect of writing modes and the effect of time constraints on writing quality, it is essential to conduct a study to investigate further if there is any interaction between writing modes and time constraints on students' writing quality. In particular, this study examines the following questions:

RQ#1: Does the effect of writing modes on students' writing quality depend on time constraint and vice versa?

RQ#2: Do writing modes affect students' writing quality?

RQ#3: Do time constraints affect students' writing quality?

RQ#4: Do the writing modes and time constraints affect the essay length?

LITERATURE REVIEW

Writing with Computer

The advancement of technology especially Information and Communication Technology, has elevated writing skill into a prominent place in human communication. Writing is not just an ordinary skill to be taken for granted but an essential one for life (Graham & Perin, 2007). Therefore, good writing skill must be fostered by the students and prompted by the teachers since writing is a skill that must be learned and can only be mastered under instruction (Brown, 2001). The teaching of writing develops hand in hand with the development of technology, as can be seen from the use of numerous technological tools, software, and applications in teaching writing ranging from the use of word processors (Bangert-Drowns, 1993; Michael Reed, 1996; Sullivan & Pratt, 1996), weblog or blog (Arslan & Şahin-Kizil, 2010; Drexler et al., 2007; Ducate & Lomicka, 2008), Facebook (Altakhaineh & Al-Jallad, 2018; Altunkaya & Topuzkanamış, 2018; Ibrahim et al., 2018) into the latest digital tools to help improve students writing quality (Ivanova et al., 2022; Kaur et al., 2023; Maghsoudi et al., 2022).

One of the most significant changes brought to writing by technology, however, is the use of computers as an alternative writing mode to handwriting. The development of computers has permeated all walks of life, including education. Computers are now a common sight in learning and are being used to replace writing by hand when taking notes, tests, and writing assignments. Governments started conducting paperless classrooms, in which all the books are digitalized and students use computers for their academic writing and even in answering exams (Meishar-Tal & Shonfeld, 2018).

The introduction of computers has enabled writers and students to produce writing in a considerably easier way. Computer permits more flexibility into the writing process compared to writing by hand. Across the phases of writing, prewriting, composing, and revising, computer enables the changing of structure and ideas. It also provides students with additional tools such as spelling-checker and grammar-checker (Ulusoy, 2006). Scholars has lauded the use of computer in writing, considering it to have changed the way people write which in turn changed the way English teacher teach writing (Ivanova et al., 2022; Williams & Beam, 2019).

Despite MacArthur's opinion on the strength of computer writing, studies on the effect of writing modes, i.e., handwriting and computer, are inconclusive. Some scholars found that computer helps improve students' writing; meanwhile, other scholars found that handwriting improves students' writing. Nineteen Norwegian EFL teacher training students were interviewed about their learning preferences in terms of writing modes. All of the students interviewed contended that they prefer to write by hand compared to

computers. The students claimed that writing by hand helps them concentrate and eventually affected their learning retention. Several students also claimed that writing prompts their creativity in learning, stating that their concentration during writing improves their flow of ideas, especially during the drafting stage of the writing (Lund, 2016). Lund concluded that handwriting is an essential tool in ELT learning, she also underlined the need to conduct further study in the affordances of different writing technology in ELT learning.

Typing vs Handwriting across Studies

A study on adults of 16 years and older in the United States of America who were taking functional writing assessments found that test takers who wrote by hand produced higher quality essays than test takers who typed with computers. Furthermore, the test takers who used computers also produced similar length of essays compared to the handwriting group. The study suggests that writing using computer may have given additional burden to the test takers considering that they took more time to answer a writing task compared to the handwriting group (Chen et al., 2011). Meanwhile, in a writing scientific texts study conducted on Slovenian six-grade students showed that despite the fact that typing produces more words than handwriting in the same amount of time, handwritten texts were found to have more information, more terminological accuracy, and, in general, more understanding of the interconnection between listed information in the text (Aberšek et al., 2018). Writing by hand is also found to produce better essays and obtain higher writing scores compared to typed essays. Students who wrote their essays by hand obtained higher writing scores in the dimensions of Ideas and Content, Linguistics expression, and Cohesion and coherence in writing (Zhu et al., 2016).

More interesting findings were reported by Wrigley (2017), who found that handwriting plays an important role in fostering students' creative paraphrasing of the information that they collected from the source text, as compared to students who type their assignments. His observation found that students who wrote their summaries produced more original content compared to those who typed their summaries due to a process of reconstructing a coherent text based on the source text. He observed that students who typed employed copy and paste technique and changed words and phrases here and there to produce his or her version of the text. On the other hand, students who write frequently stop halfway in writing a sentence and then ponder on how they best continue with the sentence. Wrigley contended that the students who write try to reformulate the information they have and create a fresh piece of writing (Wrigley, 2017).

Despite the results of studies that tend to champion handwriting as a better way to produce a piece of writing, other studies have found that writing using computer seems to produce better essay since students who type their essays using computers were found to have less grammatical mistakes, use more unique words, and eventually produce more advanced essays (Kimmons et al., 2017). In his study, Kimmons studied the essay written by students from 3 schools in the Northwestern states of the USA and found that students writing their essays with Chromebook produced higher grade-level of writing and less spelling errors, although the lower number of spelling errors may have to be attributed to the spelling checker feature of the Chromebook (Kimmons et al., 2017).

Writing using computer also seems to enable students to achieve higher writing scores in the dimensions of Ideas and Content, Linguistics expression, and Cohesion and coherence in writing (Zhu et al., 2016). In their study of 32 undergraduate students learning Chinese as a Foreign Language (CFL), they found that students who type their essay were more confident when writing using computer and that the computer assist them in writing their essay by providing alternative Chinese characters for the words they are looking for. Furthermore, students typing their essay also commented that their essay look more professional when typed rather then when written by hand. Nevertheless, we must take into account that some the results of this study were based on writing and typing essays in Chinese characters, not in English and using English alphabets (Zhu et al., 2016).

The results of a 2004 study on 42 international students taking ESL Placement Tests (EPT) at the University of Illinois at Urbana-Champaign showed the advantage of using computer in writing essay (Lee, 2004). When compared to students with handwritten essay, students with computer typed essays obtained higher scores in all the writing dimensions measured which includes organization, content, use of sources, and linguistic expressions (Lee, 2004). Another interesting finding of this study is that the raters of the essays awarded higher scores to the computer-transcribed version of an essay than the original handwritten version. The study also found that students who are accustomed to typing prefer to write using computer and cited the difficulty in editing and correcting their writing when writing by hand (Lee, 2004).

The above studies provide a picture of the conflicting landscape writing modes and writing quality. Today's writing environment tends to promote computer writing; people are more accustomed to typing than writing by hand. Instant messages, social media posts, and even letters in the form of emails are now typed. Official forms are also increasingly available in digital forms; exams are even held in a computerized environment. This situation tends to push people into computer typing, and young learners use computers from a very early age. Thus, knowing the effect of writing modes on writing quality is becoming more critical now than ever. A deeper and better understanding of this will allow English teachers to present a better learning environment for their students, especially in essay writing.

METHOD

Participants

This research involved 30 EFL university students from a state university in Indonesia who were selected based on their writing experience. The students have passed three writing courses in their undergraduate English Education program at their university and are currently enrolled in their fourth writing course. Based on their previous writing experience and a review of the previous courses' syllabuses, we assume that they are familiar with writing argumentative essays as warranted by the writing prompt used herein.

Instruments

The data on students' writing quality were obtained using an adapted version of Jacobs' ESL Composition Profile (Boardman & Frydenberg, 2008), composed of five dimensions: Content, Organization, Language Use, Vocabulary, and Mechanics. The writing prompt used in the present study instructed the participants to write a 300-word argumentative essay based on the topic provided. The prompts also guide the participants to write in at least three paragraphs. The first paragraph introduces the writer's view of the topic and mentions the participants' reasons in brief. The second paragraph should elaborate on the reasons supporting the writer's view as presented in the first paragraph. The third paragraph should restate the writer's view and summarize the supporting reasons. The prompts were first piloted and administered to a different class of 25 students from the same university, department, and semester with the participants. The participants of the prompt pilot were asked to provide inputs to improve the prompt's clarity. 25 essays of the resulting essays from the pilot were taken in random, scored by the raters and said scores were used to establish the Inter-rater Correlation Coefficient

Data Collection Procedure

Using a counterbalanced design, every participant was instructed to write four 300-word argumentative essays using two different writing modes, computer and handwriting, under two different time constraints of 30 minutes and 60 minutes. Each participant wrote four argumentative essays on the topics of "E-mail vs Telephone", "The Impact of Instant Food", "Indonesians' Dependence on Motorcycle", and "Indoor or Outdoor Sports?". The data were obtained in two meetings, as presented in Table 1. In each meeting, the data were taken during their scheduled Scientific Writing class, and the participants were told beforehand that the result of their writing would not, in any way, affect their final grade for the Scientific Writing class.

The participants willingly and voluntarily agreed to participate in the study. The participants were first divided into two groups according to their ID number, as presented in Table 1. Each group then wrote an essay based on the data collection design, using one writing mode under one-time constraint. After completing the first part of the essay writing, the participants were given a 15-minute break and then continued to write the next essay. The second meeting was held the following week, following a similar design

Table 1

Data Collection Design

		Modes of Writing		Essay Topic
		Computer Mode	Handwriting Mode	
Time Constraint	**60 Mins**	Students Number 01-17	Students Number 18-34	1
	30 Mins	Students Number 18-34	Students Number 01-17	2
	30 Mins	Students Number 01-17	Students Number 18-34	3
	60 Mins	Students Number 18-34	Students Number 01-17	4

Table 2

Descriptive Statistics of Each Writing Group

Writing Group	N	Mean	Std. Deviation
H60	30	76.30	7.22
H30	30	71.93	7.25
C60	30	77.80	6.09
C30	30	72.70	6.77

but reversing the time and mode conditions to account for prompt and order. By the end of the second meeting, the participants were asked to fill out a form with their identity and other necessary contact information.

Data Analysis

The total number of essays collected from the participants was 120. These essays were then rated by two independent raters. Both raters were university English teachers, teaching in English Education Departments with at least ten years of teaching experience. Both raters were teachers of Writing courses and are familiar and proficient in rating with Jacobs' ESL Composition Profile, and with high Intraclass Correlation Coefficient of .883, p=.000. The data obtained from each writing modes and time constraints condition were analyzed using two-way ANOVA to determine the interaction effect of the variables under investigation as well identifying the effect of each independent variable on students' writing quality.

RESULTS

Before investigating the interaction effect of Writing Modes and Time Constraints on students' Writing Quality, the data were first analyzed descriptively, as presented in Table 2.

From the data in Table 2 we can identify that, descriptively speaking, the highest writing quality score mean is obtained by the Computer 60 minutes group (M=77.80, SD= 6.09), followed by Handwriting 60 Minutes (M=76.30, SD=7.22), and then Computer 30 Minutes (M=72.70, SD=6.77), and finally Handwriting 30 Minutes (M=71.93, SD=7.25). A further breakdown of the mean scores across the dimensions of writing quality is presented in Table 3.

After determining that the data were homogeneous (the P-value on the Levene's Test was .461), the data were tested to see if they fulfill the assumption of normality. It was found that the data were normally distributed for all groups, except for the Computer 30 Minutes the p-value of the Kolmogorov-Smirnov was 0.01. Nevertheless, since two-way ANOVA is considered robust against violation of the assumption of normality especially when the sample sizes are equal (Barkaoui & Knouzi, 2018; Field, 2013; Tabachnick & Fidell, 2014)The data were analyzed using two-way ANOVA, and the results are presented in Table 4.

The statistical analysis shows that for the interaction effect of writing modes and time constraints on students' writing quality, the F-value is $F(1, 116) = .086$, $p = .770$. This result implies that there is no statistically significant interaction effect between writing modes and time constraints on students' writing quality. Meanwhile, the F-value for the main effect of writing modes on students' writing quality was $F(1, 116) = .820$, $p = .367$, this means that the F-value is not statistically significant. This result suggests that there is no statistically significant effect of writing modes on students' writing quality. As for the effect of time constraints on students' writing quality the F-value was $F(1, 116) = 14.308$, $p = <.001$ the F-value is statistically significant. This means time constraints have a statistically significant effect on students' writing quality.

The present study also found that the means of word count of the essay written in 30 minutes time in any writing modes were less than the required 300 words by the writing prompt. The means of word count were 219 and 272 for computer-typed essays and handwritten essays, respectively. Of the 30 handwritten essays written in 30 minutes, only three essays (10%) have more than 300 words, 15 (50%) have between 200 to 299 words, and 12 (40%) of these essays were even below 200 words. Meanwhile, from 30 of the computer-typed essays written in 30 minutes time, 12 (40%) essays had more than 300 words, 14 (46.67%) had between 200 to 299 words, and 4 (13%) were below 200 words. When compared to the number of words in the essay produced in 60 minutes time, the difference is clear. Of the 30 handwritten essays written in 60 minutes, 22 (73.33 %) essays had more than 300 words, with 8 (23.67%) essays had less than 300 words, but none of the essays fell below 200 words. Interestingly, from the 30 computer-typed essays written in 60 minutes time, 21 (70%) essays had more than 300 words, 7 (23.33%) had between 200 to 299 words, and 2 (6.67%) were below 200 words.

Table 3

Mean Scores Comparison Across Dimensions of Writing Quality

No	Dimension	H60	H30	C60	C30
1	Content	19.77	18.63	20.17	18.37
2	Organization	19.50	18.13	20.00	18.30
3	Language Use	18.60	17.70	19.23	18.07
4	Vocabulary	10.43	9.77	10.07	9.67
5	Mechanics	8.00	7.70	8.33	8.30
6	Word Count	329	220	342	272

Table 4
ANOVA Results

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	714.700[a]	3	238.233	5.071	.002	.116
Intercept	669312.033	1	669312.033	14247.825	.000	.992
Writingmodes	38.533	1	38.533	.820	.367	.007
TimeConstraint	672.133	1	672.133	14.308	.000	.110
Writingmodes * TimeConstraint	4.033	1	4.033	.086	.770	.001
Error	5449.267	116	46.976			
Total	675476.000	120				
Corrected Total	6163.967	119				

Table 5
Word Count Percentage Across Time and Writing Mode

Mode and Time	Word Count Percentage			Word Count Average
	<200	200-299	>300	
H60	-	26.67	73.33	329
C60	6.67	23.33	70.00	342
H30	40	50	10	220
C30	13.33	46.67	40.00	272

DISCUSSION

The present study aimed to explore the effects of writing modes and time constraints on students' writing quality, both independently and in interaction. Previous research has yielded mixed findings on whether computer-typed writing offers advantages over handwriting, particularly in time-pressured environments. By analyzing various dimensions of writing, such as content, coherence, mechanics, and vocabulary, this study provides new insights into how these factors influence writing outcomes. In the following discussion, we examine the independent effects of writing modes and time constraints, delve into the specific areas where each has an impact, and consider the broader implications for teaching and assessment in EFL contexts.

Interaction between Writing Modes and Time Constraints

The study found no statistically significant interaction effect between writing modes and time constraints on students' writing quality. This suggests that the impact of writing modes and time constraints on students' performance operated independently, without reinforcing or diminishing each other's effect. This result aligns with Lovett et al. (2010), who similarly found no interaction between time allocation and writing modes in writing performance despite differences in the time frames used. In our study, we compared 30 and 60 minutes, while Lovett explored shorter spans (10 and 15 minutes). Even with more significant time differences in this study, the lack of interaction remained consistent, reinforcing the conclusion that writing modes and time constraints independently shape writing outcomes.

Unlike earlier studies, our findings emphasize that increased time allocations do not create a dependency between writing mode and writing quality. This contradicts previous claims that technology might amplify performance under time pressure by increasing the speed and ease of writing (Kimmons et al., 2017). Thus, the general contribution of this study lies in confirming that while computer use and time matter independently, they do not interact synergistically.

Impact of Writing Modes on Writing Quality

The study found no significant main effect of writing modes on students' overall writing quality ($p = .367$). Although computer-typed essays achieved higher mean scores than handwritten essays in several writing dimensions (content, coherence, and mechanics), the differences were statistical-

ly insignificant. This outcome challenges the assumptions from earlier research, such as Zhu et al. (2016), which suggested that typing improves writing performance. The present results indicate that while computer typing offers specific advantages—such as ease of editing, rearranging ideas, and automatic correction tools—these advantages do not necessarily translate into significantly higher writing quality.

An interesting finding is the use of more varied vocabulary in handwritten essays. Students writing by hand tended to use a broader range of words and more creatively rephrased ideas. This contrasts with computer-written essays, where students frequently reused words, likely due to the ease of copy-paste functionality. This finding resonates with previous studies' findings, which found that writing by hand can deepen cognitive engagement, leading to better word choice and paraphrasing(Mueller & Oppenheimer, 2014; Wrigley, 2017).

In contrast to previous studies that argue for the superior benefits of one mode i.e. handwriting (Aberšek et al., 2018; Aragón-Mendizábal et al., 2016; Chen et al., 2011; Lund, 2016; Mueller & Oppenheimer, 2014; Zhu et al., 2016) and computer typing (Kimmons et al., 2017; Lee, 2004; MacArthur, 1988), the present study shows that both modes offer unique advantages and limitations. Computer-based writing improves speed and structure but may encourage word repetition, while handwriting enhances creativity but is more time-consuming. Therefore, it is important for educators to balance these benefits by allowing students the freedom to choose their preferred writing mode.

The Role of Time Constraints on Writing Quality

The study further revealed that time constraints significantly affected writing quality, regardless of the writing mode. Students produced better quality essays when given more time (60 minutes) compared to shorter periods (30 minutes). This finding is consistent with earlier studies (Kenworthy, 2006; Powers & Fowles, 1996; Zare et al., 2016), which emphasized that extended writing time allows students to plan, revise, and improve their essays, resulting in fewer grammatical errors and higher overall scores(Na & Yoon, 2016).

Interestingly, the findings indicate that time pressure negatively impacts students' strategies. Participants given 30 minutes were more likely to write without planning, leading to stalled progress midway through the task. On the other hand, students given 60 minutes exhibited more deliberate writing behaviors, including time for reflection and revision. These observations align with Wrigley (2017), who reported that students tend to pause and mentally organize ideas before starting their writing under relaxed time constraints.

Our results further highlight the importance of providing sufficient time for students to demonstrate their full writing potential. While timed assessments are useful for evaluating writing under pressure, ample writing time fosters creativity and allows students to engage more deeply with the writing process.

Impact of Time Constraints on Essay Length

The study also examined the relationship between time constraints and the length of essays. The findings show that students produced longer essays with higher word counts when given more time, regardless of the writing mode. In 60-minute sessions, students produced longer essays, with fewer short essays under 200 words. This finding contrasts with Lovett et al. (2010), who suggested that increased time allocation benefits students using computers more than those writing by hand. In the present study, students with handwritten essays also benefited significantly from additional time, suggesting that both writing modes require adequate time to achieve optimal results.

The findings further suggest that in terms of vocabulary, computer writing consistently yields lower scores than handwriting under the same time constraints. The vocabulary assessment focuses on range, word choice, idiom usage, appropriate register, and mastery of word forms (Boardman & Frydenberg, 2008; Winke & Lim, 2015). Repeated use of the same phrases is penalized, which might explain why handwriting students, who find it harder to rewrite sentences, were forced to be more creative and express their ideas in different ways. While computer users may write more words in less time, this does not always correlate with better essay quality. Therefore, strict word count requirements may be unfair to handwriting students, and time should be allocated to allow creative expression. Alternatively, assessments might focus more on quality rather than quantity.

In the EFL context, the results indicate that computers offer no significant advantage over handwriting. Providing students the choice of writing mode does not give them an unfair edge. As we move toward a digital age (Meishar-Tal & Shonfeld, 2018; Siddiqui & Muntjir, 2017), these findings suggest that EFL classes should allow flexibility in writing modes and offer ample time for students to develop their writing skills, especially in learning rather than testing environments.

CONCLUSION

The present study demonstrates that while writing modes do not have a statistically significant effect on students' overall writing quality, a deeper analysis of individual dimensions reveals that computer-typed essays scored higher than handwritten ones in four out of five dimensions. This suggests that computers can play a supporting role in enhancing students' writing quality. However, the results also show that neither mode offers an unfair advantage over the

other. This confirms that students, whether writing by hand or typing, perform comparably, reinforcing the notion that writing mode alone does not determine writing outcomes.

The study emphasizes the importance of time allocation in writing performance, showing that more time improves both the quality and quantity of writing, regardless of the mode. Sufficient time allows students to develop ideas more fully, enhancing their performance. EFL teachers should integrate technology, like computers, to support writing, rather than restrict students to handwriting. Allowing students to choose their preferred writing mode promotes engagement and fairness in assessment. Students must be given enough time to write to truly develop their writing and showcase their writing ability. Time allocation should also be considered wisely in writing assessment setting based on the same reason.

The present study's limitations, including specific time frames and a small sample size, suggest that future research should consider broader time constraints, additional dimensions like creativity, and larger sample sizes for more generalizable results.

Future research should explore additional variables that may moderate the effect of writing modes on performance, such as gender, essay genres, and specific writing conditions. The current study showed that time constraints have a significant impact on writing performance, but more research is needed to determine how these effects vary under different contexts. Furthermore, investigating students' writing mode preferences and their impact on performance could yield insights into whether allowing students to use their preferred methods improves writing quality.

DECLARATION OF COMPETITING INTEREST

None declared.

AUTHOR CONTRIBUTIONS

I Gusti Ngurah Agung Wijaya Mahardika: conceptualization; data curation; formal analysis; funding acquisition; investigation; methodology; project administration; resources; software; supervision; validation; visualization; writing – original draft; writing – review & editing.

IGA Lokita Purnamika Utami: conceptualization; data curation; formal analysis; funding acquisition; investigation; methodology; project administration; resources; software; validation; visualization; writing – original draft.

REFERENCES

Aberšek, M. K., Aberšek, B., & Flogie, A. (2018). Writing versus typing during science teaching: Case study in Slovenia. *Journal of Baltic Science Education*, *17*, 84–96. http://dx.doi.org/10.33225/jbse/18.17.84

Altakhaineh, A. R. M., & Al-Jallad, M. Z. (2018). The use of twitter and facebook in teaching mechanics of writing to Arabic-speaking EFL learners. *International Journal of Emerging Technologies in Learning*, *13*(9), 4–14. https://doi.org/10.3991/ijet.v13i09.8457

Altunkaya, H., & Topuzkanamış, E. (2018). The effect of using facebook in writing education on writing achievement, attitude, anxiety and self-efficacy perception. *Universal Journal of Educational Research*, *6*(10), 2133–2142. https://doi.org/10.13189/ujer.2018.061010

Amiryousefi, M. (2017). The differential effects of collaborative vs. individual prewriting planning on computer-mediated L2 writing: transferability of task-based linguistic skills in focus. *Computer Assisted Language Learning*, *30*(8), 766–786. https://doi.org/10.1080/09588221.2017.1360361

Aragón-Mendizábal, E., Delgado-Casas, C., Navarro-Guzmán, J. I., Menacho-Jiménez, I., & Romero-Oliva, M. F. (2016). A comparative study of handwriting and computer typing in note-taking by university students [Análisis comparativo entre escritura manual y electrónica en la toma de apuntes de estudiantes universitarios]. *Comunicar*, *24*(48), 101–107. https://doi.org/10.3916/C48-2016-10

Arslan, R. Ş., & Şahin-Kizil, A. (2010). How can the use of blog software facilitate the writing process of English language learners? *Computer Assisted Language Learning*, *23*(3), 183–197. https://doi.org/10.1080/09588221.2010.486575

Bangert-Drowns, R. L. (1993). The word processor as an instructional tool: a meta-analysis of word processing in writing instruction. *Review of Educational Research*, *63*(1), 69–93. https://doi.org/10.3102/00346543063001069

Barkaoui, K., & Knouzi, I. (2018). The effects of writing mode and computer ability on L2 test-takers ' essay characteristics and scores. *Assessing Writing*, *September 2017*, 0–1. https://doi.org/10.1016/j.asw.2018.02.005

Boardman, C. A., & Frydenberg, J. (2008). *Writing to communicate 2: Paragraphs and essays* (3rd ed.). Pearson Longman. https://doi.org/10.1097/00005721-199203000-00001

Brown, H. D. (2001). *Teaching by principles* (2nd ed.). Longman.

Caudery, T. (1990). The validity of timed essay tests in the assessment of writing skills. *ELT Journal*, *44*(2), 122–131. https://doi.org/10.1093/elt/44.2.122

Chen, J., White, S., McCloskey, M., Soroui, J., & Chun, Y. (2011). Effects of computer versus paper administration of an adult functional writing assessment. *Assessing Writing*, *16*(1), 49–71. https://doi.org/10.1016/j.asw.2010.11.001

Daly, J. A. (1978). Writing apprehension and writing competency. *The Journal of Educational Research*, *72*(1), 10–14. https://doi.org/10.1080/00220671.1978.10885110

Drexler, W., Dawson, K., & Ferdig, R. E. (2007). Collaborative blogging as a means to develop elementary expository writing skills. *View Publication Stats Electronic Journal for the Integration of Technology in Education*, *6*, 140–160.

Ducate, L. C., & Lomicka, L. L. (2008). Adventures in the blogosphere: From blog readers to blog writers. *Computer Assisted Language Learning*, *21*(1), 9–28. https://doi.org/10.1080/09588220701865474

Field, A. (2013). *Discovering statistics using IBM SPSS Statistics* (4th ed.). Sage.

Ghanbari, N., Karampourchangi, A., & Shamsaddini, M. R. (2015). An exploration of the effect of time pressure and peer feedback on the Iranian EFL students' writing performance. *Theory and Practice in Language Studies*, *5*(11), 2251. https://doi.org/10.17507/tpls.0511.08

Graham, S., & Perin, D. (2007). Effective strategies to improve writing of adolescents in middle and high schools. In *A report to Carnegie corporation of New York.* Alliance for Excellent Education.

Ibrahim, S., Sharina, S., Md Tahir, N., & Primsuwan, P. (2018). Promoting learners' autonomy by using facebook to enhance students' writing skill. *Journal of Creative Practices in Language Learning and Teaching*, *6*, 56–68.

Ivanova, M., Arupova, N., & Mekeko, N. (2022). Digital support for teaching punctuation in academic writing in English. *Journal of Language and Education*, *8*(3), 81–96. https://doi.org/10.17323/jle.2022.13608

Kaur, D. J., Saraswat, N., & Alvi, I. (2023). Technology-enabled language leaning: Mediating role of collaborative learning. *Journal of Language and Education*, *9*(1), 89–101. https://doi.org/10.17323/JLE.2023.12359

Kenworthy, R. (2006). Timed versus at-home assessment tests: Does time affect the quality of second language learners' written compositions? *Tesl-Ej*, *10*(1), 1–10.

Kimmons, R., Darragh, J. J., Haruch, A., & Clark, B. (2017). Essay composition across media: A quantitative comparison of 8th grade student essays composed with paper vs. chromebooks. *Computers and Composition*, *44*, 13–26. https://doi.org/10.1016/j.compcom.2017.03.001

Knoch, U., & Elder, C. (2010). Validity and fairness implications of varying time conditions on a diagnostic test of academic English writing proficiency. *System*, *38*(1), 63–74. https://doi.org/10.1016/j.system.2009.12.006

Lee, H. K. (2004). A comparative study of ESL writers' performance in a paper-based and a computer-delivered writing test. *Assessing Writing*, *9*(1), 4–26. https://doi.org/10.1016/j.asw.2004.01.001

Lovett, B. J., Lewandowski, L. J., Berger, C., & Gathje, R. A. (2010). Effects of response mode and time allotment on college students' writing. *Journal of College Reading and Learning*, *40*(2), 64–79. https://doi.org/10.1080/10790195.2010.10850331

Lund, R. E. (2016). Handwriting as a tool for learning in ELT. *ELT Journal*, *70*(1), 48–56. https://doi.org/10.1093/elt/ccv048

MacArthur, C. A. (1988). The impact of computerson the writing process. *Exceptional Children*, *54*(6), 536–542. https://doi.org/10.2307/2552790

Maghsoudi, N., Golshan, M., & Naeimi, A. (2022). Integrating digital multimodal composition into EFL writing instruction. *Journal of Language and Education*, *8*(1), 84–99. https://doi.org/10.17323/jle.2022.12021

Mangen, A., Andal, L. G., Oxborough, G. H., & Bronnick, K. (2015). Handwriting versus keyboard writing: Effect on word recall. *Journal of Writing Research*, *7*(2), 227–247. https://doi.org/10.17239/jowr-2015.07.02.1

Meishar-Tal, H., & Shonfeld, M. (2018). Students' writing and reading preferences in a paperless classroom. *Interactive Learning Environments*, *27*(7), 908–918. https://doi.org/10.1080/10494820.2018.1504306

Michael Reed, W. (1996). Assessing the Impact of Computer-Based Writing Instruction. *Journal of Research on Computing in Education*, *28*(4), 418–437. https://doi.org/10.1080/08886504.1996.10782176

Mueller, P. A., & Oppenheimer, D. M. (2014). The pen is mightier than the keyboard: Advantages of longhand over laptop note taking. *Psychological Science*, *25*(6), 1159–1168. https://doi.org/10.1177/0956797614524581

Na, S., & Yoon, H. (2016). Effects of in-class and out-of-class writing assignments on L2 writing strategy use and writing quality. *The Asia-Pacific Education Researcher*, *25*(2), 195–205. https://doi.org/10.1007/s40299-015-0250-5

Powers, D. E., & Fowles, M. E. (1996). Effects of applying different time limits to a proposed GRE writing test. *Journal of Educational Measurement*, *33*(4), 433–452. https://doi.org/10.1111/j.1745-3984.1996.tb00500.x

Siddiqui, A. T., & Muntjir, M. (2017). An approach to smart study using pen and paper learning. *International Journal of Emerging Technologies in Learning*, *12*(05), 117. https://doi.org/10.3991/ijet.v12i05.6798

Sullivan, N., & Pratt, E. (1996). A comparative study of two ESL writing environments: A computer-assisted classroom and a traditional oral classroom. *Science*, *29*(4), 491–501. https://doi.org/10.1016/S0346-251X(96)00044-9

Tabachnick, B. G., & Fidell, L. S. (2014). *Using multivariate statistics* (6th ed.). Pearson.

Ulusoy, M. (2006). The role of computers in writing process. *The Turkish Online Journal of Educational Technology*, *5*(4), 58–66. https://doi.org/10.1016/S0924-9338(97)80221-0

Williams, C., & Beam, S. (2019). Technology and writing: Review of research. *Computers and Education*, *128*, 227–242. https://doi.org/10.1016/j.compedu.2018.09.024

Winke, P., & Lim, H. (2015). ESL essay raters' cognitive processes in applying the Jacobs et al. rubric: An eye-movement study. *Assessing Writing*, *25*, 37–53. https://doi.org/10.1016/j.asw.2015.05.002

Wrigley, S. (2017). Avoiding 'de-plagiarism': Exploring the affordances of handwriting in the essay-writing process. *Active Learning in Higher Education*, *1*(13), 146978741773561. https://doi.org/10.1177/1469787417735611

Zare, M., Mohazabieh, S., & Kamali, Z. (2016). The effects of time constraints on the unity and coherence of IELTS candidates ' writing skills. *ELT Voices - International Journal for Teachers of English*, *31*(6), 24–31.

Zhu, Y., Mark Shum, S.-K., Brian Tse, S.-K., & Liu, J. J. (2016). Word-processor or pencil-and-paper? A comparison of students' writing in Chinese as a foreign language. *Computer Assisted Language Learning*, *29*(3), 596–617. https://doi.org/10.1080/09588221.2014.1000932

https://doi.org/10.17323/jle.2024.18708

Synergizing Generative Pre-Trained Transformer (GPT) Chatbots in a Process-Based Writing Paradigm to Enhance University Students' Writing Skill

Roderick Julian Robillos

Khon Kaen University, Khon Kaen, Thailand

ABSTRACT

Background: The combination of the process-based writing framework and GPT-based chatbots establishes a dynamic and interactive environment, leading students through the stages of planning, drafting, revising, and editing. This collaborative approach not only elevates writing skills but also cultivates critical thinking and self-reflection, thereby transforming the writing process into a more effective and engaging learning experience. Despite the potential synergy to revolutionize the writing process, there exists a surprising lack of research within the educational domain exploring the impact of this innovative approach.

Purpose: This study investigates the influence of a GPT-based chatbot within a process-based writing framework on university EFL students' writing skills, specifically focusing on components such as organization, content, coherence-cohesion, logical connection, and argumentation.

Method: Employing a sequential mixed methods type of research with a pre- and post-test design, 30 university EFL students were selected via purposive sampling technique. They engaged in 10 sessions that incorporated GPT-based chatbots within a process-based writing framework. Data collections were through pre-and post- writing tests, writing quizzes, and semi-structured interviews.

Results: The results highlighted substantial improvements in participants' writing performance, evident through a noteworthy increase in post-writing test scores ($\bar{x}$=17.03) in comparison to pre-writing test scores ($\bar{x}$=9.13). The study identified a progressive enhancement in four out of five writing components - organization, content, coherence-cohesion, and argumentation - across the 1st to 4th quizzes. However, the 'logical connection' component experienced a temporary decline during the 2nd and 3rd writing quizzes, rebounding significantly in the 4th quiz. Notably, the most improved writing components were 'content' and 'argumentation,' while the component related to 'logical connection' exhibited the least improved one. Qualitative findings further underscored participants' acknowledgment of the effectiveness of the strategy in facilitating their writing tasks.

Conclusion: The integration of chatbots within a writing framework was concluded as a facilitative pedagogical approach, fostering a dynamic, personalized, and effective learning experience, contributing to the multifaceted improvement of their writing skills. As educators and practitioners consider innovative approaches, this study provides a compelling case for the effective utilization of GPT-based chatbots in fostering language proficiency and a more engaging learning experience.

KEYWORDS

GPT-based applications, process-based writing, GPT-based chatbots within the process-based framework, writing skill

Citation: Robillos, R. (2024). Synergizing generative pre-trained transformer (GPT) Chatbots in a process-based writing paradigm to enhance university students' writing skills. *Journal of Language and Education, 10*(3), 79-94. https://doi.org/10.17323/jle.2024.18708

Correspondence:
Roderick Julian Robillos
E-mail: wijayamahardika@gmail.com

Received: December 20, 2023
Accepted: May 16, 2024
Published: September 30, 2024

INTRODUCTION

Effective writing skills are paramount in language learning, serving as a cornerstone for successful communication in academic and professional contexts. The significance of writing proficiency cannot be overstated, especially for EFL students navigating the challenges of tertiary education (Wonglakorn & Deerajviset, 2023). The ability to express thoughts coherently, organize ideas systematically, and construct persuasive arguments is not only a key academic requirement but also a skill with far-reaching implications for future career success (Malmir & Khosravi, 2018).

Acknowledging the profound importance of writing, this study emphasizes the central role of the process-based approach, transcending a mere set of steps to become a dynamic framework guiding learners through planning, drafting, revising, and editing. With robust support from research (Kitajroonchai et al., 2022; Robillos & Namwong, 2021; Simpson, 2013; Wutthipan, 2023), this methodology cultivates a growth mindset, encourages reflective learning, and propels continuous improvement. It nurtures creativity and exploration, enhancing students' problem-solving abilities concerning content, structure, arguments, and audience. Integrated feedback from peers, instructors, and writing tools is crucial for ongoing refinement of skills (Kitajroonchai et al., 2022), effectively reducing writing anxiety (Robillos & Namwong, 2021). This approach instills confidence through the iterative process, preparing students for future tasks with enhanced self-assurance. As real-world writing often involves multiple drafts and revisions, the process-based approach serves as a vital training ground for professional communication (Wonglakorn & Deerajviset, 2023). Moreover, it transcends traditional models, allowing students to develop skills applicable beyond the classroom, extending into real-world writing scenarios (Wuttiphan, 2023).

The rise of GPT-based chatbots, a cutting-edge language model developed by OpenAI (Open Artificial Intelligence), have sparked a paradigm shift in the field of writing assistance (Fitria, 2023; Phillips, 2022; Sinha, 2020; Su et al., 2023; Zoherey, 2023). Positioned at the forefront of technological innovation, GPT-based chatbots have garnered acclaim for their capacity to provide immediate and personalized support to users grappling with diverse writing challenges (Okonkwo & Ade-Ibijola, 2021; Zhai, 2022). Their merits encompass not only grammar and style checks, ensuring linguistic precision (Fitria, 2023), but also extend to offering insightful content development suggestions and generating thought-provoking writing prompts (Su et al., 2023). Beyond mere assistance, GPT-based chatbots introduce a novel dimension by fostering a dynamic and iterative feedback loop (Su et al., 2023; Zoherey, 2023). This continuous engagement facilitates ongoing conversations with users, enriching their understanding of the intricate nuances inherent in the writing process (Zoherey, 2023). Adaptable and scalable, GPT-based chatbots emerge as a versatile tool that not only aids in overcoming writing anxiety (Fitria, 2023) but also enhances accessibility, making the writing process more inclusive and user-friendly (Okonkwo & Ade-Ibijola, 2021). Additionally, by promoting independent learning, GPT-based chatbots empower users to not only refine their writing skills but also instill a sense of autonomy in navigating the complexities of effective communication (Zoherey, 2023).

Within the educational context of Thailand, conventional teaching approaches in writing pose challenges as they tend to inhibit students' free exploration of ideas and collaborative engagement with peers, restricting opportunities for compositional refinement. The overemphasis on rote memorization, grammar, and vocabulary accuracy further impedes the development of critical thinking, creativity, and collaborative writing skills (Robillos & Bustos, 2022; Wuttiphan, 2023). Moreover, an additional challenge arises in the reluctance of teachers to integrate technology tools into the writing process for their students. This hesitancy may stem from a lack of familiarity or confidence in utilizing such tools ((Robillos & Bustos, 2022), hindering the potential benefits technology can offer in facilitating the writing process. As a consequence, the traditional teaching methods not only impede the exploration of ideas and collaborative initiatives among students (Wuttiphan, 2023) but also fail to harness the transformative potential of technology in overcoming these limitations (Robillos, 2023). Consequently, the traditional teaching methods not only impede the exploration of ideas and collaborative initiatives among students (Robillos, 2022; Wuttiphan, 2023) but also fail to address and overcome the transformative potential of technology, leaving unaddressed the problems that students encounter in content creation, logical coherence, unity, and argumentation in their written compositions.

Recognizing the deficiencies in the current teaching practices, there emerges a pressing need for an innovative and transformative approach to teaching writing in Thai schools and universities. Traditional methods fall short in providing students with opportunities for holistic skill development, especially in articulating arguments and engaging in thoughtful analysis. To address these challenges, the integration of a robust writing framework with technological applications becomes imperative (Evmenova & Regan, 2019; Robillos, 2022; 2023). Extensive evidence supports the effectiveness of technology in various writing stages, such as prewriting, drafting, revising, proofreading, and publishing (Evmenova & Regan, 2019; Su et al., 2023; Zhang, 2021; Zoherey, 2023). The researcher leverages the capabilities of GPT-based chatbots to guide students towards a nuanced understanding of constructing and articulating arguments effectively within the process-based writing paradigm. This multifaceted approach not only addresses immediate writing challenges (Robillos & Namwong, 2021; Wuttiphan, 2023) but also cultivates foundational skills essential for academic success and beyond. As the study delves into investigating the impact of synergizing GPT-based chatbots in conjunction with the pro-

cess-based writing paradigm, the aim is not only to enhance immediate writing skills but also to underscore the potential for a transformative shift in the overall writing pedagogy. This paradigm shift marks a significant advancement in preparing students for the complex demands of contemporary communication, emphasizing empowerment and skill development in alignment with the evolving landscape of effective writing practices.

LITERATURE REVIEW

The Process Writing Approach and its Pedagogical Significance in EFL Instruction

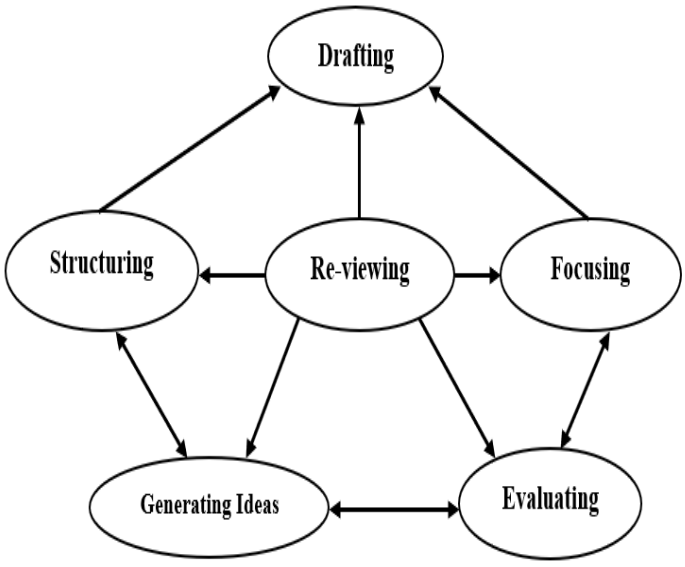
The process writing approach, recognized as a cornerstone in EFL classrooms, has garnered significant attention from researchers (Karatay, 2011; Kitajroonchai et al., 2022; Wonglakorn & Deerajviset, 2023). This writing approach places its primary emphasis on the writing process, rather than solely emphasizing the final product (White & Arndt, 1991). This widely embraced methodology has been thoroughly examined and expanded upon by scholars (Brown, 2001; Coffin et al., 2003; Robillos & Namwong, 2021; Wuttiphan, 2023), solidifying its pedagogical importance. Coffin et al. (2003) contributed an intricate eight-stage writing process, encompassing pre-writing, planning, drafting, reflecting, peer reviewing, revising, and editing. Aligning with this, Brown (2001) emphasizes the cognitive aspects of writing, highlighting prewriting, drafting, revising, and editing as pivotal stages. Additionally, Karatay (2011) delves into the nuanced stages of the process-based writing approach, expanding the spectrum to include prewriting, drafting, editing, revising, and publishing. Together, these researchers not only endorse the prominence of process writing approach but also collectively enrich our understanding of its multifaceted stages and applications within the realm of language education. In line with this perspective, the present study is grounded in the process writing approach, drawing from the framework proposed by White and Arndt (1991), which provides a concise explanation of each stage within the approach.

The initial stage, known as "focusing," involves determining the purpose and underlying motivations for writing. The subsequent stage, "structuring," revolves around organizing ideas in a manner that is coherent and comprehensible to the reader. As the writing process progresses, the "drafting" stage signifies the shift from a writer-centered perspective to one that considers the reader's perspective. Multiple drafts are generated and feedback from teachers or peers is incorporated. "Re-viewing" entails stepping back from the text and assessing it objectively, questioning its accuracy and effectiveness. Evaluation is provided throughout the process to offer continuous support to students rather than solely at the end. It should be noted that the "generating ideas" stage is particularly crucial during the initial phases of the writing process, as it involves selecting a topic and establishing a purpose

Several recent studies, including those by Wonglakorn and Deerajviset (2023), Robillos and Namwong (2021), and Kitajroonchai et al. (2022), highlight the positive impact of the

Figure 1

The Process-Based Writing Approach



Note. Adapted from White & Arndt (1991).

process writing approach. For example: Wonglakorn and Deerajviset (2023) conducted a study involving 62 Thai EFL secondary school students, using a mixed-methods research design. Their findings revealed that collaborative writing positively influenced students' skills and attitudes, with the writing process serving as a useful guide for organizing well-structured and coherent paragraphs. Similarly, Robillos and Namwong (2021) investigated the impact of a process-oriented approach on 27 Thai university students, showing significant improvements on their writing performance and positive responses to the effectiveness of the writing approach in developing compositions and enhancing writing self-regulation. These studies collectively emphasize the valuable role of the process-oriented approach in enhancing students' writing skills.

Despite the wealth of research on the process writing approach, a critical gap persists in the literature concerning the integration of innovative technologies, such as GPT-based chatbots, with this established methodology. The existing studies provide a solid foundation but fall short in exploring the potential synergies and pedagogical implications of incorporating advanced natural language processing tools into the established process-based writing framework. This gap prompts the need for further investigation into the unexplored territory of how emerging technologies can enhance and complement the process writing approach. By examining the potential benefits, challenges, and pedagogical implications of merging GPT-based chatbot with the well-established writing approach, the present study aims to contribute valuable insights into the evolving landscape of EFL instruction, paving the way for a more dynamic and personalized writing instructional paradigm in the digital age.

Affordances of GPT-Based Chatbots in Students' Writing

The integration of GPT-based chatbots into the realm of education has garnered considerable attention, particularly for its potential to enhance students' writing skills (Su et al., 2023; Shibani et al., 2017). GPT-based chat applications, driven by advanced natural language processing (NLP) algorithms, have emerged as a technological breakthrough in language learning (Moqbel & Al-Kadi, 2023; Su et al., 2023). These applications leverage extensive pre-training data to simulate human-like conversations, providing personalized language support and immediate feedback (Okonkwo & Ade-Ibijola, 2021; Shibani et al., 2017). The flexibility and adaptability of GPT-based chatbots make them promising tools for supporting various aspects of language education, with a particular focus on writing.

Prior studies have highlighted the ability of GPT-based chatbots to engage learners in interactive language exchanges, creating an environment conducive to language learning (Sinha, 2020; Su et al., 2023). The affordances lie not only in generating authentic conversations but also in providing tailored assistance for writing tasks. Learners can receive instant feedback on their written expressions, promoting continuous improvement and refinement of their writing skills (Zhai, 2022). A noteworthy affordance of GPT-based chatbots is their role in reducing writing anxiety among students (Li et al., 2017). The presence of AI-powered tools offers a less intimidating space for learners to express themselves in writing (Okonkwo & Ade-Ibijola, 2021). This aspect is particularly crucial in fostering a positive writing environment, empowering students to take ownership of their learning and become more self-directed in their writing endeavors.

However, despite these positive aspects, a comprehensive understanding of the specific benefits and challenges of GPT-based chatbots in the context of students' writing is yet to be fully explored. Existing studies tend to provide broad overviews of the potential of technology without delving into the intricacies of its application in writing instruction. That is, while studies acknowledge the overall benefits of the technology, they often lack in-depth examinations of its practical implementation and specific impact on the instructional aspects of writing. The unique affordances and limitations of GPT-based chatbots within the writing domain, especially in educational settings, warrant focused investigation. The emphasis is on the need for a more nuanced exploration of how GPT-based chatbots afford or hinder students' writing skills. The research niche lies in bridging this gap and providing insights into the specific affordances that this technology offers in the realm of writing education, contributing to a more targeted and effective approach for students in various educational contexts.

While previous research has separately explored the advantages of the process-based writing approach (Karatay, 2011; Kitajroonchai et al., 2022; Robillos & Namwong, 2021; Wonglakorn & Deerajviset, 2023; Wuttiphan, 2023) and GPT-based chat applications (Fitria, 2023; Kostka & Maliborska, 2016; Shibani et al., 2017; Sinha, 2020; Su et al., 2023), there is a scarcity of empirical studies investigating their synergistic impact on language learning, particularly in enhancing writing skills. This study aims to fill this gap by demonstrating the potential of integrating GPT-based chatbots as real-time collaborators within a process-based writing framework. Recognizing that the use of technology alone may not be as effective without a structured instructional approach, the study focuses on demonstrating how these chat applications, when integrated with a process-based approach, provide learners with instant access to language resources, vocabulary suggestions, and grammar assistance, ultimately elevating the quality of their written responses. The findings of this study seek to offer valuable insights into the efficacy of this integrated, learner-centered approach, pro-

viding practical implications for language educators looking to incorporate innovative technological tools in EFL writing instruction. Specifically, the following research questions (RQ's) are sought to be answered:

RQ#1: Is there a significant difference between the students' writing performance before and after the use of GPT–based chatbots within a process-based writing paradigm (the intervention used)?

RQ#2: Do the students' writing quiz performances improve via the implementation of the intervention in terms of organization, content, coherence-cohesion, logical connection, and argumentation?

RQ#3: What learning experiences have the students obtained in facilitating their writing tasks using the intervention?

METHOD

Design

The research design employed in the current study was a sequential mixed-method approach, drawing on both quantitative and qualitative methodologies (Creswell & Creswell, 2018). The quantitative aspect of the research involved the use of data to assess the measurable impact of integrating GPT-based chatbots within the process-based writing framework on students' writing skills. This included analyzing performance metrics such as scores in quizzes and post-writing assessments. On the other hand, the qualitative aspect delved into the subjective experiences and perceptions of participants, seeking to uncover their opinions on how the intervention influenced their writing tasks. Semi-structured interviews were utilized to capture the rich, nuanced insights of the participants, providing a more holistic understanding of the intervention's effectiveness beyond numerical metrics (Creswell & Creswell, 2028).

Furthermore, a single group of pre- and post-test design was employed to assess the impact of the intervention on the participants' writing skills. In the study, a writing pretest was administered to evaluate the participants' writing abilities and establish a baseline for comparison. The pretest aimed to provide insight into the participants' initial proficiency levels across various writing components, including organization, content, coherence-cohesion, logical connection, and argumentation. Following the intervention, a post-test was administered to the same group of participants. The post-test evaluates the impact of the intervention by measuring any changes or improvements in the participants' writing skills. This design allows for a direct comparison within the same group, offering insights into the effectiveness of the intervention over time. The adoption of a single-group design was necessitated due to the constrained availability of a larger student populace, a limitation addressed to optimize the efficacy of addressing the research inquiries.

Participants

The study involved 30 participants selected through purposive sampling (Best & Khan, 2012), comprising 10 male and 20 female university students. These participants were enrolled in a provincial university situated in the Northeastern part of Thailand and were studying the subject "Approaches to Writing". The selection of these participants in writing compositions, particularly in L2, is motivated by a pressing need to address critical deficiencies in their writing abilities. These students exhibit challenges in content creation, logical coherence, unity, and the absence of argumentation in their written compositions. The challenges observed in content creation, logical coherence, unity, and the absence of argumentation in the students' written compositions were identified through a comprehensive analysis of their course assignments.

In addition, the students' age range was between 18 and 19 years, and collectively, they exhibited an average of more than ten years of exposure to English language instruction within the Thai national education system preceding their matriculation into the university. The participants were apprised of the study's objectives, the tasks incumbent upon them, the confidential nature of their responses, their prerogative to withdraw from participation at any juncture, and the provision for elucidating any queries regarding the research procedures (Best & Khan, 2012).

Data Collection

Writing Pre-Test

A writing pre-test was employed in this study, wherein students were given one hour to develop a writing topic consisting of a minimum of 200 words. The topic was thoughtfully chosen to align with the subjects covered in their writing class. Before initiating their drafts, students engaged in preparatory activities such as question posing and brainstorming, mirroring their regular writing class practices. Subsequently, two English lecturers from the study university evaluated the written compositions using a writing scoring rubric designed by the researcher (refer to the test marking). The scoring rubric covered five aspects: organization, content, coherence-cohesion, logical connection, and argumentation. To ensure the reliability of

the ratings, the two inter-raters independently assessed ten compositions, and the correlation between their scores was calculated. The obtained inter-reliability scores for the first and second inter-raters were .87 and .84, respectively, indicating a strong level of agreement between them (Creswell & Creswell, 2018).

Writing Post-Test

The writing post-test was conducted following the implementation of the intervention. Participants were provided with a different topics (parallel in difficulty level) and were instructed to develop a composition of at least 200 words. However, they were not specifically reminded about the strategies or instruments they should employ. Revising and editing checklists were made available on the teacher's desk, but it was up to the participants whether to utilize them or not. To assess the participants' writing performance, two experienced writing experts (the same raters who had assessed the participants' written compositions during the pre-writing test) evaluated the written compositions. They are English lecturers with over five years of teaching experience in EFL writing courses at the study-university.

Students' Writing Quizzes

The writing composition tasks completed by students during the intervention comprised four distinct assignments (parallel in difficulty level) utilized throughout the program. Each composition underwent assessment using a writing rubric, evaluating aspects such as organization, content, coherence-cohesion, logical connection, and argumentation. The topics covered a range aligned with their regular writing class, ensuring relevance and stimulating student interest. Topics were intentionally broad, fostering creativity and allowing for personal experiences while providing opportunities to practice writing skills. During the writing stage, students actively collaborated to gain insights, receive constructive feedback, and refine organizational thoughts. Utilizing the GPT-based chatbot, students sought support in generating and confirming ideas, ensuring grammatical and structural accuracy, making comparisons, and developing argumentative details. The teacher remained available to address questions or concerns. In the post-writing stage, allocated for revising and editing, students autonomously scrutinized compositions with GPT-based chatbot assistance. They were responsible for employing a comprehensive revision checklist to ensure coherence and completeness and resolve lexical and organizational challenges, both individually and collaboratively. In the editing phase, students utilized an editing checklist to identify and rectify minor errors, with GPT-based chatbot support for refining details and enhancing overall quality. The writing drafts for each quiz were assessed using the same scoring rubric as their pre-and post-writing tests.

To evaluate participants' compositions, a writing scoring rubric (see Appendix A) with criteria including organization, content, word choice, and language was employed. Designed by the researcher himself, this rubric underwent review by two English lecturers at the study-university for adjustments. Additionally, a descriptive checklist ensured standardized assessment. Each criterion had a maximum score of 20 points, with subsections assigned a maximum of four points each, established through consensus among evaluators.

Semi-Structured Interviews

Semi-structured interviews were conducted to obtain comprehensive insights and detailed information regarding the students' utilization of GPT-based chatbots within the process-based writing framework. The aim was to delve into their experiences and perceptions in facilitating various components of their writing tasks, including organization, content, coherence-cohesion, logical connection, and argumentation. These interviews were carried out after the completion of the post-writing test. 15 participants voluntarily participated in the interview. Each interview session lasted approximately 30 to 45 minutes, allowing for in-depth discussions while ensuring the participants' comfort and engagement. The interviews took place in a quiet and comfortable setting within the university premises, providing a conducive environment for open and honest conversations. The participants were assured of the confidentiality of their responses, encouraging them to share their experiences freely.

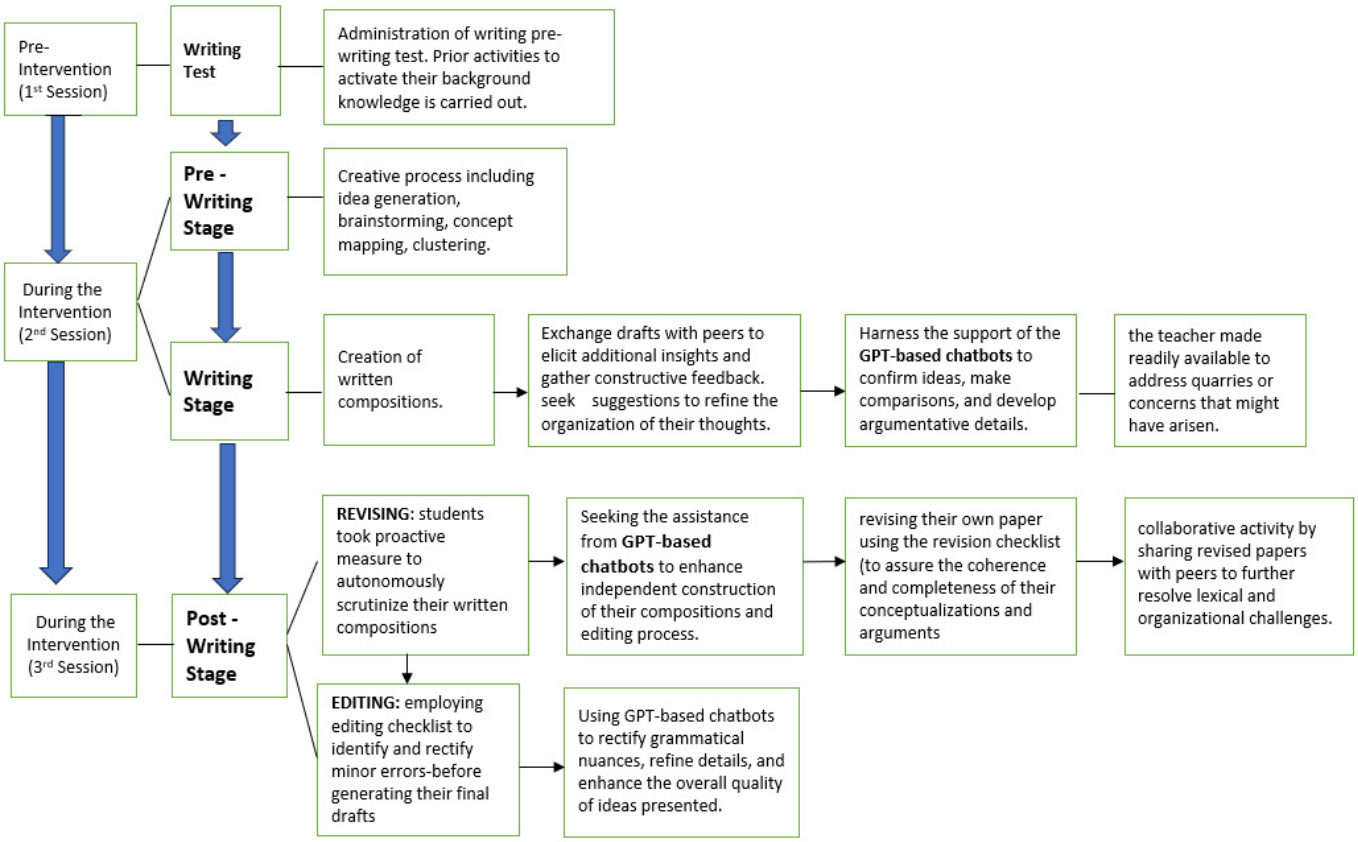
The Intervention Programme

An intervention program consisting of ten sessions was developed and implemented. Of these sessions, eight sessions were dedicated to the implementation of the process writing approach as the intervention, while one session each was allocated for the administration of the pre- and post-writing tests. The figure below presents the intervention programme implemented in the study via utilizing the GPT-based chatbot within the process-based writing framework that constitutes pre-writing, writing, and post-writing stages along with its learning activities.

In the subsequent three sessions (4^{th} – 9^{th} sessions), the students actively participated in additional composition writing exercises, delving into diverse topics (parallel in difficulty level). The primary aim of these sessions was to provide students with a comprehensive opportunity to apply the acquired strategy. This phase was strategically positioned before their involvement in the post-writing activity, slated for the 10^{th} session. Furthermore, during these composition sessions, students were encouraged to seamlessly integrate GPT-based chatbot into their writing process, leveraging its

Figure 2

The Intervention Programme



capabilities to refine and augment their compositions across varied subject matters. This integration served as an instrumental component in fostering not only writing proficiency but also in harnessing the potential of advanced linguistic support in the preparation for subsequent activities. Finally, after the intervention, (10th session), the students were tasked to compose an essay comprising at least 200 words, all completed within a 60-minute time constraint.

Data Analysis

In terms of quantitative data analysis, the study employed Descriptive Statistics, calculating and presenting measures such as frequency, mean, and Standard Deviation in a tabular format. Furthermore, a t-test statistical analysis was conducted to discern the variance between students' pre- and post-writing test results. Additionally, Analysis of Variance was applied to investigate the significant differences among quiz results. For qualitative data analysis, the responses gathered from semi-structured interviews were subjected to thorough examination. This process involved a meticulous approach known as topical coding, in line with the methodology outlined by Creswell and Creswell (2018). Texts were systematically transcribed and labeled, interpreted, and analyzed to pinpoint emerging themes within the interview responses, contributing to a comprehensive understanding of the qualitative aspects of the study.

RESULTS

Quantitative Analysis

Significant Difference between the Students' Writing Performance Before and After the Use of GPT–Based Chatbots within the Process-Based Writing Framework (RQ#1)

Overall test of difference on the participants' writing test performances

Table 1 presents the outcomes of the test aimed at comparing students' pre- and post- writing test results. The overall

mean and SD scores reveal a substantial disparity between the two sets of results. The pre-writing test performance (9.13; SD=2.619) was significantly lower than their post-writing test performance (17.03; SD=2.917). This marked difference is corroborated by a computed p-value of 0.001, which is below the significance threshold of p<0.05. This signifies that the implementation of the process writing approach had a noteworthy and positive impact on students' performance in their written assignments.

Students' Writing Quiz Performances Improve through the Integration of GPT-Based Chatbots within the Process-Based Writing Framework in Terms of Organization, Content, Coherence-Cohesion, Logical Connection, and Argumentation (RQ#2)

Table 2 displays the descriptive results for the four quizzes administered to the students. The initial quiz scores for the students commenced at a comparatively lower level. However, there was an observable improvement in the overall writing quiz scores as the intervention progressed. This improvement is reflected in the mean scores of $\bar{x}$ =12.75, $\bar{x}$ =13.86, $\bar{x}$ =14.19, and $\bar{x}$ =16.48 for Quiz 1, Quiz 2, Quiz 3, and Quiz 4, respectively. Additionally, the table reveals that four out of five writing components, namely "organization" ($\bar{x}$ =12.33; $\bar{x}$ =13.21; $\bar{x}$ =14.32; and $\bar{x}$ =16.14), "content" ($\bar{x}$ =13.14; $\bar{x}$ =14.75; $\bar{x}$ =15.96; and $\bar{x}$ =17.12), coherence-cohesion ($\bar{x}$ =12.41, $\bar{x}$ =13.89, $\bar{x}$ =15.18, and $\bar{x}$ =15.82), and "argumentation" ($\bar{x}$ =13.23, $\bar{x}$ =15.31, $\bar{x}$ =16.35, and $\bar{x}$=17.43), exhibited a gradual increase from the 1st to the 4th quizzes, respectively. One writing component (logical connection) experienced a decrease in scores during the 2nd ($\bar{x}$ =12.17) and 3rd($\bar{x}$ =12.14) quizzes but demonstrated a significant improvement in the 4th quiz ($\bar{x}$ =15.93). While the most improved writing components were "content" and "argumentation", the component related to "logical connection" showed the least improvement.

Table 3 presents the result for Repeated Measure Analysis of Variance (within subjects). It can be observed that the p-value under the *Sig.* column and sphericity assumed is less than 0.05, this indicates a significant difference among the scores in the four quizzes. The value of the ANOVA is indicated by the F column (F=97.115).

Table 4 presents the pairwise comparison of the means of the four quizzes. Here, the 1st quiz is compared to 2nd quiz, 3rd quiz, and 4th quiz. It was noticeable that the p-value was all less than 0.05 level of significance. Thus, scores in 1st quiz were significantly different to the scores in the other 3 quizzes. It was also indicated from the means found in the second table, the mean in 1st quiz is less than the mean of the other 3 quizzes, meaning, their score in 1st quiz was found significantly lower than their scores in the other quizzes. The *asterisks* from the mean scores found in the third column indicated a significant difference.

Qualitative Analysis

Learning Experiences Have the Students Obtained in Facilitating Their Writing Tasks through GPT-Based Chatbots within the Process-Based Writing Paradigm (RQ#3)

Table 1

Overall Test of Difference on the Participants' Pre- and Post- Writing Test Performances

Variables	Mean	S.D.	t-computed value	p-value
Pre-writing test performance	9.13	2.619	11.014	.001
Post-writing test performance	17.03	2.917		

Note: p<0.05

Table 2

Students' Quiz Performances

Writing Quizzes	Quiz 1		Quiz 2		Quiz 3		Quiz 4	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Organization	12.33	.56	13.21	.54	14.32	65	16.14	61
Content	13.14	.54	14.75	.67	15.96	59	17.12	.71
Coherence-cohesion	12.41	.51	13.89	.67	15.18	.78	15.82	.49
Logical Connection	12.64	.51	12.17	.67	12.14	.78	15.93	.49
Argumentation	13.23	.57	15.31	.72	16.35	.71	17.43	.76
Overall	12.75	0.53	13.86	0.67	14.19	0.71	16.48	0.63

Table 3

Repeated Measure ANOVA (Measure: MEASURE 1)

	Source	Type III sum of squares	df	Mean Square	F	Sig	Partial eta squared
Quiz	Sphericity Assumed	120.105	3	40.035	97.115	<.001	.781
	Greenhouse-Geisser	120.105	1.512	78.806	97.115	<.001	.781
	Huynh-Feldt	120.105	1.517	75.227	97.115	<.001	.781
	Lower Bound	120.105	1.010	120.10	97.115	<.001	.781
Error (Quiz)	Sphericity Assumed	33.391	81	.411			
	Greenhouse-Geisser	33.391	41.149	.810			
	Huynh-Feldt	33.391	43.102	.772			
	Lower Bound	33.391	27.010	1.229			

Note. p<0.05

Table 4

Pairwise Comparisons of the Means of the Four Quizzes. Measure: MEASURE 1

Quiz	(J) Quiz	Mean Difference (I-J)	Std Error	Sig b	95% Confidence Interval for Differenceb	
					Lower Bound	Upper Bound
1	2	-2.070*	.185	<.001	-2.597	-1.545
	3	-2.501*	.227	<.001	-3.146	-1.853
	4	-2.501*	.227	<.001	-3.146	-1.853
2	1	2.072*	.185	<.001	1.544	2.598
	3	-.428*	.140	.030	-.827	-.029
	4	-.428*	.140	.030	-.827	-.029
3	1	2.501*	.227	<.001	1.851	3.147
	2	.428*	.140	.030	.028	.828
	4	.000	.000	.000	.000	.000
4	1	2.501*	.227	<.001	1.851	3.147
	2	.428*	.140	.030	.028	.828
	3	.000	.000	.	.000	.000

Note. Based on estimated marginal means

* = the mean difference is significant at the .05 level

[b] = adjustment for multiple comparisons: Bonferroni

Theme 1: Benefits of using process-based approach on students' writing tasks

Established a strong foundational knowledge through planning activities. Prior knowledge activation is a powerful instructional approach, especially when it comes to the initial planning phase of writing. During this planning stage, students are encouraged to tap into their pre-existing mental frameworks and understanding of the subject matter (Robillos, 2021). By doing so, they establish a strong foundational knowledge base upon which to construct their written compositions. The participants emphasized the use of a specific technique known as clustering, which allowed them to explore a multitude of ideas before arranging them in a specific order. This approach helped students realize that certain words or concepts had numerous related aspects to write about. Over time, what initially appeared as random connections gradually transformed into logical patterns of thought. As one participant (P3) conveyed:

> "The clustering method enables me to examine numerous ideas before arranging them in a specific sequence." **P3**

Bolstered self-assurance through collaboration. One of the activities involved in the process-based writing approach is

collaboration activity. This activity, such as pair and group work, plays a vital role in bolstering students' self-assurance and fostering a supportive learning atmosphere (Wonglakorn & Deerajviset, 2023). Through collaboration, students can receive feedback, exchange viewpoints, and gain fresh insights, all of which contribute to a heightened confidence in expressing themselves through writing. The experiences shared by students underscore confidence and encouraging active participation in the writing process. By collaborating, students tap into their peers' collective knowledge and expertise, benefiting from constructive feedback and broadened perspectives. This collaborative approach not only enhances their writing skills but also cultivates a sense of support and motivation within the learning community. As expressed by P9:

> "Before I was hesitant to share my ideas, fearing they might be wrong, but as I shared, I realized that others also had similar thoughts, confirming my capabilities." **P9**

Empowered control of students' writing. Evaluating one's own language involves a meticulous examination of the text, scrutinizing its structure, grammar, vocabulary usage, and coherence. It necessitates students' attention to detail, identification of writing flaws or weaknesses, and consideration of ways to enhance clarity and effectiveness. This introspective practice empowers students to take control of their writing and actively participate in the process of revision and editing. The techniques of revising and editing offer students' specific strategies and guidelines for refining their writing. Drawing from P2's experiences, he/she shared his/her firsthand encounters with revising and editing techniques, indicating their value and benefits in enhancing her writing. As expressed by P2,

> "After learning techniques like revising and editing, my writing has become more meaningful and freer from grammatical errors." **P2**

Theme 2. Affordances of GPT-Based Chatbots in Facilitating Students' Writing Task/s

Writing practice. Interacting with the GPT chatbots involves writing in the target language. This provides writing practice, encouraging learners to compose questions, responses, and sentences. Regular engagement with GPT chatbots enhances writing skills, fostering fluent and accurate expression akin to real-life communication. Conversing with GPT chatbots demands critical thinking for coherent questions and responses, involving vocabulary selection, sentence structure, and grammar use. Consistent practice develops the ability to craft intricate sentences. By observing GPT chatbot's sentence construction and understanding responses, learners learn by example, gradually applying these patterns to their writing for improved fluency and quicker idea expression. As stated by P1:

> "The more I practice writing, the more I refine my skills in areas suchas sentence coherence, and overall writing structure." **P1**

Immediate feedback. GPT-based chatbots provide instant feedback on the correctness and clarity of learners' language use. Incorrect grammar, phrasing, or vocabulary can be highlighted by comparing the GPT-based chatbot's responses with the learner's input. If a learner's input contains grammatical errors or incorrect syntax, AI's response may highlight those issues. For example, if a learner writes *"I has a cat,"* the AI's response might respond with the corrected version: *"I have a cat."* This direct correction assists learners in understanding and internalizing proper grammar and sentence structure. As conveyed by P9:

> "The application improved my language skills by providing me immediate feedback on the errors on my text." **P9**

Reduced anxiety. Using a GPT-based chatbot offers a low-pressure language practice setting. Unlike interactions with native speakers, there's no immediate social judgment or pressure for perfection. Learners can freely explore, learn, and make mistakes without anxiety. This allows them to focus on improvement instead of worrying about errors. As learners practice and realize effective communication with AI, their language confidence increases. P4 expressed that:

> "I can experiment with different ways of expressing my ideas since I don't feel any pressure of a live conversation" **P4**

Theme 3: Challenge/s of Using the Strategy

Time constraints. The process writing approach places significant emphasis on engaging in each phase of the writing process, which typically includes prewriting, writing, post-writing (including revising and editing). However, within the context of this study, there arose a notable challenge for students due to time constraints that hindered their ability to complete all these stages thoroughly and within the allocated timeframe. This constraint had a tangible impact on their capacity to fully immerse themselves in each phase and deliver their written work on schedule. P6's account offers a glimpse into how these time limitations affected the students. She felt the pressure of attempting to navigate through all the writing stages within the specified timeframe. This observation underscores that the restricted time available hindered her and potentially others from fully embracing and implementing the process writing approach as it was originally intended. As articulated by P6:

> "Strategies such as planning and revising were beneficial but the time required to complete these required more time to finalize our drafts." **P6**

Misunderstandings and inaccuracies. While GPT-based models can produce coherent and contextually relevant responses, they are not infallible. That is, the AI-generated responses might contain inaccuracies or be contextually

inappropriate. This is because the model's understanding is based on patterns rather than true comprehension, and it might not always grasp the nuances of language, cultural references, or complex ideas. When learners encounter inaccurate or misunderstood responses, it can lead to confusion and misinterpretation of information. As P3 narrated:

> "Sometimes the apps gave me inaccurate details, especially in translating the words to English, which confused me and caused me misunderstanding." **P3**

Overreliance on the AI. While GPT-based chat applications offer instant answers and suggestions, some learners might start depending excessively on the AI for generating content. This overreliance can discourage the development of critical thinking skills and independent language production. Learners might prioritize convenience over engagement, missing out on the cognitive effort required to think through and formulate their own ideas. P8 expressed that:

> "I felt overly dependent on the application and this has been hindering my own way of formulating my own ideas." **P8**

Lack of personalized feedback. GPT-based models lack the capacity for personalized, detailed feedback customized to learners' individual strengths and weaknesses. In contrast, human teachers can identify areas needing improvement and provide tailored guidance. Personalized feedback is crucial for learners to comprehend progress, tackle challenges, and hone language skills effectively. The absence of this personalized touch could impede targeted improvement and hinder language proficiency advancement. P1 affirmed this sentiment.

> "Sometimes the response of the AI did not offer me detailed and personalized feedback according to my intellectual level like my teacher would give." **P1**

Theme 4: Enhanced Critical Thinking Skills

Reflective learning. Metacognition encourages students to reflect on their learning process, set goals, and evaluate their progress. By using GPT-based chatbots, students can engage in conversations and then review and analyse their interactions. This reflection helps them identify areas of strength and weakness, allowing for more targeted practice and improvement. P12 narrated that:

> " I struggled with using correct prepositions, however, I started to set a goal to focus on preposition usage next time, aiming to improve this aspect." **P12**

Reflective learning, part of metacognition, prompts learners to assess what they've learned, how they learned it, and how to enhance their learning strategies. Engaging with GPT-based chatbots mirrors real conversations. Learners can review these interactions, including their input and AI's responses, a vital metacognitive aspect. This reflection lets students recognize strengths and weaknesses, boosting motivation and confidence. P4 noted:

> "Using the applications enabled me to reflect on my language interactions which inspired me to finish my task." **P4**

Monitoring comprehension. Using GPT-based chat applications, students must continuously monitor their own comprehension of the conversation. This encourages them to pause, reflect, and ensure they are understanding the AI's responses correctly. This skill transfers to real conversations, where monitoring comprehension is essential. Monitoring comprehension is a metacognitive skill that involves assessing one's understanding of the information presented (Robillos & Bustos, 2022). When students engage in conversations with GPT-based chat applications, they actively monitor their comprehension of the AI's responses to ensure they are understanding the content correctly. If learners encounter a response from the AI that they do not fully understand, they can ask for clarification or confirmation. P3 mentioned that:

> "The intervention enabled us to compare our input with the AI's responses which could help us assess whether theintended meaning was effectively conveyed." **P3**

Error awareness and correction. When students engage with the AI, they actively look for errors in their language use and use the AI's responses to self-correct. This process trains students to be more vigilant about their language skills and to proactively correct mistakes. Metacognition underscores the importance of self-awareness and improvement, encouraging learners to actively monitor their language production. During interactions with the AI, learners develop a habit of actively seeking errors in their own language production. This heightened awareness prompts them to analyze their sentences and phrases more critically, leading to improved accuracy. P10 conveyed that:

> "After recognizing such pattern through interactions with the AI, I became more aware about using the correct preposition in real conversations." **P10**

DISCUSSION

Students' Writing Performances Before and After the Implementation of GPT-Based Chatbots within the Process-Based Writing Framework

The results presented in table 1 reveal a significant improvement in students' writing performance using GPT-based chatbots within the process-based writing framework. Prior to the intervention, the students had a mean score of $\bar{x}$=9.13, indicating a relatively lower level of performance. However, after the intervention, their average mean score increased substantially to $\bar{x}$=17.03. Results of the study underscore

a noteworthy improvement in participants' writing proficiency attributed to an innovative intervention that utilizes GPT-based chatbots within a process-based writing framework. This comprehensive method incorporates key phases of the process-based approach, including prewriting, writing, and post-writing, empowering students to approach their writing assignments strategically. The interactive GPT-based chatbots, coupled with the process-based approach, effectively fosters student engagement in composing their written pieces. Moreover, the application's provision of real-time feedback and clarifications contributes to the proactive regulation of understanding, ensuring that students receive immediate guidance during their writing endeavours (Sinha et al., 2020). This feature proves particularly beneficial during the pre-writing and writing stages, where students engage in planning and writing activities. The interactive nature of the chat application allows students to refine their ideas based on an evolving grasp of the content, promoting a dynamic and iterative approach to the writing process. The post-writing stage assumes a pivotal role in guiding students toward a deeper comprehension of their focal points and the arguments to be included in their drafts. This phase serves as a reflective platform, encouraging students to critically assess and refine their work. The evaluative process not only promotes self-directed learning but also equips students with tools to enhance their strategies in subsequent writing tasks, fostering continuous improvement. The result is in congruence with Wuttiphan's (2023) and Robillos and Namwong's (2021) studies reporting that utilizing process-based approach helped students facilitate their writing tasks and thus effectively improve their writing performance.

Students' Writing Quiz Performances Across Various Writing Aspects

It is noteworthy to emphasize that the students' writing quiz performances showed an improvement in terms of writing aspects such as organization, content, coherence-cohesion, logical connection, and argumentation. Integrating GPT-based chatbot within the process-based writing framework offers students a valuable tool for enhancing organizational skills. Through interactive discussions, students can articulate and refine their ideas, facilitating the identification of logical connections and the creation of a well-structured framework for their writing. Real-time conversations also allow for collaborative development of outlines, providing a systematic approach to organizing thoughts and ensuring a coherent structure in their written compositions (Wonglakorn & Deerajviset, 2023). Regarding 'content', the intervention enabled students to access a wealth of information in real-time, supporting them in gathering relevant and diverse content for their compositions (Fitria, 2023). The interactive nature of the chat application exposes students to a variety of perspectives, fostering a comprehensive understanding of their chosen topics. This exposure to diverse viewpoints enriches the depth of content in their writing, making their compositions more robust and well-informed.

Moreover, engaging in coherent discussions within the GPT-based chatbots (Zoherey, 2023) allow students to practice expressing ideas in a structured manner. This skill is directly transferable to their written work, where they learn to maintain logical connections between sentences and paragraphs. Shibani et al. (2017) highlighted that by capitalizing on massive pre-training data, GPT-based chatbots can comprehend the context of learners' responses and deliver appropriate language assistance, making language learning more dynamic and engaging. Real-time feedback from the chat application helps highlight inconsistencies or gaps in students' discussions, prompting them to address and enhance coherence in their written compositions.

Lastly, the application becomes a dynamic platform for students to develop and refine their argumentation skills within the process-based writing framework (Kitajroonchai et al., 2022). Through interactive debates and defending ideas in real-time, students can hone their abilities to construct compelling arguments. This practice in the virtual setting translates into more persuasive and well-supported arguments in their written compositions. Su et al. (2023) emphasized that the application's real-time feedback provides constructive insights into the strength of students' arguments, encouraging them to refine and strengthen their persuasive techniques for effective written expression. The aforementioned statements are in congruence with the responses of the participants when they were asked about the benefits they obtained from the intervention. A participant stated that the more they practice writing, the more they refine their writing skills such as content, organization, coherence, argumentation, and overall writing structure.

Conversely, the observed minimal improvement in the aspect of "logical connection" when utilizing a GPT-based chat application within a process-based writing framework could stem from a combination of factors, including the technology's primary focus on grammar and structure, potential limitations in guiding complex idea progression, and a potential misalignment between the AI's capabilities and the nuanced requirements of content coherence (Fitria, 2023; Su et al., 2023). The emphasis on grammar and the process-based writing processes might have overshadowed logical organization, and the AI's lack of explicit guidance on transitions and coherent content arrangement could have contributed to the comparatively lesser enhancement in this aspect of writing. In certain situations, the AI-generated responses might contain inaccuracies or be contextually inappropriate. When learners encounter inaccurate or misunderstood responses, it can lead to confusion and misinterpretation of information leading to illogical organization.

Learners' Learning Experiences After the Implementation of GPT-Based Chatbots within the Process-Based Writing Famework

The qualitative analysis of the study delves into the students' learning experiences (particularly in their writing skill) with the GPT-based chatbots within the process-based writing framework, were significantly enhanced. Firstly, the process-based approach provided a structured framework, allowing students to build a strong foundation through planning, drafting, and revising. Activities such as peer reviews and iterative revisions fostered collaboration and self-directed learning, leading to improved writing skills and greater confidence. The emphasis of the writing process methodology on multiple drafts and reflective practices enabled students to refine their work systematically, aligning with research that underscores the value of such methodologies (Wonglakorn & Deerajviset, 2023; Robillos & Namwong, 2021). Secondly, the GPT-based chatbots offered immediate, personalized feedback, helping students address writing issues in real-time and reducing anxiety by providing a supportive, non-judgmental space for practice. This use of advanced NLP tools facilitated continuous improvement and enhanced students' writing proficiency, reflecting findings on the benefits of technology in language education (Moqbel & Al-Kadi, 2023; Su et al., 2023; Okonkwo & Ade-Ibijola, 2021).

However, integrating GPT-based chatbots also introduced challenges, such as time constraints, inaccuracies in feedback, and the risk of overreliance on AI tools. These issues sometimes impeded the effectiveness of the writing process, emphasizing the need for a more nuanced approach to integrating technology with traditional methodologies (Zhai, 2022). Despite these challenges, the intervention effectively enhanced students' critical thinking skills, such as reflective learning and error correction, aligning with the literature on the benefits of process-based approach (Wonglakorn & Deerajviset, 2023). This highlights the importance of further research into optimizing the integration of GPT-based chatbots with established writing frameworks to address emerging challenges and improve educational practices in the digital age.

CONCLUSION

This study illuminates the transformative impact of integrating GPT-based chatbots within a process-based writing framework on learners' writing skills. The noteworthy outcomes underscore the considerable potential of this intervention in enriching language learning experiences. Participants exhibited remarkable improvements across various writing components, indicating the effectiveness of GPT-powered interactions. The consistent practice and exposure to diverse language patterns facilitated by the GPT-based chat applications contributed significantly to heightened language proficiency. Furthermore, the integration of a process-based writing approach empowered learners to self-regulate and engage in reflective practices, fostering a more effective language acquisition process. The positive reception of GPT-based chat applications by participants emphasizes their adaptability and openness to embracing technological innovations in their language learning journey. The user-friendly interface, coupled with instant feedback and accessibility, created an immersive language practice environment that resonated well with participants.

Importantly, the findings not only highlight the positive impact on writing skills but also underscore the broader implications for self-regulation, reflective learning, and the integration of technology in language education. As educators and practitioners consider innovative approaches, this study provides a compelling case for the effective utilization of GPT-based chatbots in fostering language proficiency and a more engaging learning experience. Future researchers might gain valuable insights into the potential of incorporating GPT-based chatbots within a pedagogical framework to enhance language learning outcomes.

Despite the promising outcomes, it is essential to acknowledge certain limitations warranting consideration. Firstly, the study's scope was confined to a specific context with a relatively small sample size, potentially limiting the generalizability of the findings. To address this, future research should aim to broaden the scope by incorporating diverse contexts and populations. Additionally, increasing the sample size would enhance the robustness and representativeness of the analysis. Secondly, the present study identified a less improved writing component, specifically in "logical connection." To address this in future interventions, a multifaceted approach is recommended. Begin by developing targeted instructional modules that explicitly address the intricacies of logical connections, emphasizing the use of transitional phrases and cohesive devices. Tailor the intervention to cater to individualized learning needs, recognizing varying proficiency levels among students. To strengthen the feedback mechanism of the GPT-based chatbot, provide detailed and constructive feedback on logical connections, facilitating a clearer understanding for students. Integrate explicit instruction on logical connections within the broader framework of the process-based approach, highlighting the importance of coherence in writing. Furthermore, implement a longitudinal assessment approach to track students' progression in mastering logical connections over time. A mixed-methods research design, combining quantitative measures with qualitative insights, can provide a comprehensive understanding of the factors influencing improvement. Lastly, maintain a flexible approach, continuously evaluating and adapting the intervention based on ongoing feedback and assessment results to effectively address emerging needs and challenges. This proactive strategy ensures the intervention remains responsive and impactful within evolving educational landscapes.

DECLARATION OF CONFLICTING INTERESTS

None declared.

REFERENCES

Bayat, N. (2014). The effect of the process writing approach on writing success and anxiety. *Educational Sciences: Theory & Practice, 14*(3), 1133-1141. https://doi.org/10.12738/estp.2014.3.1720

Best, J.W., & Khan, J.V. (2012). *Research in education*. Pearson/Allyn and Bacon.

Coffin, C., Curry, M., Goodman, S., Herwings, A., Lillis, M., & Swann, I. (2003). *Teaching academic writing*. Routledge. https://doi.org/10.4324/9780203994894

Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed method approaches*. SAGE Publications.

Diliduzgun, S. (2013). The effect of process writing activities on the writing skills of prospective Turkish teachers. *Eurasian Journal of Educational Research, 52*, 189-210.

Evmenova, A. S., & Regan, K. (2019). Supporting the writing process with technology for students with disabilities. *Intervention in School and Clinic*, *55*(2), 78-85. https://doi.org/10.1177/1053451219837636

Fitria, T. N. (2023). Artificial intelligence (AI) technology ino Open AI ChatGPT application: A review of chatGPT in writing English essay. In *ELT Forum: Journal of English Language Teaching*, *12*(1), 44-58. https://doi.org/10.15294/elt.v12i1.64069

Karatay, H. (2011). Process-based writing models: Planning, writing, and assessment. In M. Ozbay (Ed.), *Within writing i nstruction* (p. 21-43). Pegem Academy.

Kanlapan, T. C. E., & Velasco, J. C. (2009). Constructing a self-regulation scale contextualized in writing. *TESOL Journal, 1*(1), 79-94.

Kitajroonchai, N., Kitjaroonchai, T., & Sanitchai, P. (2022). The effects of process genre-based writing and process writing approaches on Asian EFL pre-university students' writing performance. *Journal of Language Teaching and Research, 13*(4), 860-871. https://doi.org/10.17507/jltr.1304.19

Lapan, R. T., Kardash, C. M., & Turner, S. (2002). Empowering students to become self-regulated learners. *Professional School Counselling, 5*(4), 257.

Li, Z., Dursun, A., & Hegelheimer, V. (2017). Technology and L2 writing. *The handbook of technology and second language teaching and learning* (pp. 77-92). Wiley-Blackwell.

Malmir, A., & Khosravi, F. (2018). The effect of argument mapping instruction on L2 writing achievement across writing task and writing components: A case study of Iranian EFL Learners. *Applied Research on English Language, 7*(4), 514-538. https://doi.org/10.22108/ARE.2018.111870.1318

Moqbel, M. S. S., & Al-Kadi, A. M. T. (2023). Foreign language learning assessment in the age of ChatGPT: A theoretical account. *Journal of English Studies in Arabia Felix*, *2*(1), 71-84. https://doi.org/10.56540/jesaf.v2i1.62

Morphy, P., & Graham, S. (2012). Word processing programs and weaker writers/readers: A meta-analysis of research findings. *Reading and Writing, 25*, 641–678. https://doi.org/10.1007/s11145-010-9292-5

Okonkwo, C. W., & Ade-Ibijola, A. (2021). Chatbots applications in education: A systematic review. *Computers and Education: Artificial Intelligence*, *2*, 100033. https://doi.org/10.1016/j.caeai.2021.100033

Onozawa, C. (2010). A study of the process writing approach: A suggestion for an eclectic writing approach. In *Proceedings of Kyoai Gakuen Maebashi International University* (vol. 10, pp. 153-163). Kyoai Gakuen Maebashi International University.

Pahlavani, P., & Maftoon, P. (2015). The impact of using computer-aided argument mapping (CAAM) on the improvement of Iranian EFL learners' writing self-regulation. *The Journal of Teaching Language Skills, 7*(2), 127-152. https://doi.org/10.22099/JTLS.2015.3528

Phillips, T., Saleh, A., Glazewski, K. D., Hmelo-Silver, C. E., Mott, B., & Lester, J. C. (2022). Exploring the use of GPT-3 as a tool for evaluating text-based collaborative discourse. In *Companion Proceedings of the 12th International Conference on Learning Analytics & Knwoledge* (pp.1-3). SOLAR.

Robillos, R. J. (2022). Impact of LoilooNote digital mapping on university students' oral presentation skills and critical thinking dispositions. *International Journal of Instruction, 15*(2), 501- 518. https://doi.org/10.29333/iji.2022.15228a

Robillos, R. J. & Namwong, O. (2021). Thai tertiary learners' composition writing performance and self-regulation towards EFL writing using process-oriented approach. *TESOL International Journal, 16*(7), 87-103.

Robillos, R. J., & Bustos, I. G. (2022). Learners' listening skill and metacognitive awareness through metacognitive strategy instruction with pedagogical cycle. *International Journal of Instruction, 15*(3), 393-412. https://doi.org/10.29333/iji.2022.15322a

Robillos, R. J. (2023). The impact of the FlipGrid application within the genre-based framework on students' writing skills and self-regulation of learning awareness. *Studies in Self-Access Learning Journal, 14*(4), 456–475. https://doi.org/10.37237/140404

Robillos, R. J. (2021). Learners' writing skill and self-regulation of learning awareness using computer-assisted argument mapping (CAAM). *Teaching English with Technology, 21*(4), 76-93.

Tanmongkol, N., Moonpim, R., Vimonvattaraveete, S., Suteerapornchai, T., & Kaniyoa, W. (2020). The main reason that Thailand's high school students are not adapting in the English language. *International Journal of Research and Review, 7*(6), 247-253.

Shibani, A., Koh, E., Lai, V., & Shim, K. J. (2017). Assessing the language of chat for teamwork dialogue. *Journal of Educational Technology & Society*, *20*(2), 224-237.

Sinha, S., Basak, S., Dey, Y., & Mondal, A. (2020). An educational chatbot for answering queries. In *Emerging Technology in Modelling and Graphics: Proceedings of IEM Graph 2018* (pp. 55-60). Springer Singapore.

Su, Y., Lin, Y., & Lai, C. (2023). Collaborating with ChatGPT in argumentative writing. *Assessing Writing, 57*, 100752. https://doi.org/10.1016/j.asw.2023.100752

Wang, C., Kim, D.H., Bong, M., & Ahn, H.S. (2013). Korean college students' self-regulated learning strategies and self-efficacy in learning English as a second language. *Asian EFL Journal, 15*(3), 81-112.

White, R. & Arndt, V. (1991). *Process writing*. Longman.

Wonglakorn, P., & Deerajviset, P. (2023). The effects of collaborative process writing approach on Thai EFL secondary school students' writing skills. *LEARN Journal: Language Education and Acquisition Research Network, 16*(1), 495-522.

Wuttiphan, N. (2023). Process-based writing approach: A panacea to improve Thai Chinese as a Foreign Language (TCFL) learners' writing skill and to eradicate writing block. *EDKKU Journal, 46*(1), 20-41.

Zhai, X. (2022). ChatGPT user experience: Implications for education. *Available at SSRN 4312418*. http://doi.org/10.2139/ssrn.4312418

Zhang, M. (2021). Understanding L1 and L2 interaction in collaborative writing: A lexico-grammatical analysis. *Language Teaching Research, 25*(3), 338-359. https://doi.org/10.1177/1362168819859911

Zoherey, M. (2023). ChatGPT in academic writing and publishing: A comprehensive guide. In *Artificial Intelligence in Academia, Research and Science: ChatGPT as a case Study* (1st ed). Achtago Publishing. https://doi.org/10.5281/zenodo.7803703

APPENDIX A

WRITING RUBRIC FOR QUIZZES

Components	4 marks	3 marks	2 marks	1 mark
Organization	The writing presents a clear and effective structure that enhances the overall flow and readability. Each section and paragraph seamlessly transition to the next, contributing to a strong sense of unity and purpose.	The organization is solid, with a discernible structure that supports the overall message. While some improvements could enhance the flow, the reader can easily follow the logical progression of ideas.	The organization is somewhat inconsistent, and improvements are needed to create a smoother flow. The structure may be confusing at times, impacting the overall coherence of the writing.	The writing lacks a clear organizational structure, making it difficult for the reader to follow the intended message. Significant revisions are necessary to improve overall organization.
Content	The content demonstrates a deep understanding of the topic. Relevant details are effectively incorporated, contributing to a comprehensive and engaging narrative.	The content provides a clear and sufficient exploration of the topic. While some additional depth or elaboration could enhance the overall quality, the key points are generally well-addressed.	The content is somewhat lacking, and there are notable gaps in coverage. The writer needs to provide more relevant details or information to fully address the topic.	The content is insufficient, with significant gaps in information. The writer must substantially expand and improve the content to adequately address the topic.
Coherence-Cohesion	The writing exhibits exceptional coherence and cohesion. Sentences and paragraphs are skillfully connected, creating a seamless and fluid progression of ideas. The reader can easily follow the writer's line of thought.	The coherence and cohesion are generally strong, with effective transitions between sentences and paragraphs. While there may be a few areas that could benefit from improvement, overall, the writing maintains a good flow.	The coherence and cohesion could be stronger. Some sentences or paragraphs may feel disjointed or disconnected, requiring attention to improve the overall flow of the writing.	The writing lacks coherence and cohesion, making it challenging for the reader to follow the writer's intended message. Substantial revisions are needed to create a more connected and fluid piece.
Logical Connection	The logical connections between ideas create a compelling and persuasive argument. The writer effectively builds upon each point, leading to a strong and convincing conclusion.	The logical connections are solid, with a clear progression of ideas that contributes to a persuasive argument. While some areas may benefit from additional support or development, the overall structure is effective.	The logical connections are somewhat weak, and the argument may lack sufficient support in certain areas. The writer should strengthen the logical flow to enhance the overall persuasiveness.	The logical connections between ideas are unclear or absent, resulting in a weak and unconvincing argument. Significant revisions are necessary to establish a more coherent and compelling line of reasoning.
Argumentation	The argument presents a compelling and persuasive case. It effectively anticipates and addresses counterarguments, providing thorough and convincing support for the writer's perspective. The reasoning is clear, and the overall argument is highly compelling.	The argument is solid, presenting a clear and well-supported case. While there may be room for additional depth or exploration of certain points, the writer effectively supports their perspective and addresses key counterarguments.	The argument is somewhat underdeveloped, lacking in-depth analysis or support in certain areas. The writer should provide more thorough reasoning and evidence to strengthen the overall persuasiveness of the argument.	The argument is weak, with insufficient support and limited analysis. The writer needs to substantially improve the depth and coherence of the argument to make it more persuasive and convincing.

https://doi.org/10.17323/jle.2024.18411

Academic Vocabulary Distribution in Applied Linguistics Journal Research Articles: Do SINTA Rankings Matter?

Suhandoko [1], Dian Riesti Ningrum [2], Andini Dwi Wardani [3], Ach. Nobair [1], Putroue Keumala Intan [1]

[1] UIN Sunan Ampel, Surabaya, Indonesia
[2] Universitas Airlangga, Surabaya, Indonesia
[3] Universitas Negeri Surabaya, Indonesia

ABSTRACT

Background: As a national database for indexing scientific journals, SINTA has considerable significance for the Indonesian academic community as it measures the performance of national journals and increases the visibility of Indonesian journals and researchers internationally. Although studies have been conducted to examine the role of academic vocabulary in scholarly publications, very little has investigated how academic vocabulary has been used in SINTA-indexed applied linguistics journals (SIALJ) research articles and whether there are differences in academic vocabulary coverage across SINTA rankings.

Purpose: This study examines the academic vocabulary measure of whether significant differences in academic vocabulary coverage are present in SIALJ research articles across rankings. This examination will offer insights into the linguistic expectations set by the editorial boards of the journals across rankings.

Method: Out of 8585 journals indexed by SINTA, we found 72 related to applied linguistics. We chose four journals with the highest impact factor in each ranking to ensure representativeness. We included approximately 250000 running words from each journal in each ranking and obtained 6073379 tokens in total. We used AntWordProfiler to analyse the lexical distribution with GSL and AWL as the base lists.

Results: We found that the academic vocabulary coverage in SIALJ research articles accounts for 11.01%, similar to other studies that also found that academic words typically cover at least 10% of academic texts. We also identified that the higher the journal rank, the more coverage of the academic vocabulary. However, our quantitative measurement identified no significant differences in academic vocabulary coverage in SIALJ research articles.

Conclusion: The absence of significant distribution disparities across rankings suggests a shared practice of strategies language use in SIALJ, irrespective of their rankings and challenges common assumptions about strategic language use discrepancies among journal clusters.

KEYWORDS

academic vocabulary, applied linguistics, journals, SINTA, coverage, rankings

Citation: Suhandoko, Ningrum, D.R., Wardani, A.D., Nobair, A., & Intan, P.K. (2024). Academic vocabulary distribution in applied iinguistics journal research Aaticles: Do SINTA rankings matter? *Journal of Language and Education, 10*(3), 95-107. https://doi.org/10.17323/jle.2024.18411

Correspondence:
Suhandoko
E-mail: suhandoko@uinsa.ac.id

Received: November 24, 2023
Accepted: May 16, 2024
Published: September 30, 2024

INTRODUCTION

The proliferation of open-access journals in recent years from various professional organizations, universities, independent institutions, and even bogus scientific entities has emerged extensive criticism about the publication quality. Bibliometric indices have emerged as a promising tool for identifying ethical violations in publishing practices (Gureyev & Mazov, 2022). These indices evaluate the publication status, the expertise of its authors, and the quality of their work by analysing citation frequency in the same field (Roldan-Valadez et al., 2019). While producing impactful research through scientific writing requires laborious work, it is widely acknowledged that only 10% of published work has a chance of being cited (Weinstein & Morgan, 2007). This issue is exacerbated by the language barrier, where publications written not in English often receive significantly fewer citations

(Arenas-Castro et al., 2024). Bibliometric indices allow researchers to determine the most appropriate journals for their work and monitor publication trends through self-assessment. Thus, contributing to reputable and high-quality journals can enhance the chances of their work being read and referenced (Castillo-Vergara et al., 2018; Donthu et al., 2020; Rey-Martí et al., 2016).

Several factors contribute to the assessment of a scientific journal's quality, including the rigor of the peer-review process, the reputation of the editorial board and reviewers, and the journal's impact factor. Equally important is the quality of language used to convey ideas, as effective communication can significantly influence the academic community. The use of academic vocabulary in scientific writing enhances the accuracy of communication and improves the comprehension of research findings (Awagu, 2021; Choo et al., 2017; Hyland, 2002). However, authors often prioritize linguistic simplicity over accuracy and precision (Aldawsari, 2017; Biber & Gray, 2016), which can compromise the content's accuracy (Breeze, 2008; Robbins, 2016). Therefore, writers need to strike a balance between simplicity and accuracy by carefully selecting appropriate vocabulary (Demir, 2019).

Countless studies have explored the significance of academic vocabulary in scientific writing, revealing that a good command of discipline-specific vocabulary indicates effective communication (Brun-Mercer & Zimmerman, 2015; Choo et al., 2017; Coxhead, 2012). As many scholars suggest, this practice promotes long-term academic performance (Csomay & Prades, 2018; Masrai et al., 2021). Furthermore, academic vocabulary has been found to be positively associated with journal quality (El-Omar, 2014), research impact (Pournia, 2019), and authors' credibility in the discourse practices of their scientific community (Matinparsa et al., 2022; Xodabande et al., 2022). Careful word choice establishes the author's expertise and credibility in the field, which is critical for building readers' trust and validating the argument presented (Hyland, 2013).

While previous literature has acknowledged the importance of academic vocabulary in scientific writing, there has been limited research on how academic vocabulary is used in journal articles, particularly in the context of the SINTA-indexed applied linguistics journal (SIALJ). This noticeable research gap presents us with an opportunity to examine how academic vocabulary is being used in these journals and whether there are significant differences in academic vocabulary coverage across rankings. Based on this background, the study addresses the following research questions: (1) What is the academic vocabulary coverage in SIALJ research articles? (2) How does the distribution of academic vocabulary in SIALJ research articles vary across rankings?

LITERATURE REVIEW

Open Access Journals and Bibliometric Indices

Numerous bibliometric indices are now accessible online, with both free and paid access options. Some prominent publication databases that employ accessible metrics include Scopus, Web of Science (WoS), PubMed, DOAJ, and SINTA (Science and Technology Index). These journal indexing portals provide comprehensive databases that offer bibliographic information, publication frequency, and impact factor of the journals. Although scholars have doubted the effectiveness of these measurements in determining the quality of publications, they admit that scholars often dream of having their works published in reputable and indexed journals as they tend to have a larger readership, as demonstrated by their bibliometric measurement (Garner et al., 2018; Kandi, 2016; Koushik, 2017). Beyond that, these portals play a significant role in helping researchers identify and evaluate academic literature. The significance of the portals is even more highlighted by their use in bibliometric analysis, which assesses how impactful the researchers and their works are in scientific development (Xiao et al., 2022).

In the context of Indonesia, SINTA is considered the primary national journal indexing portal that evaluates the quality of academic journals using various metrics, such as citations and h-index. In addition to assessing the quality of national journal's performance, SINTA also serves to increase the visibility of Indonesian journals and researchers in the global arena (Firmansyah & Faisal, 2019; Nandiyanto et al., 2020; Purnomo et al., 2020) as it provides academic branding to both researchers and institutions; thus, it increases their reputation and recognition in the global academic world (Ibrahim & Fadhli, 2021; Muslimin & Basthomi, 2022; Rahardja et al., 2019). Studies have reported that SINTA holds equal significance with more globally acknowledged indexing portals, like Scopus or WoS, as evidentially shown by SINTA-indexed journals' adherence to the high standard of excellence (Tamela, 2020; Wijaya & Bram, 2022; Yadira et al., 2022). Therefore, we can conclude that as a platform for disseminating knowledge and encouraging scholarly contributions, SINTA serves as a catalyst for promoting the recognition of Indonesian academia in the international arena.

Academic Vocabulary in Scientific Writing

Several factors determine the quality of a scientific journal, including the rigor of its peer-review process, the reputation of its editorial board members, the publication frequency, and the impact factor. The peer-review process is crucial to ensuring that the published articles have met high-quality standards (Wicherts, 2016). The reputation and expertise of the journal's editorial board are also important to ensuring

that the article content remains up-to-date and relevant to the journal's scope (Black et al., 1998). The publication frequency indicates the journal's quality by publishing innovative and novel research. Finally, the impact factor indicates that the published articles have a large readership, as evidenced by their citations (Koushik, 2017). However, it is important to not merely see the impact factor of a journal as the only measure of its quality. Assessing the quality of a publication should also consider how the article has conformed with the common practice, especially with the use of discipline-specific vocabulary known as academic vocabulary.

Accurate and appropriate use of academic vocabulary is a key instrument to facilitate researchers' accurate communication of their findings and ideas to their academic colleagues (Choo et al., 2017; El-Omar, 2014). Moreover, the skilful use of academic vocabulary establishes the writer's credibility in their field and increases the accessibility of their works (Awagu, 2021). The use of academic vocabulary ensures that research results are easily understandable and comparable to those of other researchers in the same field. Moreover, Asaad (2024) argues that mastering academic vocabulary, including complex and low-frequency words, is essential for producing high-quality academic writing and is strongly linked to proficient writing skills. Therefore, mastery of academic vocabulary is essential for achieving effective scientific writing and successful research and knowledge dissemination.

Academic writing is renowned for its use of discipline-specific language or academic vocabulary, which is essential in conveying complex ideas and theories. Despite its significance, the use of academic vocabulary in scholarly publications is often undervalued. Some perceive it as too complicated for common readers, leading to a tendency for authors to prioritize linguistic simplicity over precision (Aldawsari, 2017; Biber & Gray, 2016). However, such practice can compromise content accuracy and lead to a loss of credibility in scientific communication (Breeze, 2008; El-Omar, 2014; Robbins, 2016). Thus, striking a balance between technical accuracy and linguistic clarity is principal in academic writing, and writers must carefully measure their use of academic vocabulary to ensure accuracy while maintaining comprehensibility (Arianto & Basthomi, 2021; Demir, 2019; Hyland, 2009). While it may be tempting to simplify language to reach a wider audience, it is essential to maintain the specificity and precision required for scientific communication by carefully selecting the most appropriate and accurate academic vocabulary to convey their ideas effectively (Hinkel, 2003).

The crucial role of academic vocabulary in scientific discourse has prompted Coxhead to develop the Academic Word List (AWL) to standardize academic terminology (Coxhead, 2000). The study demonstrated that 570-word families comprised 10% of all academic text words. In the context of EAP, teachers can enhance the proficiency of learners in scientific communication by prioritizing this list. However, the AWL has limitations as it mainly covers written academic language and may not incorporate all relevant spoken academic vocabulary or apply to specific fields. In response to this issue, various registers and domains have been scrutinized for lexical distribution, including medical science (Chen & Ge, 2007), agriculture (Martínez et al., 2009; Muñoz, 2015), chemistry (Valipouri & Nassaji, 2013; Xodabande et al., 2023), education (Mozaffari & Moini, 2014), nursing (Yang, 2015), environmental science (Liu & Han, 2015), psychology (Safari, 2018; Xodabande & Xodabande, 2020), veterinary medicine (Özer & Akbaş, 2024), and applied linguistics (Khani & Tazik, 2013; Matinparsa et al., 2022; Shabani & Tazik, 2014; Vongpumivitch et al., 2009; Xodabande et al., 2022).

METHOD

Corpora

Our study involved the corpus compilation of research articles in applied linguistics, which was drawn from a systematic selection of SINTA-indexed journals ranked between 1 (the highest) and 6 (the lowest). In 2022, out of the 8585 journals indexed by SINTA, we identified 72 that covered applied linguistics, or a combination of applied linguistics, linguistics, and literature written in English. Within this subset, we found four journals ranked first, 22 ranked second, 26 ranked third, eight ranked fourth, five ranked fifth, and seven ranked sixth. To ensure the representativeness of the corpus, we downloaded research articles related to applied linguistics from four journals with high-impact factors to represent each rank. We then converted the articles into plain text for data analysis and removed extraneous text such as journal names, running heads, author names and affiliations, page numbers, DOIs, tables, and references. To achieve balance within the corpus, we selected and included approximately 250000 running words from their current issues. The number of articles included in each journal varies due to the differing lengths of articles across journals. Nevertheless, we continued to download articles and add them to the corpus until each sub-corpus contained approximately 250000 running words. The final corpus comprised 6073379 running words (tokens) and provided a comprehensive and diverse sample of academic texts in applied linguistics. Table 1 presents the selected journals that were included in the corpus.

Software and Base Lists for Analysis

This research utilized AntWordProfiler to analyse the use of academic vocabulary in SIALJ research articles. Ant-

Table 1
Top Four SIALJ Across Rankings

SINTA	JOURNAL TITLE	NO. OF ARTICLES	TOKENS
1	Indonesian Journal of Applied Linguistics	94	256380
	International Journal of Language Education	64	255853
	TEFLIN Journal	47	256094
	Studies in English Language and Education	49	252180
2	English Review: Journal of English Education	60	253610
	Journal on English as a Foreign Language	49	251594
	Lingua Cultura	72	255224
	Register Journal	56	252694
3	ETERNAL (English, Teaching, Learning, and Research Journal)	61	252599
	Premise: Journal of English Education and Applied Linguistics	61	253822
	Metathesis: Journal of English Language, Literature, and Teaching	64	253445
	JELTL (Journal of English Language Teaching and Linguistics	52	250485
4	Getsempena English Education Journal	66	250782
	IDEAS: Journal on ELT and Learning, Ling. and Lit.	71	251289
	Exposure Journal	74	252194
	The English-Education: Journal of English Teaching and Research	78	252374
5	Anglo-Saxon	90	254884
	Wiralodra English Journal	67	255941
	Linguistik Terapan	115	253633
	e-Journal of Linguistics	88	252160
6	ELT in Focus	44	250679
	Journal of English Language Education	90	253111
	Journal Pendidikan Bahasa Inggris Indonesia	59	251913
	Jurnal Serunai Bahasa Inggris	68	250439
	Total	***1639***	***6073379***

WordProfiler is a tool designed to profile vocabulary level and text complexity and compare the loaded corpora with a list of reference corpora. The tool includes three base word lists by default: 1000 and 2000 GSL (West, 1953) and the Academic Word List (Coxhead, 2000). Nevertheless, it is possible to evaluate texts against additional vocabulary lists, which can be manually added to the program. In this study, we decided to incorporate supplementary lists into our profiling approach because the General Service List (GSL) and the Academic Word List (AWL) only include words that function as headwords. Research in lexical frequency profiling frequently reveals that non-AWL/non-GSL words can make up more than 13% of the corpus (refer to studies by Chanasattru & Tangkiengsirisin, 2017; Matinparsa et al., 2022; Xodabande et al., 2022; Xodabande & Xodabande, 2020). However, the specific list to which these off-list words belong remains unclear. Therefore, we find it necessary to utilize an additional list, specifically, BNC-COCA lists 31-34, which we adopted from Nation (2012). This additional list enables us to more thoroughly investigate the percentage of words in English texts that are not included in standard lists. The lists encompass proper names, marginal words, transparent compounds, and acronyms.

Data Processing

After importing all sub-corpora into AntWordProfiler, the software provides an overview of vocabulary coverage,

range, and frequency in the GSL/non-AWL and AWL. Firstly, a comprehensive analysis was performed by incorporating corpora from research articles in all SIALJ to determine the overall coverage of GSL/non-AWL and AWL vocabulary. Subsequently, a comparative analysis was undertaken to investigate vocabulary coverage in each SIALJ, based on their respective rankings, to identify any discrepancies in GSL/non-AWL and AWL vocabulary coverage. Finally, we present the top 50 AWL vocabularies in SIALJ research articles and the top 25 AWL vocabularies in SIALJ in each ranking.

In order to further examine whether the distribution of academic vocabulary in SIALJ research articles varies across rankings, we conducted a one-way analysis of variance (ANOVA) because it is suitable for the specific characteristics of our dataset and research objective. Examining the difference in AWL distribution also compares the AWL distribution by comparing the means of AWL token percentage, AWL type, and AWL headword count, making ANOVA well-suited to handle proportions, counts, and measurements spanning a wide range of values. The ANOVA was conducted to test the null hypothesis (H_0), which states, "There is no significant difference in the distribution of academic vocabulary in SIALJ research articles across rankings." If the p-value was less than the chosen significance level of 0.05, we rejected the null hypothesis, indicating a significant difference in the distribution of academic vocabulary in SIALJ research articles across rankings.

RESULTS

Coverage of Lexical Items in SIALJ Research Articles

Table 2 shows the coverage of lexical items in SIALJ research articles. The table demonstrates the distribution of the 1st GSL/non-AWL in the corpus, covering approximately 4401027 tokens, which is around 72.46% of the total corpus. Combined with the 2nd GSL/non-AWL, it covers around 4703228 tokens or 77.44% of the total corpus. The Academic Word List (AWL) covers 11.01% of the total corpus, corresponding to 668500 tokens. In this case, the cumulative coverage of GSL/non-AWL and AWL accounts for 88.45% of the total corpus. Notably, 11.55% of the corpus (701651 tokens) is classified as non-AWL/non-GSL. Within this percentage, 1.89% of the words are proper names, predominantly those of authors cited in the articles. Additionally, 0.51% are marginal words, which mainly include alphabets used in bullet lists and exclamations (e.g., *hmm, uh,* and *wah*) from interview data presented in the articles. Moreover, 0.64% of the words are transparent compound nouns, such as *feedback*, *classroom*, and *teamwork*, while 0.49% comprise acronyms, including *efl*, *esl*, and *ielts*. The remaining 8.02% of the words are not found in GSL, AWL, and supplementary lists and are primarily non-English words from interviews and other discipline-specific vocabulary such as *semantic*, *syntactic*, and *guttural*. Of the 570-word families in Coxhead's AWL, 569-word families (99.82%) were found, with the word *so-called* not being found in the corpus. The absence of the word in the corpus maybe because it is categorized at level 10 in Coxhead's AWL, but it may also be because it is not a content word that expresses specific concepts; instead, it is a phrase to modify or comment on other words or concepts.

Further analysis of the above findings reveals that the ten (10) most frequently used words from the AWL account for a total of 136791 tokens, which is approximately 2.25% of the entire corpus. These words include *research* (35090 tokens), *data* (20786 tokens), *text* (15979 tokens), *analyse* (15171 tokens), *process* (14768 tokens), *strategy* (11514 tokens), *participate* (11401 tokens), *communicate* (10742 tokens), and *method* (9340 tokens). The list of the top 50 academic vocabulary words in SIALJ research articles is presented in Table 3.

The word list shown in Table 3 contains a variety of words that present the key themes and research area in applied

Table 2

Lexical Profile of the SIALJ Research Articles

List	Token	Token%	Cum Token%	Type	Headword Count
1ST GSL	4401027	72.46	72.60	3763	998
2nd GSL	302201	4.98	77.44	2819	957
AWL	668500	11.01	88.45	2671	569
BNC-COCA31	114811	1.89	90.34	2378	2266
BNC-COCA32	30938	0.51	90.85	64	32
BNC-COCA33	38972	0.64	91.49	346	270
BNC-COCA34	29724	0.49	91.98	226	222
Not in the list	487206	8.02	100.00	18190	18190
Total	***6073379***	***100.0***		***30457***	***23504***

Table 3

Top 50 Most frequent AWL Items in SIALJ Research Articles

Rank	Words	Frequency	Sub-lists	Rank	Words	Frequency	Sub-lists
1	research	34090	1	26	lecture	4631	6
2	data	19786	1	27	technique	4535	3
3	text	14979	2	28	achieve	4504	2
4	analyse	14171	1	29	technology	4473	3
5	process	13768	1	30	structure	4428	1
6	strategy	10514	2	31	factor	4419	1
7	participate	10401	2	32	category	4352	2
8	communicate	9742	4	33	positive	4303	2
9	method	8340	1	34	design	4254	2
10	culture	7483	2	35	function	4233	1
11	respond	7347	1	36	aspect	4179	2
12	conduct	6487	2	37	task	4070	3
13	motive	6412	6	38	role	3953	1
14	context	6038	1	39	theory	3908	1
15	assess	5993	1	40	error	3777	4
16	create	5905	1	41	instruct	3756	6
17	conclude	5816	2	42	approach	3679	1
18	media	5639	7	43	previous	3613	2
19	implement	5379	4	44	involve	3611	1
20	significant	5358	1	45	topic	3599	7
21	interact	5145	3	46	vary	3585	1
22	perceive	4918	2	47	item	3565	2
23	focus	4861	2	48	attitude	3548	4
24	indicate	4829	1	49	consist	3346	1
25	academy	4756	5	50	evaluate	3320	4

linguistics, including language learning and teaching (i.e., *approach, design, instruct, method, strategy, technique*), cultural influences (i.e., *attitude, context, culture, motive*), and the impact of technology (i.e., *interact, media, technology*).

Distribution of Academic Vocabulary in SIALJ Research Articles across Rankings

Table 4 shows distribution patterns of AWL and GSL/non-AWL in SIALJ across rankings. The study observed that the higher the SINTA ranking, the greater the percentage of the AWL distribution, and conversely, the higher the SINTA ranking, the lower the GSL/non-AWL percentage. The AWL percentage tends to increase with higher SINTA ranking, ranging from 12.04% for SINTA 1 to 10.56% for SINTA 6. In contrast, the GSL/non-AWL coverage increases with lower SINTA rankings, from 74.33% for SINTA 1 to 79.89% for SINTA 6, indicating that the better the quality of a journal in SINTA, the more academic vocabulary is used. Further analysis revealed that SIALJ journals with a SINTA 1 ranking contain the highest number of AWL word types, totalling 2300 types, and the highest number of word families, which is 567 words. This number gradually decreases to 1934 types and 556-word families for SINTA 6.

Further analysis also found three-word families that do not occur in SIALJ research articles with a SINTA 1 (i.e., *so-called, nuclear, offset*), five in SINTA 2 (i.e., *bulk, invoke, offset, so-called, subsidy*), eight in SINTA 3 (i.e., ce*ase, export, federal, forthcoming, invoke, offset, revenue, so-called*), ten in SINTA 4 (i.e., *adjacent, amend, commence, currency, erode, export, federal, forthcoming, nuclear, so-called*), ten in SINTA 5 (i.e., *albeit, currency, erode, federal, fee, levy, offset, regime, so-called, subsidy*), and 16 in SINTA 6 (i.e., *adjacent, aggregate,*

Table 4

Distribution of Academic Vocabulary in SIALJ Research Articles Across Rankings

Word lists	SINTA 1	SINTA 2	SINTA 3	SINTA 4	SINTA 5	SINTA 6
AWL token	122829	117523	116964	104696	104302	102186
AWL token %	12.04	11.60	11.58	10.40	10.26	10.16
AWL type	2300	2275	2126	2050	2136	1934
AWL headword	567	565	562	560	560	556

Table 5

Top 25 Most Frequent AWL Items in SIALJ Research Articles Across Rankings

SINTA 1	SINTA 2	SINTA 3	SINTA 4	SINTA 5	SINTA 6
research	**research**	**research**	**research**	**research**	**research**
participate	**data**	**data**	**data**	**data**	**text**
analyse	**analyse**	**participate**	**analyse**	**text**	**data**
text	**strategy**	**analyse**	**process**	**process**	**process**
data	**participate**	**process**	**text**	**analyse**	**analyse**
assess	**process**	**communicate**	**strategy**	**communicate**	**strategy**
strategy	**text**	**text**	**method**	**strategy**	**method**
process	culture	**respond**	**communicate**	technique	**communicate**
culture	**communicate**	**strategy**	motive	**method**	conduct
respond	context	lecture	**respond**	function	implement
context	assess	academy	error	conclude	conclude
communicate	**method**	**method**	**participate**	culture	technique
identify	media	motive	conduct	**participate**	significant
significant	**respond**	culture	conclude	create	grade
instruct	conduct	create	media	theory	media
motive	interact	implement	culture	media	create
method	lecture	conduct	achieve	context	assess
academy	perceive	technology	create	**respond**	achieve
interact	technology	perceive	implement	structure	motive
task	indicate	interact	assess	conduct	**respond**
indicate	motive	significant	factor	clause	**participate**
factor	focus	media	task	source	design
positive	item	context	focus	identify	perceive
perceive	function	focus	academy	interact	category
focus	category	indicate	context	attitude	hypothesis

albeit, bulk, confine, currency, displace, erode, nuclear, offset, so-called, subsidy, suspend, terminate). Table 5 presents the top 25 AWL in SIALJ research articles for each ranking, with words in bold type indicating that they were found in SIALJ research articles in all rankings.

Table 5 shows that 10-word families are found within the top 25 AWL in SIALJ research articles across rankings, namely, *analyse, communicate, data, method, participate, process, research, respond, strategy,* and *text*. These words naturally occur in SIALJ research articles as they are related directly to the research themes in applied linguistics, such as language

acquisition and language teaching. Specifically, these words can be found in the research method section. The consistent and prevalent occurrence of the word research emphasizes its central role as a key element in applied linguistics scholarly discourse. Additionally, words like *data* and *participate* occur consistently in SIALJ research articles, indicating their significance in applied linguistics. The word like *participate* demonstrates engagement and involvement, highlighting the significance of the active role of individuals in applied linguistics studies. Also, the frequent occurrence of the word *data* indicates the attempts to sustain its empirical methods and acknowledgment of the role of data in advancing knowledge within the field.

Furthermore, the consistent use of methodological terms like *analyse* and *text* suggests that SIALJ scholars across rankings continue to maintain rigorous analytical methods and recognize the significance of textual elements in scrutinizing language-related phenomena. In addition, the recurrent use of words like *communicate* and *strategy* in SIALJ across rankings reflects a sustained focus on effective communication and methodical approaches in the field. Finally, the prevalence of culture and context across SIALJ rankings shows the scholars' preferences in studying how languages relate to culture, signifying the field's dedication to in-depth studies across various research levels.

Further analysis also reveals that the identified words occur in the top 25-word list of SIALJ research articles belong to Coxhead's AWL sub-lists 1, 2, and 3, accounting for 3.33% (202,828 tokens) of the total corpus or 33.34% of the academic word found in SIALJ research article. However, some words occur in lower sub-lists in Coxhead's list. For instance, several words occur in sub-list 4 (i.e., *attitude, communicate, error, implement*), sub-list 5 (i.e., *academy/academic, clause*), sub-list 6 (i.e., *communicate, instruct, lecture, motive/motivate*), sub-list 7 (i.e., *grade, media*). This observation suggests that while there is substantial overlap between this AWL list and Coxhead's, the ranking and prominence of certain words differ, emphasizing the strategic lexical composition within the applied linguistics research articles in the SIALJ dataset compared to Coxhead's AWL. The high frequency of certain words within SIALJ research articles, despite their relatively low occurrences in Coxhead's AWL, suggests that these words hold particular significance or relevance within the context of applied linguistics.

In addition to the qualitative observation of how academic vocabulary is used in SIALJ research articles across rankings, we conducted a quantitative analysis to examine further whether the differences observed in our qualitative observation are also evident in quantitative measurements. In order to examine whether there is a significant difference in the coverage of academic words in SIALJ research articles across rankings, we conducted an analysis of variance (ANOVA) to test the significance of variations in academic vocabulary coverage across SINTA rankings. AWL coverage percentage, AWL types, and AWL headwords were chosen for the analysis.

Table 6 shows the results of normality tests on the three data sets using the Kolmogorov-Smirnov and Shapiro-Wilk tests, showing that the p-values for all datasets in all tests are greater than 0.05. This value suggests that all datasets have a sufficiently normal distribution, and, therefore, an analysis of variance (ANOVA) can be conducted to measure whether there is a statistically significant difference in academic vocabulary coverage in SIALJ across rankings. The result of the ANOVA calculation is presented in Table 7.

Table 6

Normality Test Results for Academic Vocabulary in SIALJ Research Articles Across Rankings

	Kolmogorov-Smirnov			Shapiro-Wilk		
	Statistic	**df**	**p-value**	**Statistic**	**df**	**p-value**
AWL Token	0.369	6	0.299	0.841	6	0.234
AWL Types	0.276	6	0.300	0.948	6	0.726
AWL Headword	0.269	6	0.300	0.968	6	0.880

Table 7

ANOVA for Academic Vocabulary in SIALJ Research Articles Across Rankings

Source of variation	Sum of squares	df	Mean Square	F	p-value	F crit
SINTA ranks	33602.39	5	6720.478	1.101877	.417176661	3.325835
Error	60991.17	10	6099.117			
Total	***14701625***	***17***				

The table shows that the p-value associated with SINTA ranks is 0.417176661, greater than the generally used significance level of 0.05. Therefore, we fail to reject the null hypothesis, suggesting no significant difference in academic vocabulary distribution in SIALJ research articles across rankings. A smaller F-statistic relative to the critical F-value (3.325835) further supports the conclusion of non-significance. The F-statistic of 1.101877 indicates that the observed variability between SINTA rankings is insignificant. In summary, the p-value and the F-statistic lead us to accept the null hypothesis and conclude that there is no statistically significant difference in the distribution of academic vocabulary in SIALJ research articles across rankings.

DISCUSSION

The importance of academic vocabulary in scholarly communication has been firmly established in the literature. Academic vocabulary is critical in ensuring that the research findings are communicated accurately and effectively within the academic community (Awagu, 2021; Choo et al., 2017). The consistent use of academic vocabulary enhances the clarity and coherence of academic discourse (Arianto & Basthomi, 2021). Moreover, the ability to adeptly use academic vocabulary is significant for establishing credibility and promoting inclusivity in scholarly communication (Matinparsa et al., 2022).

In this study, we examined the coverage and distributions of academic words in SINTA-indexed applied linguistics journal (SIALJ) research articles. We aimed to determine the extent to which academic vocabulary is utilized within these journals and to explore whether there are significant differences in vocabulary usage across different SINTA rankings.

Our findings of an 11.01% coverage of Coxhead's AWL in SIALJ research articles align with the results of other studies suggesting that academic words typically cover at least 10% of academic texts (Hyland & Tse, 2004; Khani & Tazik, 2013; Vongpumivitch et al., 2009). Regarding applied linguistics, our research shows a nearly identical AWL coverage to that found by Xodabande et al. (2022), who reported an 11.46% coverage in reputable applied linguistics journals. Further examination also found that among the top 50 AWL words identified by Xodabande et al. (2022), 24 items were also present in our study. The list of headwords includes academy, acquire, analyse, approach, assess, communicate, context, culture, data, factor, focus, instruct, interact, item, motive, participate, process, research, role, strategy, structure, task, text, and theory. Beyond this, our analysis of the Academic and Applied Linguistics Word List (ALAWL) by Xodabande et al. (2022) revealed 378 common headwords with our study. This shared academic vocabulary indicates a significant consistency within SIALJ research articles and the broader field of applied linguistics.

Our observation of an 11.01% coverage of Coxhead's AWL in SIALJ research articles aligns with the notion of a discipline-specific vocabulary, where specific terms become integral to academic discourse dependent on "contextual environments which reflect different disciplinary practices and norms" (Hyland & Tse, 2004, p. 251), irrespective of journal clusters or rankings. This terminological uniformity is deemed crucial to enhance clarity in facilitating communication within the academic community (Hyland & Tse, 2004). It pedagogically benefits English language teaching by helping learners better understand published applied linguistics academic texts they need to read (Khani & Tazik, 2013). Moreover, a consistent academic vocabulary in several lists, especially related to applied linguistics, establishes and maintains disciplinary identity (Nation, 2001) and promotes accessibility and inclusivity (Martínez et al., 2009).

The study also found that several vocabularies in SIALJ research articles mostly occur in Coxhead's AWL top three sub-lists, covering approximately 3.33%. This finding corroborates the observation of Vongpumivitch et al. (2009) that the top word lists in applied linguistics research articles accounted for 3.60%, suggesting that the top word list in SIALJ research articles also exhibits high frequency in other applied linguistics word lists. This consistency not only enhances the clarity and coherence of academic discourse within SIALJ but also reflects the commitment of Indonesian academia to linguistic convention, scholarly communication standards, and the overall quality of research. Meticulous attention to language use also implies a commitment of journal editorial boards to robust selection and editorial process, establishing it as a reputable journal in the academic landscape.

However, we also notice that some vocabulary is found in considerable numbers in SIALJ research articles despite having low frequency in Coxhead's AWL. Words such as attitude, communicate, motive, and media are particular to applied linguistics, particularly in the context of language teaching and learning. This phenomenon corroborates the assertion of scholars in diverse disciplines who challenge the conventional notion of a one-size-fits-all academic word list and emphasize the need for developing field-specific academic vocabulary (Khani & Tazik, 2013; Kwary & Artha, 2017; Xodabande & Xodabande, 2020; Yotimart, 2021).

The presence of these specialized terms may be attributed to the unique contextual environments within which applied linguistics operates, as well as the genre-specific conventions that govern academic writing in this field. As Hyland and Tse (2004) have noted, disciplinary vocabularies adapt to "a locally appropriate theoretical and methodological framework" (p. 246). In this case, Hyland's (2004) concept of genre pedagogy is highly relevant as it suggests that academic writing is required to follow genre-specific conventions that general word lists may not fully represent;

therefore, the development of field-specific lexicons is necessary (Khani & Tazik, 2013; Martínez et al., 2009; Valipouri & Nassaji, 2013; Xodabande et al., 2022). Xodabande et al. (2023) contend that a replication study, building on previous research about the development of wordlists for specific disciplines, may effectively achieve this purpose.

Further quantitative analysis reveals an absence of significant distribution disparities across various SINTA rankings, challenging the findings of Arianto and Basthomi (2021), who noted a heightened strategic language use among authors in high-quality journals. Possible explanations could be that journal editors prioritize novel contributions rather than how authors strategically present their research with accurate language. This may also suggest that all journals, regardless of their ranking in a journal database, equally emphasize the importance of strategic language use in applied linguistics research; thus, all authors are required to adhere to the notion that disciplinary vocabularies adapt to specialized needs, contributing to the identity and coherence of the field (Biber & Gray, 2016; Hyland & Tse, 2004).

This study has implications for editorial and peer review standards and EAP teaching. The study may inform editorial boards and peer reviewers that acknowledging the consistent application of strategic, discipline-specific language has challenged the notions that journals with high reputations necessarily need more refined language use. Instead, the findings suggest that editorial boards have a more egalitarian approach to linguistic expectations in their publications. While maintaining a discipline-specific register in publications is significant, the editorial boards and peer reviewers should also focus on effective communication rather than subjecting authors to prestige-based linguistic norms. Furthermore, our findings of the shared practice of academic vocabulary use across SINTA rankings suggest that EAP educators tailor their instruction by introducing academic vocabulary to ensure students are adept at scholarly discourse within their field.

While our study has shown the commitment of Indonesian academia to linguistic conventions and the meticulous attention of journal editorial boards in maintaining reputable standards, we acknowledge that the study focused on SIALJ research articles in which the findings may not fully capture the entirety of academic vocabulary use within broader applied linguistics field. Also, the study's reliance on a specific timeline for analysis may not capture potential shifts in academic vocabulary over time, especially given the evolving nature of multidisciplinary research trends, particularly considering the increasing use of AI in scientific publications. Future research should also expand the scope by including more journal samples and using larger corpora to comprehensively depict how academic vocabulary is used in SIALJ research articles. Moreover, a focused approach to examining the corpus of Indonesian authors could offer a more accurate description of how Indonesian applied linguistics researchers use academic vocabulary.

CONCLUSION

This study has examined the lexical landscape of SIALJ research articles and provides insights into the distribution and characteristics of academic vocabulary across different rankings. We identified substantial coverage of Coxhead's AWL in SIALJ research articles, aligning with the broader academic literature regarding the prevalence of academic words in scholarly texts. The substantial overlap between SIALJ and Coxhead's AWL and other applied linguistics-related word lists underscores the consistent and discipline-specific nature of academic discourse in SIALJ research articles and applied linguistics in general. Furthermore, our examination of SINTA rankings found an intriguing pattern: higher-ranked journals exhibit greater AWL distribution, not significantly, especially upon quantitative measurement. The findings challenge the notion of a discrepancy in strategic language use among journals of varying quality, suggesting that strategic language use is a shared practice across SIALJ, irrespective of their rankings. Additionally, the prevalence of specific field-related terms with lower occurrences in Coxhead's AWL highlights unique linguistic needs and preferences within applied linguistics journals, notably influenced by the current trends in research topics within the field at the time of the study. The study contributes to the ever-evolving understanding of academic vocabulary, emphasizing the significance of discipline-specific lexicon in scholarly communication. Future research should investigate the evolving nature of academic vocabulary within applied linguistics, particularly in light of the increasing prevalence of AI-generated publications, which may significantly alter the landscape of academic vocabulary. Additionally, expanding the scope to include more journals and larger corpora could provide a more comprehensive understanding of academic vocabulary use across different contexts.

ACKNOWLEDGMENTS

We would like to thank the editors and the anonymous reviewers for their insightful comments.

DECLARATION OF COMPETITING INTEREST

None declared.

AUTHOR CONTRIBUTIONS

Suhandoko: conceptualization; collected the data; analyzed the data; wrote the paper.

Dian Riesti Ningrum: collected the data; analyzed the data; reviewing and editing.

Andini Dwi Wardani: collected the data.

Ach. Nobair: collected the data.

Putroue Keumala Intan: performed statistical analysis and interpretation.

REFERENCES

Aldawsari, H. (2017). *The challenges of learning academic vocabulary among postgraduate Saudi students at New Zealand universities.* Auckland University of Technology. https://openrepository.aut.ac.nz/bitstream/handle/10292/10754/AldawsariH.pdf?sequence=4&isAllowed=y

Arenas-Castro, H., Berdejo-Espinola, V., Chowdhury, S., Rodríguez-Contreras, A., James, Raja, N. B., Dunne, E. M., Bertolino, S., Nayara Braga Emidio, Derez, C. M., Drobniak, S. M., Fulton, G. R., L. Francisco Henao-Diaz, Kaur, A., Catherine, Malgorzata Lagisz, Medina, I., Mikula, P., Narayan, V. P., & O'Bryan, C. J. (2024). Academic publishing requires linguistically inclusive policies. *Proceedings - Royal Society. Biological Sciences/Proceedings - Royal Society. Biological Sciences*, 291. https://doi.org/10.1098/rspb.2023.2840

Arianto, M. A., & Basthomi, Y. (2021). The authors' research gap strategies in ELT research article introductions: Does Scopus journal quartile matter? *Journal of Language and Linguistic Studies, 17*(4), 1743–1759. https://doi.org/10.52462/jlls.127

Asaad, H. Q. M. (2024). The role of morphological awareness in L2 postgraduates' academic writing: Is vocabulary knowledge a mediating variable? *Cogent Education, 11*(1). https://doi.org/10.1080/2331186X.2024.2327787

Awagu, I. V. (2021). Language in academic writing: Features and topical issues. *Linguistics and Literature Studies, 9*(2), 49–56. https://doi.org/10.13189/LLS.2021.090201

Biber, D., & Gray, B. (2016). *Grammatical complexity in academic English: Linguistic change in writing*. Cambridge University Press.

Black, N., Van Rooyen, S., Godlee, F., Smith, R., & Evans, S. (1998). What makes a good reviewer and a good review for a general medical journal? *JAMA, 280*(3), 231–233. https://doi.org/10.1001/JAMA.280.3.231

Breeze, R. (2008). Researching simplicity and sophistication in student writing. *International Journal of English Studies, 8*(1), 51–66.

Brun-Mercer, N., & Zimmerman, C. B. (2015). Fostering academic vocabulary use in writing. *The CATESOL Journal, 27*(1), 131–148.

Castillo-Vergara, M., Alvarez-Marin, A., & Placencio-Hidalgo, D. (2018). A bibliometric analysis of creativity in the field of business economics. *Journal of Business Research, 85*, 1–9. https://doi.org/10.1016/J.JBUSRES.2017.12.011

Chanasattru, S., & Tangkiengsirisin, S. (2017). Thw word list distribution in social science research articles. *Arab World English Journal (AWEJ), 8*(4), 412–429. https://doi.org/https://dx.doi.org/10.24093/awej/vol8no4.28

Chen, Q., & Ge, G. chun. (2007). A corpus-based lexical study on frequency and distribution of Coxhead's AWL word families in medical research articles (RAs). E*nglish for Specific Purposes, 26*(4), 502–514. https://doi.org/10.1016/J.ESP.2007.04.003

Choo, L. B., Lin, D. T. A., Singh, M. K. M., & Ganapathy, M. (2017). The significance of the academic word list among esl tertiary students in a malaysian public university. *3L: The Southeast Asian Journal of English Language Studies, 23*(4), 56–65. https://doi.org/10.17576/3L-2017-2304-05

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly, 34*(2), 213–238. https://doi.org/10.2307/3587951

Coxhead, A. (2012). Academic vocabulary, writing and English for academic purposes: Perspectives from second language learners. *RELC Journal, 43*(1), 137–145. https://doi.org/10.1177/0033688212439323

Csomay, E., & Prades, A. (2018). Academic vocabulary in ESL student papers: A corpus-based study. *Journal of English for Academic Purposes, 33*, 100–118. https://doi.org/10.1016/J.JEAP.2018.02.003

Demir, C. (2019). The needless complexity in academic writing: Simplicity vs. flowery language. *The Reading Matrix: An International Online Journal, 19*(2).

Donthu, N., Kumar, S., & Pattnaik, D. (2020). Forty-five years of Journal of Business Research: A bibliometric analysis. *Journal of Business Research, 109*, 1–14. https://doi.org/10.1016/J.JBUSRES.2019.10.039

El-Omar, E. M. (2014). How to publish a scientific manuscript in a high-impact journal. *Advances in Digestive Medicine, 1*(4), 105–109. https://doi.org/10.1016/j.aidm.2014.07.004

Firmansyah, E. A., & Faisal, Y. A. (2019). Bibliometric analysis of islamic economics and finance journals in Indonesia. *Al-Muzara'ah, 7*(2), 17–26. https://doi.org/10.29244/JAM.7.2.17-26

Garner, R. M., Hirsch, J. A., Albuquerque, F. C., & Fargen, K. M. (2018). Bibliometric indices: defining academic productivity and citation rates of researchers, departments and journals. *Journal of Neurointerventional Surgery, 1*0(2), 102–106. https://doi.org/10.1136/NEURINTSURG-2017-013265

Gureyev, V. N., & Mazov, N. A. (2022). Bibliometrics as a promising tool for solving publication ethics issues. *Heliyon, 8*(3), e09123. https://doi.org/10.1016/J.HELIYON.2022.E09123

Hinkel, E. (2003). Simplicity without elegance: Features of sentences in l1 and l2 academic texts. *TESOL Quarterly, 37*(2), 275–301. https://doi.org/10.2307/3588505

Hyland, K. (2002). Options of identity in academic writing. *ELT Journal, 56*(4), 351–358. https://doi.org/10.1093/ELT/56.4.351

Hyland, K. (2004). *Disciplinary discourses: Social interactions in academic writing*. The University of Michigan Press.

Hyland, K. (2009). *Academic discourse: English in a global context.* Continuum.

Hyland, K. (2013). Writing in the university: Education, knowledge and reputation. *Language Teaching, 4*6(1), 53–70. https://doi.org/10.1017/S0261444811000036

Hyland, K., & Tse, P. (2004). Metadiscourse in academic writing: A reappraisal. *Applied Linguistics, 25*(2), 156–177. https://doi.org/10.1093/APPLIN/25.2.156

Ibrahim, C., & Fadhli, R. (2021). Performance of Indonesia's World-Class University efficiency with bibliometrics (scientific strength) approach and data envelopment analysis. *Webology, 18*(1), 32–50. https://doi.org/10.14704/WEB/V18I1/WEB18003

Kandi, V. (2016). Journal publications, indexing and academic excellence: Have we chosen the right path. *American Journal of Infectious Diseases and Microbiology, 4*(3), 52–55. http://dx.doi.org/10.12691/ajidm-4-3-1

Khani, R., & Tazik, K. (2013). Towards the development of an academic word list for applied linguistics research articles. R*ELC Journal, 44*(2), 209–232. https://doi.org/10.1177/0033688213488432

Koushik, A. K. (2017). Indexing and impact factor: Judging the quality of a journal. *Journal of Health & Medical Informatics, 8*(4), 1000285. https://doi.org/10.4172/2157-7420.1000285

Kwary, D. A., & Artha, A. F. (2017). The academic article word list for social sciences. *MEXTESOL Journal, 41*(4), 1–11.

Liu, J., & Han, L. (2015). A corpus-based environmental academic word list building and its validity test. E*nglish for Specific Purposes, 39*, 1–11. https://doi.org/10.1016/J.ESP.2015.03.001

Martínez, I. A., Beck, S. C., & Panza, C. B. (2009). Academic vocabulary in agriculture research articles: A corpus-based study. *English for Specific Purposes, 28*(3), 183–198. https://doi.org/10.1016/J.ESP.2009.04.003

Masrai, A., Milton, J., El-Dakhs, D. A. S., & Elmenshawy, H. (2021). Measuring the contribution of specialist vocabulary knowledge to academic achievement: Disentangling effects of multiple types of word knowledge. *Asian-Pacific Journal of Second and Foreign Language Education, 6*(1), 1–15. https://doi.org/10.1186/S40862-021-00114-5/FIGURES/3

Matinparsa, H., Xodabande, I., Ghafouri, M., & Atai, M. (2022). Academic vocabulary in applied linguistics research articles: A corpus-based replication study. *Research Square*. https://doi.org/10.21203/rs.3.rs-2092705/v1

Mozaffari, A., & Moini, R. (2014). Academic words in education research articles: A corpus study. *Procedia - Social and Behavioral Sciences, 98*, 1290–1296. https://doi.org/10.1016/j.sbspro.2014.03.545

Muñoz, V. L. (2015). The vocabulary of agriculture semi-popularization articles in English: A corpus-based study. *English for Specific Purposes, 39*, 26–44. https://doi.org/10.1016/J.ESP.2015.04.001

Muslimin, A. I., & Basthomi, Y. (2022). Bibliometric analysis on an Indonesia english education department academic staffs' productivities and its potential for national accreditation stance. *Journal of Scientometric Research, 11*(2), 254–261. https://doi.org/10.5530/JSCIRES.11.2.27

Nandiyanto, A. B. D., Biddinika, M. K., & Triawan, F. (2020). How bibliographic dataset portrays decreasing number of scientific publication from Indonesia. *Indonesian Journal of Science and Technology, 5*(1), 154–175. https://doi.org/10.17509/IJOST.V5I1.22265

Nation, I. S. P. (2001). How good is your Vocabulary program. *ESL Magazine, 4*(3), 22–24.

Özer, M., & Akbaş, E. (2024). Assembling a justified list of academic words in veterinary medicine: The veterinary medicine academic word list (VMAWL). *English for Specific Purposes, 74*, 29–43. https://doi.org/10.1016/J.ESP.2023.12.002

Pournia, Y. (2019). A study on the most frequent academic words in high impact factor English nursing journals: A corpus-based study. I*ranian Journal of Nursing and Midwifery Research, 24*(1), 11. https://doi.org/10.4103/IJNMR.IJNMR_190_17

Purnomo, B. C., Munahar, S., Pambuko, Z. B., & Nasrullah, H. (2020). Biodiesel research progress in Indonesia: Data from Science and Technology Index (Sinta). *Technology Reports of Kansai University, 6*2(06), 2997-3005.

Rahardja, U., Purnama Harahap, E., Ratna Dewi, S., & Tinggi Manajemen Informatika dan Komputer Raharja Jl Jendral, S. (2019). The strategy of enhancing article citation and H-index on SINTA to improve tertiary reputation. *TELKOMNIKA (Telecommunication Computing Electronics and Control), 17*(2), 683–692. https://doi.org/10.12928/TELKOMNIKA.V17I2.9761

Rey-Martí, A., Ribeiro-Soriano, D., & Palacios-Marqués, D. (2016). A bibliometric analysis of social entrepreneurship. *Journal of Business Research, 69*(5), 1651–1655. https://doi.org/10.1016/J.JBUSRES.2015.10.033

Robbins, S. P. (2016). Finding your voice as an academic writer (and writing clearly). *Journal of Social Work Education, 52*(2), 133–135. https://doi.org/10.1080/10437797.2016.1151267

Roldan-Valadez, E., Salazar-Ruiz, S. Y., Ibarra-Contreras, R., & Rios, C. (2019). Current concepts on bibliometrics: A brief review about impact factor, Eigenfactor score, CiteScore, SCImago Journal Rank, Source-Normalised Impact per Paper, H-index, and alternative metrics. *Irish Journal of Medical Science, 188*(3), 939–951. https://doi.org/10.1007/S11845-018-1936-5/METRICS

Safari, M. (2018). Do university students need to master the GSL and AWL words: A psychology word list. *Journal of Modern Research in English Language Studies, 5*(2), 101–122. https://doi.org/10.30479/JMRELS.2019.10266.1275

Shabani, M. B., & Tazik, K. (2014). Coxhead's AWL across ESP and Asian EFL journal research articles (RAs): A corpus-based lexical study. *Procedia - Social and Behavioral Sciences, 98*, 1722–1728. https://doi.org/10.1016/J.SBSPRO.2014.03.599

Tamela, E. (2020). Move structure analysis on research article abstracts in national and international scopus indexed journals. *Proceedings of the International Conference on English Language Teaching (ICONELT 2019), 434*, 12–17. https://doi.org/10.2991/ASSEHR.K.200427.004

Valipouri, L., & Nassaji, H. (2013). A corpus-based study of academic vocabulary in chemistry research articles. *Journal of English for Academic Purposes, 12*(4), 248–263. https://doi.org/10.1016/J.JEAP.2013.07.001

Vongpumivitch, V., Huang, J. yu, & Chang, Y. C. (2009). Frequency analysis of the words in the Academic Word List (AWL) and non-AWL content words in applied linguistics research papers. *English for Specific Purposes, 28*(1), 33–41. https://doi.org/10.1016/J.ESP.2008.08.003

Weinstein, J. N., & Morgan, T. S. (2007). Threats to scientific advancement in clinical practice. *Spine, 32*(11 SUPPL.). https://doi.org/10.1097/BRS.0B013E318053D4FC

West, M. (1953). *A general service list of English words*. Longman, Green & Co.

Wicherts, J. M. (2016). Peer review quality and transparency of the peer-review process in open access and subscription journals. *PLOS One, 11*(1), 1–19. https://doi.org/10.1371/journal.pone.0147913

Wijaya, L. E., & Bram, B. (2022). Rhetorical structures of introduction sections in Sinta-indexed journals. *Jurnal Onoma: Pendidikan, Bahasa, Dan Sastra, 8*(1), 181–188. https://doi.org/10.30605/ONOMA.V8I1.1659

Xiao, Z., Qin, Y., Xu, Z., Antucheviciene, J., & Zavadskas, E. K. (2022). The journal buildings: A bibliometric analysis (2011-2021). *Buildings, 12*(1), 37. https://doi.org/10.3390/BUILDINGS12010037

Xodabande, I., Atai, M. R., Hashemi, M. R., & Thompson, P. (2023). Developing and validating a mid-frequency word list for chemistry: A corpus-based approach using big data. *Asian-Pacific Journal of Second and Foreign Language Education, 8*(32). https://doi.org/10.1186/s40862-023-00205-5

Xodabande, I., Torabzadeh, S., Qafouri, M., & Emadi, A. (2022). Academic vocabulary in applied linguistics research articles: a corpus-based study. *Journal of Language and Education, 8*(2), 154–164. https://doi.org/10.17323/JLE.2022.13420

Xodabande, I., & Xodabande, N. (2020). Academic vocabulary in psychology research articles: A corpus-based study. *MEXTESOL Journal, 44*(3).

Yadira, A., Zein, T. T., & Setia, E. (2022). The analysis schematic structure of research article from journal indexed by Sinta. *East Asian Journal of Multidisciplinary Research , 1*(5), 825–842. https://doi.org/10.55927/EAJMR.V1I5.572

Yang, M. N. (2015). A nursing academic word list. *English for Specific Purposes, 37*(1), 27–38. https://doi.org/10.1016/J.ESP.2014.05.003

Yotimart, D. (2021). Academic vocabulary in sport tourism news: A corpus-based study. *Journal of Language and Linguistic Studies, 17*(3), 1527–1535. https://doi.org/10.52462/jlls.110

https://doi.org/10.17323/jle.2024.22198

Your Article is Accepted. Academic Writing for Publication: A Deep Dive into International Research on Challenges and Strategies

Antoaneta Angelova-Stanimirova, Maya Lambovska

University of National and World Economy, Sofia, Bulgaria

ABSTRACT

Background: Academic writing for publication (AWforP) has recently come to the fore because of the critical importance of scholarly publication to academia and the internationalization of science. A review of the scientific literature on AWforP found that it is underdeveloped and lacks comprehensive frameworks and models for AWforP challenges, AWforP strategies, or both.

Purpose: To contribute to bridging these gaps, this article aims to summarize and map the AWforP challenges and AWforP strategies identified in the scholarly empirical literature.

Method: A systematic Scopus/WoS literature review was used for data collection, identifying 15 relevant sources (*n*, sample size). The review was based on the PRISMA recommendations. Data were analyzed and summarized by deduction and meta-analysis based on chi-square heterogeneity test and meta-regression, then mapped by induction and K-means clustering.

Results: First, 31 challenges to AWforP and 36 strategies for AWforP were detected. Second, an original classification of AWforP challenges was introduced. The taxonomy of academic writing strategies was expanded with AWforP strategies. Third, AWforP challenges/strategies were ranked based on their frequency of mention in the sample. Semantic difficulties were the most prevalent challenge, and attending academic writing courses was the most advised strategy. Fourth, through meta-analysis, the sample was found to be moderately statistically heterogeneous (I^2=60.97%), and the summary effect size was positive and statistically significant. Fifth, the sampled sources were mapped into five clusters based on the country of researchers studied (SSE=10.511).

Conclusion: This article conceptualizes empirical research on AWforP challenges and AWforP strategies by identifying, comprehensively systematizing, summarizing, and mapping them. Implementing the proposed taxonomy of AWforP challenges/strategies under the identified cluster specifics in the academic writing teaching and strategic research planning and control practices would improve researchers' publication activity and research management effectiveness at the university and national levels.

KEYWORDS

academic writing, writing for publication, scholarly publications, challenges and strategies, research management

Citation: Angelova-Stanimirova, A., & Lambovska, M. (2024). Your article is accepted. Academic writing for publication: A deep dive into international research on challenges and strategies. *Journal of Language and Education, 10*(3), 108-127. https://doi.org/10.17323/jle.2024.22198

Correspondence:
Maya Lambovska
E-mail: mlambovska@abv.bg

Received: May 28, 2024
Accepted: September 16, 2024
Published: September 30, 2024

INTRODUCTION

The notification 'Congratulations! Your article has been accepted for publication' is strongly desired and valuable for researchers nowadays (Agathokleous, 2022). The exceptional significance of their scholarly productivity for academia in recent years is a primary reason (Frandsen et al., 2024; Lambovska, 2023). Over the past two decades, knowledge production has become a central pillar of national (Carlsson & Wilén, 2024) and global research governance policies (Oancea, 2019). Publication activity in the Scopus/Web of Science (WoS) databases is now a key criterion for high accreditation scores (Veretennik & Okulova, 2023), university rankings (Lambovska & Todorova, 2023), and funding (Owan et al., 2023). As a result, publications in high-quality journals have become

vital to the establishment and academic success of scholars, university professors, and doctoral candidates (herein "researchers") (Habibie, 2022; Jalongo, 2024). The quality of these publications largely depends on how well the authors' ideas are expressed textually (Scholz, 2022), thereby bringing academic writing for publication (AWforP) to the forefront of the scientific literature. AWforP is a complex skill as a part of academic writing (Tikhonova et al., 2024). The complexity of academic writing stems from its linguistic and stylistic features: sentence structure, vocabulary, syntactic constructions, hedging devices, genre specificity, etc. (Biber, 2006). AWforP complexity is further amplified by its global nature, driven by the internationalization of science (Raitskaya & Tikhonova, 2020). Some authors even define AWforP as a challenging process (Du Plooy et al., 2024). The paramount importance and high complexity of AWforP for academia bring out identifying AWforP challenges and strategies as a top priority for all stakeholders, but primarily for research management bodies.

A review of the literature on AWforP shows that it is one of the emerging topics of interest in academic writing (Raitskaya & Tikhonova, 2022). Specifically regarding AWforP challenges and AWforP strategies, the following gaps in the Scopus/WoS scholarly literature were identified. First, the scholarly literature on AWforP challenges and strategies for AWforP is in its infancy. This conclusion is based on the fact that 57 records were found on the topic of this study, of which only 15 (Table 1) were relevant to it. Second, we failed to find a complete model or theoretical/conceptual framework of AWforP challenges and strategies to address them. Only the source (Ren & Hu, 2023) can be considered such an effort, but it is a community-specific case study lacking the specifics of a framework or generic model. Third, the literature review found no comprehensive systematization of or models for AWforP challenges, AWforP strategies, or both. This gap addresses both empirical and theoretical research. Source (Lim & Koay, 2024) is an exception, but only for rhetorical strategies to improve paper quality/acceptance.

To contribute to bridging the indicated gaps, albeit in a small way, this article aims to summarize and map the challenges and strategies for academic writing for publication identified in the scholarly empirical literature. Methodologically, we identify this research as a basic conceptual model of the "conceptual description" type (Meredith, 1993). A Scopus/WoS systematic literature review, deduction/induction, meta-analysis using meta-regression and chi-square heterogeneity test, and K-means cluster analysis underpin it.

The theoretical basis of this research covers the concepts of philosophy of science, linguistics, and knowledge management. Concepts of the philosophy of science were applied to specify the nature and type of this research as a conceptual model. Specifically, these are Meredith's (1993) ideas of conceptual models and frameworks forming a logical and well-structured theory. From linguistics and knowledge management, the concepts of Hyland (2021), Gillett et al. (2013), Tang et al. (2023), Teng and Yue (2023), and Bui et al.'s (2023) taxonomies were primarily used. We predominantly followed the concepts of the first three sources to classify the AWforP challenges, as we fully support the authors' views. The taxonomies of the last two sources were employed as a basis for systematizing strategies for AWforP. In our opinion, these taxonomies propose the most complete classification of academic writing strategies.

In this study, four research questions clarify our tasks:

RQ#1: What are the AWforP challenges and strategies for AWforP found in the scholarly empirical literature indexed by Scopus and Core Collection (WoS) databases before August 9, 2024?

RQ#2: Which AWforP challenges are most prevalent, and which AWforP strategies are most advised in academia?

RQ#3: What is the heterogeneity of practices in academia regarding AWforP challenges and strategies?

RQ#4: How can we reasonably cluster studies on AWforP challenges and strategies from the perspective of effective AWforP management and control?

From a linguistics and educational perspective, this article is a pioneering effort to bridge the gap between the theory and practice of AWforP by conceptualizing empirical research on AWforP challenges and strategies. To our knowledge, this is the first study to identify, comprehensively systematize, summarize, and map AWforP challenges and strategies. It also complements and extends the current taxonomy of academic writing challenges and strategies toward scholarly writing for publication. In terms of the philosophy of science, this article suggests a conceptual descriptive model. Implementing the proposed toolkit in teaching academic writing and strategic research planning and control practices would improve researchers' publication activity and research management at the university and national levels.

BASIC CONCEPTS

This section introduces the basic concepts underlying our study. Its two designations are to build conceptual foundations for the proposed taxonomy of AWforP challenges and AWforP strategies and to substantiate our inference about the nature of this study. Concepts from linguistics, knowledge management, and general management underpin the taxonomy of the AWforP challenges/strategies. We use these concepts to explain our understanding of the core of AWforP challenges/strategies and our approach to grouping them. The basic concepts in this context address the eponymous AWforP group of challenges/strategies. Con-

cepts from the philosophy of science are applied to clarify the nature of this study as a conceptual descriptive model.

In creating AWforP challenge groups, we drew primarily on Hyland's (2021) concepts of academic discourse and metadiscourse tools, Gillett et al.'s (2013) and Tang et al.'s (2023) ideas on academic writing features, and Üstünbaş's (2023) perspective on the nature of metalinguistic awareness/knowledge. The selection of concepts is mainly based on the high degree of correspondence between these authors' views and ours. Academic discourse is defined as ways of thinking about language and its use in academia (Hyland, 2021). Academic discourse covers the first AWforP challenge group of the proposed taxonomy (row 1, Table 2). Metalinguistic awareness/knowledge is viewed as cognizance of how a language works, including its structure, form and use (Üstünbaş, 2023). It covers the third AWforP challenge group here (row 3, Table 2).

In identifying AWforP strategies, we rely mostly on Teng and Yue's (2023) and Bui et al.'s (2023) taxonomies of academic writing strategies (cognitive, metacognitive, rhetorical, and social). In our view, these authors propose the most complete classification of academic writing strategies, summarizing the main prior concepts on this subject. Cognitive strategies directly address the writing process (Wischgoll, 2016), and the first AWforP strategy group here (row 1, Table 3). They cover organizing, connecting ideas, elaborating, summarizing, visualizing, inference, deducing, etc. (Bui et al., 2023; Supeno et al., 2024). Metacognitive strategies fall under metacognitive control (Teng & Yue, 2023) and facilitate aligning cognitive strategies with writing goals and the writing process monitoring (Wischgoll, 2016). They include drafting, information management, editing/revising, planning, monitoring, and evaluation (Bui et al., 2023; Rosdiana et al., 2023). Here, these strategies cover the second AWforP strategy group (row 2, Table 3). Rhetorical strategies help writers present their ideas in an understandable way (Bui et al., 2023), covering the third AWforP strategy group here (row 3, Table 3). These strategies include analogy, comparison, metadiscourse tools, formulating questions, analysis, organizing ideas, contrast, etc. (Chanamé-Chira et al., 2022). Social strategies cover seeking interplay/support from others, receiving feedback, information sharing, etc. (Bui et al., 2023; Supeno et al., 2024). They are included in the fourth AWforP strategy group, shown in row 4 of Table 3.

The basic concepts from the philosophy of science applied here address the conceptual model, the conceptual descriptive model, and the levels of conceptual models/frames. We mainly use and follow Meredith's (1993) concepts of models/frames as building blocks of coherent, logical and well-structured research methodological theory. A conceptual model is defined by Meredith (1993) as a set of concepts employed to describe or represent a process or object without explaining it. There are seven conceptual models according to Meredith (1993). Based on their explanatory power, they are grouped into three hierarchical levels (Dwayi, 2024). The first level is the lowest, covering conceptual models. The second (middle) level covers conceptual frameworks. Meta-frames/theories are at the third (highest) level. A conceptual descriptive model is a type of conceptual model that is least abstract and mostly descriptive (Meredith, 1993). It belongs to the lowest model level because of its least explanatory power (Lynn, 1976; Meredith, 1993).

METHODS

General Description of the Study

We carried out this study in three phases. In the first phase, AWforP challenges and strategies for AWforP were drawn from the literature, then systematised, and finally ranked. As a result, research questions 1 and 2 were answered. The second phase included a meta-analysis, and the third covered a cluster analysis. Research questions 3 and 4 were respectively answered in phases 2 and 3.

A systematic review of the literature (Phase 1) was employed to gather data. The data were analysed through deduction (Phase 1) and meta-analysis (Phase 2) methods. Induction (Phase 1) and K-means clustering (Phase 3) methods were applied to synthesise the findings of this paper.

The PRISMA rules were heeded in this review and meta-analysis. The latter was executed with the chi-square test of heterogeneity and meta-regression. The results were summarized in a Forest Plot diagram. IBM SPSS was applied to cluster the review data. The meta-analysis results and clusters were visualized through MS Excel.

Systematic Literature Review

PRISMA rules (Page et al., 2022) were employed for this systematic review. Using Lambovska and Raitskaya's (2022) approach, the review proceeded into five steps: identification, screening, eligibility, inclusion, and synthesis. Data statistics during the first four stages are presented in Figure 1 via a PRISMA flowchart.

In the (first) Identification step, a protocol for the review was written, and literature searches were carried out. In the review protocol (Figure 2), information (search) sources, the search phrase and strategies, and eligibility criteria (inclusion and exclusion) were described. Scopus & Core Collection of WoS scientometric databases were employed as search sources because of their scholarly nature and high coverage. We used a single search phrase obtained as a combination of keywords "academic writing", publication,

Figure 1

PRISMA flow chart

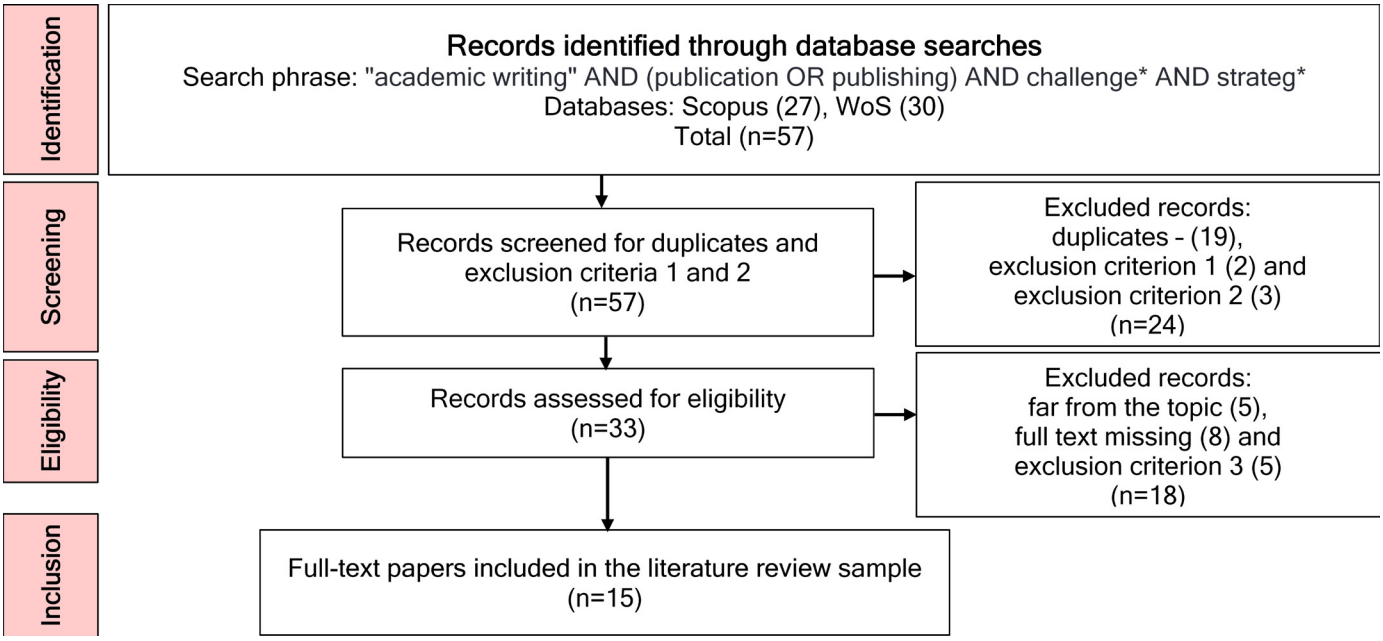
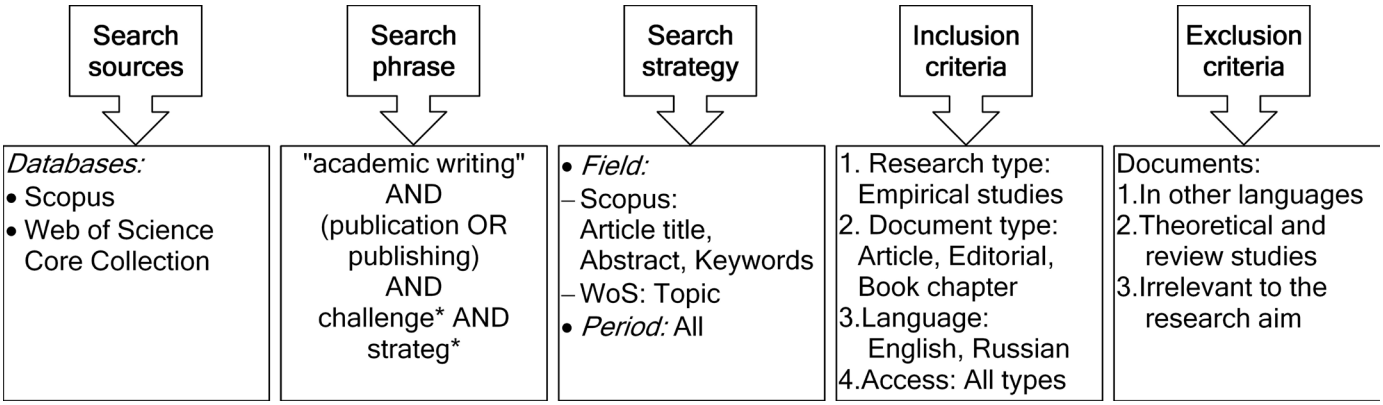


Figure 2

Protocol



publishing, challeng* and strateg*. These were chosen in line with the aim of this article. Our search strategy covered the entire indexing period of databases and the widest possible search fields, Topic (WoS) and Article Title, Abstract, and Keywords (Scopus). Eligibility criteria were mainly based on the nature of studies (theoretical/ empirical). We aimed to explore the practice of academia. Therefore, inclusion criterion 1 (research type) covered only empirical studies, and inclusion criterion 2 (document type) articles, editorials, and book chapters. Theoretical studies and reviews were excluded through exclusion criterion 2. Two languages were included in inclusion criterion 3, English and Russian. The Russian language allows covering papers from post-socialist countries. Their authors widely use Russian and are still poorly represented in Scopus & WoS.

Initial literature searches were carried out on June 10, 2024. Last updated on August 08, 2024. In the initial searches, we found 53 documents in Scopus & WoS (WoS: 28 and Scopus: 25). Later, four papers covering the search phrase were indexed in both databases (two for each). Thus, the total number of documents grew to 57. All 57 records were merged into one literature pool (herein "pool") and uploaded to Clarivate EndNote.

In the (second) Screening step, literature filtering was accomplished. First, the pool was checked for duplicate records. 19 duplicates were detected through the "find duplicates" feature. These were removed from the pool, which fell to 38 records. Then, this pool was reviewed for exclusion criteria 1 and 2. Regarding criterion 1, we found one record in Arabic and one in German. Three records were found un-

der criterion 2 (theoretical article, review and note). All five records were excluded for further review. As a result, the pool dropped to 33 records.

In the (third) Eligibility step, study selection was performed. The 33 records were analyzed for relevance to our research topic. Five documents were conflicting with the topic and were removed from the pool. We checked if the full text of the remaining 28 records was available. Of these, eight papers were not open access. Their full texts were not found in other databases or by queries to the authors, and these articles were excluded from the review. So, 20 papers are left in the pool. These papers were analyzed in detail according to exclusion criterion 3, using deduction and expert judgment methods. Five papers were found irrelevant to our research aim and removed from the pool. Each of the authors then performed a second in-depth review. Thus, the eligibility of the remaining 15 papers for this study was confirmed.

In the (fourth) Inclusion step, the sample of this study was formed. It covers the 15 eligible sources (Table 1). The remaining 23 sources (57 reduced by 19 duplicates and 15 eligible – "sample") are shown in the Appendix below. Of these, two sources fall under criterion 1 (other languages), three under exclusion criterion 2 (theoretical sources), five were far from this topic, eight with full text missing, and five under exclusion criterion 3 (irrelevant to the research aim).

The (fifth) Synthesis step covered data extraction, systematization, and ranking. First, data extraction was done for AWforP challenges and AWforP strategies based on an in-depth analysis of the selected sources. We tabulated these results, thus answering question 1. Second, applying the induction method, the AWforP challenges (Table 2) and AWforP strategies (Table 3) were systematised into groups based on their nature and concepts from the previous section. The country of participants under study (here "researchers") was also identified (Table 1). We use the term "undefined country" for the country of top-tier journal editors (sources [3], [8] and [9], Table 1) because their experience is not country-specific, but the term "international" for the sources [1], [4], and [7], (Table 1) addressing researchers from two or more countries. Third, AWforP challenges/strategies were ranked based on their total frequency of mention in the sample (Tables 2 and 3, columns Total). On this basis, in response to question 2, the most prevalent/advised AWforP challenges/strategies were identified. The top-ranked ones are shown in Figure 3.

Meta-Analysis

We ran a meta-analysis to estimate the statistical heterogeneity of the sample, thus answering research question 3. A systematic review and statistical estimation of its summary results are always covered by the meta-analysis (Higgins et al., 2023). Therefore, this meta-analysis was held as a follow-up to our systematic review, in four steps under Milani et al.'s recommendations (2024) on the meta-analysis procedure. We adhered to the established methodology throughout the process. Meta-analysis variables were defined in the first step. We used two variables: the total mentions of AWforP challenges (AWPC) and the total mentions of AWforP strategies (SAWP). Their values are shown in the Source Total rows in Tables 2 and 3, respectively. Statistical heterogeneity (I^2) was assessed in the second step through the chi-square test. When evaluating it, we followed the notion of Andrade (2020) that heterogeneity is low when I^2 is below 50%, moderate when I^2 is between 50% and 75%, and high when I^2 exceeds 75%. The third step was choosing a method. We decided to use the meta-regression method because it can simultaneously model the effects. In addition, our sample size meets the requirements of more than 10 items (Andrade, 2020). In the fourth step, we assessed the summary effect size using Carlson et al.'s Rule 8 (2023). The summary results were pictured by a Forest Plot (Figure 4) portraying the summary effect size estimate, its confidence intervals (CI), and the variations between the sampled sources. The Forest Plot was created in MS Excel after Neyeloff et al.'s (2012) guidelines on using Excel for this purpose. We deliberately decided not to conduct the sensitivity analysis, as it would reduce the variety of AWforP challenges/strategies.

Cluster Analysis

The K-means method was applied to cluster the sampled sources. Research question 4 was thus answered. The data were processed in 10 iterations by IBM SPSS. AWPC (Country Total row, Table 2) and SAWP (Country Total row, Table 3) were used as cluster analysis variables. The sampled sources were clustered using the country of researchers studied (Table 1) and the two variables. We decided to group the sources based on their country coverage (the countries of origin of the researchers studied) into five clusters (Figure 5). Our decision was founded on a relatively balanced country distribution by clusters (Table 4) and a low value of the sum of squared errors – SSE (Blömer et al., 2016). As a rule, low SSE denotes high compactness of the clusters (Selmi et al., 2024). Here, we tried to balance the results of these criteria. Trials were conducted for two, three, four, and five clusters. The results for the five clusters were the most satisfactory in terms of these two criteria. We visualised the clusters and their centroids through MS Excel.

RESULTS

Sample of the Study

This subsection presents the sample of this study (Table 1). The sample covers 15 sources. The authors of the sources, their years of publication, document type, respondents and the respondents' countries are also shown in Table 1.

Table 1

Research Sample Used

№	Source	Document type	Respondents	Country of Respondents
[1]	Bakla & Karakaş (2022)	Article	English-speaking researchers	International
[2]	Giraldo (2019)	Article	Columbian university professors	Columbia
[3]	Good & Pullins (2024)	Editorial	Top-tier journal editors	Undefined
[4]	Gupta et al. (2022)	Article	Non-native English-speaking (NNES) doctoral students (DS)and their faculty supervisors	International
[5]	Harvey et al. (2020)	Article	Australian health practitioners in clinical services	Australia
[6]	Langum & Sullivan (2020)	Article	Norwegian DS	Norway
[7]	Lillis & Curry (2022)	Article	NNES researchers	International
[8]	Lim & Koay (2024)	Editorial	Top-tier journal editors	Undefined
[9]	Martín (2017)	Editorial	Top-tier journal editor	Undefined
[10]	Niemelä & Naukkarinen (2021)	Article	Finish DS	Finland
[11]	Ren & Hu (2023)	Article	Chinese DS	China
[12]	Rezaei & Seyri (2019)	Article	Iranian DS	Iran
[13]	Shehata & Eldakar (2018)	Article	Egyptian researchers	Egypt
[14]	Subaveerapandiyan & Sinha (2024)	Article	Zambian university librarians	Zambia
[15]	Zhigalev et al. (2022)	Article	Russian DS	Russian Federation (RF)

Challenges and Strategies for Academic Writing for Publication

In response to question 1, we found 31 AWforP challenges and 36 AWforP strategies in the sample. These and their total, by-source and by-country, mentions are shown in Tables 2 and 3. In both tables, each mention of a challenge/strategy is counted once per source and is marked with the symbol √. The Total columns show the total mentions for each challenge/strategy (per row). The Source Total rows include the total mentions of all challenges/strategies in each source (per column), representing the values of the AWPC/SAWP variables used in the meta-analysis. The Country Total rows capture the total mentions of all challenges/strategies for each country, representing the values of the AWPC/SAWP variables used in the cluster analysis. For the undefined country, the AWPC/SAWP values reflect the total mentions in source columns [3], [8] and [9], and for the "international" item, they reflect those in columns [1], [4] and [7].

Further, AWforP challenges and strategies for AWforP were systematised into five and four groups, respectively, based on their nature and the basic concepts used here (see Basic Concepts section). The AWforP challenge groups (Table 2) are about academic discourse features, centre-periphery relations, (meta)linguistic knowledge, researcher behaviour, and research environment. The AWforP strategy groups (Table 3) are for cognitive, metacognitive, rhetorical, and social strategies. Using the same principles of systematization, the AWforP challenge group of academic discourse features and all AWforP strategic groups were subdivided.

The Most Prevalent Challenges and Advised Strategies for Academic Writing for Publication

In response to question 2, the most prevalent (top-ranked) three AWforP challenges and the most advised AWforP strategies were found (Figure 3). The rankings are based on the total mentions of AWforP challenges and AWforP strategies, shown in the Total columns of Tables 2 and 3.

The top-ranked AWforP challenges were semantic difficulties (rank 1, 10 mentions – m.), lack of English language proficiency, difficulties with writing conventions in English, difficulty deciding on research structure (rank 2, 8 m.), grammar problems, and difficulty organizing texts (rank 3, 5 m.). The most advised strategies for AWforP in academia were attending academic writing courses (rank 1, 10 m.), ongoing support from superior/university, formal training at universities (rank 2, 7 m.), information management through corpus tools, providing access to resources, using peer feedback, and requesting proofreading/feedback (rank 3, 5 m.).

Figure 3

Top-Ranked Challenges and Strategies for Academic Writing for Publication

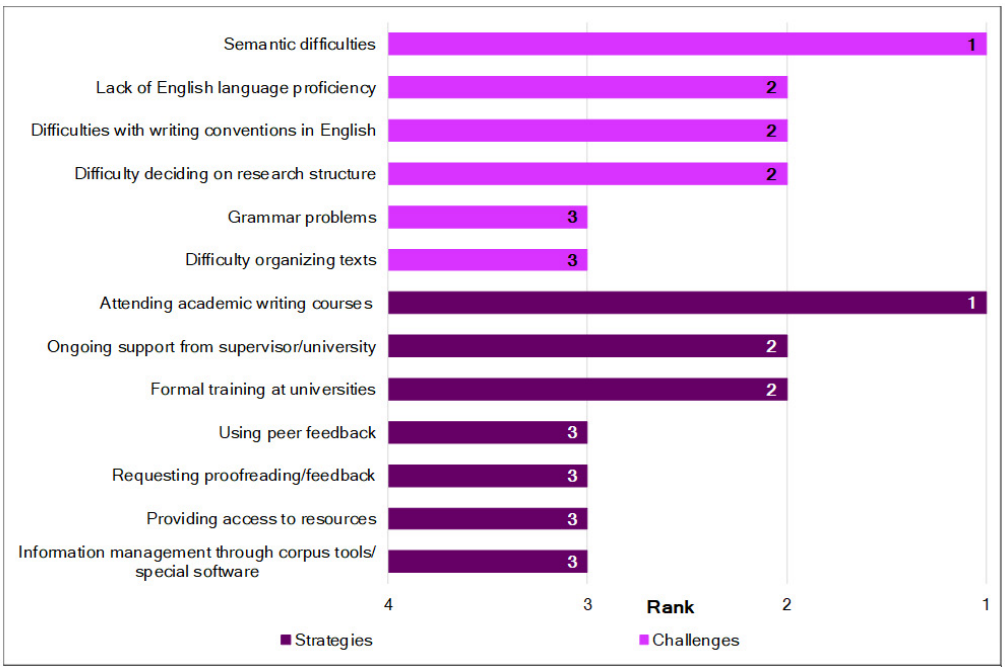


Table 2

Challenges to Academic Writing for Publication

	Challenges to (AWPC)/ Sources and Countries	[1], International	[2], Columbia	[3], Undefined	[4], International	[5], Australia	[6], Norway	[7], International	[8], Undefined	[9], Undefined	[10], Finland	[11], China	[12], Iran	[13], Egypt	[14], Zambia	[15], RF	Total
1	Academic discourse features																
1.1	*Complexity*																
a	Difficulties in developing concepts				√		√										2
b	Difficulty conceiving research											√			√	√	3
1.2	*Writing structure and organization*																
a	Difficulty deciding on the manuscript structure	√			√	√	√		√			√			√	√	8
b	Difficulty text organizing			√	√		√				√		√				5
c	Difficulties with writing the Introduction/ Discussion			√		√	√										3
1.3	*Writing style*																
a	Concise writing difficulties	√			√				√								3
b	Difficulty finding the author's "personal voice»						√										1
1.4	*Writing patterns in English*																

	Challenges to (AWPC)/ Sources and Countries	[1], International	[2], Columbia	[3], Undefined	[4], International	[5], Australia	[6], Norway	[7], International	[8], Undefined	[9], Undefined	[10], Finland	[11], China	[12], Iran	[13], Egypt	[14], Zambia	[15], RF	Total
a	Difficulties with conventions for writing (incl. the IMRaD model)				√		√			√	√	√	√	√	√		8
b	Difficulty translating some special expressions							√									1
c	Difficulty understanding genres of academic writing	√									√	√	√				4
2	Centre-periphery relations																
2.1	Difficulties due to national (academic) culture									√	√		√			√	4
2.2	Difficulty conveying local debates in mainstream journals									√							1
2.3	Dissatisfaction with the English language hegemony							√					√				2
2.4	Idiosyncratic forms of writing (in terms of international journals)									√			√				2
2.5	Lack of culture in international publishing													√	√		2
3	(Meta)linguistic knowledge																
3.1	Difficulty paraphrasing others' ideas	√					√										2
3.2	Gaps in linguistic terminology										√	√				√	3
3.3	Grammar problems				√						√	√			√	√	5
3.4	Lack of English language proficiency		√			√					√	√	√	√	√	√	8
3.5	Problems with academic metadiscourse tools										√	√				√	3
3.6	Semantic difficulties	√			√	√	√	√			√	√	√		√	√	10
4	Researcher behaviour																
4.1	Competing priorities at work and home					√											1
4.2	Experiencing writer's block	√															1
4.3	Fear of international publishing													√		√	2
4.4	Lack of discipline		√			√											2
5	Research environment																
5.1	Difficulties with peer feedback	√				√					√						3
5.2	Lack of (clear) communication with supervisors				√									√			2
5.3	Lack of instructions (for writing academic papers)												√	√	√		3
5.4	Lack of resources		√			√										√	3
5.5	Lack of support (financial, rewards, etc.)					√								√	√		3
5.6	Political decisions of some journals												√				1

Challenges to (AWPC)/ Sources and Countries	[1], International	[2], Columbia	[3], Undefined	[4], International	[5], Australia	[6], Norway	[7], International	[8], Undefined	[9], Undefined	[10], Finland	[11], China	[12], Iran	[13], Egypt	[14], Zambia	[15], RF	Total
Source Total	***7***	***3***	***2***	***8***	***9***	***8***	***3***	***2***	***4***	***10***	***9***	***10***	***7***	***9***	***10***	***101***
Country Total	***18[a]***	***3***	***8[b]***		***9***	***8***				***10***	***9***	***10***	***7***	***9***	***10***	***101***

Note. Sources match those in Table 1. [a] AWPC of the "international" item is the sum of columns [1], [4], and [7], [b] AWPC of the undefined country is the sum of columns [3], [8], and [9]

Table 3

Strategies for Academic Writing for Publication

Strategies (SAWP)/ Sources and Countries	[1], International	[2], Columbia	[3], Undefined	[4], International	[5], Australia	[6], Norway	[7], International	[8], Undefined	[9], Undefined	[10], Finland	[11], China	[12], Iran	[13], Egypt	[14], Zambia	[15], RF	Total
1 Cognitive strategies																
1.1 About the writing process																
a Applying novel research designs/methods											√					1
b Translating only the data necessary for the target publication							√									1
c Using dictionaries	√										√					2
1.2 About learning and training																
a Attending academic writing courses	√			√	√	√			√	√		√	√	√	√	10
b Formal training at universities				√					√	√		√	√	√	√	7
c Self-Study	√										√	√			√	4
2 Metacognitive strategies																
2.1 Planning																
a Leaving enough time for/between writing and proofreading	√				√											2
b Pre-writing activities	√	√														2
2.2 Drafting																
a Drafting the manuscript in the native language before its translation											√					1
2.3 Evaluation																
a Using a proofreading checklist				√							√	√				3
2.4 Information management through:																
a Corpus tools/special software	√			√	√						√			√		5

	Strategies (SAWP)/ Sources and Countries	[1], International	[2], Columbia	[3], Undefined	[4], International	[5], Australia	[6], Norway	[7], International	[8], Undefined	[9], Undefined	[10], Finland	[11], China	[12], Iran	[13], Egypt	[14], Zambia	[15], RF	Total
b	Interactive technologies															√	1
c	Using reading matrices		√										√				2
2.5	*Revising/Editing*	√			√							√					3
3	Rhetorical strategies																
3.1	*Improving writing style*																
a	Borrowing linguistic expressions from published articles											√					1
b	Balancing academic and literary styles						√										1
c	Using metadiscourse tools	√										√					2
3.2	*Improving paper quality/acceptance*																
a	«Tell with the Title» (select a short, informative, and engaging title)			√					√								2
b	«Grasp attention with the Abstract» (write a brief but comprehensive summary)			√					√								2
c	"Craft the Keywords" (choose relevant keywords, incl. theoretical and methodological terms, and geographical/contextual features)								√								1
d	«Sell the study in the Introduction» (hook the reader; highlight the topic's importance, literature gaps, and contributions of the study)								√								1
e	"Build the ground with the Literature" (lay the theoretical/conceptual foundation)			√					√								2
f	"Clarify the Methodology" (detail the sample, data collection procedures/methods)			√					√								2
g	"Frame the findings with the Results" (visualize findings and interpret results)			√					√								2
h	«Dazzle with the Discussion» (write a meaningful discussion, clarify contributions)			√					√								2
i	"Culminate in the Conclusion" (summarize key contributions and results, highlight significance and impact of the study, discuss future work and limitations)								√								1
j	"Strike with the References" (include seminal & recent, credible & relevant sources)								√								1
4	Social strategies																
4.1	*University social strategies*																
a	Creating writing centres at the university				√	√					√					√	4
b	Ongoing support from the supervisor or university					√	√				√		√	√	√	√	7
c	Promotion/reward system modification					√								√			2

	Strategies (SAWP)/ Sources and Countries	[1], International	[2], Columbia	[3], Undefined	[4], International	[5], Australia	[6], Norway	[7], International	[8], Undefined	[9], Undefined	[10], Finland	[11], China	[12], Iran	[13], Egypt	[14], Zambia	[15], RF	Total
d	Providing access to resources				√							√		√	√	√	5
e	Providing an appropriate environment/ culture					√										√	2
4.2	*Researchers' social strategies*																
a	Requesting proofreading/feedback	√				√	√					√	√				5
b	Using peer feedback		√		√	√					√	√					5
c	Using professional editing/proofreading services	√			√							√	√				4
d	Using writing retreats				√							√					2
	Source Total	***10***	***3***	***6***	***10***	***9***	***4***	***1***	***10***	***2***	***5***	***14***	***8***	***5***	***5***	***8***	***100***
	Country Total	***21a***	***3***	***18b***		***9***	***4***				***5***	***14***	***8***	***5***	***5***	***8***	***100***

Note. Sources match those in Table 1. [a] SAWP of the "international" item is the sum of columns [1], [4], and [7], [b] SAWP of the undefined country is the sum of columns [3], [8], and [9].

Results of the Meta-Analysis

In response to question 3, the following results from the meta-analysis were found (Figure 4). First, the statistical heterogeneity (I^2) assessed by the chi-square test was 60.97%, a moderate value.

Second, the sample summary outcome was 1.29 (Summary column, Outcome row, Figure 4 legend), calculated by combining the effect sizes of the sampled sources, with a standard error (SE) of 0.54 - average for this sample (Summary column, SE row, Figure 4 legend).

Third, a Forest Plot diagram was created to depict the results of each source and the overall heterogeneity. It is shown in Figure 4. Each horizontal line in Figure 4 corresponds to the individual confidence interval of a particular source. The horizontal lines of four sources intersect the ordinate, thus increasing the heterogeneity of the sample. These are sources [2] (R=100, CI=[-13.16;213.16]), [3] (R=33.33, CI=[-12.86;79.53]), [7] (R=300, CI=[-39.48;639.48]), and [8] (R=20, CI=[-7.72;47.72]), where R is the effect size in %.

Fourth, the sample summary effect size, depicted by the summary diamond on the line closest to the abscissa, was R=129.17 (Summary column, Rate row, Figure 4 legend), CI=[24.11;234.24].

Clustering the Sample on the Challenges and Strategies for Academic Writing for Publication

Five clusters were generated in response to question 4 (Figure 5 and Table 4). The clusters and their relative distribution in the sample by country are visualized in Figure 5a. Figure 5b depicts the cluster centres (centroids) and the location of their elements (countries covered). The content of the clusters (country coverage and sources included) and their SSE are shown in Table 4. Clusters were named after their values of total mentions of AWforP challenges (AWPC) and AWforP strategies (SAWP).

The clusters have the following features. Cluster 1 (labelled "medium–high") covered sources/countries reporting a medium value of total mentions of AWforP challenges (AWPC) and a high value of total mentions of AWforP strategies (SAWP). These are two items: the undefined country and China. The former corresponds to sources [3], [8], and [9] (Table 1), and the latter to [11]. The centroid coordinates of cluster 1 were (8.5;16), and its SSE was 4.124. Cluster 2 (labelled "maximum–maximum") included sources reporting AWPC and SAWP maximum values. These are three editorials ([1], [4], and [7]) combined into one item labelled "international". The centroid coordinates of cluster 2 were (18;21), and its SSE was zero (0). Cluster 3 (labelled "medium–medi-

um") comprised sources reporting AWPC and SAWP medium values. These are sources [5], [12], and [15] addressing researchers from Australia, Iran, and the Russian Federation, respectively. The centroid coordinates of cluster 3 were (9.67;8.33), and its SSE was 1.885. Cluster 4 (labelled "minimum–minimum") covered only one source - [2], the one reporting AWPC and SAWP minimum values. This source explores researchers from Colombia. The centroid coordinates of cluster 4 were (3;3), and its SSE was zero. Cluster 5 (labelled "medium-low") comprised sources/countries reporting medium AWPC and low SAWP values. These are sources [6], [10], [13], and [14], respectively, addressing researchers from Norway, Finland, Egypt, and Zambia. The centroid coordinates of cluster 5 were (8.5;4.75), and its SSE was 4.502.

The summary SSE of the cluster map was 10.511. Three levels of SSE were observed for this cluster map: zero, medium, and higher. The SSE of clusters 2 and 4 is zero because they cover only one country/item whose AWPC and SAWP define the cluster centroid. Cluster 3 SSE (1.885) is medium for this cluster map. Clusters 1 and 5 have higher SSE values of 4.124 and 4.502, respectively. One possible reason is that these two clusters encompass the most sources - four each.

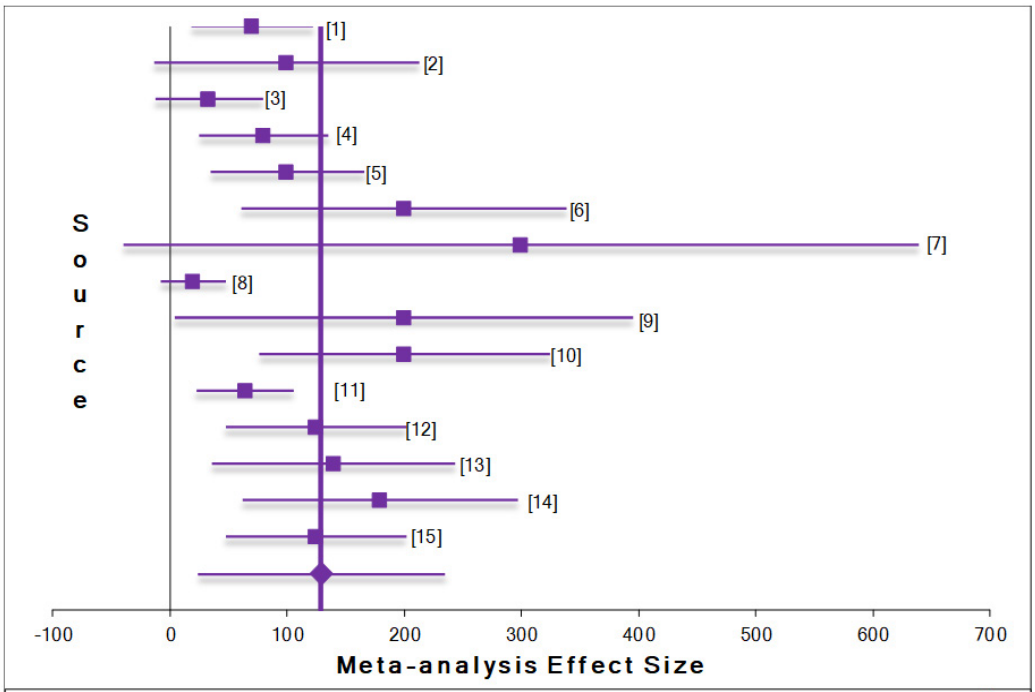
DISCUSSION

This research summarizes and maps AWforP challenges and strategies for AWforP by conceptualizing the empirical literature indexed by Scopus/WoS before August 09, 2024.

From a linguistics and education perspective, this research is a pioneering effort to address the absence of a comprehensive systematization, model or conceptual framework for AWforP challenges and AWforP strategies in the scholarly literature. In particular, it introduces a taxonomy of AWforP challenges and AWforP strategies. The first part of this taxonomy covers an original classification of AWforP challenges based on Hyland's (2021), Gillett et al.'s (2013), and Tang et al.'s (2023) concepts and our experience as researchers. The second part of the taxonomy, regarding AWforP strategies, can be considered an extension of Bui et al.'s (2023) and Teng & Yue's (2023) taxonomies on academic writing

Figure 4

Forest Plot of the Sample



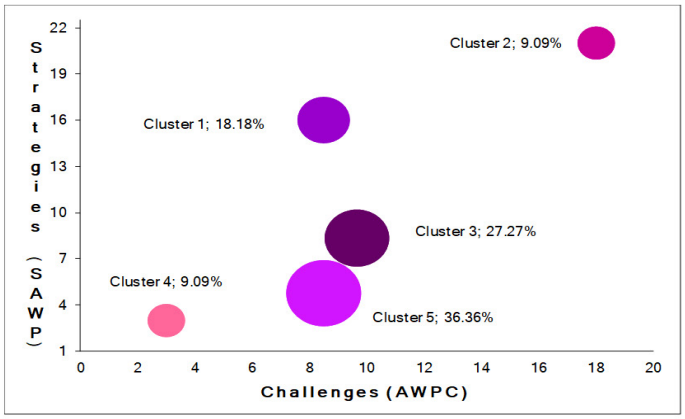
Source	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]	[14]	[15]	Summary
Outcome[a]	0.7	1	0.33	0.8	1	2	3	0.2	2	2	0.64	1.25	1.4	1.8	1.25	1.29
SE[b]	0.26	0.58	0.24	0.28	0.33	0.71	1.73	0.14	1	0.63	0.21	0.4	0.53	0.6	0.4	0.54
Rate (*R*)[c]	70	100	33.33	80	100	200	300	20	200	200	64.29	125	140	180	125	129.17
95% *CI*[d]																
CI lower	18.14	-13.16	-12.86	24.56	34.67	61.41	-39.48	-7.72	4	76.04	22.29	47.52	36.29	62.4	47.52	24.11
CI upper	121.86	213.16	79.53	135.44	165.33	338.59	639.48	47.72	396	323.96	106.29	202.48	243.71	297.6	202.48	234.24

Note. The source number is shown in Table 1.
[a] Outcome = Effect Size. [b] *SE* = Standard Error. [c] *R* = Effect Size in %. [d] *CI* = Confidence Interval.

Figure 5

Clusters to the AWforP Challenges and Strategies

(a) Cluster Map of the Sample



(b) Cluster Centroids

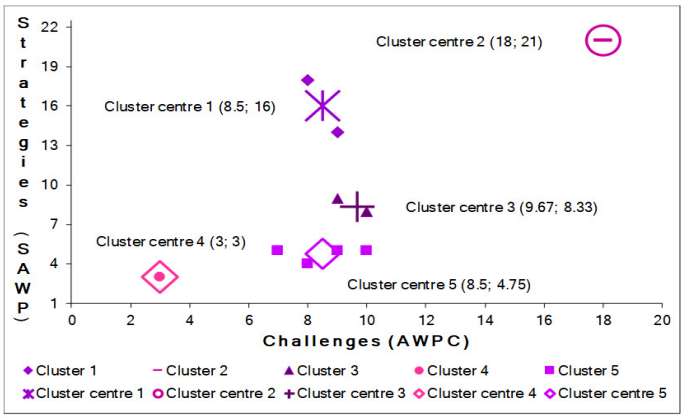


Table 4

Results of the Cluster Analysis

Indicator	Cluster 1 ("Medium-High")	Cluster 2 ("Maximum–Maximum")	Cluster 3 ("Medium–Medium")	Cluster 4 ("Minimum–Minimum")	Cluster 5 ("Medium-Low")	Total
SSE	4.124	0	1.885	0	4.502	10.511
Content of the clusters						
Country coverage	Undefined, China	International	Australia, Iran, Russia	Columbia	Norway, Finland, Egypt, Zambia	11
Literature source[a]	[3], [8], [9], [11]	[1], [4], [7]	[5], [12], [15]	[2]	[6], [10], [13], [14]	15

Note. [a] Sources match those in Table 1.

strategies towards AWforP. This part also integrates the ideas of Chanamé-Chira et al. (2022), Rosdiana et al. (2023), Supeno et al. (2024) and Wischgoll (2016), and summarizes the concepts of all authors from the sampled sources. We would like to highlight the expansion of the group of rhetorical strategies in the proposed taxonomy with the subgroup for improving paper quality/acceptance based on Good and Pullins's (2024) and Lim and Koay's (2024) views.

From the point of view of the philosophy of science, it can be categorized methodologically as a basic conceptual model of the "conceptual description" type under Meredith (1993) or at a classificatory level under Lin (1976). The primarily descriptive and low abstract nature of this study presupposes this categorization.

From a managerial perspective, this study directly addresses activities and tools of strategic research planning and control at both university and national levels. Specifically, choosing a strategy is a strategic planning procedure, identifying challenges is an element of risk assessment, and taking corrective action against challenges/inappropriate strategies is a control influence known as "regulation" (Nedyalkova, 2020). The last two activities fall under strategic control (Zhelev & Kostova, 2024), while the third closes the control feedback loop (Nedyalkova, 2024) in strategic management.

Research Question 1

Our interpretation of the main findings on AWforP challenges/strategies follows: To begin with the AWforP challenges (31 items, Table 2). The group of AWforP challenges to the academic discourse features was the most numerous (10 items). Given our topic, this is a logical outcome, as this group covers key challenges to AWforP rather than academic writing in general. Another important point concerns the AWforP strategies (36 items, Table 3). They were systematized into four groups: cognitive, metacognitive, rhetorical, and social. Within the rhetorical strategy group, we separated the improving paper quality/acceptance subgroup, covering strictly specific strategies for writing scientific publications. We titled these strategies based on the leading ideas of their authors (Good & Pullins, 2024; Lim & Koay, 2024),

editors of top-tier journals. In the context of the research topic, this subgroup and the rhetorical strategy group are logically the most numerous, with 10 and 12 items, respectively. Finally, the large number of AWforP challenges and AWforP strategies discovered through this research, particularly strategies to improve paper quality/acceptance, is evidence that academia has been excited about the topic and has been working on it in recent years.

Research Question 2

Regarding the ranking of AWforP challenges (Figure 3), the most prevalent ones (ranked 1st to 3rd, including semantic difficulties, grammar problems, difficulties with writing conventions in English, etc.) fall into the groups of challenges to academic discourse features and (meta)linguistic knowledge – three per group. Furthermore, these two groups were the most mentioned, with 38 and 31 mentions, respectively. The groups of research environment (15), centre-periphery relations (11), and researcher behaviour (six) follow. These results are reasonable because most researchers studied in this sample were non-native English-speaking or doctoral students (Table 1) and were not experienced enough in AWforP, including in international journals.

Regarding the ranking of AWforP strategies (Figure 3), the most advised strategies (ranked 1st and 2nd) were two cognitive learning/teaching strategies (attending academic writing courses and formal training at universities) and one social (ongoing support from superior/university). Three social strategies (providing access to resources, requesting proofreading/feedback and using peer feedback) and one metacognitive (information management through corpus tools) were ranked 3rd. Because of their nature, we view the top-ranked strategies as highly suitable for overcoming the most prevalent challenges. With 36 mentions, the social strategy group was the most advised. The groups of cognitive (25), rhetorical (20) and metacognitive strategies (19) follow. Notably, rhetorical strategies were not strongly recommended, including those to improve paper quality/acceptance. Given the most prevalent challenges and the researchers studied, these results are entirely logical and expected by us. The same goes for most AWforP strategies.

Research Question 3

Generally, the meta-analysis's main advantage is that it aggregates the results of multiple studies, thus providing a more reliable summary estimate than an individual study. In our research, the statistical heterogeneity of the sample (I^2=60.97%) was moderate. This level of heterogeneity is typical of meta-analyses in the social sciences because most studies do not have identical empirical settings (Hansen, 2022). According to the results, four studies ([2], [3], [7], and [8]) from our sample increased its heterogeneity to the level of 60.97%. These studies have statistically insignificant results because their horizontal lines intersect the Forest Plot ordinate (Figure 4). We intentionally did not eliminate these four sources through sensitivity analysis (Milani et al., 2020) as we aimed for a "deep dive" into this topic. The summary results of the meta-analysis give reason to conclude that the sample effect size is positive and statistically significant.

Research Question 4

Based on the cluster analysis results (Figure 5 and Table 4), the following interpretations can be made about the features of the cluster map: First, two types of clusters can be recognized in the cluster map based on the level of correspondence between the total mentions of AWforP challenges (AWPC, Table 2, Country Total row) and those of AWforP strategies (SAWP, Table 3, Country Total row). Clusters with high similarity in these indicators belong to the first type. These are clusters with two identical title elements, namely clusters 2 ("maximum–maximum"), 3 ("medium–medium") and 4 ("minimum–minimum"). There is no similarity between AWPC and SAWP for the clusters of the second type. These are clusters with two different title elements, namely clusters 1 ("medium–high") and 5 ("medium–low"). The second point is that two clusters stand out, the features of which differ significantly from the others in the cluster map. These are clusters 2 ("maximum–maximum") and 4 ("minimum–minimum"). They address only one country and have zero SSE and similar extreme centroid coordinates of AWPC and SAWP as their titles suggest. Our next conclusion is that there are three clusters with similar centroid medium AWPC values but quite different SAWP. These are clusters 1 (8.5;16), 3 (9.67;8.33) and 5 (8.5;4.75). Clusters 1 and 5 even have the same AWPC value (8.5). Finally, the total SSE of the cluster map (10.511, Table 4) is relatively low for this moderately heterogeneous sample. Therefore, our findings regarding clustering can be deemed reliable.

Limitations

The main limitation of this research concerns the review protocol applied (Figure 2), namely its search sources, search phrase and inclusion and exclusion criteria (language, research type, and document type). Furthermore, this research does not consider the field of study of the sampled sources. Exploring the hidden effects and causes of correlations between the two variables of total mentions of AWforP challenges and AWforP strategies (AWPC and SAWP), including by source and country, is also beyond the scope of this study.

CONCLUSION

This article provides a "deep dive" (in-depth study) into international research on challenges and strategies for academic writing for publication. These were identified, systematised, summarized, and mapped therein, thus answering the research questions raised.

From a linguistics and education perspective, this research adds to academic writing theory by expanding knowledge about academic writing for publication. In particular, it proposes a taxonomy of challenges and strategies for academic writing for publication, thus complementing and extending the existing eponymous taxonomy in academic writing. In addition, this taxonomy enriches the toolbox of pedagogy, specifically the teaching of academic writing and the training of doctoral students. As far as the available literature suggests, this is the first study to identify and summarize the challenges and strategies of academic writing for publication. This summarization can be considered a more significant contribution to the topic than the contribution of the studies published to date. Further, the present research bridges the gap between the theory and practice of academic writing for publication by conceptualizing empirical studies on its challenges and strategies. From a management perspective, the proposed taxonomy expands the strategic research planning and control toolkit, especially that of the regulatory process as part of strategic research control, thus enriching research management knowledge.

In terms of practice, the following main implications of the present study can be outlined. First, the parties concerned (researchers, teachers, and university research managers) can choose appropriate strategies for writing academic publications from the taxonomy suggested here to increase publication quality and activity and, as a result, enhance university rankings and scores. Second, the parties concerned can borrow strategies from their or other cluster sources/countries. Third, the proposed taxonomy of academic writing challenges and publishing strategies can be incorporated into curricula of the academic writing and research methodology disciplines, taught to doctoral students. Fourth, university/government management can integrate the proposed taxonomy of challenges and strategies into their strategic research planning and control systems. This would be particularly beneficial to the research regulatory systems. In conclusion, the results of this study can become a pillar of a road map to enhance the publication activity of researchers and research management effectiveness at the university and national levels.

We suggest future research on the topic be conducted in several directions. The first is to periodically perform analogous empirical literature studies to expand and complement the proposed taxonomy of challenges and strategies for academic writing for publications. Another direction is to develop and implement university-, community-, and country-specific taxonomies of this type. If they exist, we propose analyses of their features and implementation problems to be carried out and good practices to be promoted. The next direction covers developing methodologies for evaluating the effectiveness of strategies to overcome the challenges of academic writing for publications. Last but not least, studies could be conducted to explore the effects of applying such taxonomies on the researchers' publication activity and research management/governance effectiveness.

DECLARATION OF COMPETITING INTEREST

None declared.

AUTHOR CONTRIBUTIONS

Antoaneta Angelova-Stanimirova: conceptualization; data curation; formal analysis; funding acquisition; investigation; methodology; project administration; resources; software; supervision; validation; visualization; writing – original draft; writing – review & editing.

Maya Lambovska: methodology; visualization; writing – original draft; writing – review & editing.

REFERENCES

Agathokleous, E. (2022). Mastering the scientific peer review process: Tips for young authors from a young senior editor. *Journal of Forestry Research*, *33*(1), 1-20. https://doi.org/10.1007/s11676-021-01388-8

Andrade, C. (2020). Understanding the basics of meta-analysis and how to read a Forest plot: As simple as it gets. *Journal of Clinical Psychiatry*, *81*(5), Article 20f13698. https://doi.org/10.4088/JCP.20F13698

Bakla, A., & Karakaş, A. (2022). Technology and strategy use in academic writing: Native, native-like versus non-native speakers of English. *Iberica*, *2022*(44), 285–314. https://doi.org/10.17398/2340-2784.44.285

Biber, D. (2006). *University Language: A Corpus-based Study of Spoken and Written Registers*. J. Benjamins.

Blömer, J., Lammersen, C., Schmidt, M., & Sohler, C. (2016). Theoretical analysis of the k-means algorithm – A survey. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 9220 LNCS, pp. 81–116). https://doi.org/10.1007/978-3-319-49487-6_3

Bui, H. P., Nguyen, L. T., & Nguyen, T. V. (2023). An investigation into EFL pre-service teachers' academic writing strategies. *Heliyon*, *9*(3), Article e13743. https://doi.org/10.1016/j.heliyon.2023.e13743

Carlson, R. B., Martin, J. R., & Beckett, R. D. (2023). Ten simple rules for interpreting and evaluating a meta-analysis. *PLoS Computational Biology*, *19*(9), Article e1011461 https://doi.org/10.1371/journal.pcbi.1011461

Carlsson, V., & Wilén, E. J. (2024). "It is controlling, but you don't really care." Researchers' perceptions of legitimation of research policy. *Science and Public Policy*, *51*(4), 609–617. https://doi.org/10.1093/scipol/scae004

Chanamé-Chira, R., Santisteban-Chévez, D., Manayay-Tafur, M., Solano-Cavero, J. K., Villón-Prieto, R. D., Villón-Prieto, C. R., & Quintana-Marreros, C. (2022). Discursive and rhetorical strategies: A problem of academic writing. *RISTI - Revista Iberica de Sistemas e Tecnologias de Informacao*, *2022*(E53), 137–154.

du Plooy, B., Albertyn, R., Troskie-de Bruin, C., & Belcher, E. (2024). Academic writing for publication: The experience and facilitation of liminality for developing higher levels of scholarliness. *Innovations in Education and Teaching International*, 1-13. https://doi.org/10.1080/14703297.2024.2363899

Dwayi, V. V. M. (2024). Reimagining how the critical realist ways of methodological triangulation might allow for resolving paradigm in/commensurability in research methodologies. *European Conference on Research Methodology for Business and Management Studies, 23*(1), 63–70. https://doi.org/10.34190/ecrm.23.1.2464

Frandsen, T. F., Lamptey, R. B., & Borteye, E. M. (2024). Promotion standards to discourage publishing in questionable journals: A follow-up study. *Journal of Academic Librarianship*, *50*(5), Article 102895. https://doi.org/10.1016/j.acalib.2024.102895

Gillett, A., Hammond, A., & Martala, M. (2013). *Inside Track to Successful Academic Writing*. Pearson Education Limited.

Giraldo, F. (2019). An English for research publication purposes course: Gains, challenges, and perceptions. *GiST Education and Learning Research Journal*, *18*, 198–219. https://doi.org/10.26817/16925777.454

Good, V., & Pullins, E. B. (2024). The nine habits of highly effective researchers: Strategies for strengthening scholarly submissions. *Journal of Personal Selling and Sales Management*, *44*(2), 101-107. https://doi.org/10.1080/08853134.2024.2324883

Gupta, S., Jaiswal, A., Paramasivam, A., & Kotecha, J. (2022). Academic writing challenges and supports: Perspectives of international doctoral students and their supervisors. *Frontiers in Education*, *7*, Article 891534. https://doi.org/10.3389/feduc.2022.891534

Habibie, P. (2022). Writing for scholarly publication in an interconnected disjunctured world. *Journal of Second Language Writing*, *58*, Article 100933. https://doi.org/10.1016/j.jslw.2022.100933

Hansen, C., Steinmetz, H., & Block, J. (2022). How to conduct a meta-analysis in eight steps: A practical guide. *Management Review Quarterly*, *72*(1), 1-19. https://doi.org/10.1007/s11301-021-00247-4

Harvey, D., Barker, R., & Tynan, E. (2020). Writing a manuscript for publication: An action research study with allied health practitioners. *Focus on Health Professional Education: A Multi-Professional Journal*, *21*(2), 1–16. https://doi.org/10.11157/fohpe.v21i2.397

Higgins, J. P. T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. A. (2023). *Cochrane Handbook for Systematic Reviews of Interventions* version 6.4 *(updated August 2023).* Cochrane.

Hyland, K. (2021). Academic discourse. In Hyland, K., Paltridge, B. & Wong, L. (Eds.) *The Bloomsbury Handbook of Discourse Analysis* (2nd ed., pp. 125-138). Bloomsbury Publishing Plc.

Jalongo, M. R. (2024). Scholarly publication during doctoral candidature: Obstacles, benefits, and strategies for success. *Early Childhood Education Journal*, 1-12. https://doi.org/10.1007/s10643-024-01724-7

Lambovska, M. (2023). What is behind the shine? The dark side of research evaluation: A conceptual framework. *TEM Journal*, *12*(4), 2552–2563. https://doi.org/10.18421/TEM124-67

Lambovska, M., & Raitskaya, L. (2022). Specificity of the motivation for high-quality publications in Russia. *TEM Journal*, *11*(3), 1205–1212. https://doi.org/10.18421/TEM113-28

Lambovska, M., & Todorova, D. (2023). Striving for high-quality publications: Motivational profiles of management within a Bulgarian university cluster. *TEM Journal*, *12*(2), 1100–1109. https://doi.org/10.18421/TEM122-56

Langum, V., & Sullivan, K. P. H. (2020). Academic writing, scholarly identity, voice and the benefits and challenges of multilingualism: Reflections from Norwegian doctoral researchers in teacher education. *Linguistics and Education*, *60*, Article 100883. https://doi.org/10.1016/j.linged.2020.100883

Lillis, T., & Curry, M. J. (2022). The dynamics of academic knowledge making in a multilingual world Chronotopes of production. *Journal of English for Research Publication Purposes*, *3*(1), 109–142. https://doi.org/10.1075/jerpp.22002.lil

Lim, W. M., & Koay, K. Y. (2024). So you want to publish in a premier journal? An illustrative guide on how to develop and write a quantitative research paper for premier journals. *Global Business and Organizational Excellence*, *43*(3), 5–19. https://doi.org/10.1002/joe.22252

Lin, N. (1976). *Foundations of social research*. McGraw-Hill.

Martín, E. (2017). Current Sociology and the challenges of inequality in academia: 65 years forging spaces of intelligibility. *Current Sociology*, *65*(3), 327–335. https://doi.org/10.1177/0011392117694516

Meredith, J. (1993). Theory building through conceptual methods. *International Journal of Operations & Production Management*, *13*(5), 3–11. https://doi.org/10.1108/01443579310028120

Milani, A., Dessi, F., & Bonaiuto, M. (2024). A meta-analysis on the drivers and barriers to the social acceptance of renewable and sustainable energy technologies. *Energy Research and Social Science*, *114*, Article 103624. https://doi.org/10.1016/j.erss.2024.103624

Nedyalkova, P. (2020). Quality of internal auditing in the public sector. Perspectives from the Bulgarian and international context. In *Contributions to Management Science*. Springer Nature. https://doi.org/10.1007/978-3-030-29329-1

Nedyalkova, P. (2024). Concepts of the nature and development of control. In A. Derbali (Ed.), *Recent developments in financial management and economics* (pp. 14-25). IGI Global. https://doi.org/10.4018/979-8-3693-2683-1.ch002

Neyeloff, J. L., Fuchs, S. C., & Moreira, L. B. (2012). Meta-analyses and forest plots using a Microsoft Excel spreadsheet: Step-by-step guide focusing on descriptive data analysis. *BMC Research Notes*, *5*, Article 52. https://doi.org/10.1186/1756-0500-5-52

Niemelä, H., & Naukkarinen, J. (2021). On the rocky road to academia: Stumbling blocks for Finnish engineering students with English as a second language. *International Journal of Engineering Pedagogy*, *10*(6), 36–56. https://doi.org/10.3991/IJEP.V10I6.14559

Oancea, A. (2019). Research governance and the future(s) of research assessment. *Palgrave Communications*, *5*(1), Article 27. https://doi.org/10.1057/s41599-018-0213-6

Owan, V. J., Bassey, B. A., & Ubi, I. O. (2023). Construction and standardisation of an instrument measuring lecturers' persistence to publish in Scopus-indexed journals. *Journal of Applied Learning and Teaching*, *6*(2), 158–171. https://doi.org/10.37074/jalt.2023.6.2.37

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., … Moher, D. (2022). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Revista Panamericana de Salud Publica/Pan American Journal of Public Health*, *46*, Article e112. https://doi.org/10.26633/RPSP.2022.112

Raitskaya, L., & Tikhonova, E. (2022). An in-depth glimpse into research on academic writing. *Journal of Language and Education*, *8*(2), 5–18. https://doi.org/10.17323/jle.2022.14586

Raitskaya, L., & Tikhonova, E. (2020). Seven deadly sins: Culture's effect on scholarly editing and publishing. *Journal of Language and Education*, *6*(3), 167–172. https://doi.org/10.17323/jle.2020.11205

Ren, S., & Hu, G. (2023). Two Chinese medical doctors' English scholarly publishing practices: Challenges, contradictions and coping strategies. *Iberica*, *2023*(45), 289–315. https://doi.org/10.17398/2340-2784.45.289

Rezaei, S., & Seyri, H. (2019). Iranian doctoral students' perceptions of publication in English: Motives, hurdles, and strategies. *Journal of Applied Research in Higher Education*, *11*(4), 941–954. https://doi.org/10.1108/JARHE-02-2019-0040

Rosdiana, L. A., Damaianti, V. S., Mulyati, Y., & Sastromiharjo, A. (2023). The role of metacognitive strategies in academic writing skills in higher education. *International Journal of Learning, Teaching and Educational Research*, *22*(6), 328–344. https://doi.org/10.26803/ijlter.22.6.18

Scholz, F. (2022). Writing and publishing a scientific paper. *ChemTexts*, *8*(1). Article 8. https://doi.org/10.1007/s40828-022-00160-7

Selmi, A. T. E., Zerarka, M. F., & Cheriet, A. (2024). Enhancing K-Means clustering with post-redistribution. *Ingenierie Des Systemes d'Information*, *29*(2), 429–436. https://doi.org/10.18280/isi.290204

Shehata, A. M. K., & Eldakar, M. A. M. (2018). Publishing research in the international context: An analysis of Egyptian social sciences scholars' academic writing behaviour. *Electronic Library*, *36*(5), 910–924. https://doi.org/10.1108/EL-01-2017-0005

Subaveerapandiyan, A., & Sinha, P. (2024). Assessing scholarly communication competence in Zambian library professionals. *Global Knowledge, Memory and Communication*. https://doi.org/10.1108/GKMC-09-2023-0351

Supeno, Sundari, H., & Yohanna, L. (2024). Willingness to write among EFL university students: A case of a virtual writing course in Indonesia. *Taiwan Journal of TESOL*, *21*(1), 37–78. https://doi.org/10.30397/TJTESOL.202404_21(1).0002

Tang, X., Zhou, H., & Li, S. (2023). Predictable by publication: Discovery of early highly cited academic papers based on their own features. *Library Hi Tech*, *42(*4), 1366-1384. https://doi.org/10.1108/LHT-06-2022-0305

Teng, M. F., & Yue, M. (2023). Metacognitive writing strategies, critical thinking skills, and academic writing performance: A structural equation modelling approach. *Metacognition and Learning*, *18*(1), 237–260. https://doi.org/10.1007/s11409-022-09328-5

Tikhonova, E. V., Kosycheva, M. A., & Mezentseva, D. A. (2024). Ineffective strategies in scientific communication: Textual wordiness vs. clarity of thought in thesis conclusion section. *Integration of Education*, *28*(2), 249–265. https://doi.org/10.15507/1991-9468.115.028.202402.249-265

Üstünbaş, Ü. (2023). Machine translation use in language learning: Learner characteristics, beliefs, and ethical concerns. In *Transforming the Language Teaching Experience in the Age of AI (*pp. 141 - 159*)*. https://doi.org/10.4018/978-1-6684-9893-4.ch009

Veretennik, E., & Okulova, O. (2023). Of performance and impact: How AACSB accreditation contributes to research in business schools. *Higher Education Policy*, *36*(4), 758–780. https://doi.org/10.1057/s41307-022-00284-y

Wischgoll, A. (2016). Combined training of one cognitive and one metacognitive strategy improves academic writing skills. *Frontiers in Psychology*, *7,* Article 187. https://doi.org/10.3389/fpsyg.2016.00187

Zhelev, Z., & Kostova, S. (2024). Investigating the application of digital tools for information management in financial control: Evidence from Bulgaria. *Journal of Risk and Financial Management*, *17*(4), Article 165. https://doi.org/10.3390/jrfm17040165

Zhigalev, B. A., Belorukova, M. V., Ganyushkina, E. V., & Zolotova, M. V. (2022). Effective strategies forming L2 professional communicative competency in postgraduate groups. *Yazyk I Kultura-Language and Culture*, *59*, 202–226. https://doi.org/10.17223/19996195/59/11

APPENDIX

LITERATURE SOURCES NOT INCLUDED IN THE SAMPLE

Afros, E. (2014). Replying/responding to criticism in language studies. *English for Specific Purposes*, *34*(1), 79–89. https://doi.org/10.1016/j.esp.2013.09.003

Akpan, I. J., Kobara, Y. M., Owolabi, J., Akpan, A. A., & Offodile, O. F. (2024). Conversational and generative artificial intelligence and human–chatbot interaction in education and research. *International Transactions in Operational Research*, *0*, 1-31. https://doi.org/10.1111/itor.1352

Al Lily, A. E., Ismail, A. F., El-Deeb, W. M., Alghamdi, A. K., Aldalbahi, A., Alnajjar, A. O., & El-Lateef, H. M. A. (2020). 33 Tips for Arabs who wish to publish in Scopus and Clarivate-indexed journals. *Scientific Journal of King Faisal University Basic and Applied Sciences*, *21*(2), i–vii. https://doi.org/10.37575/B/SCI/1155

Alstete, J. W., & Flavian, H. (2024). Advice from the editor's desk: an introductory guide to success in education publishing. *Quality Assurance in Education*, *32*(2), 232–245. https://doi.org/10.1108/QAE-09-2023-0161

Binns, C., & Low, W. Y. (2019). Publish or the population perishes: The challenges of regional publishing in public health. *Asia-Pacific Journal of Public Health*, *31*(5), 396–403. https://doi.org/10.1177/1010539519849966

Bozkurt, A. (2024). GenAI et al.: Cocreation, authorship, ownership, academic ethics and integrity in a time of generative AI. *Open Praxis*, *16*(1), 1–10. https://doi.org/10.55982/openpraxis.16.1.654

Brunner, D. D. (1991). Who owns this work?: The question of authorship in professional/academic writing. *Journal of Business and Technical Communication*, *5*(4), 393–411. https://doi.org/10.1177/1050651991005004004

Gnutzmann, C., & Rabe, F. (2014). "That's the problem of making it sound like a Native Speaker has written it." Academic writing and publishing in the foreign language English. *Fachsprache-Journal of Professional and Scientific Communication*, *36*, 31–52.

Gnutzmann, C., Jakisch, J., & Rabe, F. (2015). Resources for publishing in English as a foreign language: Strategies, peers and techniques. In *English as a Scientific and Research Language: Debates and Discourses: English in Europe, Volume 2* (vol. 3, pp. 59-84). Walter de Gruyter.

Imani, A., & Habil, H. (2012). NNS postgraduate students' academic writing: Problem-solving strategies and grammatical features. In W. Mansor, M. H. Zakaria, A. A. Samad, & N. A. Ibrahim (Eds.), *Procedia - Social and Behavioral Sciences*, *66*, 460–471. https://doi.org/10.1016/j.sbspro.2012.11.290

Jalongo, M. R. (2013). Professional wisdom and writing for publication: Qualitative interviews with editors and authors in early childhood education. *Early Childhood Education Journal*, *41*(1), 65–79. https://doi.org/10.1007/s10643-012-0569-y

Janke, K. K., Wilby, K. J., & Zavod, R. (2020). Academic writing as a journey through "chutes and ladders": How well are you managing your emotions? *Currents in Pharmacy Teaching and Learning*, *12*(2), 103–111. https://doi.org/10.1016/j.cptl.2019.11.001

Kohl, K. E. (2011). Fostering academic competence or putting students under general suspicion? Voluntary plagiarism check of academic papers by means of a web-based plagiarism detection system. *ALT-J: Research in Learning Technology*, *19*(SUPPL.1), 90–99. https://doi.org/10.3402/rlt.v19s1/7611

Kornhaber, R., Cross, M., Betihavas, V., & Bridgman, H. (2016). The benefits and challenges of academic writing retreats: an integrative review. *Higher Education Research and Development*, *35*(6), 1210–1227. https://doi.org/10.1080/07294360.2016.1144572

McAnulla, S. J., Ball, S. E., & Knapp, K. M. (2020). Understanding student radiographer attrition: Risk factors and strategies. *Radiography*, *26*(3), 198–204. https://doi.org/10.1016/j.radi.2019.12.001

Murray, R. (2012). Developing a community of research practice. *British Educational Research Journal*, *38*(5), 783–800. https://doi.org/10.1080/01411926.2011.583635

Mustapha, M. R., Ahamad, F., Hazahari, N. Y., & Samsudin, N. (2024). Ethical issues in the halal food supply chain: A systematic bibliometric review. *Journal of Islamic Marketing*. https://doi.org/10.1108/JIMA-07-2023-0210

Oliver, S. (2015). Spanish authors dealing with hedging or the challenges of scholarly publication in English L2. In *R*. Plo Alastrué & C. Pérez-Llantada (Ed.), *English as a scientific and research language: Debates and discourses. English in Europe* (vol. 2, pp. 141-158). De Gruyter Mouton. https://doi.org/10.1515/9781614516378-009

Roulston, K. (2023). Writing with pleasure: A review. *The Qualitative Report*, *28*. https://doi.org/10.46743/2160-3715/2023.6431

Salamonson, Y., Koch, J., Weaver, R., Everett, B., & Jackson, D. (2010). Embedded academic writing support for nursing students with English as a second language. *Journal of Advanced Nursing*, *66*(2), 413–421. https://doi.org/10.1111/j.1365-2648.2009.05158.x

Seyedalhosseini, S. M., Habib, F., & Majedi, H. (2012). Interactional approach in scales and levels of urban design in urban planning process. *Bagh-E Nazar*, *9*(22), 23–32.

Sukirman, & Kabilan, M. K. (2023). Indonesian researchers' scholarly publishing: An activity theory perspective. *Higher Education Research and Development*, *42*(8), 2030–2047. https://doi.org/10.1080/07294360.2023.2209522

Zhezhera, E. (2017). Research communication skills for academic staff: Lessons of experience. *INTED2017 Proceedings*, *1*, 4421–4429. https://doi.org/10.21125/inted.2017.1048

https://doi.org/10.17323/jle.2024.23747

Text Redundancy in Academic Writing: A Systematic Scoping Review

Elena Tikhonova [1], **Daria Mezentseva** [2], **Peter Kasatkin** [3]

[1]RUDN University, Moscow, Russia
[2]Russian Biotechnological University, Moscow, Russia
[3]MGIMO University, Moscow, Russia

ABSTRACT

Background: The aim of academic writing is to effectively communicate and disseminate new knowledge and discoveries through the clear and concise expression of scientific ideas, highlighting the importance of being both brief and thorough in academic writing. The quality of this type of writing is under question. There are various sources that degrade the clarity and quality of writing. One of these aspects is redundancy, there are studies examining redundancy in written texts, however, redundancy in academic writing has received little attention. So far, there is no common understanding of the problem in academic writing, nor a common classification, nor a clear description of the causes of this phenomenon and its effects on the quality of academic texts.

Purpose: To map the existing literature on text redundancy, exploring its definitions and types, investigate the factors contributing to redundancy in academic writing, Furthermore, the article seeks to assess the impact of text redundancy on the clarity, coherence, and overall quality of academic communication.

Method: The Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR) guidelines and the "PCC" mnemonic (Population, Context, Concept) were applied for inclusion and exclusion criteria were utilized. A literature search was carried out in June 2024. Employing a detailed search strategy, the review engaged two electronic databases – Scopus and Google Scholar, initially identifying 252 studies.

Results: 65 English-language studies addressing the text redundancy were included in the review. The synthesis of the selected research revealed that redundancy is perceived differently: as a phenomenon that reduces the quality and comprehension of the text; and a strategy that makes the text understandable and explicit. Different classification of redundancy were presented: by mode of redundancy expression and repetition, by nature, and by its role and impact. The functions and impact on academic written communication redundancy were reviewed.

Conclusion: This review explores the dual nature of text redundancy in communication, particularly within academic writing. It highlights that redundancy can enhance comprehension by reinforcing key ideas or hinder communication through excessive repetition. The study classifies redundancy into three categories: functional (beneficial), wordiness (excessive), and contextual redundancy, providing a framework for writers to manage redundancy effectively. The article emphasises the importance of balancing necessary repetition with conciseness to maintain clarity and reader engagement, as excessive redundancy may lead to reader fatigue.

KEYWORDS

text redundancy, academic writing, readability, text comprehension, concise writing

Citation: Tikhonova, E., Mezentseva, D., & Kasatkin, P. (2024). Text redundancy in academic writing: A systematic scoping review. *Journal of Language and Education, 10*(3), 128-160. https://doi.org/10.17323/jle.2024.23747

Correspondence:
Elena Tikhonova
E-mail: etihonova@hse.ru

Received: June 12, 2024
Accepted: September 16, 2024
Published: September 30, 2024

INTRODUCTION

Conciseness is widely recognized as an essential characteristic of academic writing, as it enhances readability and ensures efficient communication of information (Lynn, 2016). By allowing readers to engage with critical content more easily, concise writing can increase the reach and influence of academic work

(Lynn, 2016). Alongside objectivity, formality, and clarity, conciseness forms one of the foundational pillars of effective academic writing (Chauhan, 2022). It maintains reader engagement and promotes a clear conveyance of meaning (Mu & Lim, 2022). The more explicit and straightforward the language, the easier it becomes to read and understand (Schlesinger, 1966; Baten, 1981). According to Shannon (1948), the predictability of a word within a given context affects how much information it carries; a highly predictable word contributes less new information, rendering it redundant. Therefore, it is the responsibility of the writer to craft a text that is as clear and comprehensible as possible for its intended audience (Demir, 2019; Soltani & Kuhi, 2022). Failure to do so can lead to the perception that the writer lacks expertise or confidence (Every, 2017). Effective writing requires authors to learn how to be precise and economical in their language, using only as many words as needed to convey their ideas (Abdollahi-Guilani et al., 2012; Prasetyo, 2015; Cao & Zhuge, 2022).

Writers must also anticipate the points at which readers may need additional guidance to fully comprehend the text and should incorporate supportive signals throughout (Kuhi, 2017; Dhivya & Koperundevi, 2024). This emphasis on reader comprehension aligns with the core aim of academic writing, which is to communicate knowledge to as wide an audience as possible (Demir, 2019). In this context, understanding the role and implications of redundancy in academic writing becomes critical.

The concept of redundancy in communication goes beyond simple textual repetition; it also encompasses multimodal and visual elements. However, redundancy at the textual level is often underexplored, leading to uninformative writing and ineffective communication. Existing research indicates that redundancy levels in academic writing can be significant, with studies reporting estimates of 50-75% redundancy in printed English texts (Newman & Waugh, 1960; Tuinman & Gray, 1972; Guerrero, 2009; Bazzanella, 2011; Yang, 2021). Understanding this issue is crucial for both authors and readers, as the presence of redundancy can markedly affect the efficiency and impact of communication (Bensoussan, 1990; Dhivya & Koperundevi, 2024). While all languages inherently contain redundancy, which can support successful communication, particularly in natural and imperfect conditions (Trudgill, 2009), its degree can vary based on the type and intent of the message (Marinashvili, 2020).

Recognizing text redundancy requires identifying and removing excessive information to improve the accuracy and efficiency of communication (Thadani & McKeown, 2008; Alontseva & Ermoshin, 2019; Rahman & Borah, 2021). This recognition process operates on both macro and micro levels. At the micro level, readers decode individual words and sentences, while at the macro level, they draw on prior knowledge to derive overall meaning from the text (De Beaugrande, 1980; Lotfipour & Sarhady, 2000). Texts that are easier to read may either be more effective in their micro-level cues or present challenges when redundancy at the macro level is unrecognized (Schlesinger, 1977; Bensoussan, 1990).

Despite the emphasis on clarity and conciseness in academic writing, text redundancy remains a significant and often overlooked barrier to effective communication. This challenge is exacerbated by the lack of a comprehensive review addressing how redundancy manifests in academic writing, leading to inconsistent understandings, terminologies, and classifications (Horning, 1979; Lotfipour, 1982; Xue & Hwa, 2014; Yang, 2021; Leufkens, 2023). Such inconsistencies make it difficult for writers to eliminate redundancy effectively and for readers to engage with academic texts.

This scoping review aims to address these gaps by providing a detailed analysis of text redundancy in academic writing, its characteristics, causes, and effects on the quality of communication..

Research questions

RQ#1: To study thoroughly the term “redundancy”, how researchers approach the description of this phenomenon, and identify the key characteristics of it.

RQ#2: To find in the identified sources and analyse the reasons, functions and classifications of text redundancy in academic writing.

RQ#3: To identify the impact of redundancy on academic text, its informativeness, clarity and coherence.

RQ#4: To suggest a structuralsed course on redundancy reduction in academic writing for doctoral students.

METHOD

Transparency Statement

To address our research question, we conducted a scoping review to outline the current literature. This review sought to define the extent of research done, recognize new evidence, and identify gaps in study, thus adding to the discussion in research and educational policy. We followed the PRISMA-ScR protocol. Before starting the research, a set of guidelines was established. The authors ensure that this manuscript presents an accurate, thorough, and comprehensive report of the conducted research; it addresses all important aspects of the study; and any deviations from the initial plan are properly acknowledged and justified.

Eligibility Criteria

This review was executed through a structured process that included: (1) formulating the research question; (2) identifying relevant literature; (3) selecting appropriate studies for inclusion; (4) extracting the key data from these studies; and (5) summarising and presenting the findings. The selection criteria for the literature were divided into three categories, based on the suggested mnemonic (Population, Concept, and Context) for framing research questions in scoping reviews, with an additional focus on the language, time period, geographical affiliation and type of publication (see Table 1). The research materials encompassed a range of document types, including original papers, book chapters, conference materials, editorials, unpublished doctoral and PhD dissertations, all discussing the concept of text redundancy.

Information Sources and Search Strategy

The literature search was carried out in two databases: Scopus and Google Scholar. Figure 1 illustrates the adherence to the PRISMA-ScR Protocol. First, a preliminary search was conducted in Google Scholar to identify studies relevant to the topic of interest. This search facilitated the identification of key terms and index terms closely related to the topic, focusing on aspects such as the definition of text redundancy, its characteristics, types, causes and consequences for both authors and readers. Using these terms, a comprehensive search strategy was developed and carried out on June 12, 2024.

The search terms obtained and refined by consulting relevant publications connected with the topic of interest of this study were combined using Boolean operators (OR and AND) and truncation symbols. In both Scopus and Google Scholar, the search entries were the following: “text redundancy”, “academic OR scientific AND text redundancy”, “academic writing AND redundancy”.

For Google Scholar, only the first 50 results from each query were reviewed, based on the observation that the subsequent entries are increasingly less relevant and consistent with the focus of the review. 102 studies were found in the Scopus database. Additionally, the reference lists of selected studies were examined to uncover further pertinent research. From this search, 252 studies were retrieved (as shown in Figure 1).

Selection of Sources of Evidence

The sourced references’ titles were organized in a Zotero library, and duplicate entries were eliminated using a reference management tool. The library, which includes all relevant titles, was systematically examined by two reviewers in separate phases: (1) screening based on title and abstract, (2) evaluation of the full text. Consensus meetings were conducted at each phase to discuss studies that met the inclusion criteria. Any disagreements among reviewers were addressed through consultation with a third reviewer.

Through the initial screening of titles and abstracts, 101 studies were excluded based on the predefined criteria. The review of the remaining 151 studies led to the removal of those unavailable or duplicated and exclusion of 80 studies that did not meet the inclusion criteria, leaving 55 studies for incorporation into the survey. An additional 10 sources were identified during the reference list screening process. Finally, 65 sources were included in the scoping review (see Appendix 1).

Data Charting Process

Data extraction was conducted by two independent reviewers, making the process more objective and ensuring that nothing of substance was overlooked. The complete data set extracted by one reviewer was then cross-checked by the other to ensure accuracy and consistency. Any differences between reviewers were addressed and resolved through consensus meetings. To systematically organise the extracted data, a standardised Excel spreadsheet was created. This spreadsheet captured a range of data points critical for our analysis, including: the name of the institution involved in the study; geographical coverage of the document; publication year of the document; objectives and a brief description of the document; the target population addressed by the study; definitions of text redundancy; characteristics of text redundancy; types of text redundancy; factors influencing text redundancy; consequences of text redundancy for both authors and readers; tools and strategies for reducing text redundancy. This structured approach facilitated a comprehensive and systematic review of the literature, enabling the authors to identify and synthesise key findings related to text redundancy in academic writing.

Summarising and Reporting the Results

Following the data charting phase, the same reviewers synthesised information concerning each aspect of the text redundancy phenomenon identified during the charting stage. The terminological ambiguity encountered during the source selection phase necessitated a detailed analysis of the identified definitions of text redundancy to extract their core characteristics. These characteristics were essential for establishing a consensus definition of text redundancy. The identified definitions of term “text redundancy” were organised in sequentially numbered Microsoft Word documents. The coding process, conducted by the first and second authors, followed the methodology proposed by Braun and Clarke (2006). Initially, the first author examined the text to generate a preliminary set of codes. These initial codes were then reviewed and refined in collaboration, leading to the development of potential themes. Subsequently,

Table 1
Eligibility Criteria

Criterion	Inclusion	Exclusion	Justification
Population	All the studies describing text redundancy in academic context, such research may involve university teachers, students, research staff, university administration and educational programmes (compulsory and elective).	All the studies outside the defined field.	This scoping review focuses on text redundancy and all participants in academic writing and text comprehending.
Concept	The concept of redundancy in the current review includes the study of redundant information in different forms (texts, educational programmes, university administration / organisation activities) and its impact on the effectiveness of academic communication.	Studies which do not relate to the concept of text redundancy.	The research on duplication of information in various forms and its impact on cognitive load, comprehension and quality of interaction in educational and research environments.
Context	The context of this review is higher education and research activities in higher education institutions. The studies cover the discourse in different countries and consider aspects such as academic writing, courses on its development, its representation in academic texts of different genres.	Studies outside writing context.	The focus of the research is text redundancy in academic writing. Anything beyond that would be unmanageable with the resources available for verification.
Language	English	Any other languages	English is the international language of scientific communication.
Time period	1948-2024	None	The aim is to get all the information of the 20th and 21st centuries. There is little information available about text redundancy, and even less that can be classified as academic writing.
Types of sources	Any types	Unavailable sources	Gathering all the sources possible
Geographical affiliation	Any location	None	Getting international perspective

each researcher independently performed a thematic analysis using these codes, a step critical for ensuring a thorough and impartial evaluation of the data.

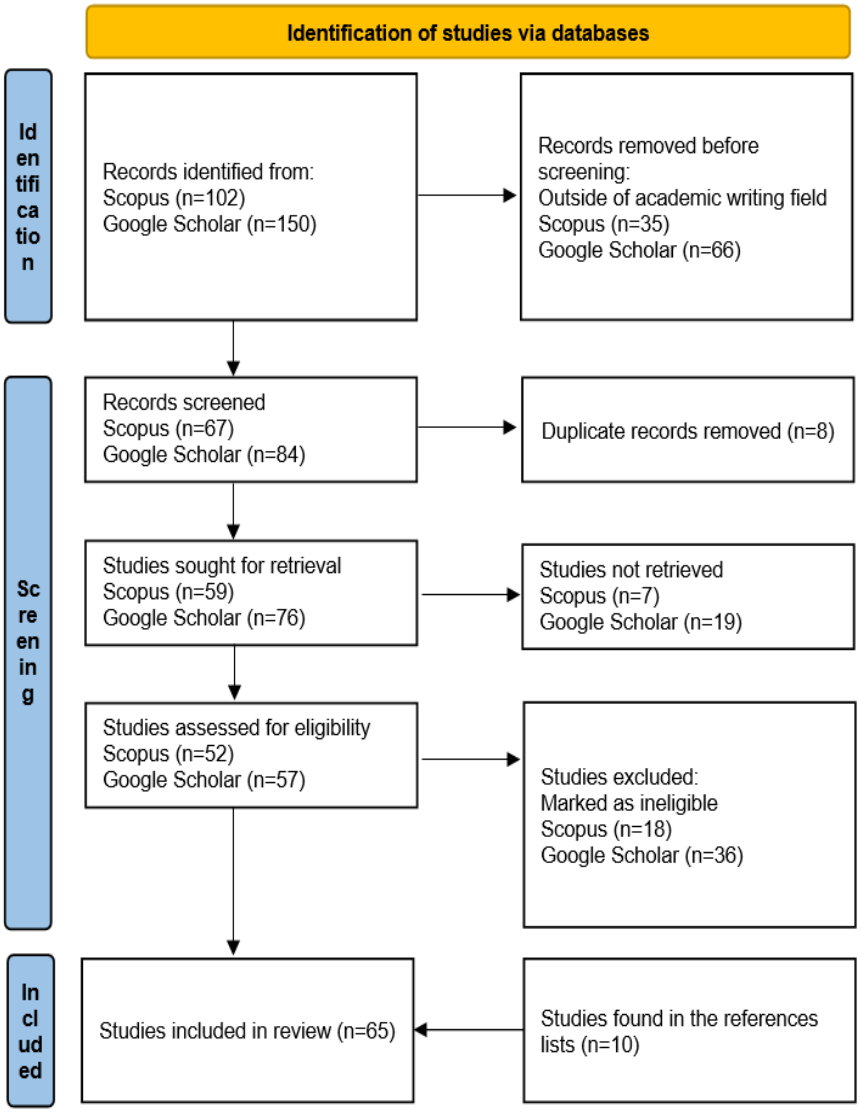
Discussions between the coders achieved over 91% consensus on the themes, codes, and references, indicating a high degree of inter-coder reliability. Any differences were addressed through detailed comparison and dialogue, which led to the modification of some codes and the reclassification of certain themes. A second round of coding was conducted based on these adjusted themes, further refining the analysis. Similarly, the factors leading to and the implications arising from the excessive incorporation of superfluous words in scholarly texts, as identified by the researchers of the studies encompassed in this review, were systematically coded and conceptualised.

Data Visualisation

The metadata of the articles included in the review were processed using VOSviewer, a software tool for constructing and visualising bibliometric networks, was utilized to process the metadata of the articles in the review. This software makes it easier to visualize connections between the sources being analyzed and helps to find clusters in the research. With VOSviewer, it is possible to visually identify the primary research directions, displaying the connections between topics and indicating which areas have received more attention.

Using the VOSviewer software we defined the co-occurrence of keywords related to the topic of «redundancy»(see Figure 2). The nodes represent different keywords, and the edges depict the number of times they appear together in the same document. The thicker the edge, the more often the keywords appear together. Even though in the review

Figure 1
PRISMA-ScR Protocol



we included the studies only focused on text redundancy, the keywords reveal that other types of redundancy are also common, as shown by the clusters we identified. The keywords are divided into four groups marked by different colours: (1) Concepts directly related to redundancy, such as "cognitive load theory" and "multimedia learning" (green); (2) Keywords related to experimental design or research methodology, such as "accommodation" and "experimental pragmatics" (blue); (3) Keywords related to reading comprehension and attention, such as "split attention", "visual display", and "spatial contiguity" (yellow); (4) Keywords related to text processing and communication, such as "genre", "source text", "ellipsis", and "strategies" (red). It suggests that redundancy is a complex concept that is related to a variety of topics.

Figure 3 displays the research trends on redundancy throughout the years. The nodes correspond to various ideas linked to redundancy, while the edges indicate how many publications mention both concepts. The edge's color reflects the publication year, with blue indicating topics being in a research scope longer and yellow indicating more recent research trends. It indicates that the research focus on redundancy has been changing over the years. The studies examining how redundancy impacts reading and visual presentation have been studied much longer, and the studies concentrating on how redundancy is involved in multimedia learning and cognitive load theory has appeared recently.

Figure 4 shows a word cloud that visualizes the most frequently discussed concepts in a body of research related to redundancy. The size of each word indicates how frequently that concept was mentioned in the identified sources. The largest cluster of words is centered around the key concept in this research. The other clusters of words suggest that redundancy is often discussed in relation to text processing ("source text ellipsis", "genre", "strategies", and "media headlines"), cognitive load ("cognitive load theory", "multimedia learning", "split-attention", "reading", "spatial contiguity", "visual display", and "a cognitive theory of multimedia") and pragmatic ("accommodation" and "experimental pragmatics"). No cluster has been formed specifically for text redundancy, and only one of the main clusters presented by the VOSviewer programme is close to text redundancy, indicating that it has not been studied enough.

An Overview of the Selected Articles

We analised demographic characteristics of the sources chosen for the review. Figure 5 displays a visualised net of the year-wise and Figure 6 summarises country-wise distribution of the included sources. Among the included sources, 27 studies were published in the 20th century, starting from 1948, and 38 studies were released during the last 24 years. In the middle of the 20th century there was little interest in the subject. The interest in the topic of redundancy began to grow in the 1970s. The interest increased significantly in the 21st century, particularly in the 2010s, when 16 studies were published, nearly a quarter of all papers found. This suggests that the academic community has been focusing on redundancy and how it affects scientific communication for the last few decades.

The topic received contributions from a combined 23 countries (see Figure 5). Approximately one-third of the research papers (n=23) were released in the USA. 9 countries exhibited comparable and relatively low engagement in the text redundancy with 2-7 studies. The other countries, such as Bahrain, Belgium, Colombia, France, Hungary, Indonesia, Iran, Poland, The Netherlands, Turkey, Ukraine and Uzbekistan, all had an equal contribution (1 study each), indicating a very minimal interest in the issue of redundancy.

RESULTS

Definitions of "Text Redundancy"

Identified Definitions

First, we want to address the problem of terminology. There are different interpretations of the term "redundancy" by scholars, and there is no unanimous understanding of the term in the research. The analysed sources identified that redundancy can be observed from two perspectives. Two concepts of this term are frequently mentioned by researchers with the opposite connotations: (1) excessive redundancy (see Table 2); and (2) beneficial redundancy (see Table 3).

After a thorough analysis of the definitions of excessive redundancy, we can define that it is the unnecessary repetition of information beyond what is needed for clarity, often involving extra words or repeated expressions. Excessive redundancy can make communication awkward and inefficient and be identified when removing certain words or phrases doesn't change the sentence's meaning. While beneficial redundancy (see Table 3) ensures information is still

Figure 2

Co-Occurrence of Keywords and Clusters

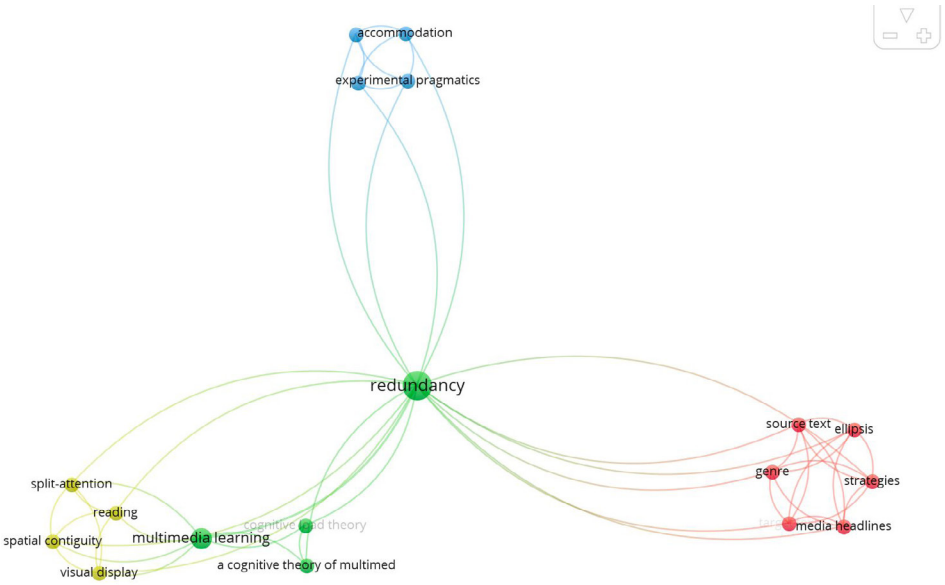


Figure 3

Research Trend on Redundancy Over Time

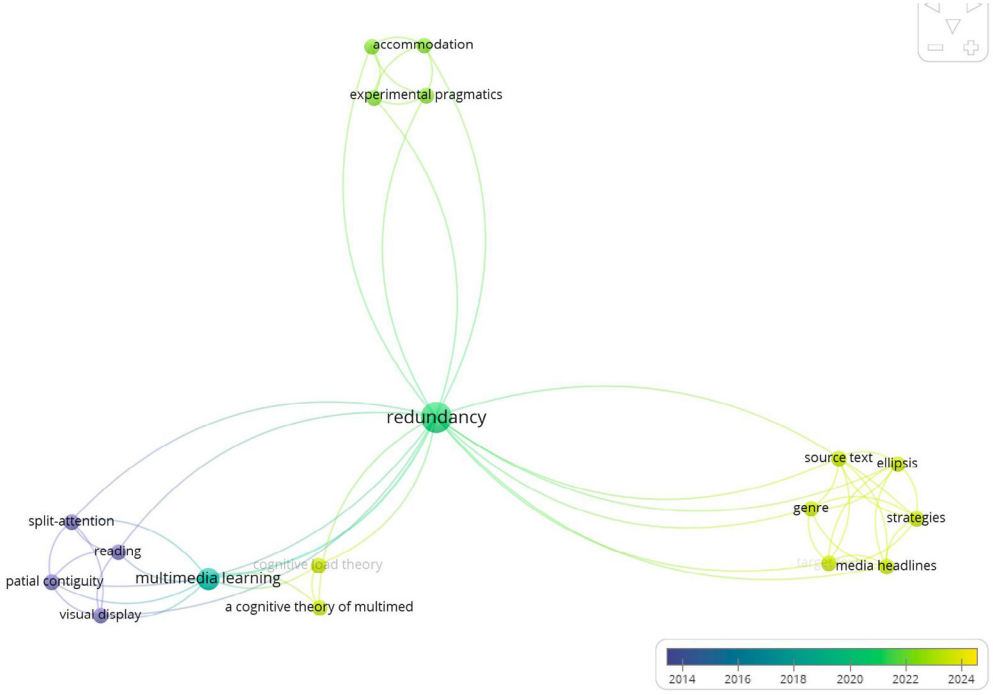


Figure 4

Density of Text Redundancy Concepts

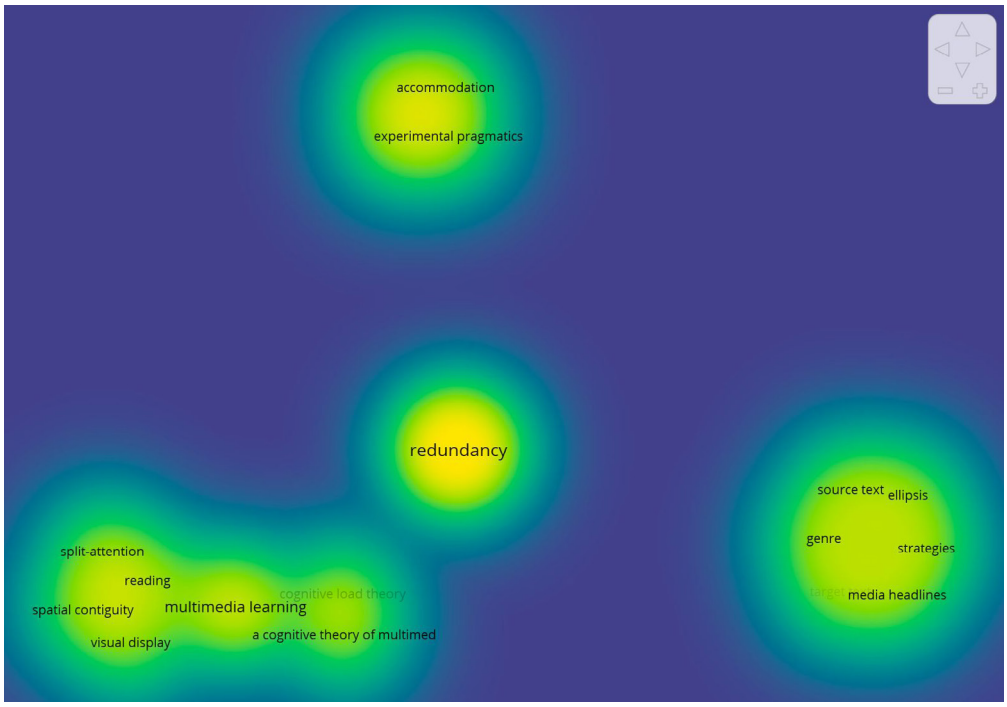


Figure 5

Publication Years of the Included Studies

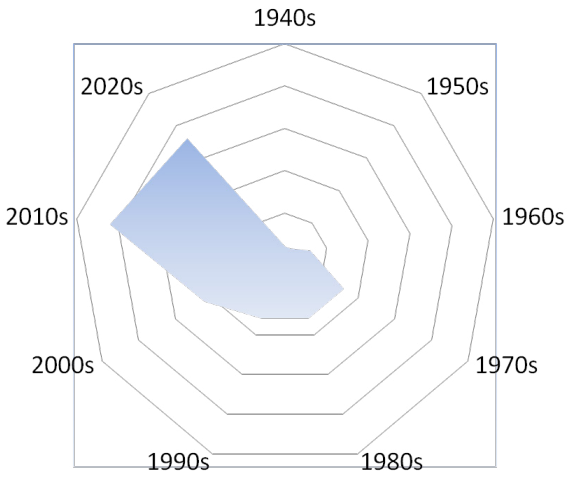
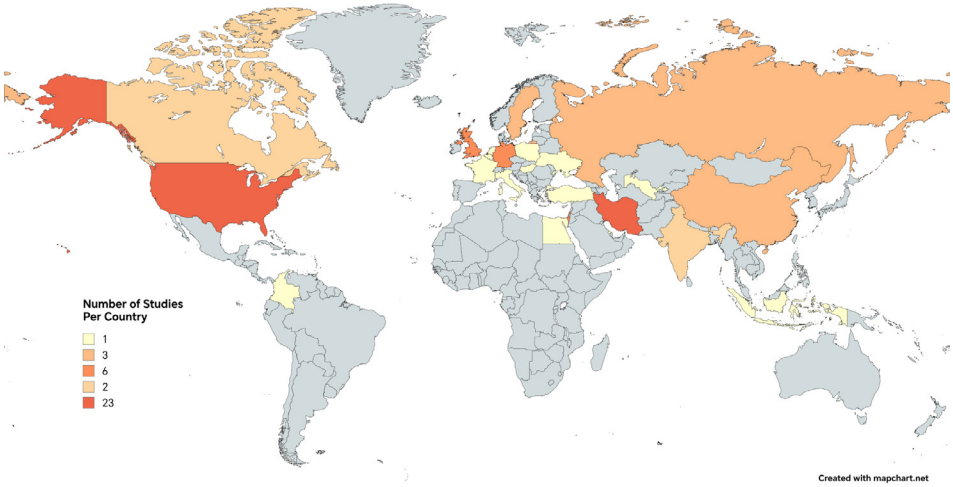


Figure 6

Geographic Affiliation of the Authors



conveyed if parts are lost. It repeats concepts through various channels to ensure message clarity and reliability. This redundancy compensates for reader inattention, enhances text predictability, and facilitates information processing. By providing more information than minimally necessary, redundancy ensures effective communication and strengthens the connection between readers and writers.

Redundancy Characteristics

Beneficial redundancy ensures information is conveyed even if parts are lost, but excessive redundancy can make text awkward and impede efficient knowledge transfer. It is often identified when removing certain words or phrases still results in a sentence that conveys the same meaning. We identified the characteristics that collectively illustrate how excessive redundancy complicates communication, detracts from clarity, and may hinder effective information transmission (see Table 4).

When information is duplicated across multiple sources or cue systems, it leads to unnecessary overlap that complicates rather than clarifies communication. Excessive redundancy poses significant challenges that can undermine effective information exchange, while beneficial redundancy can play a crucial role in reinforcing communication and enhancing clarity. The characteristics that highlight the multifaceted role of redundancy in enhancing communication effectiveness and ensuring message clarity are presented in Table 5.

Beneficial redundancy is essential for enhancing communication, as it reinforces key messages through repetition.

Table 2

Definition of Excessive Redundancy

Source	Definition
Shannon, 1948, 1951	"mutual information"; "the difference between the entropy of the messages actually transmitted and the maximum entropy that the channel could transmit"
Klare, 1963	"the extent to which a given unit of language is determined by nearby units"
McGarry, 1975	"the complement of entropy (uncertainty), i.e. as entropy rises, redundancy decreases, and vice versa"
Horning, 1979	"redundancy exists whenever information is duplicated by more than one source, and in the case of reading, information is duplicated by at least four sources or cue systems"
Darian, 1979	"information whose meaning may be predicted or limited by other information in the discourse"
Hunnicutt, 1985	"the information in a complete sentence over and above that which is essential"
Smith,1971	"there is redundancy whenever the same alternatives can be eliminated in more than one way"; "whenever information is duplicated by more than one source"
Smith, 1978	"the reader's prior knowledge … the reader could not perceive the redundancy in a written text unless it reflected knowledge already present in the reader's mind"
Rosie, 1973	"anything other than the minimum required to represent or transmit information is considered redundant"
Forlini et al., 1982	"the unnecessary repetition of an idea"
Dawson, 1992	"the use of more words than are necessary to express a thought, especially the use of two expressions that mean the same thing"
Grant-Davie, 1995	"a kind of linguistic cholesterol, clogging the arteries of our prose and impeding the efficient circulation of knowledge"
Lehmann, 2005	"a message is redundant if it contains such elements which contribute nothing to the information not already conveyed by the rest of the message"; "repeating an utterance"
Every, 2017	"the excessive use of unnecessary words"; "the repetition of the same idea in different words, or tautology (in a narrower sense)"
Wolf et al., 2023	"an information-theoretic measure that quantifies the amount of information obtained about one random variable (e.g., prosody) by observing the other random variable (e.g., text)"
Lotfipour & Sarhadi, 2000	"a feature is redundant if its presence is apparently unnecessary"; "a piece of information is redundant if it is reiterated (in any mode) in relation to features or pieces of information occurring before it in the text"
Bazzanella, 2011	"redundancy measures how much the information transmitted from the source differs from the maximum possible information, given the same set of symbols"
Trudgill, 2011	"multiple expressions of a single meaning within the same phrase or clause"
Xue & Hwa, 2014	"some extraneous word or phrase that do not add to the meaning of the sentence but possibly make the sentence more awkward to read"; "consider a word or a phrase to be redundant if deleting it results in a fluent English sentence that conveys the same meaning as before"
Heltai, 2018	"the expression of the same information or meaning component more than once, or the overt expression of a piece of information or a meaning component that is considered self-evident and best left implicit, to be inferred from other linguistic items in the given piece of discourse or from the situation and/or general world knowledge"
Alontseva & Ermoshin, 2019	"how much the length of a text in a given language can be reduced without losing any part of the information"
Rasulov & Artikov, 2023	"the expression of simple content in compound sentences"; " a language unit that does not have any semantic load in the text, that is, does not perform any task"

Source	Definition
Marinashvili, 2020	"the excessive information (in other words repeated or unnecessary information), defined as percentage content of excessive information in the texts of a given language";
	"information may be discarded from the text without the harm to its meaning and easily restored as it is determined by the structure of the language itself"
Yang, 2021	"the part of the information that is more than the minimum required is redundant"
Kravtchenko & Demberg, 2022	"more information is provided than needed to recover the intended meaning or world state"
Dhivya & Koperundevi, 2024	"the unnecessary repetition of words or elements within a sentence that do not contribute substantially to its meaning"

Table 3

Definition of Beneficial Redundancy

Source	Definition
Darian, 1979	"a method of reinforcing communication, by repetition of concepts through linguistic or nonlinguistic channels"
Zola, 1981	"a measure of certainty"
Lotfipour, 1982	"redundancy as a textual strategy functions in two ways: compensating for the attention failure of the reader and neutralizing the linearity of the text"
Horning, 1991	"the characteristic of written language that helps ensure that the reader gets the message, so it is another means by which readers and writers can connect in text"
Lehmann, 2005	"additional explanation in demanding communicative conditions, or a poetic function"
Bazzanella, 2011	"the repetition of relevant information"
Marinashvili, 2020	"one of the factors increasing the reliability of received information"
Yang, 2021	"excessive information provided in information transmission beyond a minimum amount to ensure the effectiveness of communication"
Leufkens, 2023	"the expression of a single meaning by means of both a lexical and a grammatical element";
	"facilitates processing and acquisition, as it increases the predictability and robustness of the sentence"

This redundancy increases the reliability of the information conveyed, fostering trust and clarity in the exchange. However, it is crucial to strike a balance between beneficial redundancy and excessive redundancy to maintain effective communication. By doing so, we can ensure that messages remain clear, concise, and meaningful.

Classifications of Redundancy

We have not been able to capture a unified perception regarding the types of redundancy. In this review we observe different approaches to classifying text redundancy in a language in general which can be applied to academic writing as well. We identified redundancy classifications based on different methods which can be divided into three groups: (1) by mode of redundancy expression and repetition; (2) by nature; (3) by its role and impact. To the first category we can assign the following: Yang (2021) classifies redundancy by parts of speech (see Table 6); Grant-Davie (1995) differentiates this phenomenon by redundant phrases it appears in (see Table 7); Lehmann (2005) divides it into four types by repetition types (see Table 8).

Grant-Davie (1995) classified redundancy into 5 types of redundant phrases it appears in: redundant pairs, redundant modifiers, redundant categories, phrases used where words would do, and empty sentence openings.

Lehmann (2005) identified 4 types of redundancy: pleonasms, tautology, repetition and hypercharacterization.

The second category of redundancy classification is based on the nature principle. Some agreement has been achieved among researchers (Wit & Gillette, 1999; Xu, 1984; Rasulov & Artikov, 2023), classifying it into linguistic (grammatical) redundancy and non-linguistic (contextual) redundancy, researchers assume a dichotomy based on the causes for emergence whether the redundancy is generated from grammatical rules or the contextual need (see Table 9). Grammatical redundancy is internal to the language system, is systematic and obligatory, whereas contextual redundan-

Table 4

Characteristics of Excessive Redundancy

Characteristic	Meaning
Multiple Sources of Information	Excessive redundancy occurs when information is duplicated across multiple sources or cue systems, resulting in unnecessary overlap (Horning, 1979; Smith, 1971).
Predictable and Non-contributory Information	Redundant information may be predictable or limited by surrounding discourse, indicating a lack of new value. It includes elements that exceed what is essential for conveying meaning, complicating communication (Rosie, 1973; Darian, 1979; Hunnicutt, 1985; Lehmann, 2005).
Unnecessary Repetition	Excessive redundancy involves the unnecessary repetition of ideas, phrases, or words that do not enhance overall meaning, including tautological expressions (Forlini et al., 1982; Dawson, 1992; Grant-Davie, 1995; Xue & Hwa, 2014; Every, 2017).
Length Reduction	The extent of redundancy can be quantified by how much a text's length can be reduced without losing essential information, indicating excessive information (Alontseva & Ermoshin, 2019; Marinashvili, 2020).
Semantic Load	Redundant language units lack semantic weight and do not contribute meaningfully to the text, sometimes providing more information than needed to recover the intended meaning (Rasulov & Artikov, 2023; Kravtchenko & Demberg, 2022).

Table 5

Characteristics of Beneficial Redundancy

Characteristic	Meaning
Reinforcement of Communication	Redundancy serves as a method to reinforce communication by repeating concepts, which can occur through both linguistic and nonlinguistic channels (Darian, 1979; Lehmann, 2005).
Measure of Certainty	It acts as a measure of certainty, helping to ensure that the intended message is conveyed clearly (Zola, 1981; Horning, 1991; Bazzanella, 2011).
Compensation for Attention Failures	Redundancy functions as a textual strategy that compensates for potential attention failures of the reader, helping to maintain comprehension (Lotfipour, 1982).
Increased Reliability of Information	It contributes to the reliability of the information received, making it more trustworthy and clear (Marinashvili, 2020; Yang, 2021).
Facilitation of Processing and Acquisition	Redundancy facilitates the processing and acquisition of information by increasing predictability and robustness in sentences (Leufkens, 2023).

cy is voluntary. Contextual redundancy involves the judgment of the speaker concerning the receptor's background or it may simply be used to achieve a certain rhetorical effect (Wit & Gillette, 1999).

One more redundancy classification based on nature is represented by the study of Albers et al. (2023) who divided redundancy into content redundancy and modal redundancy (see Table 10).

In the third type of redundancy classifications Lotfipour & Sarhady (2000) distinguish text redundancy based on their role in the text and their impact on understanding and perception of information (see Table 11).

Text redundancy classifications encompass the various methods by which information can be repeated or reinforced within a text. By understanding these classifications, writers can analyze language use more effectively, allowing them to identify instances of unnecessary repetition or, conversely, purposeful reinforcement. This awareness not only aids in improving the clarity of the text but also enhances overall communication effectiveness, ensuring that information is conveyed in a manner that is both engaging and easily comprehensible for the reader.

Sources of Redundancy

Redundancy in texts and communication is a multi-level phenomenon that can manifest at various levels of language and under different circumstances. Researchers identify several key sources of redundancy that are important to consider when analyzing written and spoken texts. First, there is visual information, which refers to the text itself as it appears on the page. Redundancy at this level can manifest through the overuse of headings, subheadings, or repeated formatting elements that do not add meaning but may confuse the reader or complicate the perception of the

Table 6

Redundancy Classification Based on Parts of Speech It Appears in

Type	Example
Noun Redundancy	"I should like to apply for *a secretary job*."
Pronoun Redundancy	"Students have to pay *their* tuition fees *themselves*."
Verb Redundancy	"There are other interesting and special means of transportation *remained* in China."
Adjective Redundancy	"There are *all kinds of different* Zongzi all over the country."
Adverb Redundancy	"A lot of people *compete together* for one job, of course some of them will lose."
Adverb Redundancy	"In China, our situation is *relatively better.*"
Coordinating Conjunction Redundancy	"She worked hard *yet* without complaining."
Subordinating Conjunction Redundancy	"Most of the students *who* studying abroad did not want to come back."
Preposition Redundancy	"Some students will even quit the jobs they just find and *to* seek another."
Article Redundancy	"If you have courage of facing *the* all kinds of the difficulties, studying abroad is a good opportunity for you."
Auxiliary Verb Redundancy	"Many people *are* regret that they didn't get degrees when they were students."

Note. The table is based on the information from Yang (2021).

Table 7

Redundancy Classification by Redundant Phrases It Appears in

Type	Example
Redundant pairs	*"benefits and advantages"*
Redundant modifiers	*"mandatory requirement"*
Redundant categories	*"rectangular in shape"*
Phrases used where words would do	*"at this point in time» instead of «now"*
Empty sentence openings	*"There is a strong likelihood of rain tomorrow."*

Note. The table is based on the information from Grant-Davie (1995).

information (Smith, 1971; Bartell et al., 2006; McCrudden et al., 2013).

Second, orthographic information pertains to how readers rely on their knowledge of spelling and letter sequences to predict the next elements of a text. For example, in languages with regular orthographic rules, such as English, knowledge of typical letter combinations helps to anticipate words, which can create predictability and, thus, redundancy (Smith, 1971; Chetail, 2015; Staub, 2015). Syntactic information is related to how sentence structure creates redundancy. Grammatical rules allow us to predict the next element in a sentence, especially in languages with a fixed word order. This can lead to the repetition of information at the syntactic level when the word or phrase structure does not add new information but merely confirms already known data (Smith, 1971, Wit & Gillette, 1999; Berdicevskis, 2015).

At the semantic level, redundancy occurs when the context provides sufficient cues for the readers to easily predict the meanings of words or phrases, rendering additional explanation superfluous. This phenomenon is particularly prevalent in academic texts, where authors might feel compelled to elaborate on concepts that are already clear from the surrounding context. For instance, unnecessary elaboration can take the form of reiterating obvious points that the target audience is likely to understand without further clarification. Additionally, the use of synonyms in place of a single, precise term can create a sense of redundancy that detracts from the overall clarity of the writing. Such practices not

Table 8

Redundancy Classification by Repetition Types

Type	Explanation	Example
Pleonasms	In general, a pleonastic expression contains constituents – typically two – one of which implies – technically: entails – the other. Thus, the meaning of the latter constituent is part of the meaning of the former; the meaning of one constituent entails the meaning of the other without being identical to it.	*"return back"*
Tautology	It refers to a proposition that is always true independently of the truth values of its constituents.	*"each and every", "necessary and unnecessary", "null and void", "enough is enough", "business is business", "It will rain or it will not rain."*
Repetition	The synonymous elements are identical.	*"This is totally impossible - totally impossible."*
Hypercharacterization (reinforcement)	The focal component is expressed by an inflectional or derivational morpheme.	*"more easier"*

Note. The table is based on the information from Lehmann (2005).

only dilute the impact of the message but also contribute to a more cumbersome reading experience, potentially frustrating readers who seek concise and direct communication (Smith, 1971; Bodenreider, 2003). Moreover, excessive semantic redundancy can hinder the flow of arguments, making it challenging for readers to follow the author's line of reasoning. In academic writing, where precision and clarity are paramount, recognizing and minimizing semantic redundancy is essential for effective communication and ensuring that the core ideas are conveyed with maximum impact.

Redundancy also exists at the phonetic and morphological levels. For example, the addition of affixes that convey the same information already present in the root of a word can create unnecessary repetition of meaning. Additionally, certain morphological constructions may include multiple affixes that, while grammatically correct, do not contribute new information and can clutter the communication (Stanley, 1967; Darian, 1979; Caballero, 2014).

At the discourse level, redundancy can be observed through the repetition of the same thoughts or ideas in different parts of a text, which can significantly impact the overall coherence and effectiveness of the writing. This phenomenon is particularly common in classroom contexts, where students may reiterate points made earlier in their presentations or written assignments, often in an attempt to emphasize their arguments. Similarly, in academic articles, authors may inadvertently duplicate their thesis or central arguments without providing additional insights or perspectives. This redundancy not only reduces the overall effectiveness of the writing but can also make the text more challenging to process for readers, who may struggle to discern the key contributions of the work amidst the repetition. When authors reiterate the same points without adding depth or nuance, they risk losing the attention of their audience and undermining the impact of their arguments (Darian, 1979; Lyster, 1998; Freywald, 2018). Moreover, excessive redundancy at the discourse level can create a sense of monotony, making it difficult for readers to maintain their focus and interest. In academic writing, where clarity and precision are essential, it is crucial for authors to be vigilant about avoiding unnecessary repetition.

Another important source of redundancy is stylistic errors, which can manifest in various forms, such as the excessive use of linking words or referential expressions. Linking words, or transition phrases, are essential for guiding readers through the flow of ideas within a text; however, when they are overused, they can create a convoluted narrative that hinders comprehension. Violations of the formal connection between sentences can exacerbate these issues, leading to a disjointed reading experience. In scientific texts, where precision and clarity are paramount, authors may inadvertently create gaps in logic or coherence by failing to establish clear relationships between their statements. This can occur when sentences are poorly structured or when the logical flow is disrupted by the inclusion of superfluous information. Such stylistic errors not only reduce the readability of the text but can also undermine the credibility of the authors, as readers may perceive them as less rigorous in their writing (Buscail & Saint-Dizier, 2009; Alontseva & Ermoshin, 2019).

Redundancy can also be closely related to content, particularly in the use of redundant modifiers or descriptions that do not contribute any new meaning but merely reiterate ideas that have already been expressed. For example, this can be seen in excessive explanations, where the meaning of the modifiers completely or partially overlaps with the meaning of the main word (Grant-Davie,1995; Lehmann, 2005; Yang,

Table 9

Redundancy Classification by Causes for Emergence

Type	Grammatical redundancy	Contextual redundancy
Definition	The internal systematicity and rule governed behavior of a language in which two or more of its features serve the same function. It is internal to the language in the sense that it is generated from grammatical rules and is independent of situational, contextual and nonlinguistic considerations; it is truly redundant since it serves only to repeat information already given by another feature.	This repetition consists of the reproduction of identical elements of information or of elements that are only apparently identical. Contextual redundancy is not systematically generated by grammatical rules, although nongrammatical circumstances may suggest or require its use. Such circumstances include sociolinguistic and psycholinguistic factors. Unlike grammatical redundancy there is not one kind of contextual redundancy and a subcategorization can be made on the basis of the structure of the redundant expression.
Categories	The English –s	Identical or synonymous repetition
1.	English requires the morpheme -s to mark third person singular verbs in the present tense. Since English is not a 'pro-drop' language, the presence of an expressed subject makes the -s morpheme redundant. That morpheme, nevertheless, is obligatory. According to the grammatical rules of English, the speaker may not use the -s in some contexts and omit it in others. The -s morpheme is semantically superfluous since it offers no more information than is already expressed by the subject of the sentence	This kind of redundancy occurs when the expression contains two (or more) identical or synonymous words or subexpressions. *"Last year I visited the Eiffel Tower, the tallest steel construction in the center of Paris."* From the examples it is clear that the redundant expressions often do carry a semantic goal.
2	Questions Most sentences have at least two features that indicate the interrogative nature of the expression. English clearly has a backup system for ensuring that certain utterances are understood as questions. Information questions: (1) (a) *"How is your mother?"* The interrogative markers in this sentence are: (1) Interrogative word: "how." (2) Subject-predicate inversion. (b) *"Where did you buy that car?"* In this example the interrogative markers are: (1) Interrogative word. (2) Introduction of the auxiliary, "did." (3) Subject-auxiliary inversion.	Isolating, salient repetition An isolating redundant expression contains at least two subexpressions, of which one implicitly contains one or more features or characteristics of the other. *"I love the salty sea."*
3	Spelling The rules of spelling function typically operate in written expressions by conforming to a pattern of expectation in the reader a uniform spelling increases the redundant coding of an expression and thereby increases the comprehensibility of the utterance or written text. *"evereewan shoot edher too the saim spellin" = "Everyone should adhere to the same spelling."*	Contrasting repetition Contrasting redundancy occurs when two (or more) words or expressions that semantically constitute a contrast are repeated or in some other way redundantly coded. *"Although his parents are Asian, his eyes are blue and not dark."*
4	Word order Although it may be harder to recognize word order as a form of redundancy, the word order of a sentence constitutes one of the most important linguistic coding systems besides the words and expressions themselves. *"Her book the he gives."* It does present the information (i.e., what is the subject, what is the indirect object, etc.) in a more accessible manner, simply by conforming to the expectations that the receptor has of a sentence	Distinguishing, differentiating repetition A form of repetition of information in a context of differentiating one object from another. Many words or expressions that are not ambiguous in one context, may be ambiguous in another. In order for contextual redundancy to occur it requires a context with possible alternatives besides the one being singled out in the expression. *"I am looking at the monkey in the group with the red boundary."*

Type	Grammatical redundancy	Contextual redundancy
5	Double negotiation Double negatives introduce a redundancy in the sentence that reduce the possibility of a mistake. *"I can't give you no money"*	
6	Concordance of adjectives and articles with noun in gender and number (does not apply to the English language)	
7	Indirect object pronoun redundancy (does not apply to the English language)	

Note. The table is based on the information from Wit & Gillette (1999).

Table 10

Redundancy Classification by Nature

Type	Explanation	Function
Content redundancy	Content redundancy occurs when the same information is presented more than once. This includes any situation in which multiple sources present the same information, irrespective of the combination of sources, such as animation and written text, animation and narration, or written text and narration	Content redundancy enhances learning and decreases cognitive load.
Modal redundancy	Modal redundancy occurs when multiple information is concurrently presented in the same mode (auditory or visual), resulting in an excessive load in either the auditory or the visual channel. Since modal redundancy does not presume a contentual overlap, it can occur in combination with content redundancy or on its own. As an example, modal redundancy occurs whenever animation or narration is accompanied by written text, irrespective of its content	Modal redundancy harms learning and increases cognitive load.

Note. The table is based on the information from Albers et al. (2023).

2021). When authors include redundant modifiers, they risk diluting the impact of their message, as readers may become bogged down by the repetition and lose sight of the core ideas being presented.

Finally, interference between languages can be a significant source of redundancy, particularly for second-language learners (L2). When individuals are acquiring a new language, they often rely heavily on their native language as a reference point. This reliance can lead to the practice of literal translations, where phrases or sentences are directly translated without considering the nuances and idiomatic expressions of the target language. Such an approach can result in awkward phrasing and redundancy, as learners may inadvertently replicate structures or expressions that are common in their native language but do not convey the same meaning or efficiency in the new language (Heltai, 2018; Yang, 2021; Al-Qaddoumi & Ageli, 2023).

Reasons for Text Redundancy

Redundancy in texts can arise from various functional needs and cognitive considerations, making it a crucial component of both written and spoken communication. Scholars have identified several key reasons for the presence of redundancy in texts, particularly in academic and technical writing. Redundancy plays an important role in compensating for readers' attention lapses and addressing the linear nature of text. Lotfipour-Saedi (1982) emphasizes that since human cognitive capacity is limited, readers may struggle to focus on all concepts in a text. As a result, writers often repeat key ideas in different ways. This repetition ensures that even if the reader misses some information, it can be recovered later in the text. By providing more information than strictly necessary, redundancy protects against comprehension failure and helps readers process complex or dense material more effectively.

Table 11

Redundancy Classification by Its Role and Impact

Type	Explanation	Example
Exact Repetition	Not confined to mere words, but they include exact repetition of groups, the same patterns, clauses, clause complexes (sentences). While exact repetition as a textual strategy in general and as one manifestation of redundancy in particular may not be tolerated in some texts, it can engrave some vital effects such as expressive and aesthetic aspects of communication, emphasis, rousing the interest of the addressees, and excitement on the receivers in some others.	"A muscle [*contracts*] extremely rapidly when it *contracts* against no load - to a state of full contraction in approximately 0.1 seconds for the average muscle" (Gayton, 1985).
Functors	Grammatical words without any meaning by themselves, those which replace the presupposed items in the context. This group of words occur with the highest frequency in all different texts and this can be related to the economy principle in language. The functors have the potentiality to replace a word, a group, a clause, a sentence, etc. Reiteration of the words without any modification may make a text boring and overredundant; hence, the use of function words can counteract this effect. The references cannot be effective beyond some limited spans in text; otherwise, they result in confusion and ambiguity, i.e., the more the distance between the presupposing and presupposed items, the more restrictions on the use of references.	Pronouns and demonstratives, bound morphemes (e.g., *"re-"*), and definite articles.
Semantic Redundancy		
Grammatically Undeletable Redundancy(GUR)	Reiteration is fulfilled by content words including synonyms, antonyms, general words, different parts of speech, comparisons, and different codes. One synonymous lexical item can refer back to another, to which it is related by having a common referent.	"For instance, they contain a large quantity of carbonic anhydrase, which [catalyzes] the reaction between carbonic dioxide and water, *increasing* the rate of this reaction many thousand fold" (Gayton, 1985). "Before treating a patient for any disease, the physician must [find out] what the disease is. In other words, he must diagnose the disease"(Gayton, 1985).
Grammatically Deletable Redundancy (GDR)	The elaboration of a piece of information as distinct from what we have presented so far. It seems to be right in assuming that the elimination of the previous modes of redundancy is not permissible in terms of both grammar and meaning. Most of the realizations of GDR can be omitted without impairing the text grammatically.	
Grammatically Deletable Redundancy (GDR)		
Reiteration by paraphrase	Explaining the meaning of a word, a phrase, etc. by using other words in an attempt to make the meaning easier and more clear to understand. This is mostly realized in clause or sentence forms, and they appear immediately following the presupposed items whereas this is not the case for the latter.	The word "paraphrase" is so general that it can subsume some other semantic reiterations including exemplification, clarification, appositive, and relative clauses. These versions of paraphrase do not have the same value in terms of their overt/covert realizations.
a) Relative Clause Relative	Relative clauses are considered to be redundant due to their function in reiterating an item preceding them. These redundant elements are so crucial in unfolding the meaning of their presupposed items that their length sometimes trespasses the whole main clause of a sentence. Eliminating all the relative clauses of a text makes it writer-specific or restricts it to a particular group of readers, while the sole purpose of writing is to communicate the message in the best way possible. The relative clauses make sentences self-contained, i.e.. the reader does not have to search around the text to recover their meanings. The omission of the relative clauses, on the one hand, may make the text under-redundant for some readers, and presenting them in main clauses. On the other hand, it may make the text over-redundant.	

Type	Explanation	Example
b) Appositives	Appositive refers to words, phrases, or clauses in a sentence having the same reference.	"*especially", "particularly", "for example", "e.g.", "such as", "i.e.", "in other words", "that is", etc*. "[Infectious diseases which spread by contact or touch are called contagious diseases. *Malaria smallpox, diphteria, and colds are examples of infectious diseases* (Guyton, 1985)" "Thus far, all the acquired immunity that we have discussed has been [active immunity]. That is, the person's body develops either antibodies or sensitized lymphocytes in response to invasion of the body by a foreign antigen (Guyton, 1985)"
Reiteration by Intertextuality	Two kinds of intertextual relationships, i.e. relationships existing between elements of a given text (passive intertextuality), and relationships existing between distinct texts (active intertextuality). Intertextuality is taken as redundant due to the fact that it duplicates the preceding or following information: hence, its presence is apparently unnecessary despite the fact that the use of intertextuality is discoursally motivated.	
Reiteration by Cross-References	The devices by which one can keep track of references retrospectively or prospectively in the unfolding discourse. These redundant elements provide links between the discoursal themes and reiterate them in various ways so that they make the process of reading more possible.	*"in the following chapter";* *"as mentioned above"; "it will be discussed elsewhere"*
Redundancy and Predictor Signals	Predictor signals are enumerations and words that inform readers retrospectively and prospectively in a text. They serve as warnings of what the writer is about to produce, or what he has already produced.	*There are a few stages: (1)...(2)...(3)..., first,... second, ..., third,...*
Redundancy and Summary	In scientific texts, the more the reader comes to the end of the text, the more redundant elements are crystalized. It seems that the density of redundancy reaches its climax in the summary of a text because the writer without anything new only reviews the main points of the text. This part can be presented either covertly (without any marker signifying the summary) or overtly. The writer's awareness of readers' memory limitations helps him keep step with the readers, i.e., he does not hasten to overload readers with a lot of information in a short span. These condensed parts of texts have important cognitive effects on the reader's comprehension and recall, and they can act as feedback whether the reader has extracted the intended message or not.	*"in sum"; "to close up the text"; "I should now conclude by summarizing my arguments"*
Redundancy and Conjunctions	The lack of these tools (1) conjunctives, (2) conjunctive adverbs, (3) correlatives; (4) coordinators, (5) subordinators) does not seriously damage comprehension because readers are usually able to make bridging inferences.	(1) *"so that", "as long as",* etc.; (2) *"however", "therefore",* etc.; (3) *"either or", "both... and",* etc.; (4) *"but", "and, ...";* (5) *"because", "when",* etc.

Type	Explanation	Example
Redundant Collocations	It includes any pair of lexical items that stand to each other in some recognizable lexico-semantic (word meaning) relation. Accordingly, such pairs of words can be (1) synonyms, (2) antonyms, (3) series, (4) hyponyms, and (5) paranomy occuring freely both within a sentence or across sentence boundaries.	(1) *"climb and ascent";* (2) *"like and hate";* (3) *"north and east";* (4) *"chair and table";* (5) *"car and brakes".* The water gain occurs only through the application of [water] *droplets* to the soil surrounds or directly to the plant; this may me through rainfall or other forms of participation, irrigation or flooding or dew formation (Criffitlis, 1975)

Note. The table is based on the information from Lotfipour-Saedi & Sarhady (2000).

In professional scientific and technical writing, redundancy often manifests as a strategy to increase efficiency by condensing information. Gengshen (1990) notes that experts frequently use abbreviated terms, contracted forms, or shorthand expressions to streamline communication. This approach assumes that the intended audience shares common background knowledge, enabling more concise communication. In this case, redundancy serves to optimize communication within a specialized group, allowing for focus on critical content while reducing the need for explanations that would be necessary for a less informed audience.

Another reason for redundancy is to enhance cohesion and clarity within the text. As Bazzanella (2011) points out, the repetition of certain words or structures can reinforce the logical connections between ideas, making the text more coherent. This is particularly important in complex academic writing, where clarity is critical for the reader's understanding of intricate arguments. Redundancy through repeated keywords or phrases can help maintain focus on core ideas and ensure that the reader is continuously reminded of the main thesis or key concepts.

Additionally, redundancy can also serve as a pedagogical tool. In educational contexts, especially in instructional materials, repetition of key points or ideas is a common method to reinforce learning. This is especially true for students who may be encountering complex ideas for the first time (Watkowska, 2021). By intentionally including redundancy, educators aim to facilitate deeper understanding and long-term retention of information (Darian, 1979).

Finally, redundancy may arise as a consequence of cultural and linguistic differences in communication styles. In some languages or cultural contexts, repetition is a valued rhetorical device that signals emphasis or politeness. For instance, in certain Asian languages, redundancy can be used to ensure clarity or avoid miscommunication, reflecting a different approach to information density compared to Western academic writing (Yang, 2021). In multilingual academic environments, this can lead to the transfer of redundant structures from one language to another, particularly among non-native speakers of English.

Redundancy Impact on Text Comprehension

Redundancy in a text occurs when it presents or invokes information that readers already possess, either because they knew it beforehand or because it was previously introduced in the text. According to Grant-Davie (1995) and Every (2017), this repetition can lead to inefficiency, as the reader is provided with cues that are not necessarily new. Smith (1971) highlights that this redundancy has two key effects: it offers repetitive cues to the reader and narrows the range of possible language elements that can occupy certain positions in the sentence. This process can aid in comprehension by guiding the reader's expectations about upcoming information but can also lead to reduced engagement with the material if overused (Lotfipour & Sarhady, 2000).

Redundancy can also negatively impact writing by diminishing the clarity, efficiency, and overall impact of ideas. Forlini et al. (1982), Grant-Davie (1995) and Marinashvili (2020) note that when ideas are repeated without purpose, they lose their sharpness, which weakens the text's ability to effectively convey its message. This is especially problematic in technical writing, where readers do not have immediate access to the writer for clarification. In such contexts, unnecessary redundancy can lead to misunderstandings that may have serious consequences.

Wang (2021) points out that unnecessary redundancy in writing does not contribute new information nor serve a rhetorical or literary purpose, which significantly affects the precision and readability of the content. When readers encounter repeated or irrelevant information, the flow of the text is disrupted, and comprehension is hindered, leading to frustration and a potential loss of interest in the material (Demir, 2019).

However, some degree of redundancy is essential for ensuring coherence and readability in a text. Horning (1993) and Leufkens (2023) argue that redundancy in the form of overlapping information or cohesive ties between sentences contributes to the coherence of a text. This adherence to the given-new contract, where new information is presented in the context of already familiar or «given» information, helps readers relate new concepts to their existing knowledge. Grant-Davie (1995) and Bazzanella (2011) state that the easier it is for readers to make connections between new and prior information, the better they can understand and retain that information. This overlap is crucial for readability, as it allows readers to form a continuous thread of understanding throughout the text.

Redundancy is also embedded in the structure of language itself. Smith (1978), Wit & Gillette (1999) and Berdicevskis (2015) point out that readers use the inherent redundancy of language, relying on alternative cues such as the shape of words, their sounds, and their syntactic and semantic contexts, to recognize meaning. Even in writing that is perceived as concise, there is functional redundancy, allowing competent readers to draw on multiple sources of information to comprehend the text. Grant-Davie (1995) and Dasril et al. (2019) add that beginning readers and writers struggle because they have not yet mastered the use of these redundant pathways to meaning. Therefore, redundancy is not only a common feature of language but also a vital tool for comprehension, especially for more experienced readers.

Functions of Text Redundancy

Redundancy in text serves several important functions across different genres and contexts, playing a pivotal role in enhancing communication effectiveness. Firstly, redundancy can provide clarity and emphasis, reinforcing key ideas and ensuring that critical information is not overlooked (Rathjens, 1985; Bazzanella, 2011). In educational texts, for instance, reiterating concepts can aid in comprehension and retention, allowing learners to grasp complex subjects more thoroughly.

In narrative genres, redundancy can contribute to character development and thematic depth. By echoing certain phrases or motifs, authors can create a sense of rhythm and cohesion, drawing attention to significant emotional or narrative arcs (Baten, 1981). This technique not only enriches the reader's experience but also deepens their engagement with the text.

Moreover, in technical and instructional writing, redundancy can serve a practical purpose by ensuring that essential details are communicated clearly (Horning, 1991; Lotfipour & Sarhady, 2000). Instructions may repeat critical steps or warnings to minimize the risk of misunderstanding, thereby enhancing user safety and effectiveness. Similarly, in legal, formal or scientific documents, redundancy can provide clarity and precision, ensuring that terms are clearly defined and understood, reducing the potential for ambiguity.

In diverse contexts, redundancy can also accommodate varying levels of reader knowledge (Bazzanella, 2011). For example, in scientific writing, where audiences may range from experts to laypersons, repeating foundational concepts can bridge the knowledge gap, making complex information accessible to a broader audience. Additionally, redundancy can foster a sense of connection and familiarity in persuasive writing. By reiterating key arguments or values, authors can strengthen their appeal and resonate more deeply with their audience, making their message more memorable.

Overall, while redundancy is often viewed as a stylistic flaw, its functions across different genres and contexts reveal its potential to enhance clarity, engagement, and understanding, ultimately contributing to more effective communication. Supporting the idea that redundancy improves quality and coherence of a text, Wit & Gillette (1999) distinguished 6 functions of text redundancy: comprehensibility, resolving ambiguity, isolating a feature, contrasting elements, emphasizing or intensifying, creating poetic effect (see Table 12). While text redundancy can serve these functions, it is essential for academic writers to strike a balance, ensuring that redundancy enhances rather than detracts from the clarity and effectiveness of their writing.

Redundancy Reduction

Reducing redundancy is crucial for enhancing the clarity, conciseness, and overall impact of communication. Redundancy, or the unnecessary repetition of ideas, words, or phrases, can obscure the intended message and reduce its effectiveness (Marinashvili, 2020; Dhivya & Koperundevi, 2024). By eliminating non-essential elements without altering the meaning of the text, writers can create concise and focused messages that are easier for the reader to understand. This practice is especially valued in professional and academic writing, where precision and efficiency are key to effective communication (Rathjens, 1985).

Concise and focused communication not only improves clarity but also reflects a high level of professionalism. Reducing redundancy ensures that the audience remains engaged and that the message is delivered in a sharp and purposeful manner. In fields such as academia and technical writing, where complex information is often presented, excessive redundancy can dilute the core ideas, leading to confusion or disengagement. Therefore, mastering the skill of crafting concise and clear messages is essential for maintaining audience interest and ensuring that the intended message is conveyed effectively (Dhivya & Koperundevi, 2024).

However, while eliminating redundancy is typically desirable, some level of repetition is often necessary to ensure communication reliability. In some contexts, reducing re-

Table 12

Functions of Text Redundancy

	Function	Explanation	Examples
1	Comprehensibility	Language from one point of view can be regarded as a communication process. It serves to communicate a message or a feeling to a (potential) audience.	*"Last year I visited the Dar es Salaam, the capital of Tanzania.";* *"I like that marine-colored, blue dress, that hangs over there.";* *"I can't give you money – no."*
2	Resolving ambiguity	In many official occasions precision of expression is needed.	*"the truth, the whole truth and nothing but the truth.";* *"I live in Carbondale, Pennsylvania."*
3	Isolating a feature	In their speech and writing people frequently want to focus on a salient characteristic of a certain object.	*"I love the salty sea."*
4	Contrasting elements	Sometimes, what seems redundant actually contrasts two elements in the expression.	*"I like coffee and you don't."*
5	Emphasizing or intensifying	The redundant feature intensifies the meaning of the expression.	*"The green, green grass of home.";* *"I am completely and entirely crazy about her.";* *"I had a blue, blue Christmas."*
6	Creating poetic effect	It encapsulates all uses of redundancy with no clear semantic purpose, but with an intention to shock, to please, to horrify, to move, etc.	*"Warning. Danger. Stay out."*

Note. The table is based on the information from Wit & Gillette (1999).

dundancy too much can decrease the effectiveness of a message, especially when dealing with complex or technical subjects. Shorter, more concise messages may transmit information effectively, but key points could be lost or not adequately emphasized, making the message less reliable (Bazzanella, 2011). A balance between conciseness and necessary repetition is therefore crucial, as redundancy can help emphasize critical information and ensure it is understood (Tuinman & Gray, 1972).

Redundancy can also serve as a cognitive aid, providing cues that help readers navigate complex or unfamiliar content. When dealing with challenging material, redundancy reinforces key concepts, aiding in comprehension and retention (Albers et al., 2023). In this way, redundancy makes dense information more accessible without overwhelming the reader. This is particularly important in technical writing, where misunderstandings can have significant consequences, and redundancy can act as a safeguard against miscommunication (Lotfipour & Sarhady, 2000).

Excessive redundancy, on the other hand, can negatively affect the readability of a text. Redundant language can slow down the reader and make the text feel repetitive or tedious, ultimately harming its overall effectiveness (Lehmann, 2006; Xue & Hwa, 2014). Teaching students to eliminate unnecessary redundancy, particularly at the syntactic level, not only enhances the fluency of their writing but also helps them communicate their ideas more clearly. By removing unnecessary repetitions, writers can create sharper, more engaging messages where every word contributes meaningfully to the overall flow of the text (Dawson, 1992; Grant-Davie, 1995).

Clear and logical writing is key to minimizing redundancy and improving comprehension. Writers should avoid vague expressions and redundant phrases, as these can slow down the reader's understanding and reduce interest in the text (Yang, 2021). Using specific language, minimizing generic words, and ensuring that key terms are consistently repeated in the same format can help avoid confusion. Furthermore, cutting unnecessary synonyms and repeated constructions helps streamline the text, making it more coherent and readable, which enhances the reader's focus on the main ideas (Wallwork & Southern, 2020).

Eliminating redundant elements also preserves the original meaning of the text while enhancing fluency. Writers should focus on ensuring that each word in a sentence serves a specific purpose. This approach helps maintain the reader's

attention without overwhelming them with unnecessary information, making the text more effective in academic and professional settings. Careful word choice and cutting out redundant phrases contribute to the overall clarity and precision of the message (Xue & Hwa, 2014; Tikhonova & Mezentseva, 2024).

Despite the general consensus that reducing redundancy improves text quality, redundancy can play a positive role in certain contexts. Redundancy can actually enhance comprehension by increasing predictability and robustness in language. In academic writing, for example, repeating key concepts can help ensure that readers fully grasp the material, particularly in complex or technical texts. In these cases, redundancy serves to reinforce understanding and prevent ambiguity, acting as a tool to improve communication (Leufkens, 2023).

Balancing redundancy is particularly important when the writer is unsure of the reader's level of background knowledge. Writers often employ strategies like «audience unspecificity allowance» by repeating key ideas to ensure that readers with varying levels of expertise can follow the text. This redundancy allows even less knowledgeable readers to engage with the material while providing additional cues that enhance understanding for more experienced readers. As a result, redundancy bridges the gap between different levels of reader comprehension (Lotfipour & Sarhady, 2000; Yang, 2021).

Moreover, redundancy helps compensate for the limits of human cognitive capacity. Readers cannot always focus on every concept in a text, so by repeating key ideas in different ways, writers ensure that critical information is emphasized and easier to recall. This method of using redundancy to safeguard against cognitive overload is particularly useful in academic writing, where dense information can easily overwhelm the reader if not properly reinforced (Lotfipour & Sarhady, 2000; Albers et al., 2023).

While redundancy can make language harder to process in some cases by violating linguistic economy, it also simplifies language by increasing predictability. Redundancy helps readers anticipate and understand content more easily, making texts more accessible while maintaining clarity. This balance between economy and predictability is a key consideration in academic writing, where managing redundancy is essential for producing clear, precise, and comprehensible texts (Leufkens, 2023).

DISCUSSION

The purpose of this review was to explore the phenomenon of text redundancy, its implications for comprehension, and its role in academic and professional writing. The results demonstrate that redundancy is a complex and multifaceted concept, serving both functional and detrimental purposes depending on how it is applied. By analyzing the various definitions and uses of redundancy across the literature, we have identified key themes that help clarify its impact on communication, as well as the fine line between necessary repetition and wordiness.

The findings of this review reveal a dual perception of redundancy. It affects the text in opposite ways. On one hand, redundancy is often viewed negatively as excessive redundancy which leads to inefficiency and cognitive overload. Studies such as those by Forlini et al. (1982) and Every (2017) argue that unnecessary repetition in a text can obscure the intended message and diminish its clarity. This form of redundancy, commonly referred to as *wordiness*, involves the excessive use of synonyms, filler phrases, or repeated constructions that do not add value to the content. The result is a text that slows down the reader, causing frustration and reducing engagement (Wallwork & Southern, 2020).

On the other hand, redundancy can serve a crucial role in enhancing readability and ensuring comprehension, particularly in complex or technical writing. Smith (1971) and Horning (1991) suggest that when used strategically, beneficial redundancy helps reinforce key ideas, allowing readers to better process and retain information. In this context, redundancy acts as a cognitive aid, particularly for readers who may not grasp complex material upon first reading. This form of *functional redundancy* is essential in academic writing, where precision and clarity are paramount, and repeating important concepts helps reduce ambiguity and misinterpretation (Lotfipour & Sarhady, 2000).

Consolidating Definitions of Redundancy

The diverse definitions of redundancy across the literature suggest a need for a more consolidated understanding of the term. Redundancy is not simply a negative trait in writing; rather, it has both positive and negative aspects depending on its use. Based on the results of this review, the most appropriate definition of redundancy would be: the strategic or unintentional repetition of information, ideas, or structures within a text, which can either enhance clarity and comprehension or lead to inefficiency and cognitive overload, depending on context and audience needs. This definition captures the dual nature of redundancy, recognizing that it can either aid or hinder the reading experience depending on how it is applied.

Moreover, the concept of *wordiness* should be recognized as a specific type of redundancy. Wordiness refers to excessive, unnecessary repetition that does not contribute to the text's meaning or purpose. It detracts from the clarity and efficiency of the message by introducing irrelevant or repetitive elements that slow down comprehension. Thus, wordiness can

be classified as a subcategory of redundancy that focuses on the negative aspects of excessive repetition.

Refining the Concept of Redundancy

The review also emphasizes the importance of understanding redundancy within the broader context of communication. In academic writing, redundancy is not always a sign of poor style; rather, it can be an intentional strategy to ensure the reader fully understands the material. This is particularly important in technical texts or educational materials, where the complexity of the content often requires multiple forms of reinforcement. However, when redundancy crosses into wordiness - where repetition no longer serves a purpose - it undermines the effectiveness of communication. Writers need to be aware of this balance and apply redundancy thoughtfully to enhance, rather than detract from, the text.

Additionally, the findings highlight the challenges that arise from audience variability. Lotfipour and Sarhady (2000) point out that writers cannot always predict the reader's background knowledge or cognitive capacity. As a result, some readers may find a text overly redundant, while others may struggle with under-redundancy. This audience-specific challenge underscores the importance of adaptive redundancy - where writers consciously repeat key concepts to accommodate a broader range of readers. This strategy ensures that readers of varying expertise levels can access and comprehend the material effectively.

Classification of Redundancy

The review proposes a classification system for redundancy based on its function and impact on communication:

(1) Functional Redundancy or Beneficial Redundancy: Intentional repetition that enhances clarity and comprehension by reinforcing key ideas. This type of redundancy is particularly useful in complex or technical writing, where precision is crucial, and repeating concepts ensures understanding.

(2) Wordiness or Excessive Redundancy: A negative form of redundancy characterized by unnecessary repetition that does not contribute meaningfully to the text. Wordiness detracts from clarity and efficiency, leading to cognitive overload and reduced reader engagement.

(3) Contextual Redundancy: Redundancy used to address the varying levels of reader expertise. Writers may repeat key ideas to ensure that less knowledgeable readers can follow the text, while more experienced readers can still benefit from the additional cues provided by redundancy.

This classification allows for a more nuanced understanding of redundancy and its role in communication, recognizing that not all repetition is detrimental. It also provides a framework for writers to evaluate when and how to apply redundancy effectively. The results of this review have important implications for academic and professional writing. Writers should be mindful of the role redundancy plays in shaping the reader's comprehension. While functional redundancy can be beneficial, particularly in technical or educational texts, excessive redundancy (wordiness) should be minimized to maintain clarity and reader engagement. This balance between repetition and conciseness is critical for producing high-quality texts that are both accessible and efficient. Additionally, writing instructors and educators should focus on teaching students how to identify and eliminate unnecessary redundancy while recognizing when repetition can serve a strategic purpose. By understanding the dual nature of redundancy, writers can craft texts that are clear, precise, and tailored to their audience's needs.

Implications. How to overcome redundancy

The outcomes of this review on text redundancy found direct application in the development of pedagogic materials aimed at improving doctoral students' academic writing across disciplines. In this section, we focus on the «Avoiding Text Redundancy in Academic Writing» course, which adopts a genre-based approach informed by corpus linguistics, academic writing studies, and rhetorical analysis (see Table 13). This course is designed for L1 and L2 speakers of English, offering practical strategies to identify and reduce redundancy in their academic writing. The course emphasizes the balance between clarity and conciseness, aiming to eliminate unnecessary repetition that often leads to wordiness.

This course incorporates corpus-based descriptions to highlight redundancy patterns and hands-on analysis of authentic academic texts to demonstrate how different disciplines handle repetition. Course materials are made available to students through the Moodle learning platform, allowing for flexible access to readings, exercises, and discussion forums. These materials include annotated corpus texts and instructional videos that introduce the key concepts of redundancy and rhetorical intent, with exercises designed to solidify knowledge through peer discussion and self-evaluation.

Corpus-Based Materials and Tasks

The course materials are organized around a set of core principles, including the move/step model of academic discourse, which serves as a foundational framework for understanding how redundancy operates within different sections of academic writing. The readings offer compre-

hensive descriptions of rhetorical moves and their functions, illustrating how redundancy may serve either to reinforce or obscure key ideas. Each reading is accompanied by short video lectures in which instructors use excerpts from authentic academic texts to demonstrate redundancy patterns, explain rhetorical intent, and offer strategies for streamlining writing without losing meaning.

Additionally, the course includes a series of corpus-based exercises that guide students through the process of analyzing authentic academic texts. These exercises focus on identifying patterns of redundancy and distinguishing between necessary and excessive repetition. Through this hands-on engagement, students are encouraged to apply the theoretical principles discussed in the readings to their own writing.

To promote deeper understanding and collaborative learning, group discussions are a core component of the course. Students engage in peer review sessions where they critique each other's writing, offering feedback on redundant elements and suggesting improvements. These peer-to-peer interactions, modeled after the approaches advocated by Flowerdew (2008), foster an environment of shared learning and critical reflection. The feedback provided in these sessions helps students refine their writing by reducing redundancy and enhancing the clarity of their arguments. The course's structured peer review process is complemented by detailed guidelines on how to evaluate redundancy in academic writing. These guidelines encourage students to assess whether their writing and that of their peers are clear, whether key ideas are effectively communicated without excessive repetition, and whether the argumentation is sufficiently concise. By focusing on these key elements, students learn to apply critical thinking to their writing, improving both their own texts and those of their peers.

Swales (2014) emphasizes that tasks should be designed to be sequenced and goal-directed, drawing upon a range of cognitive and communicative procedures. In alignment with this view, the course integrates differentiated activities that scaffold learning from initial identification of redundancy to the practical application of revision strategies. The tasks are structured to progress from the analysis of self-compiled corpora to the revision of drafts, allowing students to gradually build their skills and apply them in context.

The course also uses annotated corpora, made accessible through tools like Callisto, to support students in conducting a detailed analysis of redundancy in authentic academic texts. These annotated texts highlight common linguistic patterns that contribute to redundancy, such as excessive use of synonymous terms or unnecessary phrases. By exploring the annotated corpus, students learn to identify redundancy across various disciplines and develop strategies for avoiding it in their own writing.

By the end of the course, students are expected to apply the corpus-based techniques they have learned to revise their own research article drafts, reducing unnecessary repetition and improving clarity. Through the combination of theoretical instruction, hands-on corpus exploration, peer feedback, and iterative revisions, the course equips students with the tools necessary to produce clear, concise, and effective academic writing. The goal is to help students reduce redundancy while maintaining the coherence and clarity of their arguments (see Table 14).

Mapping of Materials and Tasks to Specific Course Modules

The course "Avoiding Text Redundancy in Academic Writing" is structured into five comprehensive modules, each targeting a specific aspect of identifying, analyzing, and eliminating redundancy in academic writing. The course is designed for doctoral students across disciplines and offers a detailed, hands-on approach to improving clarity and conciseness in research papers and other academic texts. The pedagogical approach integrates theoretical instruction

Table 13

Course Module Description and Focus

Module	Topics and Focus
Module 1: Understanding Redundancy	Introduction to redundancy and its aspects in academic writing. Corpus compilation and analysis of redundancy patterns.
Module 2: Reducing Wordiness	Identifying unnecessary repetition and applying strategies to reduce wordiness in academic texts.
Module 3: Corpus-Based Exploration	Hands-on analysis of annotated corpora to examine the function of redundancy across different disciplines.
Module 4: Enhancing Argumentation	Balancing repetition with clarity in academic argumentation. Revising drafts for conciseness and coherence.
Module 5: Peer Review and Final Project	Peer review sessions focused on redundancy. Final project submission with revisions based on feedback.

with practical tasks, corpus-based analysis, peer review, and iterative revisions, allowing students to apply the concepts directly to their own writing. Below is a detailed description of each module, including its focus, activities, and materials.

Module 1. Understanding Redundancy in Academic Writing

The first module introduces students to the concept of redundancy in academic writing, explaining different types of redundancy, its reasons, functions, tools and strategies to detect and overcome. This part of the course sets the stage for the subsequent modules. The main focus is to help students recognize the different forms of redundancy that can appear in research articles, such as unnecessary repetition of ideas, redundant phrases, or the overuse of similar transitions.

Students begin by compiling a self-curated corpus of 10-15 research articles from reputable journals in their field of study. These articles are required to be recently published and written by different authors. This task encourages students to become familiar with the structure and style of contemporary academic writing in their disciplines while paying close attention to patterns of redundancy. After compiling their corpus, students conduct a preliminary analysis, noting redundant elements, transitions, and section structures in the articles.

Following the individual analysis, students participate in group discussions where they compare their findings with peers. They analyze how different authors handle redundancy and discuss when repetition may serve a rhetorical or functional purpose, such as reinforcing key ideas, and when it may become excessive or unnecessary. Through these discussions, students gain a deeper understanding of how redundancy can either enhance or hinder clarity and how authors in their field manage repetition to create more effective texts.

Materials used: self-compiled corpus of research articles, readings on redundancy in academic writing, group discussion materials and notes on redundancy patterns.

Module 2. Identifying and Reducing Wordiness

The second module focuses on the practical skills necessary to identify and reduce wordiness, which is a specific form of redundancy in academic writing. Wordiness occurs when unnecessary words, phrases, or overly verbose expressions are used, making the text difficult to follow and less efficient.

Students are provided with sample academic texts that include redundant expressions, verbose phrases, or unnecessary synonyms. The first task is to identify and analyze these examples of wordiness. Students compare the original texts with revised versions that have been edited for conciseness, observing how the removal of redundant words improves the clarity, flow, and readability of the text without sacrificing the intended meaning. Once students understand the principles of identifying wordiness, they are tasked with applying these techniques to their own writing. Each student selects a section from their research paper, thesis, or dissertation draft and revises it to eliminate redundant elements. This exercise not only improves the clarity of their writing but also helps them develop the skills to critically evaluate their own work.

The module also includes peer review sessions, where students provide feedback on each other's revisions, identifying remaining instances of wordiness and offering suggestions for further improvement. This peer-to-peer interaction encourages reflection and collaborative learning.

Materials: sample academic texts with redundant elements, revised versions of these texts for comparison, student's own research drafts, peer review guidelines and rubrics for evaluating conciseness.

Table 14

Types of Materials and Assignments in the Course

Materials	Tasks
Self-compiled corpus	Writing and revising research article draft
Video lectures, readings	Identifying redundancy and discussing strategies to reduce it
Corpus-based exercises, rhetorical analysis handouts	Group work and discussion on redundancy patterns in academic texts
Annotated corpus, Callisto	Rhetorical analysis of redundancy patterns in research articles
Annotated corpus, Callisto	Language use analysis focused on redundancy; online discussion forum
Self-compiled corpus, Callisto	Annotation of redundant elements
Student research article draft, Callisto	Peer review and annotation of first draft for redundancy
Move/step model-based guidelines	Revision of research article drafts focusing on reducing redundancy

Module 3. Corpus-Based Exploration of Redundancy

In this module, students move into more advanced, data-driven analysis using corpus-based tools. The goal of this module is to deepen students' understanding of redundancy by allowing them to explore how repetition functions across different disciplines and research texts. Using these tools, students can identify where redundancy serves a clear rhetorical or functional role and where it becomes excessive.

Students are introduced to corpus-based tools such as Callisto and concordancers, which allow them to analyze the annotated corpora provided for the course. The annotated corpus includes examples from a range of academic disciplines and highlights linguistic and structural redundancy. Students are asked to explore different sections of their self-compiled corpus (from Module 1) using these tools to investigate how redundancy manifests in introductions, literature reviews, results, and discussions. They focus on identifying linguistic markers of redundancy, such as overuse of certain phrases, excessive synonyms, or unnecessary repetition of concepts.

In group sessions, students share their findings, discussing which aspects of redundancy are common in their field and how authors typically manage them. This exploration helps students recognize discipline-specific patterns of redundancy and apply them to their own writing. The corpus-based analysis also provides students with insights into how other researchers have effectively reduced redundancy without losing clarity or emphasis on key points.

Materials: annotated corpora from various disciplines, corpus analysis tools (e.g., Callisto, concordancer software), self-compiled research corpus from Module 1, guidelines for corpus-based analysis tasks.

Module 4. Enhancing Academic Argumentation

Building on the skills developed in the previous modules, the fourth module shifts the focus towards refining academic argumentation. The objective here is to teach students how to maintain a balance between necessary repetition for emphasis and conciseness for clarity. Academic argumentation often requires a careful approach to redundancy, where key points must be reinforced without over-explaining or becoming redundant.

Students begin by analyzing the argumentation in their self-compiled corpus, paying particular attention to how repetition is used in the structure of introductions, methodologies, discussions, and conclusions. They identify instances where authors repeat important points to emphasize their arguments and note how this is done effectively, without overloading the reader with unnecessary repetition.

Following this, students apply these insights to their own writing. They are tasked with revising the argumentation in their drafts, reducing redundancy while ensuring that their key arguments remain clear and coherent. The focus is on identifying sections where repetition has been overused and finding alternative ways to express or emphasize the same points more concisely. Instructor feedback is provided to each student, with personalized suggestions on how they can improve the logical flow and clarity of their argumentation.

Materials: annotated corpus examples highlighting argumentation patterns, student research article drafts, instructor-provided feedback templates on redundancy in argumentation.

Module 5. Peer Review and Final Project

The final module of the course involves peer review sessions and the completion of a final project, allowing students to put all the skills they've acquired throughout the course into practice. This module emphasizes collaborative learning through peer feedback and the refinement of writing based on constructive critique.

Students engage in structured peer review sessions, using a detailed rubric to evaluate each other's drafts for redundancy, clarity, and conciseness. The peer review focuses on whether unnecessary repetition is present and whether the text can be improved by reducing wordiness. Students provide detailed feedback on how to streamline writing and enhance the overall clarity of the argumentation. After receiving peer feedback, students revise their drafts for a final submission. The final project requires students to submit a fully revised research article draft, demonstrating improved conciseness and clarity, along with a reflective report. The report asks students to reflect on the revision process, outlining how they applied the redundancy reduction techniques learned throughout the course and what impact these revisions had on the overall quality of their writing.

Materials: peer review rubric for evaluating redundancy and clarity, student research article drafts, reflective report template for final project submission.

Course Limitations

The course presents several practical challenges. One limitation is the time-intensive nature of corpus-based tasks, which may not be feasible for all educational settings. The course also relies heavily on access to annotated, discipline-specific corpora, which may not be available to all students. Moreover, the course requires students to have a working familiarity with corpus tools, which could pose a challenge for those without prior experience in linguistics

or academic writing analysis. Despite these limitations, the course structure provides a comprehensive framework for doctoral students to develop their writing skills by minimizing redundancy and enhancing the clarity of their academic discourse.

In some contexts, teachers may need to adjust the course's interactive and corpus-based activities to fit within tighter schedules. For example, corpus exploration tasks could be assigned as homework, while classroom time is devoted to discussions of the findings. As corpus representativeness grows over time through student annotations, future iterations of the course will have access to a broader range of examples, allowing for more precise analysis of redundancy patterns across disciplines.

CONCLUSION

This review examined the dual nature of text redundancy and its role in communication, especially in academic and professional writing. The findings suggest that redundancy can either improve comprehension by reinforcing key ideas or detract from communication when it results in excessive repetition or wordiness. The review consolidates redundancy as the strategic or unintentional repetition of information or structures, which can either aid or hinder communication based on its context and purpose.

The classification of redundancy into functional (beneficial) redundancy, wordiness (excessive redundancy), and contextual redundancy provides a useful framework for writers to manage repetition more effectively. Understanding these categories allows writers to make informed decisions about when to employ redundancy for emphasis and when to avoid it to maintain clarity. Balancing necessary repetition with conciseness is crucial for maintaining clarity and reader engagement, as excessive redundancy can lead to reader fatigue and disengagement, while well-placed repetition can enhance retention and understanding.

The limitations of this review include its primary focus on written communication, leaving the role of redundancy in other mediums, such as spoken or digital communication, underexplored. While the nuances of written texts have been examined, the dynamics of redundancy in oral communication, where tone, inflection, and immediacy play significant roles, remain largely unaddressed. Additionally, the review provides general insights into academic writing without delving deeply into the varying uses of redundancy across different academic fields or genres, which can exhibit distinct conventions and expectations regarding repetition.

Future studies should focus on exploring redundancy in digital communication, including social media, email, and online academic platforms, where brevity is often prioritized. The unique characteristics of these platforms may influence how redundancy is perceived and utilized, offering a rich area for investigation. Additionally, research on how redundancy functions in oral communication, presentations, or instructional settings would provide valuable insights into the effectiveness of spoken repetition and its impact on audience engagement and comprehension. Understanding redundancy in these contexts could lead to more effective communication strategies tailored to specific audiences and situations.

Finally, discipline-specific studies on redundancy within different academic fields could reveal more nuanced ways in which repetition supports or detracts from the clarity and effectiveness of communication in specialized contexts. Different fields may have varying tolerance levels for redundancy based on their conventions, audience expectations, and the complexity of the subject matter. By examining these aspects, future research can contribute to a more comprehensive understanding of redundancy as a multifaceted phenomenon that plays a critical role in effective communication across diverse contexts. Such insights could ultimately inform best practices for writers and speakers, enhancing their ability to convey ideas clearly and effectively in both academic and professional settings.

DECLARATION OF COMPETITING INTEREST

None declared.

AUTHOR CONTRIBUTIONS

Elena Tikhonova: conceptualization; data curation; formal analysis; investigation; methodology; project administration; resources; software; supervision; validation; visualization; writing – original draft; writing – review & editing.

Daria Mezentseva: data curation; formal analysis; investigation; software; visualization; writing – original draft.

Petr Kasatkin: data curation; investigation; methodology; software; visualization; writing – original draft; writing – review & editing.

REFERENCES

Abdollahi-Guilani, M., Mirzaeifard, S., Aghaei, K., & Khojastehrad, S. (2012). Clashes of conciseness and wordiness between English and Persian verbs. *Asian Social Science, 8*(10), Article 10. https://doi.org/10.5539/ass.v8n10p118

Al-Qaddoumi, K., & Ageli, N. (2023). Redundancy and ellipsis in the translation into English of selected Arabic media texts. *International Journal of Translation and Interpretation Studies, 3*(2), 16–23. https://doi.org/10.32996/ijtis.2023.3.2.2

Albers, F., Trypke, M., Stebner, F., Wirth, J., & Plass, J. L. (2023). Different types of redundancy and their effect on learning and cognitive load. *British Journal of Educational Psychology, 93*(S2), 339–352. https://doi.org/10.1111/bjep.12592

Alontseva, N. V., & Ermoshin, Y. A. (2019). The problem of language redundancy on the example of a scientific text. *RUDN Journal of Language Studies, Semiotics and Semantics, 10*(1), 129–140. https://doi.org/10.22363/2313-2299-2019-10-1-129-140

Bartell, A. L., Schultz, L. D., & Spyridakis, J. H. (2006). The effect of heading frequency on comprehension of print versus online information. *Technical Communication, 53*, 416–425.

Baten, L. (1981). *Text comprehension: The parameters of difficulty in narrative and expository prose texts: A redefinition of readability* [Unpublished doctoral dissertation]. University of Illinois.

Bazzanella, C. (2011). Redundancy, repetition, and intensity in discourse. *Language Sciences, 33*(2), 243–254. https://doi.org/10.1016/j.langsci.2010.10.002

Bensoussan, M. (1990). Redundancy and the cohesion doze. *Journal of Research in Reading, 13*(1), 18–37. https://doi.org/10.1111/j.1467-9817.1990.tb00320.x

Berdicevskis, A. (2015). Estimating grammeme redundancy by measuring their importance for syntactic parser performance. In R. Berwick, A. Korhonen, A. Lenci, T. Poibeau, & A. Villavicencio (Eds.), *Proceedings of the Sixth Workshop on Cognitive Aspects of Computational Language Learning* (pp. 65–73). Association for Computational Linguistics. https://doi.org/10.18653/v1/W15-2410

Bodenreider, O. (2003). Strength in numbers: Exploring redundancy in hierarchical relations across biomedical terminologies. *AMIA Annual Symposium Proceedings, 2003*, 101–105.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Buscail, L., & Saint-Dizier, P. (2009). Textual and stylistic error detection and correction: Categorization, annotation, and correction strategies. *2009 Eighth International Symposium on Natural Language Processing*, 205–210. https://doi.org/10.1109/SNLP.2009.5340918

Caballero, G., & Kapatsinski, V. (2014). Perceptual functionality of morphological redundancy in Choguita Rarámuri (Tarahumara). *Language, Cognition and Neuroscience, 30*(9), 1134–1143. https://doi.org/10.1080/23273798.2014.940983

Cao, M., & Zhuge, H. (2022). Automatic evaluation of summary on fidelity, conciseness, and coherence for text summarization based on semantic link network. *Expert Systems with Applications, 206*. https://doi.org/10.1016/j.eswa.2022.117777

Chauhan, P. (2022). Fundamentals of academic writing: A literature review. *Journal of NELTA, 27*(1-2), 161–180. https://doi.org/10.3126/nelta.v27i1-2.53201

Chetail, F. (2015). Reconsidering the role of orthographic redundancy in visual word recognition. *Frontiers in Psychology, 6*, 645. https://doi.org/10.3389/fpsyg.2015.00645

Darian, S. (1979). The role of redundancy in language and language teaching. *System, 7*(1), 47–59. https://doi.org/10.1016/0346-251X(79)90022-8

Dasril, Zaim, M., & Ningsih, K. (2019). Coherence and unity of students' writing on background of the study of research proposals. In *Proceedings of the 1st International Conference on Education Social Sciences and Humanities*. Atlantis Press. https://doi.org/10.2991/icesshum-19.2019.65

Dawson, J. H. (1992). Avoid redundancy in writing. *Weed Technology, 6*(3), 782. https://doi.org/10.1017/S0890037X0003623X

De Beaugrande, R. (1980). Text discourse in European research. *Discourse Processes, 3*, 287–300.

Demir, C. (2019). Writing intelligible English prose: Conciseness vs. verbosity. *Söylem Filoloji Dergisi, 4*(2), 482–505. https://doi.org/10.29110/soylemdergi.617184

Dhivya, P., & Koperundevi, E. (2024). TBLT: Avoiding redundancy to reduce complexity. *Forum for Linguistic Studies, 6*(1), 567–576.

Every, B. (2017). Writing economically in medicine and science: Tips for tackling wordiness. *Medical Writing, 26*, 17–20.

Flowerdew, J. (2008). Scholarly writers who use English as an additional language: What can Goffman's "Stigma" tell us? *Journal of English for Academic Purposes, 7*, 77–86. https://doi.org/10.1016/j.jeap.2008.03.002

Forlini, G., Bauer, M. B., & Prentice-Hall, I. (1982). *Prentice Hall Grammar and Composition: Level 4*. Prentice Hall PTR.

Freywald, U., & Finkbeiner, R. (2018). Exact repetition or total reduplication? Exploring their boundaries in discourse and grammar. In R. Finkbeiner & U. Freywald (Eds.), *Exact Repetition in Grammar and Discourse* (pp. 3–28). De Gruyter Mouton. https://doi.org/10.1515/9783110592498-001

Gengshen, H. (1990). An exploration into sci-tech interpretations: Abstract interpreting approach. *Babel, 36*(2), 85–96. https://doi.org/10.1075/babel.36.2.03gen

Grant-Davie, K. (1995). Functional redundancy and ellipsis as strategies in reading and writing. *JAC: A Journal of Composition Theory, 15*(3), 455–469.

Guerrero, F. G. (2009). A new look at the classical entropy of written English (arXiv:0911.2284). arXiv. https://doi.org/10.48550/arXiv.0911.2284

Heltai, P. (2018). Explicitation, redundancy, ellipsis, and translation. In K. Károly & A. Fóris (Eds.), *New trends in translation studies* (pp. 45-74). Akadémiai Kiadó.

Horning, A. S. (1979). On defining redundancy in language: Case notes. *Journal of Reading, 22*(4), 312–320.

Horning, A. S. (1991). Readable writing: The role of cohesion and redundancy. *Journal of Advanced Composition, 11*(1), 135–145.

Horning, A. S. (1993). *The psycholinguistics of readable writing: A multidisciplinary exploration*. Ablex Publishing Corporation.

Hunnicutt, S. (1985). Intelligibility versus redundancy - Conditions of dependency. *Language and Speech, 28*(1), 47–56. https://doi.org/10.1177/002383098502800103

Kravtchenko, E., & Demberg, V. (2022). Informationally redundant utterances elicit pragmatic inferences. *Cognition, 225*, 105159. https://doi.org/10.1016/j.cognition.2022.105159

Kuhi, D. (2017). Hybridity of scientific discourses: An intertextual perspective and implications for ESP pedagogy. *The Journal of Applied Linguistics and Applied Literature: Dynamics and Advances, 5*(2), 61–80. https://doi.org/10.22049/JALDA.2018.26150.1048

Lehmann, C. (2005). Pleonasm and hypercharacterisation. In G. Booij & J. Van Marle (Eds.), *Yearbook of Morphology 2005* (pp. 1–13). Springer. https://doi.org/10.1007/1-4020-4066-0_5

Leufkens, S. (2023). Measuring redundancy: The relation between concord and complexity. *Linguistics Vanguard, 9*(s1), 95–106. https://doi.org/10.1515/lingvan-2020-0143

Lotfipour-Saedi, K. (1982). Applying an analysis of writer-reader discourse processes to a pedagogy of reading [Unpublished doctoral dissertation]. University of Lancaster.

Lotfipour-Saedi, K., & Sarhady, T. (2000). Redundancy, its discoursal function, and textual realizations in different genres. *The International Journal of Humanities, 7*(1), 31–41.

Lynn, M. (2016). Conciseness is critical. *Cornell Hospitality Quarterly, 57*(4), 346–347. https://doi.org/10.1177/1938965516665740

Lyster, R. (1998). Recasts, repetition, and ambiguity in L2 classroom discourse. *Studies in Second Language Acquisition, 20*(1), 51–81. https://doi.org/10.1017/S027226319800103X

Marinashvili, M. (2020). Average word length and text redundancy variability: French texts case study. *Polonia University Scientific Journal, 38*, 67–75. https://doi.org/10.23856/3849

McCrudden, M. T., Hushman, C. J., & Marley, S. C. (2013). Exploring the boundary conditions of the redundancy principle. *The Journal of Experimental Education, 82*(4), 537–554. https://doi.org/10.1080/00220973.2013.813368

McGarry, J. (1975). Redundancy and entropy in language. *The Journal of General Psychology, 93*(1), 101–106.

Mu, W., & Lim, K. (2022). Revision for concision: A constrained paraphrase generation task. *arXiv*. https://doi.org/10.48550/arXiv.2210.14257

Newman, E. B., & Waugh, N. C. (1960). The redundancy of texts in three languages. *Information and Control, 3*(2), 141–153. https://doi.org/10.1016/S0019-9958(60)90731-2

Prasetyo, Y. (2015). Sentence conciseness in thesis abstracts of English department students. *Jurnal Edukasi, 1*(1), 71–80.

Rahman, N., & Borah, B. (2021). Redundancy removal method for multi-document query-based text summarization. *2021 International Symposium on Electrical, Electronics and Information Engineering* (pp. 568-574) . ACM. https://doi.org/10.1145/3459104.3459197

Rasulov, Z., & Artikov, A. (2023). The principle of redundancy in compound sentences. *Integration Conference on Integration of Pragmalinguistics, Functional Translation Studies and Language Teaching Processes* (vol. 2, pp. 48–51). Bukhara State University.

Rathjens, D. (1985). The seven components of clarity in technical writing. *IEEE Transactions on Professional Communication, 28*(4), 42–46. https://doi.org/10.1109/TPC.1985.6448848

Rosie, A. M. (1973). *Information and communication theory*. Van Nostrand Reinhold Company.

Schlesinger, I. M. (1966). *The influence of sentence structure on the reading process* (Technical Report No. 21). U.S. Office of Naval Research Information Systems Branch.

Schlesinger, I. M. (1977). *Production and comprehension of utterances*. Lawrence Erlbaum Associates.

Schüler, A., Scheiter, K., & Gerjets, P. (2013). Is spoken text always better? Investigating the modality and redundancy effect with longer text presentation. *Computers in Human Behavior, 29*(4), 1590–1601. https://doi.org/10.1016/j.chb.2013.01.047

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal, 27*(3), 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

Shannon, C. E. (1951). Prediction and entropy of printed English. *The Bell System Technical Journal, 30*(1), 50–64. https://doi.org/10.1002/j.1538-7305.1951.tb01366.x

Skorikova, T. P. (2017). Linguo-pragmatic aspects of the study of Russian public academic speech genres. *Russian Linguistic Bulletin, 4*(12). https://doi.org/10.18454/RULB.12.05

Smith, F. (1971). *Understanding reading: A psycholinguistic analysis of reading and learning to read*. Holt, Rinehart, and Winston.

Smith, F. (1978). *Understanding reading* (2nd ed.). New York: Holt, Rinehart, and Winston.

Soltani, K., & Kuhi, D. (2022). Writer responsibility across sections of research articles: Recycling of directional determinants by English and Iranian academics. *Journal of English for Academic Purposes, 60*. https://doi.org/10.1016/j.jeap.2022.101191

Stanley, R. (1967). Redundancy rules in phonology. *Language, 43*(2), 393–436. https://doi.org/10.2307/411542

Staub, A. (2015). The effect of lexical predictability on eye movements in reading: Critical review and theoretical interpretation. *Language and Linguistics Compass, 9*, 311–327. https://doi.org/10.1111/lnc3.12151

Swales, J. (2014). *Genre analysis: English in academic and research settings*. Cambridge University Press. https://doi.org/10.1075/z.184.513swa

Thadani, K., & McKeown, K. (2008). A framework for identifying textual redundancy. In *Proceedings of the 22nd International Conference on Computational Linguistics* (pp. 873–880). ACM. https://doi.org/10.3115/1599081.1599191

Tikhonova, E. V., & Mezentseva, D. A. (2024). Wordiness in academic writing: A systematic scoping review. *Research Result. Theoretical and Applied Linguistics, 10*(1), 133–157. https://doi.org/10.18413/2313-8912-2024-10-1-0-8

Trudgill, P. (2009). Sociolinguistic typology and complexification. In G. Sampson, D. Gil, & P. Trudgill (Eds.), *Language complexity as an evolving variable* (pp. 98–109). Oxford University Press. https://doi.org/10.1093/oso/9780199545216.003.0007

Trudgill, P. (2011). *Sociolinguistic typology: Social determinants of linguistic complexity*. Oxford University Press.

Tuinman, J. J., & Gray, G. (1972). The effect of reducing the redundancy of written messages by deletion of function words. *The Journal of Psychology, 82*(2), 299–306. https://doi.org/10.1080/00223980.1972.9923820

Wallwork, A., & Southern, A. (2020). Chapter 5: Readability and avoiding redundancy. In A. Wallwork & A. Southern (Eds.), *100 tips to avoid mistakes in academic writing and presenting* (pp. 39–74). Springer International Publishing. https://doi.org/10.1007/978-3-030-44214-9_5

Wang, F. X. (2001). *English-Chinese comparative semantics*. Foreign Language Press.

Watkowska, D. (2021). Redundancy in ELF: A corpus-based study on negative and modal concord. *Anglica: An International Journal of English Studies, 30*(2), 71–86. https://doi.org/10.7311/0860-5734.30.2.04

Wit, E. C., & Gillette, M. (1999). *What is linguistic redundancy?* [Technical Report]. https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.714.5763&rep=rep1&type=pdf

Wolf, L., Pimentel, T., Fedorenko, E., Cotterell, R., Warstadt, A., Wilcox, E., & Regev, T. (2023). *Quantifying the redundancy between prosody and text. arXiv*. https://doi.org/10.48550/arXiv.2311.17233

Xu, S. H. (1984). The redundancy of language. *Modern Foreign Languages, 2*, 3–8. https://doi.org/CNKI:SUN:XDWY.0.1984-02-000

Xue, H., & Hwa, R. (2014). Redundancy detection in ESL writings. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing* (pp. 683–691). ACL. https://doi.org/10.3115/v1/E14-1072

Yang, Y. (2021). The investigation on redundancy errors in writing of Chinese English learners with different proficiency. *Proceedings of the 2021 International Conference on Education, Language and Art*. Atlantis Press. https://doi.org/10.2991/assehr.k.220131.025

Zola, D. (1981). *The effect of redundancy on the perception of words in reading* [Technical Report]. University of Illinoys.

APPENDIX 1

Al-Qaddoumi, K., & Ageli, N. (2023). Redundancy and ellipsis in the translation into English of selected Arabic media texts. *International Journal of Translation and Interpretation Studies, 3*(2), 16–23. https://doi.org/10.32996/ijtis.2023.3.2.2

Albers, F., Trypke, M., Stebner, F., Wirth, J., & Plass, J. L. (2023). Different types of redundancy and their effect on learning and cognitive load. *British Journal of Educational Psychology, 93*(S2), 339–352. https://doi.org/10.1111/bjep.12592

Alontseva, N. V., & Ermoshin, Y. A. (2019). The problem of language redundancy on the example of a scientific text. *RUDN Journal of Language Studies, Semiotics and Semantics, 10*(1), 129–140. https://doi.org/10.22363/2313-2299-2019-10-1-129-140

Bartell, A. L., Schultz, L. D., & Spyridakis, J. H. (2006). The effect of heading frequency on comprehension of print versus online information. *Technical Communication, 53*, 416–425.

Bazzanella, C. (2011). Redundancy, repetition, and intensity in discourse. *Language Sciences, 33*(2), 243–254. https://doi.org/10.1016/j.langsci.2010.10.002

Bensoussan, M. (1990). Redundancy and the cohesion doze. *Journal of Research in Reading, 13*(1), 18–37. https://doi.org/10.1111/j.1467-9817.1990.tb00320.x

Berdicevskis, A. (2015). Estimating grammeme redundancy by measuring their importance for syntactic parser performance. In R. Berwick, A. Korhonen, A. Lenci, T. Poibeau, & A. Villavicencio (Eds.), *Proceedings of the Sixth Workshop on Cognitive Aspects of Computational Language Learning* (pp. 65–73). Association for Computational Linguistics. https://doi.org/10.18653/v1/W15-2410

Bodenreider, O. (2003). Strength in numbers: Exploring redundancy in hierarchical relations across biomedical terminologies. *AMIA Annual Symposium Proceedings, 2003*, 101–105.

Buscail, L., & Saint-Dizier, P. (2009). Textual and stylistic error detection and correction: Categorization, annotation, and correction strategies. *2009 Eighth International Symposium on Natural Language Processing* (pp. 205–210). IEEE. https://doi.org/10.1109/SNLP.2009.5340918

Caballero, G., & Kapatsinski, V. (2014). Perceptual functionality of morphological redundancy in Choguita Rarámuri (Tarahumara). *Language, Cognition and Neuroscience, 30*(9), 1134–1143. https://doi.org/10.1080/23273798.2014.940983

Chetail, F. (2015). Reconsidering the role of orthographic redundancy in visual word recognition. *Frontiers in Psychology, 6*, 645. https://doi.org/10.3389/fpsyg.2015.00645

Darian, S. (1979). The role of redundancy in language and language teaching. *System, 7*(1), 47–59. https://doi.org/10.1016/0346-251X(79)90022-8

Dasril, Zaim, M., & Ningsih, K. (2019). Coherence and unity of students' writing on background of the study of research proposals. *Proceedings of the 1st International Conference on Education Social Sciences and Humanitie*s. Atlantis Press. https://doi.org/10.2991/icesshum-19.2019.65

Dawson, J. H. (1992). Avoid redundancy in writing. *Weed Technology, 6*(3), 782. https://doi.org/10.1017/S0890037X0003623X

De Beaugrande, R. (1980). Text discourse in European research. *Discourse Processes, 3*, 287–300.

Demir, C. (2019). Writing intelligible English prose: Conciseness vs. verbosity. *Söylem Filoloji Dergisi, 4*(2), 482–505. https://doi.org/10.29110/soylemdergi.617184

Dhivya, P., & Koperundevi, E. (2024). TBLT: Avoiding redundancy to reduce complexity. *Forum for Linguistic Studies, 6*(1), 567–576.

Every, B. (2017). Writing economically in medicine and science: Tips for tackling wordiness. *Medical Writing, 26*, 17–20.

Forlini, G., Bauer, M. B., & Prentice-Hall, I. (1982). *Prentice Hall Grammar and Composition: Level 4*. Prentice Hall PTR. https://books.google.ru/books?id=9KBIjqiYUu8C

Freywald, U., & Finkbeiner, R. (2018). Exact repetition or total reduplication? Exploring their boundaries in discourse and grammar. In R. Finkbeiner & U. Freywald (Eds.), *Exact repetition in Ggammar and discourse* (pp. 3–28). De Gruyter Mouton. https://doi.org/10.1515/9783110592498-001

Grant-Davie, K. (1995). Functional redundancy and ellipsis as strategies in reading and writing. *JAC, 15*(3), 455–469.

Guerrero, F. G. (2009). A new look at the classical entropy of written English (arXiv:0911.2284). arXiv. https://doi.org/10.48550/arXiv.0911.2284

Heltai, P. (2018). Explicitation, redundancy, ellipsis, and translation. In K. Károly & A. Fóris (Eds.), *New trends in translation studies* (pp. 45-74). Akadémiai Kiadó.

Horning, A. S. (1979). On defining redundancy in language: Case notes. *Journal of Reading, 22*(4), 312–320.

Horning, A. S. (1991). Readable writing: The role of cohesion and redundancy. *Journal of Advanced Composition, 11*(1), 135–145. http://www.jstor.org/stable/20865767

Horning, A. S. (1993). *The psycholinguistics of readable writing: A multidisciplinary exploration*. Ablex Publishing Corporation.

Hunnicutt, S. (1985). Intelligibility versus redundancy - Conditions of dependency. *Language and Speech, 28*(1), 47–56. https://doi.org/10.1177/002383098502800103

Kravtchenko, E., & Demberg, V. (2022). Informationally redundant utterances elicit pragmatic inferences. *Cognition, 225*, 105159. https://doi.org/10.1016/j.cognition.2022.105159

Lehmann, C. (2005). Pleonasm and hypercharacterisation. In G. Booij & J. Van Marle (Eds.), *Yearbook of Morphology 2005* (pp. 1–13). Springer. https://doi.org/10.1007/1-4020-4066-0_5

Leufkens, S. (2023). Measuring redundancy: The relation between concord and complexity. *Linguistics Vanguard, 9*(s1), 95–106. https://doi.org/10.1515/lingvan-2020-0143

Lotfipour-Saedi, K. (1982). Applying an analysis of writer-reader discourse processes to a pedagogy of reading [Unpublished doctoral dissertation]. University of Lancaster.

Lotfipour-Saedi, K., & Sarhady, T. (2000). Redundancy, its discoursal function, and textual realizations in different genres. *The International Journal of Humanities, 7*(1), 31–41.

Lyster, R. (1998). Recasts, repetition, and ambiguity in L2 classroom discourse. *Studies in Second Language Acquisition, 20*(1), 51–81. https://doi.org/10.1017/S027226319800103X

Marinashvili, M. (2020). Average word length and text redundancy variability: French texts case study. *Polonia University Scientific Journal, 38*, 67–75. https://doi.org/10.23856/3849

McCrudden, M. T., Hushman, C. J., & Marley, S. C. (2013). Exploring the boundary conditions of the redundancy principle. *The Journal of Experimental Education, 82*(4), 537–554. https://doi.org/10.1080/00220973.2013.813368

McGarry, J. (1975). Redundancy and entropy in language. *The Journal of General Psychology, 93*(1), 101–106.

Newman, E. B., & Waugh, N. C. (1960). The redundancy of texts in three languages. *Information and Control, 3*(2), 141–153. https://doi.org/10.1016/S0019-9958(60)90731-2

Rahman, N., & Borah, B. (2021). Redundancy removal method for multi-document query-based text summarization. *2021 International Symposium on Electrical, Electronics and Information Engineering*. IEEE. https://doi.org/10.1145/3459104.3459197

Rasulov, Z., & Artikov, A. (2023). The principle of redundancy in compound sentences. *Integration Conference on Integration of Pragmalinguistics, Functional Translation Studies and Language Teaching Processes* (vol. 2, pp. 48–51). Bukhara State University.

Rathjens, D. (1985). The seven components of clarity in technical writing. *IEEE Transactions on Professional Communication, 28*(4), 42–46. https://doi.org/10.1109/TPC.1985.6448848

Rosie, A. M. (1973). *Information and communication theory*. Van Nostrand Reinhold Company.

Schlesinger, I. M. (1966). *The influence of sentence structure on the reading process* [Technical Report No. 21]. U.S. Office of Naval Research Information Systems Branch.

Schlesinger, I. M. (1977). *Production and comprehension of utterances*. Lawrence Erlbaum Associates.

Schüler, A., Scheiter, K., & Gerjets, P. (2013). Is spoken text always better? Investigating the modality and redundancy effect with longer text presentation. *Computers in Human Behavior, 29*(4), 1590–1601. https://doi.org/10.1016/j.chb.2013.01.047

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal, 27*(3), 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

Shannon, C. E. (1951). Prediction and entropy of printed English. *The Bell System Technical Journal, 30*(1), 50–64. https://doi.org/10.1002/j.1538-7305.1951.tb01366.x

Skorikova, T. P. (2017). Linguo-pragmatic aspects of the study of Russian public academic speech genres. *Russian Linguistic Bulletin, 4*(12). https://doi.org/10.18454/RULB.12.05

Smith, F. (1971). *Understanding reading: A psycholinguistic analysis of reading and learning to read*. Holt, Rinehart, and Winston.

Smith, F. (1978). *Understanding reading* (2nd ed.). Holt, Rinehart, and Winston.

Stanley, R. (1967). Redundancy rules in phonology. *Language, 43*(2), 393–436. https://doi.org/10.2307/411542

Staub, A. (2015). The effect of lexical predictability on eye movements in reading: Critical review and theoretical interpretation. *Language and Linguistics Compass, 9*, 311–327. https://doi.org/10.1111/lnc3.12151

Thadani, K., & McKeown, K. (2008). A framework for identifying textual redundancy. In *Proceedings of the 22nd International Conference on Computational Linguistics* (pp. 873–880). ACM. https://doi.org/10.3115/1599081.1599191

Tikhonova, E. V., & Mezentseva, D. A. (2024). Wordiness in academic writing: A systematic scoping review. *Research Result. Theoretical and Applied Linguistics, 10*(1), 133–157. https://doi.org/10.18413/2313-8912-2024-10-1-0-8

Trudgill, P. (2009). Sociolinguistic typology and complexification. In G. Sampson, D. Gil, & P. Trudgill (Eds.), *Language complexity as an evolving variable* (pp. 98–109). Oxford University Press. https://doi.org/10.1093/oso/9780199545216.003.0007

Trudgill, P. (2011). *Sociolinguistic typology: Social determinants of linguistic complexity*. Oxford University Press.

Tuinman, J. J., & Gray, G. (1972). The effect of reducing the redundancy of written messages by deletion of function words. *The Journal of Psychology, 82*(2), 299–306. https://doi.org/10.1080/00223980.1972.9923820

Wallwork, A., & Southern, A. (2020). Chapter 5: Readability and avoiding redundancy. In A. Wallwork & A. Southern (Eds.), *100 tips to avoid mistakes in academic writing and presenting* (pp. 39–74). Springer International Publishing. https://doi.org/10.1007/978-3-030-44214-9_5

Wang, F. X. (2001). *English-Chinese comparative semantics*. Foreign Language Press.

Watkowska, D. (2021). Redundancy in ELF: A corpus-based study on negative and modal concord. *Anglica: An International Journal of English Studies, 30*(2), 71–86. https://doi.org/10.7311/0860-5734.30.2.04

Wit, E. C., & Gillette, M. (1999). *What is linguistic redundancy?* [Technical Report]. https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.714.5763&rep=rep1&type=pdf

Wolf, L., Pimentel, T., Fedorenko, E., Cotterell, R., Warstadt, A., Wilcox, E., & Regev, T. (2023). *Quantifying the redundancy between prosody and text*. *arXiv*. https://doi.org/10.48550/arXiv.2311.17233

Xu, S. H. (1984). The redundancy of language. *Modern Foreign Languages, 2*, 3–8. https://doi.org/CNKI:SUN:XDWY.0.1984-02-000

Xue, H., & Hwa, R. (2014). Redundancy detection in ESL writings. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing* (pp. 683–691). https://doi.org/10.3115/v1/E14-1072

Yang, Y. (2021). The investigation on redundancy errors in writing of Chinese English learners with different proficiency. *Proceedings of the 2021 International Conference on Education, Language and Art.* Atlantis Press. https://doi.org/10.2991/assehr.k.220131.025

Zola, D. (1981). *The effect of redundancy on the perception of words in reading* [Technical Report]. University of Illinoys.

https://doi.org/10.17323/jle.2024.18806

Research Trends on Students' Writing Skills: A Bibliometric Analysis Using Scopus Database

Jaja

Universitas Swadaya Gunung Jati Cirebon, Indonesia

ABSTRACT

Introduction: Students' writing skills positively impact their overall academic ability. Writing allows students to gain knowledge while enhancing their intelligence, initiative, courage, and willpower. Moreover, it fosters creativity, inspiring students to think outside the box. Therefore, it can be said that teaching students to write is crucial to the learning process in the classroom, even though, in reality, teaching students to write is an activity that neither teachers nor students enjoy. It seems contradictory that writing is both necessary and undesirable simultaneously.

Purpose: This study investigates research trends in students' writing skills through bibliometric analysis recorded on the Scopus database from 2014 to 2023.

Method: The study identified 7650 publications in the Scopus database when searching for the keyword "students, writing, and skills." Using the Prisma model, it selected 2214 publications as the sample. The data were analyzed using bibliometric analysis.

Results: The analysis has revealed that Graham *(f= 105)*, Hwang *(f= 102)*, and Kent *(f= 84)* were the top three authors who investigated students' writing skills in their studies, showcasing the diversity of research in our field. Meanwhile, Reading and Writing, Asian EFL Journal, and International Journal of Instruction are the top three publishers with the most publications, further highlighting the breadth of our field. At the country level, the United States is the most productive country, with Arizona State University being the most productive university in research about students' writing. The study also found that writing in higher education contexts is the most widely researched area.

Conclusion: More importantly, the bibliometrics analysis has also revealed that argumentative writing skills, collaborative writing, EFL learners, and writing attitude are some areas that need more research.

KEYWORDS

bibliometric analysis, research trends, Scopus database, students writing skills

INTRODUCTION

The four essential skills that students must develop and teachers should teach in language learning in an integrated manner are listening, speaking, reading, and writing (Yusuf et al., 2019). Writing is one of the four essential language skills students must master. It involves using language to convey ideas, emotions, or intentions through written expression in the form of text (Putri & Aminatun, 2021). Each text, be it a narrative, a poem, or an academic paper, carries unique characteristics, functions, purposes, structures, and guidelines (Jaja et al., 2019). From a viewpoint that treats texts as independent entities, students' compositions are viewed as langue-meaning they reflect the writer's understanding of structure and their awareness of the rules governing text creation (Hyland, 2002). Therefore, effective writing is not just about precision but about embracing the diversity and richness of language by communicating the writer's intended meaning. In 2004, a survey by *the American National Commission on Writing* on 120 large American companies employing eight million people indicated that writing is a threshold skill that offers a ticket into a professional career (James, 2007).

Citation: Jaja. (2024). Research trends on students' writing skills: A bibliometric analysis using Scopus database. *Journal of Language and Education, 10*(3), 161-182. https://doi.org/10.17323/jle.2024.18806

Correspondence:
Jaja
E-mail: jaja@ugj.ac.id

Received: December 28, 2023
Accepted: May 16, 2024
Published: September 30, 2024

The development of science and technology changed how people write. To deal with it, teachers must find different strategies for students to learn to write (Bugis, 2018; Chevalier, 2011; Davis & Davis, 2013; Hung et al., 2012; Lipschutz, 2010) from an early age (Tolchinsky, 2016). Guidance and instructions in writing and knowledge about existing literary works can influence students' writing interests and motivation. As literature has repeatedly documented, writing can be considered a motor activity that involves mental processes because writing is a medium to convey what the writer feels (Muhanif et al., 2021; Nugroho, 2014). Writing is also a productive language activity that produces work from creative and critical thinking processes (Karim & Mustapha, 2020). Several studies show that many students experience difficulties in writing, both in terms of technique and content (Whai et al., 2013). Ceylan (2019) also identified that difficulty expressing ideas contributes to low writing skills. Therefore, teachers and schools must pay sufficient attention to developing students' writing skills by fostering their learning motivation. Jaja and Rahayu (2021) claim that exposing students to the benefits of writing could foster their motivation to learn to write.

In line with motivation, effective instructional strategies in writing are also critical. Sakkir and Dollah (2019) suggested that developing students' writing skills can be started by training in proper writing techniques, such as good grammar, punctuation, and writing structures. Furthermore, many studies have also recommended the use of various activities such as essaying, journaling, short stories, and poetry writing as effective strategies for teaching writing (Coleman & Willis, 2015; Iftanti, 2016; Kottacheruvu, 2023; Lee & Deakin, 2016). In addition, students also need to be taught writing strategies, such as brainstorming, outlining, and revision (Sadiku, 2015). By mastering good writing techniques and methods, students can write more effectively and efficiently in determining future students' academic and workplace success (Barone, 2010).

However, writing skills are not only related to writing techniques and strategies; the content and ideas conveyed are also critical in writing. Writing down concepts offered in class or a text helps students grasp and retain information (Bangert-Drowns et al., 2004; Graham & Hebert, 2011; Graham & Perin, 2007). In addition, students need to get constructive feedback from teachers and classmates (Elbow, 1973). This feedback can help students improve their writing quality (DaCrema & Stout, 2012; Stout, 2014). In this case, the teacher can provide specific and clear feedback regarding aspects that need improvement, such as grammar, writing structure, and content.

Students' writing skills have a positive impact on their overall academic ability. Students can gain knowledge by writing while also enhancing their intelligence, initiative, and creativity, as well as their courage and willpower (Baker, 2016; Li et al., 2022; Myhill, 2018). Therefore, teaching students to write is crucial to the learning process in the classroom, even though, in reality, teaching students to write is an activity that neither teachers nor students enjoy. It seems contradictory that writing is both necessary and undesirable simultaneously. The problems in learning to write in schools like that can be studied using bibliometric analysis.

Bibliometric analysis regarding students' writing skills is still minimal. Although several researchers have carried out this study, such as Crosthwaite et al. (2022) and Sun and Lan (2023), the bibliometric analysis carried out only focuses on the L2 written corrective feedback aspect, meanwhile Hyland and Jiang (2022) also only focuses on the interaction aspect in written texts. In this research, the bibliometric analysis focuses on various writing skills that have positive and significant implications for students' writing skills at multiple levels of education. Thus, this research aims to look at research trends and find important aspects of students' writing skills based on bibliometric analysis from the Scopus database in 2014-2023. It is hoped that the results of this research will be helpful and become a reference for teachers and researchers to improve students' writing skills in the future.

LITERATURE REVIEW

Writing Theory

A Cognitive Process Theory of Writing

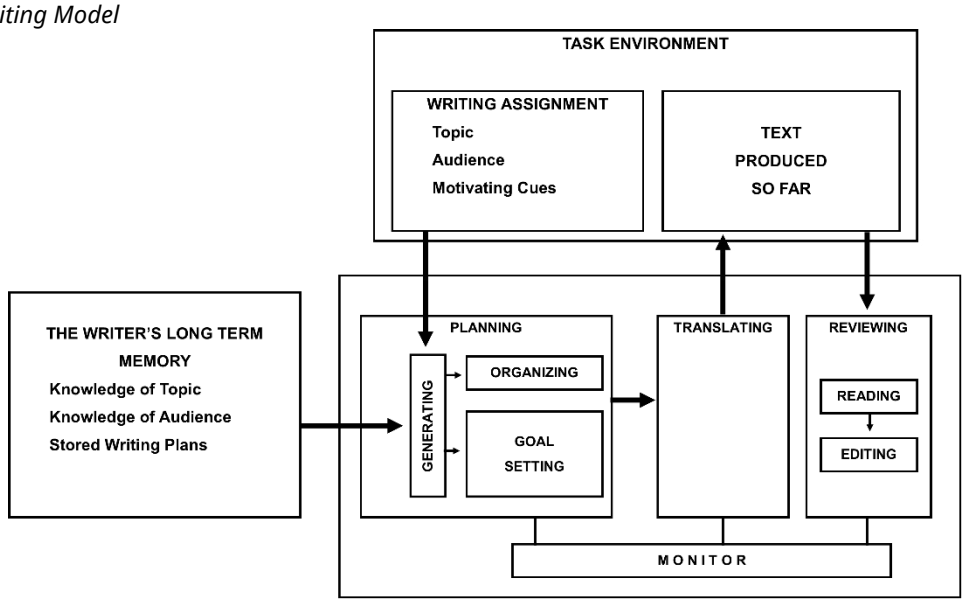
Flower and Hayes (1981) aimed to present a formal model of the writing process. Their theory is built upon four hypotheses. Firstly, they propose that the writing process is most effectively comprehended as a sequence of distinct cognitive processes that the writer organizes or regulates while actively engaged in writing. Secondly, these processes exhibit a hierarchical and closely interconnected structure, wherein any specific process can be intricately embedded within other processes. Thirdly, composing itself is depicted as a goal-directed cognitive process shaped by the author's network of fundamental goals. Lastly, writers formulate their objectives through two primary methods: formulating overarching goals and corresponding sub-goals and occasionally modifying the principal goal or establishing new goals based on acquired knowledge. Figure 1 illustrates the primary components of the writing process, as Flower and Hayes (1981) proposed.

Figure 1 delineates the primary constituents of the writing process. The composing process model encompasses three key components:

(1) The task environment, encompassing all external elements beyond the writer's physical boundaries, includes the rhetorical problem or assignment and extends to incorporate the evolving text.

Figure 1

Structure of the Writing Model



(2) The writer's long-term memory serves as the repository for the writer's accumulated knowledge, encompassing not only the subject matter but also understanding the audience and various writing strategies.

(3) The writing processes themselves, specifically the fundamental processes of Planning, Translating, and Reviewing, all of which fall under the supervision of the Monitor.

Social Constructivist Theory of Writing

The social constructivist theory of writing suggests that writing is essentially a social activity shaped by interactions within a community and cultural environment. According to this theory, writing is not a solitary endeavor but a collaborative process where knowledge and meaning are developed through social interaction. Hyland (2002) points out that writing is deeply social, with ideas being formed and reformed through engagement with others. The dynamics between the writer and their audience and the norms of their discourse community significantly influence the writing process. Bazerman (1988) adds that writing is embedded in a social context, where the writer's relationship with their readers and the community's expectations play crucial roles. This viewpoint highlights that writing involves negotiating meaning within a social framework, where the writer and the audience actively create and interpret.

Genre Theory of Writing

Genre theory suggests that writing is influenced by recurring social situations and the specific conventions associated with different genres. Writers develop their expressive skills by learning and adhering to the rules, structures, and expectations unique to each genre. According to Miller (1984), genres are not merely forms; they represent ways of acting, interacting with others, and making sense of the world. This viewpoint highlights that writing involves personal creativity and responding to the social contexts in which it occurs. Devitt (2004) builds on this idea by explaining that writers utilize genres as tools to engage with a community, achieve specific goals, and address particular audiences. Therefore, genre theory emphasizes the significance of understanding the rhetorical situation and social purpose behind each writing act.

Expressivist Theory of Writing

The expressivist theory of writing highlights writing as a tool for personal expression, creativity, and self-development. It sees writing as a way to uncover one's thoughts and feelings, emphasizing the importance of individuality and authenticity. According to Expressivist views, writing should not be constrained by strict rules to allow for the development of a writer's distinct style. Elbow (1973) observes that writing can lead to new insights and thoughts that were not initially apparent. This approach promotes exploring personal ideas and experiences, helping writers discover their unique voices. As Murray (1985) suggests, effective writing results from a writer's self-exploration and engagement with language and the world. The theory supports writing as a reflective and investigative activity, enabling writers to gain deeper self-awareness and perspective through their writing practice.

Types of Students' Writing Skills

Students' writing skills can be categorized into several types, each reflecting different aspects of writing proficiency.

Narrative Writing Skills

Narrative writing skills are crucial for enabling writers to convey stories in an engaging and organized manner. Narrative writing involves crafting a story that includes elements such as characters, plot, setting, and theme, all of which contribute to developing a cohesive narrative. According to Graham and Perin (2007), narrative writing skills help students organize and structure ideas into a coherent story that readers can follow. They also note that narrative writers should be able to use various techniques, such as description, dialogue, and characterization, to create a pleasurable and impactful reading experience. Furthermore, Applebee and Langer (2011) indicate that narrative writing skills are essential for enhancing students' ability to communicate effectively and creatively across different contexts. Thus, developing narrative writing skills is a key aspect of language education that supports students in expressing their ideas more efficiently and imaginatively.

Expository Writing Skills

Expository writing skills are crucial for conveying information clearly and systematically. Expository writing aims to explain or describe a topic objectively without influencing the reader with personal opinions. According to Anson and Schwegler (2012), expository writing involves presenting information in an organized manner supported by relevant evidence, which helps readers better understand complex topics. They state that expository writing requires arranging information logically and clearly, allowing readers to follow arguments and understand the material presented. Mastering these skills is essential in academic and professional contexts as it enables individuals to communicate ideas and information effectively and systematically.

Persuasive Writing Skills

Persuasive writing skills are a crucial aspect of effective communication, where the writer aims to influence the reader's opinions or actions through convincing arguments. According to Hyland (2005), persuasive writing involves conveying information, constructing solid arguments, and using appropriate rhetorical strategies to affect the audience. These skills include the ability to identify and understand the audience, formulate logical and data-based arguments, and address objections convincingly. For instance, Tannen (1998) suggests that the ability to present arguments persuasively can enhance communication effectiveness across various contexts, from academic writing to professional communication. Mastering these skills makes writers more effective in persuading readers and achieving their communication goals.

Descriptive Writing Skills

Descriptive writing skills involve creating vivid and detailed imagery through written language, allowing readers to visualize and experience scenes, characters, and events as though they were present. This skill requires using rich, sensory details and precise language to craft a mental image for the reader. Langan (2001) highlights that descriptive writing appeals to the reader's senses and emotions, offering a more immersive and memorable experience. This writing style is vital across various genres, including fiction, memoir, and non-fiction, as it helps establish atmosphere, develop characters, and improve storytelling. Williams (2006) adds that effective descriptive writing offers a clear depiction, stirs emotions, and sets the narrative tone. Mastery of descriptive writing is crucial for captivating readers and enhancing the text's overall impact.

Creative Writing Skills

Creative writing skills involve the ability to write imaginatively and originally, encompassing character development, plot construction, and expressive language. According to Burroway (2019), these skills require a deep understanding of narrative structure and writing techniques that allow the writer to create a touching and engaging reading experience. King (2000) explains that creative writing involves a reflective process where writers refine their ideas and translate them into authentic and compelling prose. These skills are not only about creativity but also about mastering writing techniques that enable effective self-expression.

Analytical Writing Skills

Analytical writing skills are crucial abilities that involve critical and systematic thinking in crafting well-structured arguments and analyses. According to Brown (2014), these skills are essential for assessing arguments, logically organizing thoughts, and presenting evidence effectively. This ability not only aids in creating coherent writing but also in constructing in-depth, evidence-based arguments. Furthermore, Chaffee (2015) highlights that effective analytical writing requires a clear understanding of the topic, the ability to break down complex ideas, and the capacity to present findings in a structured manner. With these skills, writers can produce informative and persuasive work, thereby contributing to broader academic and professional discussions.

Academic Writing Skills

Academic writing skills are a crucial component of higher education, involving organizing, structuring, and presenting ideas clearly and systematically in an academically recognized format. According to Leki (1998), academic writing

is the process used in scholarly contexts to communicate complex ideas and arguments clearly and coherently. This skill encompasses knowledge of grammar and spelling and the ability to construct arguments, conduct research, and format writing according to academic standards. Flowerdew and Peacock (2001) further emphasize that effective academic writing entails engaging with and contributing to ongoing scholarly discussions within a discipline. Thus, academic writing skills are essential for academic and professional success, enabling individuals to participate in intellectual discourse actively and produce work that significantly impacts their field of study.

Technical Writing Skills

Technical writing skills refer to the ability to convey complex information in a clear, structured, and understandable manner. This type of writing is commonly used in documents such as user manuals, technical reports, and software guides. According to Markel (2018), technical writing involves presenting information clearly and accurately while considering the audience and communication goals. This skill encompasses a deep understanding of the technical topic, the use of appropriate language, and the ability to organize information logically. Technical writers must be able to create content that is not only informative but also engaging and accessible to readers from diverse backgrounds (Gerson & Gerson, 2018). As technology and system complexity evolve, technical writing skills become increasingly crucial in helping users understand and utilize products or services effectively.

Research Focus

In this study, the term "students" includes individuals at various levels of education, including pupils, students, and college students. Meanwhile, the main focus of this study on writing skills is on the development of overall writing skills, including techniques and strategies, and how current studies address and support various aspects of writing skills. The "skills" referred to in this study are multiple types of students writing skills (such as narrative, expository, persuasive, descriptive, creative, analytical, academic, and technical writing skills) which include various components of writing skills, such as specific techniques (e.g., grammar and sentence structure), strategies (e.g., brainstorming and revision), and process approaches (e.g., drafting and editing).

These "skills" encompass all these aspects, providing a comprehensive picture of how writing skills are developed and learned. Thus, it can be said that the context of students writing skills in this study involves various types of students writing skills (such as narrative, expository, persuasive, descriptive, creative, analytical, academic, and technical writing skills) that students need to have and develop in various aspects, such as writing techniques, strategies, and process approaches at multiple levels of education, based on research from the Scopus database in 2014-2023.

Research Questions

This study aims to find trends and novelties in students writing skills research in the last ten years (2014-2023). The previous ten years were chosen for investigation since more science and technology were used in education during this time. The advancement of science and technology in language teaching and learning makes it easier for students to create and publish their writings in various electronic media, including social media, journals, and other supporting media. Some questions underlie research on students' writing skills by retrieving data from the Scopus publication database.

RQ#1: What is the frequency of research on students writing skills in 2014-2023?

RQ#2: What keywords are the most widely used in the title of research articles on students' writing skills?

RQ#3: What words are the most widely used in the abstract of students' writing skills research articles?

RQ#4: Who is the author with the most citations in the students' writing skills articles?

RQ#5: What publisher publishes the most articles on students' writing skills?

RQ#6: Which countries and organizations are the most productive in students' writing skills research?

RQ#7: Which level of education is the most widely used in research on students' writing skills?

RQ#8: What are the future writing skills research opportunities?

METHOD

This study aims to visualize research trends using bibliometric analysis. Bibliometric visualization uses *Vos Viewer* software, which is free and easy to use to analyze trends, opportunities, or relationships between keywords in research.

Sample

The sample of this research is 2214 research articles from the Scopus database. The search uses a Scopus account

with an institution login because by using a personal Scopus account, researchers cannot search articles in the database but can only search for authors. The settings for searching the results of the research are *TITLE-ABS-KEY (students AND writing AND skills) AND PUBYEAR > 2013 AND PUBYEAR < 2024*.

Indicators

The indicators used in this study are the publication between 2014 and 2023 (the last ten years), the trend of keywords in research titles, popular words in abstracts, authors with the most significant number of citations, most productive countries, most writing skills, level of education, and writing skills research opportunities in the future. The indicator refers to the various types of students' writing skills (such as narrative, expository, persuasive, descriptive, creative, analytical, academic, and technical writing skills) that students need to have and develop in various aspects, such as writing techniques, strategies, and process approaches at multiple levels of education.

Data Collection and Analysis

The data collection procedure is done by specifying the Scopus database as the data source. Next, a search for research results articles was conducted in the database using the keywords explained in the sample described above. Search results on the Scopus database in the last ten years (2014-2023) for research on students' writing skills obtained 7629 research articles. The results are presented in Figure 2 below.

The search results are exported into CSV (comma-separated value) format for manual analysis and visualization using *Vos Viewer software*. To tidy up the research article data in CSV form, open the file using *Microsoft Excel software,* then set the file using the open *Microsoft Excel procedure* > After it is open > click the data tab > select *from text/CSV* > select the location of the *export CSV database file* > *load* > wait a few moments for the CSV data to be fully loaded. After using these settings, the CSV file will be easier to read and can be analyzed to select research results. The display conversion is presented in Figure 3.

There were 7650 research articles, which were analyzed manually using *Microsoft Excel*. At this stage, 3156 research articles were obtained unrelated to students' writing skills; thus, the search results that matched the keywords totaled 4494 research articles only. Then, 4494 research articles were re-selected (manually using Microsoft Excel) to ignore research results not sourced from *journals, conferences, books, book chapters, reviews, editorials, erratum, letters, notes, and reviews*. From this step, 2280 articles were irrelevant. Thus, the final results were 2214 research articles from journal publications that match the keywords of students' writing skills, which can be used for bibliometric analysis. The selection procedure for the research results used PRISMA models sourced from Moher et al. (2009), presented in Figure 4.

Data were analyzed using the *Vos Viewer software* bibliometric analysis. The VosViewer software settings use the setting defined by Syahid and Qodir (2021) with slight modifications, as presented in Table 1.

Figure 2

Search Queries in the Scopus Database

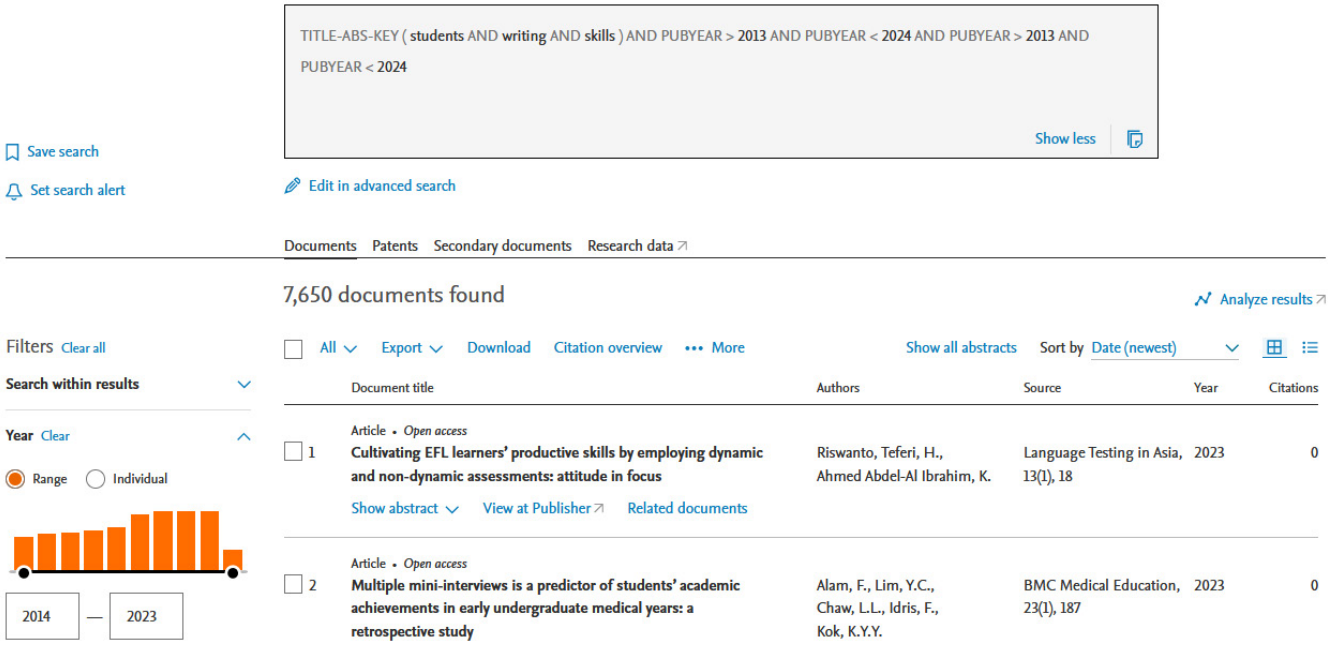
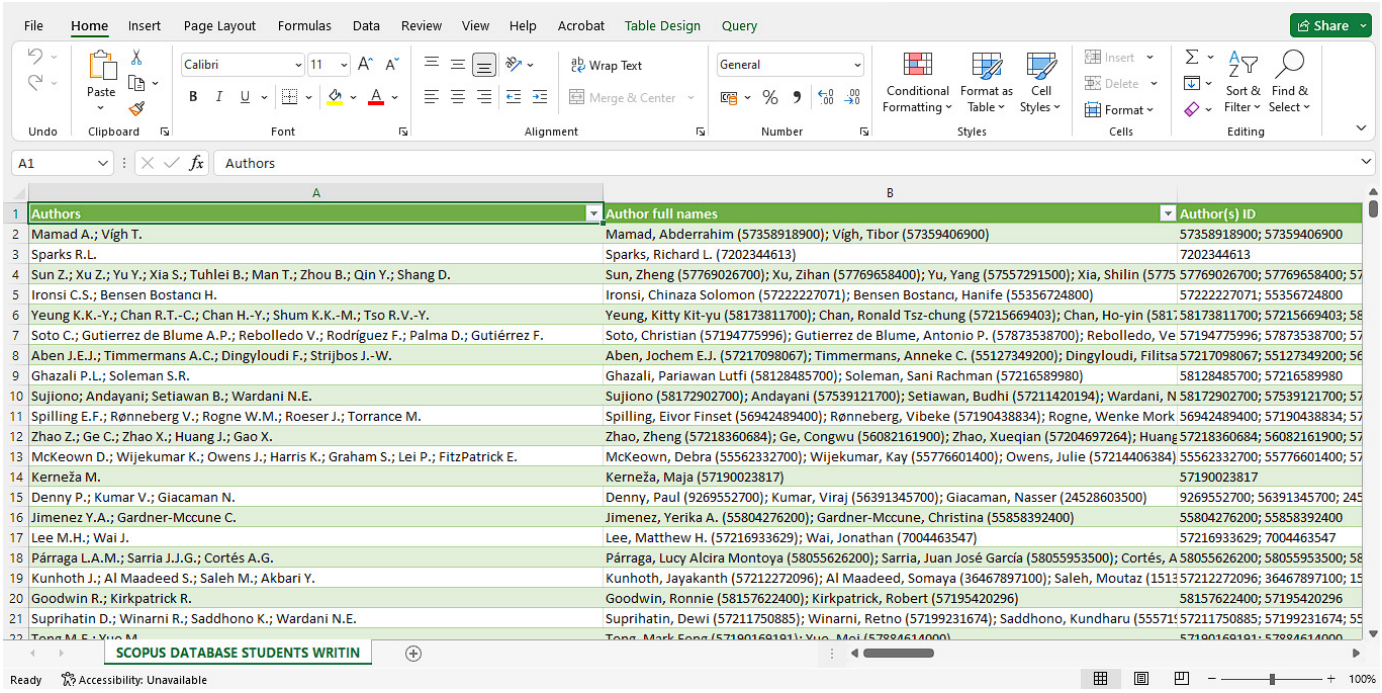


Figure 3

CSV Layout Format into Microsoft Excel View



	A	B	
1	Authors	Author full names	Author(s) ID
2	Mamad A.; Vígh T.	Mamad, Abderrahim (57358918900); Vígh, Tibor (57359406900)	57358918900; 57359406900
3	Sparks R.L.	Sparks, Richard L. (7202344613)	7202344613
4	Sun Z.; Xu Z.; Yu Y.; Xia S.; Tuhlei B.; Man T.; Zhou B.; Qin Y.; Shang D.	Sun, Zheng (57769026700); Xu, Zihan (57769658400); Yu, Yang (57557291500); Xia, Shilin (5775	57769026700; 57769658400; 57
5	Ironsi C.S.; Bensen Bostancı H.	Ironsi, Chinaza Solomon (57222227071); Bensen Bostancı, Hanife (55356724800)	57222227071; 55356724800
6	Yeung K.K.-Y.; Chan R.T.-C.; Chan H.-Y.; Shum K.K.-M.; Tso R.V.-Y.	Yeung, Kitty Kit-yu (58173811700); Chan, Ronald Tsz-chung (57215669403); Chan, Ho-yin (581	58173811700; 57215669403; 58
7	Soto C.; Gutierrez de Blume A.P.; Rebolledo V.; Rodríguez F.; Palma D.; Gutiérrez F.	Soto, Christian (57194775996); Gutierrez de Blume, Antonio P. (57873538700); Rebolledo, Ve	57194775996; 57873538700; 57
8	Aben J.E.J.; Timmermans A.C.; Dingyloudi F.; Strijbos J.-W.	Aben, Jochem E.J. (57217098067); Timmermans, Anneke C. (55127349200); Dingyloudi, Filitsa	57217098067; 55127349200; 56
9	Ghazali P.L.; Soleman S.R.	Ghazali, Pariawan Lutfi (58128485700); Soleman, Sani Rachman (57216589980)	58128485700; 57216589980
10	Sujiono; Andayani; Setiawan B.; Wardani N.E.	Sujiono (58172902700); Andayani (57539121700); Setiawan, Budhi (57211420194); Wardani, N	58172902700; 57539121700; 57
11	Spilling E.F.; Rønneberg V.; Rogne W.M.; Roeser J.; Torrance M.	Spilling, Eivor Finset (56942489400); Rønneberg, Vibeke (57190438834); Rogne, Wenke Mork	56942489400; 57190438834; 57
12	Zhao Z.; Ge C.; Zhao X.; Huang J.; Gao X.	Zhao, Zheng (57218360684); Ge, Congwu (56082161900); Zhao, Xueqian (57204697264); Huang	57218360684; 56082161900; 57
13	McKeown D.; Wijekumar K.; Owens J.; Harris K.; Graham S.; Lei P.; FitzPatrick E.	McKeown, Debra (55562332700); Wijekumar, Kay (55776601400); Owens, Julie (57214406384)	55562332700; 55776601400; 57
14	Kerneža M.	Kerneža, Maja (57190023817)	57190023817
15	Denny P.; Kumar V.; Giacaman N.	Denny, Paul (9269552700); Kumar, Viraj (56391345700); Giacaman, Nasser (24528603500)	9269552700; 56391345700; 245
16	Jimenez Y.A.; Gardner-Mccune C.	Jimenez, Yerika A. (55804276200); Gardner-Mccune, Christina (55858392400)	55804276200; 55858392400
17	Lee M.H.; Wai J.	Lee, Matthew H. (57216933629); Wai, Jonathan (7004463547)	57216933629; 7004463547
18	Párraga L.A.M.; Sarria J.J.G.; Cortés A.G.	Párraga, Lucy Alcira Montoya (58055626200); Sarria, Juan José García (58055953500); Cortés, A	58055626200; 58055953500; 58
19	Kunhoth J.; Al Maadeed S.; Saleh M.; Akbari Y.	Kunhoth, Jayakanth (57212272096); Al Maadeed, Somaya (36467897100); Saleh, Moutaz (1513	57212272096; 36467897100; 15
20	Goodwin R.; Kirkpatrick R.	Goodwin, Ronnie (58157622400); Kirkpatrick, Robert (57195420296)	58157622400; 57195420296
21	Suprihatin D.; Winarni R.; Saddhono K.; Wardani N.E.	Suprihatin, Dewi (57211750885); Winarni, Retno (57199231674); Saddhono, Kundharu (55571	57211750885; 57199231674; 55

Figure 4

Prisma Flow for Sample Identification

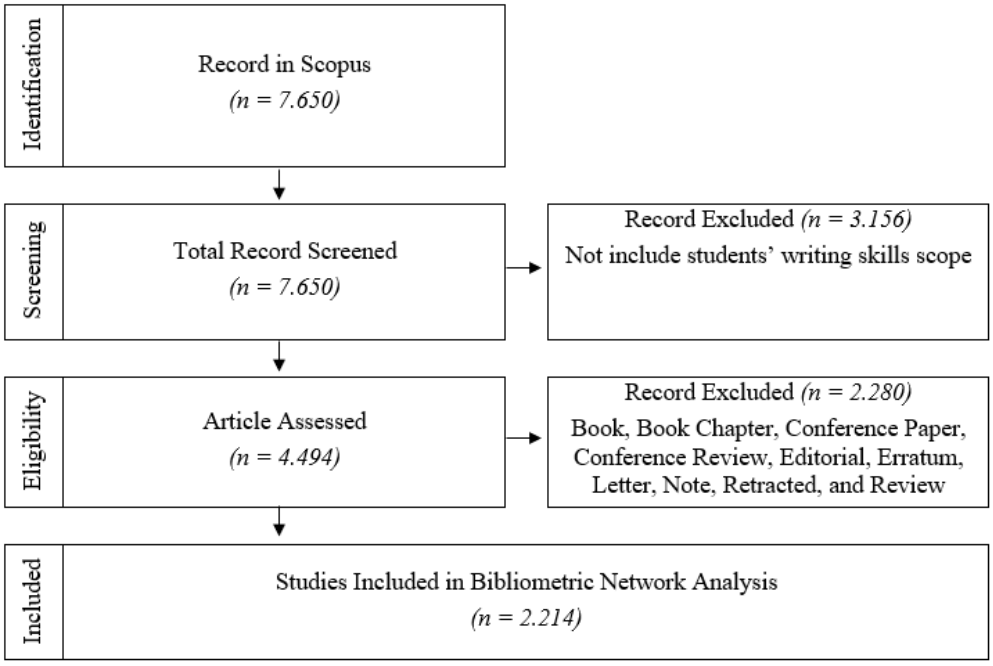


Table 1
Vos Viewer Software Retrieval Setting to Bibliometric Analysis

Type and Unit of Analysis	Counting Method	Threshold	A Large Number of Analysis Unit	Selection
Co-Authorship				
Authors	Full	1 Document	Not Ignored	Maximum
Organizations	Full	1 Document	Not Ignored	Maximum
Countries	Full	1 Document	Not Ignored	Maximum
Co-Occurrence				
Keywords	Full	5 Occurrences	Ignored	281
Abstract Words	Full	10 Occurrences	Ignored	721
Citation				
Authors	N/A	1 Document	Not Ignored	Maximum
Documents	N/A	10 Citations	N/A	Maximum

RESULTS

Based on the findings of a bibliometric analysis of 2214 research articles in the last ten years (2014-2023) on the *Scopus database* with *Vos Viewer software*, this study obtained data related to the keywords that were most widely used in titles and abstracts, authors with the highest number of citations, publishers with the highest number of publications, the most productive countries, and organizations in research. The study also collected data about the writing skills of the students under study, their level of education, and opportunities or *novelties* for future research. These results are presented in the visualization image and table below.

Research Frequency Related to Students' Writing Skills 2014-2023

The frequency of students' writing skills research results from the Scopus database in 2014-2023 was analyzed using Microsoft Excel software CSV data. Sorting is done to group publications by year of publication. Based on this analysis, data was obtained that research on students' writing skills had increased in 2014-2020 and fluctuated in 2021-2023. During this period, the highest number of publications occurred in 2022 (340 publications). The research frequency data is presented in Figure 5.

The Most Widely Used Keywords in Students' Writing Skills Research Articles

The stages in making *a map based on bibliographic data* for *the most widely used keywords* in the study were carried out using *co-occurrence* analysis based on the author's *keywords*. *The minimum number of occurrences* is automatically set at five, and the number of keywords is 281. The results are presented in Figure 6. Based on this analysis, the data obtained are 10 clusters with the most keywords: ***writing*** *(f= 304)*. These results are relevant to other keywords, such as ***writing skills*** *(f= 166)* and ***academic writing*** *(f=150)*, which are the most widely used in students' writing skills research. This finding indicates that most of the research focuses on students' writing skills in the academic field.

The Words That Are Most Widely Used in Abstract Research Articles on Students' Writing Skills

The analysis of the most widely used words in the abstract is carried out by *creating a map based on text data* by selecting abstract *fields* to get data. The calculation uses *full counting* with *a minimum number of occurrences* of 10 and *several terms* of 721. The visualization results are presented in Figure 7. The results show 10 clusters with the word ***effect*** *(f= 616)*, the most used in the abstract. Other words that are widely used are ***test*** *(f= 459)*, ***control group*** *(f= 369)*, ***experimental group*** *(f= 344),* and ***grade*** *(f= 344)*. These results indicate that most research articles on students' writing skills focus on the influence or effect caused by the research group *(experimental and control)* on learning outcomes (*grade*) after a series of learning activities and tests.

The Writer with The Most Number of Citations in Students' Writing Skills

The total number of authors of research articles on students' writing skills based on the *Scopus database* is 2158 authors from 2214 publications. To find the authors with the

Figure 5

Publication Frequency on Students Writing Skills Researchers by Year

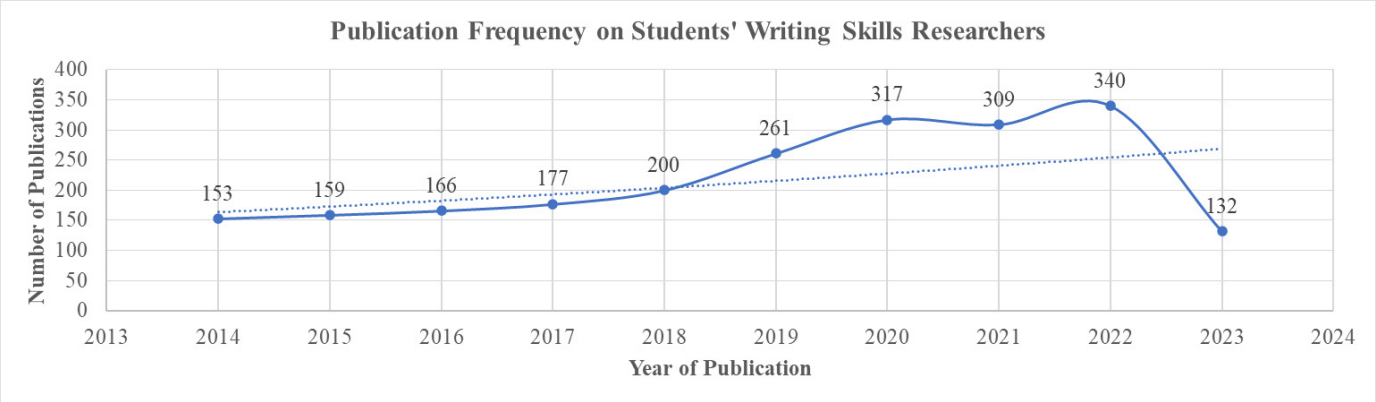
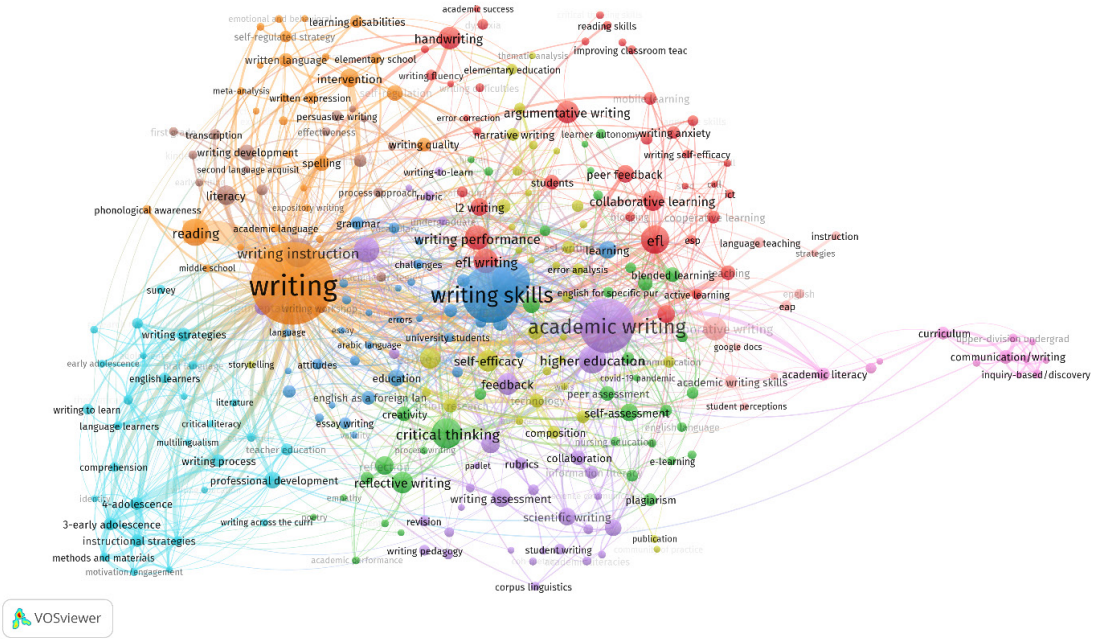


Figure 6

The Most Widely Used Keywords in Articles Relating to the Research on Students Writing Skills



highest number of citations, this study analyzed *co-authorship* with *the minimum number of documents* set as one and the maximum *number of authors* set *(f= 2158)*. Based on this analysis, ten authors with the highest number of citations were obtained. The top three authors are ***Graham*** (2019) with 105 citations, ***Hwang*** (2014) with 102 citations, and ***Kent*** (2014) with 84 citations. The complete data are presented in Table 2.

Based on Table 2, writing skills are an essential aspect that can be applied at various levels of education, individually and collaboratively. In addition, learning writing skills can also be integrated using media such as *game-based learning* and *augmented reality*.

The Publishers with The Most Publication in Students' Writing Skills Research Articles

Several publishers of students' writing skills research articles with the most publications were obtained through *CSV export data* from the *Scopus database*. The authors used CSV data to obtain publisher data and then sorted it manually using *Microsoft Excel* software to obtain 794 publishers. This analysis process has identified three publishers with the most publications of research articles on students' writing skills: *Reading and Writing (f= 65)*, *Asian EFL Journal (f= 49)*, and *International Journal of Instruction (f= 37)*. Complete data is presented in Table 3.

Figure 7

The Word That is Most Widely Used in an Abstract Article

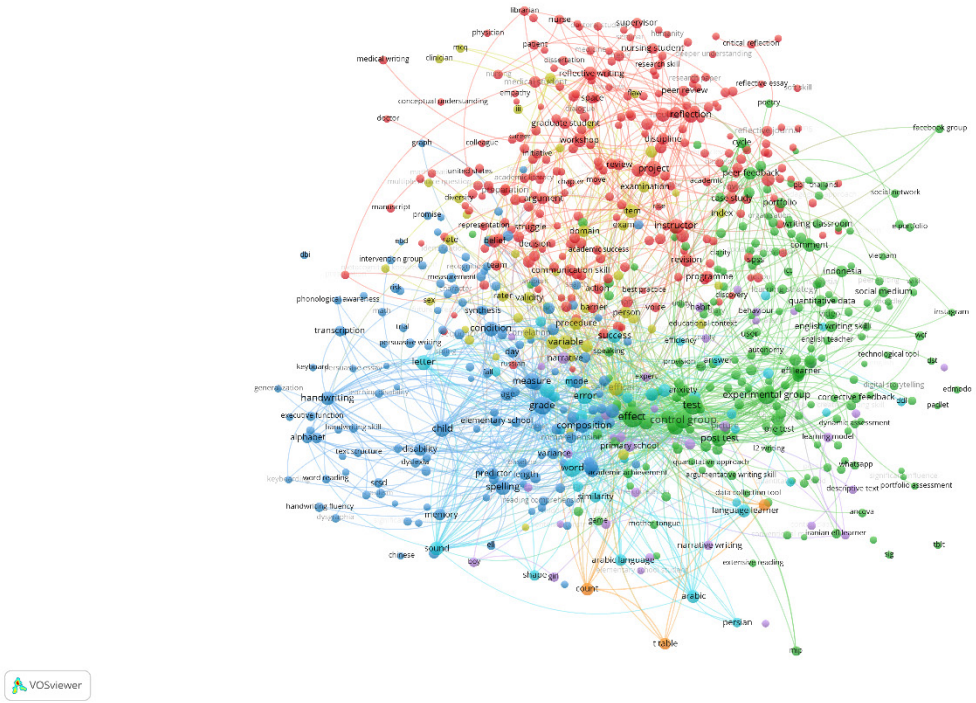


Table 2

The Most Cited Author

Author and Year	Title	Cited
Graham (2019)	Changing how writing is taught	105
Hwang et al. (2014)	Improving English as a foreign language writing in elementary schools using mobile devices in familiar situational contexts	102
Kent et al. (2014)	Writing fluency and quality in kindergarten and first grade: The role of attention, reading, transcription, and oral language	84
Charon et al. (2016)	Close reading and creative writing in clinical education: Teaching attention, representation, and affiliation	79
Koster et al. (2015)	Teaching children to write: A meta-analysis of writing intervention research	79
Wang (2014)	Promoting collaborative writing through wikis: A new approach for advancing innovative and active learning in an ESP context	76
Lin et al. (2018)	A flipped contextual game-based learning approach to enhancing EFL students' English business writing performance and reflective behaviors	71
Wang (2017)	Exploring the effectiveness of integrating augmented reality-based materials to support writing activities	67
Naber & Wyatt (2014)	The effect of reflective writing interventions on the critical thinking skills and dispositions of baccalaureate nursing students	67
Graham et al. (2018)	Effectiveness of literacy programs balancing reading and writing instruction: A meta-analysis	66

The Most Productive Countries and Organizations in Students' Writing Skills Research Articles

The Most Productive Country

Data from the most productive countries were obtained from the bibliographic database from the Scopus database, which included the *type of analysis co-authorship* and unit of analysis *countries*. The minimum document of a country is set at one, and the number of countries is set at 114. This analysis shows that the *United States* ranks first as the most productive country, with 642 documents. Meanwhile, *Indonesia* ranks second with 172 papers, and *Malaysia* is third with 119 papers. The complete data are presented in Table 4.

The Most Productive Organization

Data of the most productive organizations were obtained *from the bibliographic database* from the *Scopus database* through the *type of analysis co-authorship* and *unit of analysis organizations*. The minimum document of a country is set at one, and the maximum *number of countries* is set at 3655. The results show that *Arizona State University* is the most productive organization, with nine documents, followed by *Georgia State University* and the *University of Delaware*, ranked second and third, respectively, with seven papers each. The complete data are presented in Table 5.

Table 3

Top 10 Publishers on Students' Writing Skills Research Articles

Publisher	Publication	Percentage (%)
Reading and Writing	65	2.95
Asian EFL Journal	49	2.22
International Journal of Instruction	37	1.68
Journal of Language Teaching and Research	28	1.27
Theory and Practice in Language Studies	22	1.00
Computer-Assisted Language Learning	20	0.91
Reading and Writing Quarterly	20	0.91
Frontiers in Psychology	19	0.86
Journal of Writing Research	18	0.82
Assessing Writing	16	0.73

Table 4

The Most Productive Country for Students' Writing Skills Research Articles

Country	Documents	Cited	Total Link Strength
United States	642	5.173	122
Indonesia	172	433	15
Malaysia	119	477	34
Turkey	111	506	8
United Kingdom	89	699	56
Iran	88	518	28
China	82	482	48
Australia	81	750	50
Spain	74	431	19
Saudi Arabia	72	179	36

Table 5
The Most Productive Organization for Students' Writing Skills Research Articles

Organization	Country	Documents	Cited	Total Link Strength
Arizona State University	United States of America	9	208	17
Georgia State University	United States of America	7	122	6
University of Delaware	United States	7	55	9
Utrecht University	Netherlands	6	175	6
University of Amsterdam	Netherlands	5	122	7
Prince Sattam Bin Abdulaziz University	Saudi Arabia	5	18	7
Universidad Técnica Particular de Loja	South America	5	9	0
University of Porto	Portugal	4	147	1
Educational Testing Service	United States	3	75	3
Texas A&M University	United States	2	44	7

The Level of Education with The Most Widely Used in Students' Writing Skills Research Articles

This study analyzed bibliographic database files to determine the level of education that becomes the context of studies on students' writing (elementary school, primary school, high school, and higher education). The fields used are titles and abstracts with the *whole counting* method. *The minimum number* of occurrences is set at five, and the maximum number of terms is set at 2807 terms. Based on this analysis, the results show that the level of education with the most widely used for research on students writing skills is *higher education (f= 121)*, followed by *secondary school (f= 71)*, *elementary school (f= 43)*, and *high school (f= 39)*. The complete data is presented in Figure 8.

Opportunities and Novelty for Future Writing Skills Research

Opportunities for further research and novelties identified in this study based on the bibliometric analysis of students' writing skills research from the Scopus database in 2014-2023 are presented in Figure 9.

Based on Figures 9 A, B, C, and D, several *novelty* studies on students writing skills were found, such as research on *argumentative writing*, *collaborative writing*, *EFL learners*, and *writing attitude* as a topic that still has a chance and has not been widely carried out during the last ten years according to the *Scopus database bibliometric* analysis.

DISCUSSION

In today's educational environment, students' writing skills are undeniably important. Mastery in writing is crucial for achieving academic excellence and effective communication in both professional and personal settings. Studies show that solid writing skills are closely associated with better academic results and enhanced cognitive skills (Graham & Perin, 2007). Proficient writing supports critical thinking and problem-solving skills, essential for handling complex issues (Graham, 2006). Additionally, in the current digital age, the skill to express ideas clearly and convincingly is increasingly necessary for success across various domains (National Commission on Writing, 2004). Therefore, prioritizing writing instruction is fundamental to equipping students for future challenges.

Some important findings obtained based on bibliometric analysis in the study of students' writing skills are the most widely used keywords in students' writing skills, the most used word in the abstract on students' writing research, the total number of authors of research articles on students writing skills, the most popular publisher, the most productive countries and organization, the level education, and the opportunities.

The Frequency of Research on Students' Writing Skills in 2014-2023 (RQ1)

Based on the findings of a bibliometric analysis of 2214 research articles in the last ten years on the Scopus database, this study obtained data related to the keywords that

Figure 8

The Level of Education with the Most Widely Used in Students' Writing Skills Research Articles

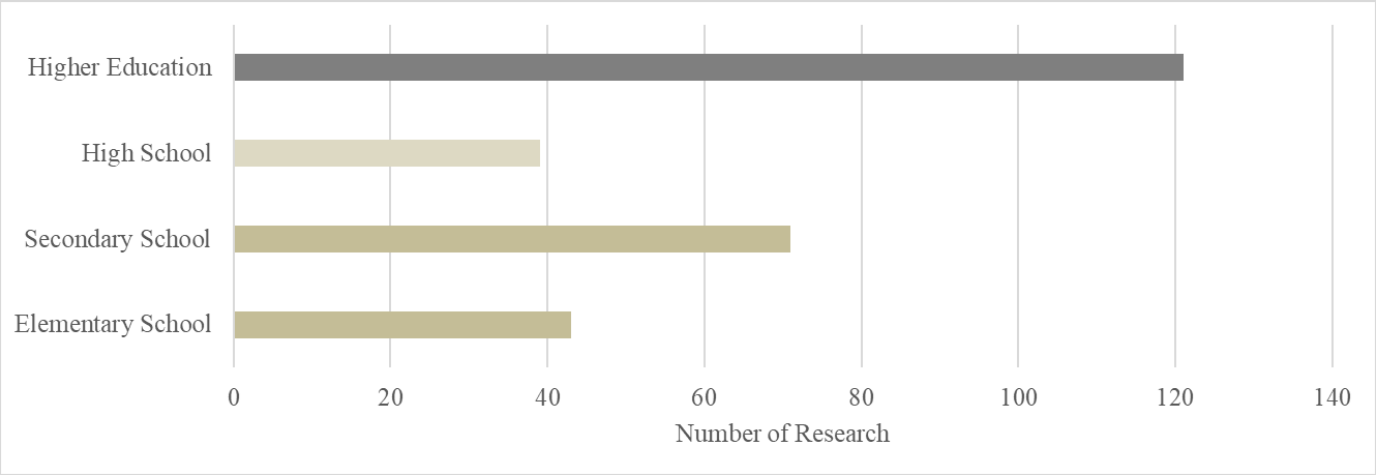
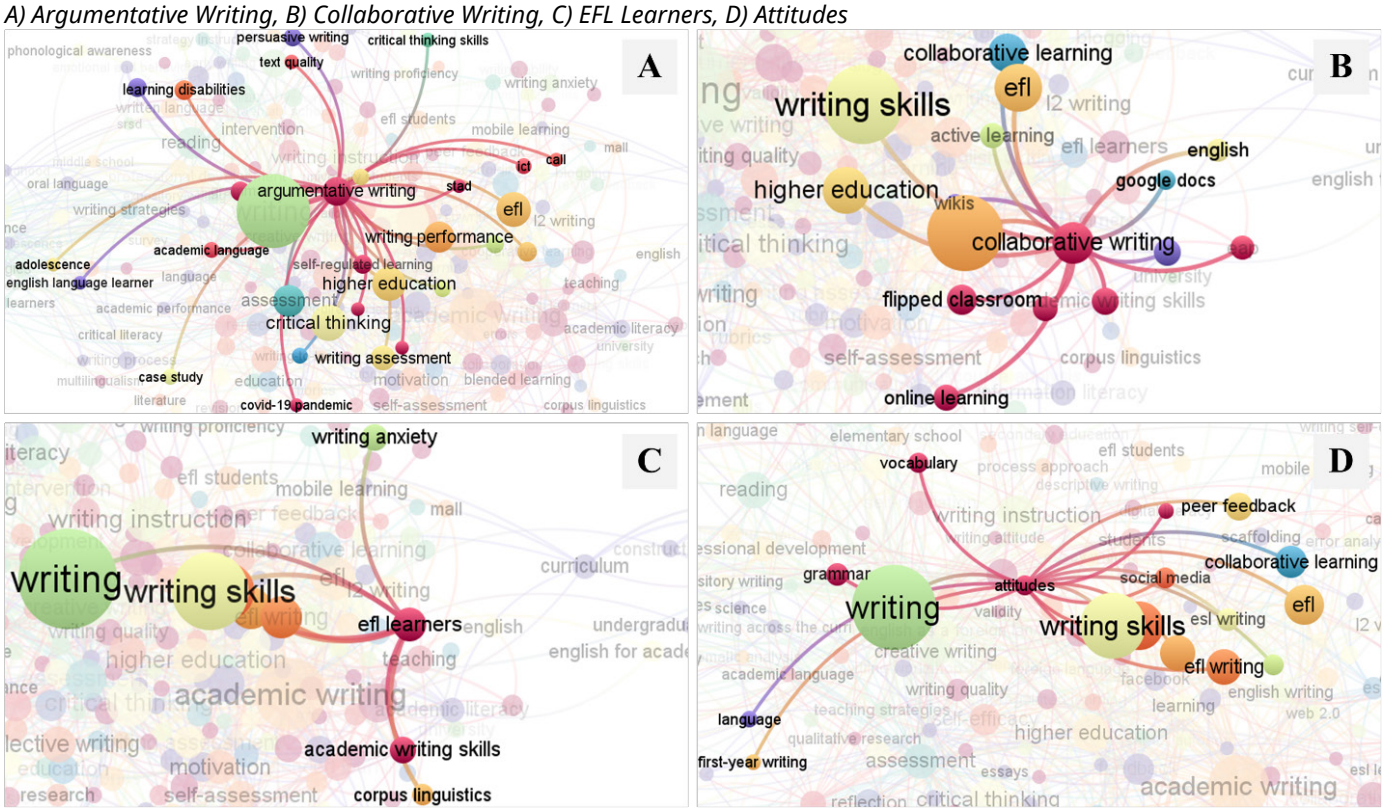


Figure 9

A) Argumentative Writing, B) Collaborative Writing, C) EFL Learners, D) Attitudes



were most widely used in titles and abstracts, authors with the highest number of citations, publishers with the highest number of publications, the most productive countries, and organizations in research. The study also collected data about the writing skills of the students under study, their level of education, and opportunities or novelties for future research. From the findings of a bibliometric analysis on the *Scopus database* for 2014-2023, research on students' writing skills has increased from 2014 to 2020 and fluctuated from 2021 to 2023. During this period, the highest number of publications occurred in 2022.

The Most Widely Used Keywords in Students' Writing Skills Research Articles (RQ2)

The most widely used keywords in students writing skills research articles were obtained in 10 clusters with the most keywords: *"writing."* These results are relevant to other key-

words, such as *"writing skills"* and *"academic writing,"* the most widely used in writing skills research (see Figure 6). Writing skills were the search terms that were most frequently utilized in this survey. This finding indicates that most of the research focuses on writing skills in the academic field. For instance, Wale and Bogale (2021) stated that students who participated in inquiry-based writing instruction improved their academic writing skills. Meanwhile, Shulgina et al. (2024) state that obtaining additional comments during online peer editing is an effective way to enhance students' writing performance further because peer editing is crucial in developing students' academic writing.

The Most Commonly Used Words in the Abstracts of Research Articles on Students' Writing Skills (RQ3)

Analyzing the most widely used words in the abstract gets as much data as possible in 7 clusters with the word *"effect,"* the word most used in the abstract. Other words that are widely used are *"test," "control group," "experimental group,"* and *"grade"* (see Figure 7). These results indicate that most research articles on students' writing skills focus on the influence or effect caused by the research group *(experimental and control)* on learning outcomes *(grade)* after a series of learning activities and tests. For instance, Arici and Kaldirim (2015) observed that writing instruction courses, which involved Turkish pre-service teachers using activities based on a process-oriented writing approach, significantly improved the written expression skills of the experimental group compared to the control group. Ali and Ulker (2020) mentioned that the pre-test and post-test results indicated a significant improvement in students' writing skills through the application of Inquiry-based Learning.

The Author with the Highest Number of Citations in the Student' Writing Skills Articles (RQ4)

The total number of authors of research articles on students' writing skills based on the Scopus database is 2158 authors from 2214 publications. The three authors in this field of study with the most citations are Graham, Hwang, and Kent. Graham (2019) stated that if students want to be successful at school, at work, and in everyday life, students must learn to write. Writing cannot happen naturally but requires practice and proper instruction, especially in formal education. To practice writing skills, Hwang et al. (2014) train students with the help of mobile devices to be more flexible and develop their writing skills anywhere and anytime. As a result, students are not only able to write more sentences, but they are also able to describe the objects they write about clearly and thoroughly. Kent et al. (2014) carried out a study on a framework for the progression of students writing, encompassing a collaborative focus on reading abilities, spelling, fluency in handwriting, and oral language elements. The findings indicated that the mindfulness approach demonstrated a more suitable fit than the one solely based on language and literacy elements. Writing skills are an essential aspect that can be applied at various levels of education, individually and collaboratively.

In addition, learning writing skills can also be integrated using media such as *game-based learning* and *augmented reality*. Wang (2014) states that language learning can be done collaboratively using social networking platforms. This is one of the new approaches to teaching that has emerged due to the influence of technological advances. Wang (2017) also uses technology as a tool in students writing learning with Augmented Reality media. Wang stated that AR is a reference that teachers can use to encourage students to have good writing skills.

The other frequently cited topics include research by Charon et al. (2016), which explored the significance of creative writing for students; Koster et al. (2015), which examined effective writing interventions for elementary students; Lin et al. (2018), which demonstrated how flipped contextual game-based instruction can improve students' English writing performance; Naber & Wyatt (2014), who investigated the impact of reflective writing interventions on critical thinking skills and student attitudes; and Graham et al. (2018), which assessed the effectiveness of literacy programs that combine reading and writing instruction. Thus, based on the ten articles with the highest citation counts, research on students' writing skills from 2014 to 2023 covers aspects of the writing process, strategies, and techniques that can facilitate students to improve their writing skills.

Authors with many citations often play a crucial role in research. Their publications can lay the groundwork for further studies, advance scientific disciplines, and offer valuable references for other researchers. However, the impact of a highly cited author is determined by the number of citations and the quality and context of those citations. Research shows that not all citations are necessarily positive, as some may critique or challenge the original work (Huang et al., 2022; Xu et al., 2022).

Articles with many citations can shape other researchers' perceptions of topics or fields. A study on the role of highly cited articles indicates that such papers can be markers of scientific productivity and intellectual impact (Plomp, 1990). Nonetheless, it is important to recognize that citations might sometimes reflect the popularity of a topic rather than the inherent quality or positive influence of the research. Thus, while high citation counts are generally seen as a good sign, the actual impact of a highly cited author is determined by the contribution their work makes to scientific advancement and discussion.

The Publisher Publishes the Most Articles on Students' Writing Skills (RQ5)

Regarding the publishers that publish the most articles on students' writing skills, *Reading and Writing, Asian EFL Journal, and International Journal of Instruction* are the top three publishers with the most publications based on the number of documents identified. These three journals can be a reference for teachers to find out and apply the latest research results to improve students' writing skills in this area. Also, for researchers, these journal references can make it easier to find gaps or novelties that have been carried out by other researchers worldwide and as a reference for publishing their research articles. However, it should be underlined that the level of popularity of these research findings is based on the quantity of publications. Meanwhile, it does not mean that other journals that do not have a large number of publications are not popular or of good quality. Just because a journal such as "The Asian EFL Journal" publishes monthly does not mean it is better than other journals that publish less often.

Beyond the top three journals listed in Table 3, which rank as the most popular journals for publications on students' writing skills, several other journals between ranks 11 and 20 can also be considered for publishing articles in this field. These journals include *Frontiers in Education, Indonesian Journal of Applied Linguistics, Asian ESP Journal, International Journal of Language Education, Journal of Language and Linguistic Studies, TESOL International Journal, International Journal of Applied Linguistics and English Literature, Journal of Second Language Writing, Journal of Language and Education,* and *International Journal of English Language and Literature Studies*.

Some popular journals may be influenced by factors such as publication frequency, acceptance rate, speed of publication, special issues, publication fee, and the journal's indexing rating. For authors, before publishing an article in a journal, it is recommended to check all the details about the journal, especially the indexing status of those journals through platforms like SCImago Journal & Country Rank (SJR) or by checking the Scopus source page. For instance, although the International Journal of Instruction is one of the most popular journals that publishes a lot of content about students' writing skills, it has recently been discontinued from Scopus indexing.

Countries and Organizations Most Productive in Research on Students' Writing Skills (RQ6)

The Most Productive Countries in Research on Students' Writing Skills

Over the past decade, the most productive countries in research on students' writing skills, based on the number of identified documents, are the United States in first place, Indonesia in second place, and Malaysia in third place (see Table 4). At the same time, the countries ranked from fourth to tenth in research on students' writing skills are Turkey, the United Kingdom, Iran, China, Australia, Spain, and Saudi Arabia. They are nations from diverse continents, including the Americas, Asia, and Europe. Regarding article citations, the United States leads, with Australia in second place and the United Kingdom in third. Additionally, regarding the overall strength of article links, the United States ranks first, followed by the United Kingdom in second and Australia in third. This indicates that the United States excels not only in the volume of research article publications but also in the number of citations and the strength of interconnected article links. Conversely, although Indonesia and Malaysia have published more articles than Australia and the United Kingdom, their citation counts and article link strengths do not surpass those of Australia and the United Kingdom.

The Most Productive Organization in Research on Students' Writing Skills

Arizona State University is the leading organization in the United States for publishing research on students' writing skills, as evidenced by the number of documents published (see Table 5). Other prominent organizations based on document counts include Georgia State University, the University of Delaware, Educational Testing Service, and Texas A&M University in the U.S.; Utrecht University and the University of Amsterdam in the Netherlands; Prince Sattam Bin Abdulaziz University in Saudi Arabia; Universidad Técnica Particular de Loja in South America; and the University of Porto in Portugal. In terms of citations, Arizona State University leads with 208 citations, followed by Utrecht University with 175 citations in second place, and the University of Porto with 147 citations in third place. This indicates that Arizona State University is not only the most prolific in publishing research on students' writing skills but also significantly influential, as reflected in the high number of citations, demonstrating the impact and relevance of its publications.

Education Level is the Most Commonly Studied Factor in Research on Students' Writing Skills (RQ7)

Most studies about students' writing took place in higher education contexts, followed by secondary school, elementary school, and high school (see Figure 8). In this bibliometric analysis visualization, one of the keywords identified as the most widely used in research on students' writing skills is "Academic Writing" (see Figure 7). These visualization findings support the findings in Figure 8. Swales and Feak (2012) describe academic writing as a unique genre marked by its formal tone, organized structure, and adherence to specific disciplinary norms. It is a key method for communicating and evaluating scholarly ideas in higher education.

Mastering academic writing requires more than just understanding language mechanics; it demands a deep grasp of the disciplinary context and the ability to meet complex academic expectations. This complexity highlights the need for specialized instruction and support to help students develop strong writing skills (Gillett et al., 2009). Historically, the focus on enhancing writing skills in higher education stems from the belief that these skills are essential for academic and professional success, which has shaped the direction of research (Hyland, 2003).

The Future Writing Skills Research Opportunities (RQ8)

The areas that need further research and exploration are *argumentative writing, collaborative writing, EFL learners,* and *attitude writing* (see Figure 9 A, B, C, D). An explanation of these aspects is presented in the following paragraph.

Argumentative Writing

Argumentative rhetoric aims to persuade others to adopt the writer's or reader's desired beliefs and behaviors by manipulating their attitudes and opinions (Ka-kan-dee & Kaur, 2015; Keraf, 2010). In other words, argumentative writing is when the writer presents reasons and evidence to support their argument and convince the reader of the position's validity (Harmer, 2004). An effective scientific argumentative essay should adhere to specific standards. This requires ensuring that the statement is supported by solid reasoning, evidence, or data and, if applicable, incorporating counterarguments to strengthen the author's position (Mulyati & Hadianto, 2023).

Studies in the cognitive tradition, like those by Graham et al. (2018), show that argumentative writing development in students' is influenced by factors like writing goals and knowledge of persuasion genres. These studies often employ interventions to improve students' argumentative writing skills. While there is existing research, the ongoing exploration into the factors that influence argumentative writing suggests that this area still requires more nuanced investigation, mainly from various educational and cultural perspectives.

Collaborative Writing

Collaborative writing, on the other hand, is a potent technique that promotes critical thinking, peer learning, and active engagement in the final output. The requirement to deal with structuring, negotiating, and merging ideas in written form led to introducing more conventional approaches into the classroom, making collaborative writing a relatively young discipline compared to other practices (Dobao, 2012). In collaborative learning, students should be organized into study clusters, facilitating collective effort, the opportunity to engage in conversations with peers, a strong inclination to educate group members, and utilizing students' interaction within these collaborative groups for their advantage (Ghufron et al., 2023).

Research on collaborative writing, especially in educational settings, indicates that it serves as an effective bridge for enhancing argumentative writing skills. Studies have explored its role in fostering peer discourse, improving writing synthesis, and integrating conflicting information from multiple sources. While collaborative writing has been studied, understanding its full potential and diverse applications in different learning contexts (such as EFL environments) remains an open research area (Granado et al., 2019; Matos, 2021).

EFL Learners

In the context of EFL, further study is required. EFL refers to individuals who study English in nations that do not speak the language. Japanese citizens, for instance, who study English, are EFL learners (Iwai, 2011). Harmer (2007) has a similar definition, limiting EFL to English teaching in which students take short courses in English or learn English in their home country. Proficiency in writing is crucial for English as a Foreign Language (EFL) students, particularly in the contemporary era where effective written communication is gaining heightened significance (Kusumaningrum et al., 2019).

There is increasing interest in how technologies like chatbots assist EFL students in argumentative writing. However, since these technological interventions are relatively new, there are gaps in understanding how these tools affect students' writing skills, attitudes, and learning processes. Further studies could explore long-term effects and the best practices for incorporating such technologies into the EFL classroom (Guo et al., 2021).

Writing Attitudes

Regarding students' writing attitudes, Ni'mah et al. (2017) state that it considerably impacts students' writing skills. Students' writing skills are significantly affected by writing attitudes (Graham & Perin, 2007). Here, kids with a positive mindset demonstrate superior writing abilities to those with a bad attitude. Negative attitudes can hamper an effective writing process. This mentality can be detrimental to writing because it is a complex endeavor requiring much work. Attitudes, such as cognitive involvement, influence writing skills. For instance, students who write more diligently and with more effort than students with fewer positive views do so (McKenna et al., 1995). Writing attitudes have been a topic of interest, but there is an ongoing need for more comprehensive studies to understand its impact on writing proficiency and its interaction with other variables (Graham & Perin, 2007).

Based on previous research, this bibliometric analysis reveals that argumentative writing, collaborative writing, EFL learners, and writing attitudes influence students' writing skills. Murtadho (2021) underscores the significance of argumentative writing as a critical area for researchers seeking to enhance college-level writing. Additionally, Pham (2021) highlights the substantial improvement in writing fluency observed in students who engage in collaborative writing in group and individual assignments. Furthermore, Chen (2022) emphasizes that EFL learners generally perceive mastering effective writing as one of the most challenging and frustrating aspects of language learning. Similarly, Özen and Duran (2021) discuss the importance of writing attitude, emphasizing its close relationship to the practical implementation of skills in language learning programs. These skills include writing short texts and composing stories. Teachers can support and guide students to cultivate a positive attitude towards writing at every stage of writing development, serving as mentors and participants.

Students' writing proficiency is significantly influenced by their exposure to different writing theories and genres. Effective writing instruction incorporates theoretical frameworks that address both the cognitive aspects of writing and the structural elements of various genres, leading to improved writing outcomes (Bereiter & Scardamalia, 1987). The quality of writing is significantly affected by the processes and techniques writers use and their motivation levels. Effective writing instruction should address these elements to enhance students' writing skills and outcomes (Applebee & Langer, 2011). Furthermore, successful writing comes from various strategies and techniques supporting the writing process, such as organization, coherence, and language use. Motivation also plays a crucial role, affecting writers' persistence and commitment to refining their work (Graham & Perin, 2007).

Historically, several researchers have studied the topic of students' writing skills using bibliometric analysis, including Crosthwaite et al. (2022), Hyland and Jiang (2022), and Sun and Lan (2023). However, these studies only focus on aspects such as L2 written corrective feedback or the interaction aspect in written texts. In contrast, our bibliometric analysis takes a broader approach, considering the general context of students' writing skills. Our goal is to identify trends and opportunities related to the writing skills that students should possess. Furthermore, we include data from all countries, regardless of whether English is their primary language. Therefore, it is crucial to investigate further the essential aspects of students' writing skills across various countries and languages. By doing so, we can improve students' writing skills at different educational levels.

The expected outcome of this analysis is to identify future trends and opportunities for students' writing skills. However, the findings of this bibliometric analysis cover essential aspects such as argumentative writing, collaborative writing, EFL learners, and writing attitude. Research on students' writing skills has decreased in intensity from 2021 to 2023. In today's technological era, writing skills are crucial for students, especially in technology and digital literacy. The advancement of science and technology should encourage teachers, researchers, and students to improve their writing skills actively. Nowadays, students are expected to be critical in various areas, not just science.

Additionally, students need to be able to collaborate on various tasks, including writing. Moreover, students from countries where English is not the primary language must learn foreign languages to support their knowledge. Furthermore, since writing is an effort to convey information, students' must also use proper ethics in writing so that the meaning of their text aligns with the purpose of the writing. Research on students' writing skills, including argumentative writing, collaborative writing, EFL learners, and writing attitude, is precious for improving students' writing skills. Not only in the academic realm but also in students' writing skills will help them achieve tremendous success in their professional careers.

Limitations

The limitation of this research is that it only uses Scopus database. Additionally, the research parameters are limited to analyzing research frequency, commonly used words in articles and abstracts, authors with the highest citation counts, prominent publishers, the productivity of countries and organizations, and educational levels. In the future, conducting bibliometric analysis using multiple data sources such as Web of Science (WOS), EBSCO, Eric, Microsoft Academic, Google Scholar, and various other databases will produce better results. Furthermore, other indicators, such as content analysis, research methods, and data analysis, can also be used to enrich the data analysis results. Moreover, this type of bibliometric analysis will provide more valuable information for readers, researchers, authors, teachers, and students.

CONCLUSION

This study aims to find trends and novelties in students' writing skills research in the last ten years (2014-2023) using bibliometric analysis from the Scopus database. The results of this bibliometric analysis demonstrate that since 2014, there has been a rise in research on students' writing abilities. Several international journals that can be used as references in this research are *Reading and Writing, Asian EFL Journal*, and *International Journal of Instruction*. From this bibliometric analysis, *argumentative writing*, *collaborative writing*, *EFL learners*, and *writing attitudes* have not been widely practiced in the last ten years. These four topics can be used as one of the research themes for students writing skills at various levels of education for future researchers.

Several innovations in this field have been applied in various research over the last ten years, such as structured training and integrating technology such as mobile devices, games, augmented reality, social networks, methods, and learning models at various levels of education. Findings from this analysis, such as *argumentative writing skills*, *collaborative writing*, *EFL learners*, and *writing attitudes,* are interesting topics to research and develop in the future. Teachers or researchers can analyze, research, or integrate these rare topics to improve students' writing skills at various levels of education. These findings hold significance within writing skills, serving as valuable points of reference and suggestions for forthcoming investigations and educational endeavors. By incorporating these revelations into their work, educators and scholars can leverage these insights to uncover diverse crucial facets of writing skills that may have eluded prior research efforts.

The limitation of this research is that it only uses Scopus database. Furthermore, the research parameters are confined to analyzing research frequency, commonly utilized words in articles and abstracts, authors with the highest citation counts, prominent publishers, the productivity of countries and organizations, and educational levels. Although the analysis of this topic is limited and only obtained from the Scopus database, we hope these findings can become valuable recommendations for teachers, researchers, and students to enrich their knowledge in the field of language, especially in terms of writing skills.

Recommendations for further research are that researchers can use various databases such as Scopus, Web of Science, Ebsco, Eric, and other databases to obtain more complex data. Besides that, broader parameters can be used in research, such as type of research, methods, data analysis, impact, and other variables, to deepen the scope of study in this field and provide valuable information for readers.

DECLARATION OF COMPETITING INTEREST

None declared.

REFERENCES

Abd Karim, R., & Mustapha, R. (2020). Students' perception on the use of digital mind map to stimulate creativity and critical thinking in ESL writing course. *Universal Journal of Educational Research, 8(12A)*, 7596-7606. https://doi.org/10.13189/UJER.2020.082545

Ali, H. F., & Ulker, V. (2020). *The effect of inquiry-based approach on development of reading and writing skills of university EFL students* [Preprint]. SSRN. https://doi.org/10.2139/ssrn.3621259

Anson, C. M., & Schwegler, R. (2012). *The longman handbook for writers and readers* (6th ed.). Pearson.

Applebee, A. N., & Langer, J. A. (2011). *Writing instruction that works: Proven methods for middle and high school classrooms*. Teachers College Press.

Arici, A. F., & Kaldirim, A. (2015). The effect of the process-based writing approach on writing success and anxiety of pre-service teachers. *The Anthropologist*, *22*(2), 318-327. https://doi.org/10.1080/09720073.2015.11891883

Baker, K. M. (2016). Peer review as a strategy for improving students' writing process. *Active Learning in Higher Education, 17*(3), 179–192. https://doi.org/10.1177/1469787416654794

Bangert-Drowns, R. L., Hurley, M. M., & Wilkinson, B. (2004). The effects of school-based writing-to-learn interventions on academic achievement: A meta-analysis. *Review of Educational Research*, *74*(1), 29–58. https://doi.org/10.3102/00346543074001029

Barone, D. M. (2010). Engaging young ELLs with reading and writing. In G. Li & P. Edwards, (Eds.), *Best practices in ELL instruction* (pp. 84-102). The Guildford Press.

Bazerman, C. (1988). *Shaping written knowledge: The genre and activity of the experimental article in science*. University of Wisconsin Press.

Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition.* Lawrence Erlbaum Associates.

Bugis, Y. (2018). *Creating digital stories with Saudi Arabian pre-service teachers: using the analysis, design, development, implementation, and evaluation model to promote lesson plan development* [Doctoral Dissertation]. University of Northern Colorado. https://www.proquest.com/docview/2103953544

Burroway, J. (2019). *Writing fiction: A guide to narrative craft* (11th ed.). University of Chicago Press.

Brown, H. D. (2014). *Principles of language learning and teaching.* Pearson Education.

Ceylan, N. O. (2019). Student perceptions of difficulties in second language writing. *Journal of Language and Linguistic Studies*, *15*(1), 151-157. https://doi.org/10.17263/jlls.547683

Chaffee, J. (2015). *Thinking critically.* Cengage Learning.

Charon, R., Hermann, N., & Devlin, M. J. (2016). Close reading and creative writing in clinical education: Teaching attention, representation, and affiliation. *Academic Medicine: Journal of the Association of American Medical Colleges*, *91*(3), 345–350. https://doi.org/10.1097/ACM.0000000000000827

Chen, A. H. (2022). The effects of writing strategy instruction on efl learners' writing development. *English Language Teaching, 15*(3), 29-37. https://doi.org/10.5539/elt.v15n3p29

Chevalier, R. D. (2011). When did ADDIE become addie? *Performance Improvement*, *50*(6), 10–14. https://doi.org/10.1002/PFI.20221

Coleman, D., & Willis, D. S. (2015). Reflective writing: The student nurse's perspective on reflective writing and poetry writing. *Nurse Education Today*, *35*(7), 906–911. https://doi.org/10.1016/J.NEDT.2015.02.018

Crosthwaite, P., Ningrum, S., & Lee, I. (2022). Research trends in L2 written corrective feedback: A bibliometric analysis of three decades of Scopus-indexed research on L2 WCF. *Journal of Second Language Writing*, *58*, 100934. https://doi.org/10.1016/J.JSLW.2022.100934

DaCrema, J. J., & Stout, D. (2012). *Ten tips for improving business writing. White paper prepared for ACC6902- Management Accounting Systems*. Youngstown State University.

Davis, A. L., & Davis, A. L. (2013). Using instructional design principles to develop effective information literacy instruction: The ADDIE model. *College & Research Libraries News*, *74*(4), 205–207. https://doi.org/10.5860/crln.74.4.8934

Devitt, A. J. (2004). *Writing genres*. Southern Illinois University Press.

Elbow, P. (1973). *Writing without teachers*. Oxford University Press.

Ferretti, R. P., & Graham, S. (2019). Argumentative writing: Theory, assessment, and instruction. *Reading and Writing*, *32*, 1345-1357. https://doi.org/10.1007/s11145-019-09950-x

Fernández Dobao, A. (2012). Collaborative writing tasks in the L2 classroom: Comparing group, pair, and individual work. *Journal of Second Language Writing*, *21*(1), 40–58. https://doi.org/10.1016/J.JSLW.2011.12.002

Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication*, *32*(4), 365-387. https://doi.org/10.2307/356600

Flowerdew, J., & Peacock, M. (2001). Research perspectives on English for academic purposes. Cambridge University Press.

Gerson, S. J., & Gerson, S. M. (2018). *Technical communication: Process and product* (10th ed.). Pearson.

Ghufron, S., Nafiah, Syahruddin, Kaswadi, & Mustofa. (2023). The effect of STAD-type cooperative learning based on a learning tool on critical thinking ability in writing materials. *International Journal of Instruction*, *16*(1), 61–84. https://doi.org/10.29333/iji.2023.1614a

Gillett, A., Hammond, A., & Martala, J. (2009). *Inside track: Writing successfully in science.* Open University Press.

Graham, S. (2006). Writing as a goal of instruction. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 187-207). Guilford Press.

Graham, S. (2019). Changing how writing is taught. *Review of Research in Education*, *43*(1), 277–303. https://doi.org/10.3102/0091732X18821125

Graham, S., & Hebert, M. (2011). Writing to read: A meta-analysis of the impact of writing and writing instruction on reading. *Harvard Educational Review*, *81*(4), 710–744. https://doi.org/10.17763/haer.81.4.t2k0m13756113566

Graham, S., Liu, X., Aitken, A., Ng, C., Bartlett, B., Harris, K. R., & Holzapfel, J. (2018). Effectiveness of literacy programs balancing reading and writing instruction: A meta-analysis. *Reading Research Quarterly*, *53*(3), 279–304. https://doi.org/10.1002/RRQ.194

Graham, S., & Perin, D. (2007). *Writing next: Effective strategies to improve writing of adolescents in middle and high schools. A Report to Carnegie Corporation of New York*. Alliance for Excellent Education.

Granado-Peinado, M., Mateos, M., Martín, E., & Cuevas, I. (2019). Teaching to write collaborative argumentative syntheses in higher education. *Reading and Writing*, *32*(8), 2037-2058. https://doi.org/10.1007/s11145-019-09939-6

Guo, K., Li, Y., Li, Y., & Chu, S. K. W. (2024). Understanding EFL students' chatbot-assisted argumentative writing: An activity theory perspective. *Education and Information Technologies*, *29*(1), 1-20. https://doi.org/10.1007/s10639-023-12230-5

Harmer, J. (2004). *How to teach writing.* Pearson Longman.

Harmer, J. (2007). *The practice of English language teaching*. Longman.

Huang, H., Zhu, D. & Wang, X. (2022). Evaluating scientific impact of publications: combining citation polarity and purpose. *Scientometrics, 127*, 5257–5281. https://doi.org/10.1007/s11192-021-04183-8

Hung, C.-M., Hwang, G.-J., & Huang, I. (2012). A project-based digital storytelling approach for improving students' learning motivation, problem-solving competence and learning achievement. *Educational Technology & Society*, *15*(4), 368–379.

Hwang, W. Y., Chen, H. S. L., Shadiev, R., Huang, R. Y. M., & Chen, C. Y. (2014). Improving English as a foreign language writing in elementary schools using mobile devices in familiar situational contexts. *Computer Assisted Language Learning*, *27*(5), 359–378. https://doi.org/10.1080/09588221.2012.733711

Hyland, K. (2002). *Teaching and researching writing*. Longman.

Hyland, K., & Jiang, F. K. (2022). Interaction in written texts: A bibliometric study of published research. *Studies in Second Language Learning and Teaching*, *13*(4), 903-924. https://doi.org/10.14746/ssllt.40220

Hyland, K. (2003). Genre-based pedagogies: A social response to process. *Journal of Second Language Writing*, *12*(1), 17-29. https://doi.org/10.1016/S1060-3743(02)00124-8

Hyland, K. (2005). *Metadiscourse: Exploring interaction in writing*. Continuum.

Iftanti, E. (2016). Improving students' writing skills through writing journal articles. *Jurnal Bahasa Lingua Scientia*, *8*(1), 1–14. https://doi.org/10.21274/LS.2016.8.1.1-14

Iwai, Y. (2011). The effects of metacognitive reading strategies: Pedagogical implications for EFL/ESL teachers. *The Reading Matrix: An International Online Journal, 11*(2), 150-159.

Jaja, J., Mudopar, M., Kurnia, M. D., & Muliawati, H. (2019). Representation of linguistic aspects in the genre of text in junior high school's bahasa Indonesia textbooks in the 2013 curriculum. In *International Symposium on Social Sciences, Education, and Humanities (ISSEH 2018)* (pp. 75-79). Atlantis Press. https://doi.org/10.2991/ISSEH-18.2019.18

Jaja, & Rahayu, I. (2021). *Apresiasi dan kajian prosa fiksi, menuju proses kreatif menulis* [Appreciation and analysis of fictional prose, towards the creative writing process]. Media Edukasi Indonesia.

James, N. (2007). *Writing at work: How to write clearly, effectively, and professionally*. Allen & Unwin.

Ka-kan-dee, M., & Kaur, S. (2015). Teaching strategies used by Thai EFL lecturers to teach argumentative writing. *Procedia - Social and Behavioral Sciences*, *208*, 143–156. https://doi.org/10.1016/J.SBSPRO.2015.11.191

Kent, S., Wanzek, J., Petscher, Y., Al Otaiba, S., & Kim, Y. S. (2014). Writing fluency and quality in kindergarten and first grade: The role of attention, reading, transcription, and oral language. *Reading and Writing*, *27*(7), 1163–1188. https://doi.org/10.1007/S11145-013-9480-1

Keraf, G. (2010). *Argumentasi dan narasi* [Argumentation and narrative]. Gramedia.

King, S. (2000). *On writing: A memoir of the craft*. Scribner.

Koster, M., Tribushinina, E., de Jong, P. F., & van den Bergh, H. (2015). Teaching children to write: A meta-analysis of writing intervention research. *Journal of Writing Research*, *7*(2), 249–274. https://doi.org/10.17239/JOWR-2015.07.02.2

Kottacheruvu, N. (2023). Developing writing skills through english short stories: A case study in the classroom. *International Journal of Language and Literary Studies*, *5*(1), 287–299. https://doi.org/10.36892/IJLLS.V5I1.1243

Kusumaningrum, S. R., Cahyono, B. Y., & Prayogo, J. A. (2019). The effect of different types of peer feedback provision on EFL students' writing performance. *International Journal of Instruction*, *12*(1), 213–224. https://doi.org/10.29333/iji.2019.12114a

Langan, J. (2001). *College writing skills with readings* (5th ed.). McGraw-Hill.

Lee, J. J., & Deakin, L. (2016). Interactions in L1 and L2 undergraduate student writing: Interactional metadiscourse in successful and less-successful argumentative essays. *Journal of Second Language Writing*, *33*, 21–34. https://doi.org/10.1016/J.JSLW.2016.06.004

Leki, I. (1998). *Academic writing: Exploring processes and strategies.* Cambridge University Press.

Li, M., Chen, Y., Zhang, L., Wu, X., & Huang, C. (2022). Investigating learners' engagement and chinese writing learning outcomes with different designs of SVVR-based activities. *Sustainability, 14*(8), 4767. https://doi.org/10.3390/SU14084767

Lin, C.-J., Hwang, G.-J., Fu, Q.-K., & Chen, J.-F. (2018). A flipped contextual game-based learning approach to enhancing efl students' english business writing performance and reflective behaviors. *Educational Technology & Society*, *21*(3), 117–131.

Lipschutz, B. D. (2010). *The use of digital storytelling to improve the effectiveness of social and conflict resolution skills training for elementary students* [Doctoral Dissertation]. Temple University. http://dx.doi.org/10.34944/dspace/1728

Markel, M. (2018). *Technical communication* (12th ed.). Bedford/St. Martin's.

Matos, F. (2021). Collaborative writing as a bridge from peer discourse to individual argumentative writing. *Reading and Writing, 34(5)*, 1321-1342. https://doi.org/10.1007/s11145-020-10117-2

McKenna, M. C., Kear, D. J., & Ellsworth, R. A. (1995). Children's attitudes toward reading: A national survey. *Reading Research Quarterly*, *30*(4), 934. https://doi.org/10.2307/748205

Miller, C. R. (1984). Genre as social action. *Quarterly Journal of Speech, 70(2)*, 151-167. https://doi.org/10.1080/00335638409383686

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *BMJ*, *339*(7716), 332–336. https://doi.org/10.1136/BMJ.B2535

Muhanif, M., Suhartono, S., & Juhana, J. (2021). Pengaruh kedisiplinan dan kreativitas terhadap keterampilan menulis siswa sekolah dasar [The influence of discipline and creativity on the writing skills of elementary school students]. *Edukatif: Jurnal Ilmu Pendidikan*, *3*(4), 1962–1973. https://doi.org/10.31004/EDUKATIF.V3I4.1046

Mulyati, Y., & Hadianto, D. (2023). Enhancing argumentative writing via online peer feedback-based essay: A quasi-experiment study. *International Journal of Instruction*, *16*(2), 195–212. https://doi.org/10.29333/iji.2023.16212a

Murray, D. M. (1985). *A writer teaches writing*. Houghton Mifflin.

Murtadho, F. (2021). Metacognitive and critical thinking practices in developing EFL students' argumentative writing skills. *Indonesian Journal of Applied Linguistics*, *10*(3), 656-666. https://doi.org/10.17509/ijal.v10i3.31752

Myhill, D. (2018). Grammar as a meaning-making resource for language development. *L1-Educational Studies in Language and Literature*, *18, Running Issue* (Running Issue), 1–21. https://doi.org/10.17239/L1ESLL-2018.18.04.04

Naber, J., & Wyatt, T. H. (2014). The effect of reflective writing interventions on the critical thinking skills and dispositions of baccalaureate nursing students. *Nurse Education Today*, *34*(1), 67–72. https://doi.org/10.1016/J.NEDT.2013.04.002

National Commission on Writing. (2004). *Writing: A ticket to work... or a ticket out: A survey of business leaders.* College Entrance Examination Board.

Ni'mah, U., Kadarisman, A. E., & Suryati, N. (2017). The roles of writing attitudes and writing apprehension in EFL learners' writing performance. *LET: Linguistics, Literature and English Teaching Journal*, *6*(2), 1-14. https://doi.org/10.18592/LET.V6I2.1459

Nugroho. (2014). *Menjadi penulis kreatif* [Becoming a creative writer]. Notebook.

Özen, N. E., & Duran, E. (2021). Developing a story writing attitude scale for secondary school students. *International Journal of Contemporary Educational Research, 8*(1), 84-93. https://doi.org/10.33200/ijcer.788115

Parmawati. (2013). *The efffectiveness of Think-Pair-Share (TPS) to teach writing viewed from students creativity* [Unpublished doctoral dissertation]. Universitas Negeri Surakarta. https://digilib.uns.ac.id/dokumen/detail/35709/The-efffectiveness-of-Think-Pair-Share-TPS-to-Teach-Writing-Viewed-From-Students-Creativity

Pham, V. P. H. (2021). The effects of collaborative writing on students' writing fluency: An efficient framework for collaborative writing. *SAGE Open, 11*(1), 1-11. https://doi.org/10.1177/2158244021998363

Plomp, R. (1990). The significance of the number of highly cited papers as an indicator of scientific prolificacy. *Scientometrics 19*, 185–197. https://doi.org/10.1007/BF02095346

Putri, N., & Aminatun, D. (2021). Using facebook to practice writing skill: What do the students think? *Journal of English Language Teaching and Learning, 2*(1), 45-50. https://doi.org/10.33365/jeltl.v2i1.852

Sadiku, L. M. (2015). The importance of four skills reading, speaking, writing, listening in a lesson hour. *European Journal of Language and Literature*, *1*(1), 29–31. https://doi.org/10.26417/EJLS.V1I1.P29-31

Sakkir, G., & Dollah, S. (2019). Measuring students' writing skills using Facebook group application in EFL context. *International Journal of Humanities and Innovation*, *2*(3), 69–74. https://doi.org/10.33750/IJHI.V2I3.43

Shulgina, G., Costley, J., Shcheglova, I., Zhang, H., & Sedova, N. (2024). Online peer editing: the influence of comments, tracked changes and perception of participation on students' writing performance. *Smart Learning Environments*, *11*(1), 30. https://doi.org/10.1186/s40561-024-00315-8

Stout, D. E. (2014). A business communication module for an MBA managerial accounting course: A teaching note. *Accounting Education*, *23*(2), 155–173. https://doi.org/10.1080/09639284.2013.847330

Sun, Y., & Lan, G. (2023). A bibliometric analysis on L2 writing in the first 20 years of the 21st century: Research impacts and research trends. *Journal of Second Language Writing*, *59*, 100963. https://doi.org/10.1016/J.JSLW.2023.100963

Syahid, A., & Qodir, A. (2021). Journal of Language and Linguistic Studies: A fifteen-year bibliometric quest for a bigger impact. *Journal of Language and Linguistic Studies*, *17*(1), 290–314.

Swales, J. M., & Feak, C. B. (2012). *Academic writing for graduate students: Essential tasks and skills.* University of Michigan Press.

Tannen, D. (1998). *The argument culture: Moving from debate to dialogue.* Random House.

Tolchinsky, L. (2016). From text to language and back again: The emergence of written language. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (2nd ed., pp. 144–159). Guilford Press.

Wale, B. D., & Bogale, Y. N. (2021). Using inquiry-based writing instruction to develop students' academic writing skills. *Asian-Pacific Journal of Second and Foreign Language Education*, *6*, 1-16. https://doi.org/10.1186/s40862-020-00108-9

Wang, Y. C. (2014). Promoting collaborative writing through wikis: a new approach for advancing innovative and active learning in an ESP context. *Computer Assisted Language Learning, 28*(6), 499–512. https://doi.org/10.1080/09588221.2014.881386

Wang, Y. H. (2017). Exploring the effectiveness of integrating augmented reality-based materials to support writing activities. *Computers & Education*, *113*, 162–176. https://doi.org/10.1016/J.COMPEDU.2017.04.013

Whai, M. K. G., Wei, W., & Man, C. K. (2013). *Writing difficulties faced by politeknik kuching sarawak commerce diploma students in doing their assignments* [Writing difficulties faced by politeknik kuching sarawak commerce diploma students in doing their assignments]. *The Asian Journal of English Language & Pedagogy*, *1,* 90-101. https://ojs.upsi.edu.my/index.php/AJELP/article/view/783

Williams, J. (2006). *Style: Lessons in clarity and grace* (8th ed.). Pearson.

Xu, L., Ding, K. & Lin, Y. (2022). Do negative citations reduce the impact of cited papers? *Scientometrics, 127*, 1161–1186. https://doi.org/10.1007/s11192-021-04214-4

Yusuf, Q., Jusoh, Z., & Yusuf, Y. Q. (2019). Cooperative learning strategies to enhance writing skills among second language learners. *International Journal of Instruction, 12*(1), 1399-1412. https://doi.org/10.29333/iji.2019.12189a

https://doi.org/10.17323/jle.2024.22312

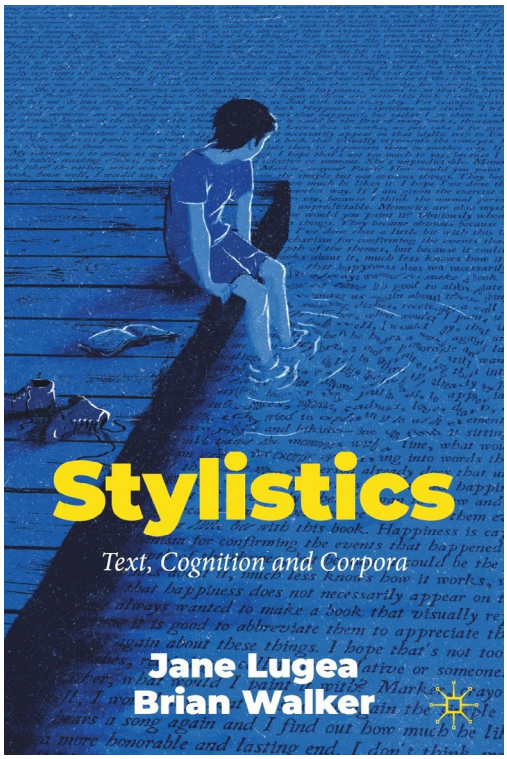
Stylistics: Text, Cognition and Corpora: A Book Review

Apri Wardana Ritonga [1], **Ayu Desrani** [2]

[1]State Islamic University of Maulana Malik Ibrahim Malang, Indonesia
[2] Yogyakarta State University, Indonesia

STYLISTICS: TEXT, COGNITION AND CORPORA, BY JANE LUGEA AND BRIAN WALKER, CHAM: PALGRAVE MACMILLAN CHAM, 2023, 296 PP. ISBN: 978-3-031-10422-0, 32,09 € (EBOOK).

This book introduces the three main pillars of contemporary stylistic research: text, cognition, and corpus, and shows how these three pillars work together to reveal the beauty and complexity of style in a text. Traditionally, stylistics has always centered on text, but now it also considers the role of cognition and the use of corpora to provide an empirical basis and increase objectivity. This book fills the void of comprehensive explanations for Stylistics scholars by combining detailed textual analysis, cognitive processes in textual interpretation, and corpus methodology. According to Mahlberg (2017), cognitive and corpus approaches complement each other and support the goals of contemporary Stylistics. Each chapter provides an analytical framework informed by Cognitive Science and corpus methodology to examine the influence of literary texts on readers' understanding.



This book consists of ten main chapters comprehensively organized by Lugea and Walker. In the opening chapter, the authors invite readers to explore the world of stylistics by introducing the concept of style. The authors not only discuss what style is but also explain how it is formed within a text. Each chapter presents in-depth empirical data related to stylistics, characterization, and the features of novel characters within the text. Furthermore, each chapter is supported by theoretical approaches and qualitative methodologies that contextualize the analysis more thoroughly, providing readers with a more comprehensive understanding.

Chapter 1 of this book serves not only as an introduction but also as an invitation to understand the worlds created by the texts we read. In this chapter, the authors carefully establish a solid foundation for further exploration of the book's content. Authors explain that by understanding and integrating texts, cognition, and corpora, we can uncover the ways in which style is formed within texts, thereby allowing readers to appreciate the beauty and uniqueness of each literary work. The explanations in this chapter begin by introducing the idea that a text creates a "world," a meta-

Citation: Ritonga, A.W., & Desrani, A. (2024). Stylistics: Text, cognition and corpora: Book review. *Journal of Language and Education, 10*(3), 183-185. https://doi.org/10.17323/jle.2024.22312

Correspondence:
Apri Wardana Ritonga
E-mail: 230104310018@student.uin-malang.ac.id

Received: August 23, 2023
Accepted: September 16, 2024
Published: September 30, 2024

phor that illustrates how a text interacts with its discourse context and how the reader's understanding of the text is influenced by that context.

Chapter 2 in particular provides a detailed guide to the World Stylistics approach, using Text World Theory (TWT) by Werth (1999). This theory teaches that discourse participants, such as readers, use clues from the text and existing knowledge to form a mental representation of the discourse, called the text world. With the help of Cognitive Science, TWT provides tools to better analyze and describe discourse. To show how TWT works, the authors apply it to the opening paragraph of Tess of the D'Ubervilles by Thomas Hardy (1893), revealing how this novel has expressed its concern about social inequality from the start. However, they also note that TWT is less practical for the analysis of longer texts. Therefore, authors suggest using a corpus to help find entry points into the text for further analysis.

Chapter 3 of this book invites readers to dive into the world of narrative point of view by combining various studies from leading experts such as Short (1996) and Simpson (2000). The authors introduce a new way of thinking about narrative, narration, and story, showing how narrative texts enable writers to communicate with readers. Using textual cues and existing knowledge, readers form an impression of the world of the text, which consists of narratives and stories. The narrator acts as a storyteller who tells the story from his or her point of view, adding characters and having the option to use the character's point of view. The authors then focus on a concrete example from Hemingway's (1926) The Sun Also Rises (TSAR), which is known for its objective style and lack of subjectivity. Corpus analysis shows that Hemingway's narrator uses few adverbs, evaluative adjectives, and modal auxiliary verbs, supporting the assessment that the novel has a journalistic style. However, point-of-view indicators such as 'let's' and 'ought to' were found to be redundant in the characters' direct speech, not in the narrative.

In Chapter 4 of this book, the reader is invited to understand the basic differences between narrative and discourse when talking about character. The authors explore various techniques for presenting a character's Speech, Writing and Thought Presentation (SW&TP). To explain how SW&TP can be analyzed using a corpus, the authors start by manually annotating these categories in Evelyn Waugh's (1903-1966) short story entitled 'Mr. Loveday's Little Outing'. This process is guided by the model SW&TP, which helps create a more detailed and structured analysis. Then, the authors go on to compare the results of this analysis with the fictional portion of Lancaster's corpus SW&TP (LancFic). Through a combination of a stylistic framework and a corpus approach, the author shows that analysis that is usually complicated can become easier and more focused. In addition, this analysis also involves a comparison between 'norms' SW&TP in fiction and researched texts.

Chapter 5 of this book invites us to delve deeper into the analysis of dialogue in drama through a combination of Pragmatics and Conversation Analysis (CA) models. By taking a quote from the British film I, Daniel Blake produced in 2016 (p. 116), the authors show how Gricean Conversational Implicature, Speech Acts, CA, and (im)politeness complement each other in understanding character interactions. A corpus approach is used to study dialogue in sitcoms, compare the characters' dialogue linguistically, and reveal prominent patterns of language use from the main characters. These patterns were further analyzed using a pragmatic (im)politeness framework, which revealed the meaning behind their speaking styles. This case study shows how quantitative deviations in texts relate to our social intuitions about the norms of oral communication, providing deep insight and analytical focus for understanding narrative style more comprehensively.

On the journey to understanding characters in fiction, Chapter 6 of this book invites readers to explore the concepts of character and characterization. The author introduces a character analysis checklist from Culpeper's (2001) work, then expands it to apply to characters in prose and drama. With detailed analysis, the authors show how keywords can be used to analyze the characters in Shaw's (1916) Pygmalion. At first these keywords were just a list of words, but they become full of meaning when viewed through the Cognitive Stylistic Framework, which emphasizes the importance of text, cognition, and corpus.

Moving on to Chapter 7, the authors take us deeper into the world of figurative language in fiction. They explain that understanding metaphor, metonymy, and simile requires a textual and cognitive approach supported by a corpus approach. Using the idea of Encyclopaedic Knowledge from Langacker's (2008) Cognitive Grammar, the authors help us understand the cross-domain mappings underlying metaphor and simile, as well as the domain mapping in metonymy. This chapter enriches our understanding of figurative language in fiction, showing how an integrated approach can reveal hidden depths of meaning in texts.

Chapter 8 takes us through research on thought style as an accumulation of linguistic features carried out over fifty years. The authors map various lexical and grammatical features that have been related to thought style while introducing pragmatic and cognitive approaches to deepen understanding of fictional thought. The chapter culminates in the innovative use of corpus methods and Cognitive Grammar to revisit Benjy's classic study of mind style in Faulkner's (1929) The Sound and the Fury. With these new techniques, the authors re-examine existing understandings of the text, illustrating Stylistics' commitment to a rigorous and replicable methodology. This book shows how new approaches can reinvigorate the analysis of classic texts, enriching our understanding of thought and character in literature.

Chapter 9 invites readers to delve deeper into the linguistic world of Humour Studies. By explaining the main characteristics of humorous discourse: ambiguity, incongruity, and resolution, the authors introduce the General Theory of Verbal Humour (GTVH) was proposed by Raskin and Attardo (1991). The GTVH is particularly relevant for stylistic analysis because the idea of script opposition helps analyze irony, and Knowledge Resources (KRs) structure the construction of jokes at various levels. The authors then incorporate TWT into joke analysis, defining the KRs situation as the world of joke texts, thereby integrating text and context in humour discourse. This modification connects GTVH to Stylistics research, showing the application of concepts and models from the previous chapter to humour. This chapter also shows how corpus methods can be used to validate claims about inappropriate or unexpected language use, providing new insights into humour analysis.

Chapter 10 concludes the discussion of this book by making the practice of Stylistics explicit. The authors provide a detailed guide to completing a Stylistics project, summarizing key principles such as rigor, replicability, and objectivity. This final chapter is a practical guide for readers who wish to explore and apply Stylistic analysis to their own work. Reading this book in full is like putting together a puzzle, where each piece combined reveals a complete and clear picture. The two authors of this book have succeeded in presenting the study of Stylistics coherently and interestingly and are equipped with the results of fifty years of research, making the contents of this book empirical and comprehensive. The study of Stylistics still needs to be followed up and developed to provide stronger tools for scholars to explore the effects and meaning of literary works.

Overall, this book provides a thorough and innovative exploration of contemporary stylistics, integrating text, cognition, and corpus approaches. Its strengths lie in the clear presentation of theoretical concepts and the practical tools it offers for analyzing literary texts. However, the book's reliance on corpus analysis may be daunting for readers unfamiliar with these methods, and the application of TWT may prove less practical for analyzing longer texts. Despite these limitations, the book's combination of cognitive science and corpus methodologies makes it a valuable resource for scholars seeking to deepen their understanding of stylistics.

ACKNOWLEDGMENTS

The authors would like to express his deepest gratitude to the Indonesia Endowment Fund for Education Agency (LPDP) for sponsoring our doctoral studies and supporting the publication of this article.

REFERENCE

Mahlberg, M. (2017). Corpus Stylistics. In M. Burke (Ed.), *The Routledge handbook of stylistics*. Routledge. https://doi.org/10.4324/9781315795331

CONTENT

EDITORIAL

RESEARCH PAPERS

REVIEW PAPERS

BOOK REVIEWS