Check for updates

Titles in research articles and doctoral dissertations: cross-disciplinary and cross-generic perspectives

Jialiang Hao[1,2]

Abstract

Although titles are often regarded as a minimal aspect of academic discourse, they play a crucial role in knowledge construction across various disciplines and genres. This study examined four features of titles, namely, title length, punctuation usage, structure, and content information, with a corpus comprising 1600 titles of research articles (RAs) from top journals and doctoral dissertations (DDs) from prestigious universities across four soft and hard science disciplines. The results confirm disciplinary and generic variations within the titles of these two critical academic genres. Titles in linguistics and medicine are generally longer than those in economics and computer science (CS). Slightly more titles in hard disciplines contain punctuation than do those in soft disciplines. The average title length of RAs is longer than that of DDs, and more RA titles than DD titles have punctuation in all four disciplines, with no apparent difference in the punctuation variety across the two genres, except for CS titles. Nominal group titles and compound titles are the two most common types, and prepositional phrase titles are the least common in all four disciplines and genres. The content information in titles is different in each discipline and genre. These findings are partially congruent with those of previous studies, indicating the significance of further investigating titles across disciplines and genres.

Keywords Titles · Research articles · Doctoral dissertations · Disciplines

Introduction

Research articles (RAs) and doctoral dissertations (DDs) constitute two important academic texts that most disciplinary communities use for the "dissemination and ratification of knowledge" (Koutsantoni, 2006, p. 20). RAs are major academic publications that are considered a professional genre (Kawase, 2015) through which new research results and contributions are reported and shared with the scientific community. Thus, RAs, written by more seasoned writers, constitute one of the most researched academic genres (Soler, 2007). A dissertation or thesis, terms that are derived from the difference between US and

✉ Jialiang Hao
hao_scarlett@yeah.net

1   Xi'an International Studies University, Xi'an, Shaanxi, China

2   Weinan Normal University, Weinan, Shaanxi, China

UK conventions[1] (Kawase, 2018), is the "highest level of advanced academic literacy" (Thompson, 2005, p. 307) and the most advanced educational genre (El-Dakhs, 2018; Kawase, 2015). Specifically, DDs constitute a genre positioned on a continuum between student-generated coursework papers and RAs subject to rigorous peer review (Xiao & Sun, 2020) and are characterized by specific features and conventions inherent to the genre (El-Dakhs, 2018; Kawase, 2015; Koutsantoni, 2006; Swales, 2004). Previous studies have compared specific linguistic features of the two genres, such as hedges (Koutsantoni, 2006) and stances (Qiu & Ma, 2019), as well as certain sections, such as abstracts (El-Dakhs, 2018; Hyland & Tse, 2005) and introductions (Bunton, 2002; Kawase, 2015; Taş, 2008), revealing the existence of genre variations that impact the writing of each section. Therefore, variations within the same sections of RAs and DDs need to be examined in terms of their context (El-Dakhs, 2018).

Titles are an essential component of academic texts, including RAs and DDs. Titles represent a "highly condensed" (Wang & Bai, 2007, p. 397) small subgenre that plays a pivotal role in indicating the contents of a document, attracting the correct groups of readers, and showing the expertise and professional identities of the authors. These terms are "front matter and summary matter" and cover the key information in a text (Swales, 1990, p. 179). Scanning titles in databases, references, and catalogs is an "early decision point" (Swales, 1990, p. 222) for academics to decide whether to read further. After Swales (1990) claimed that titles had not been thoroughly studied as an academic genre, interest in RA titles increased dramatically in various aspects (e.g., Cheng et al., 2012; Haggan, 2004; Hartley, 2007a, 2007b; Kerans et al., 2020; Lewinson & Hartley, 2005; Pearson, 2020, 2021; Soler, 2007, 2011), including structure (e.g., Wang & Bai, 2007; Xie, 2020), content (e.g., Goodman et al., 2001; Sahragard & Meihami, 2016), length (e.g., Bramoullé & Ductor, 2008; Jiang & Hyland, 2023), and lexical density (e.g., Li & Xu, 2019; Nagano, 2015); additionally, titles in a single discipline (e.g., Fox & Burns, 2015; Kerans et al., 2016) and multiple disciplines (e.g., Ball, 2009; Moslehi & Kafipour, 2022) and of synchronic (e.g., Nair & Gibbert, 2016; Méndez et al., 2014) and diachronic (e.g., Chen & Liu, 2023; Guo et al., 2015; Jiang & Jiang, 2023; Whissell, 2012, 2013; Yang, 2019) studies have been a focus. A plethora of related research has demonstrated the prominence of titles in RA academic writing. An RA may be one of many publications that a researcher has produced, whereas the vast majority of scholars traditionally write only one dissertation (Finlay et al., 2012). PhD candidates, as academic beginners, are confronted with pressure to publish to sustain employment and progress in their professional trajectories. Recognizing the generic similarities and distinctions between DDs, which constitute an educational genre, and RAs, which belong to a professional genre, can yield greater benefits for their future scholarly publications. Compared with RA titles, DD titles have received limited attention, and cross-generic comparative studies of RA and DD titles have rarely been performed. Thus, whether the titles of academic works differ across genres has not been addressed (Gesuato, 2008; Soler, 2011).

Although brief, titles for DDs and RAs in reputed journals are challenging to write. They should be attractive, informative (Ball, 2009; Hartley, 2005; Pearson, 2021) and concise (Bahadoran et al., 2019; Jiang & Jiang, 2023; Kerans et al., 2020; Wang & Bai, 2007). Moreover, scholars must construct titles under specific disciplinary academic community conventions, as title variations across disciplines have been identified (Ball, 2009; Diao,

---

[1] This paper follows US conventions.

2021; Haggan, 2004; Hartley, 2007a; Hyland & Zou, 2022). Thus, "before advice can be given on title writing, extensive research needs to be conducted to determine the discourse conventions within and across different disciplines and fields" (Anthony, 2001, p. 193). To this end, in this paper, we attempt to fill this gap by performing an exploratory comparison between RA and DD titles in various disciplines and providing preliminary research on the length, punctuation usage, structure, and content information of titles. This study will act as a reference for PhD candidates, supervisors, journal reviewers and editors, as well as EAP researchers, contributing to the existing literature on titleology.

Literature review

A good title is a "concise statement of the main topic of the research and should identify the variables or theoretical issues under investigation and the relationship between them" (American Psychological Association, 2020, p. 31); thus, writers may spend an "inordinate amount of time, discussion and mental effort" (Swales, 1990, p. 222) on making their titles appropriate and attractive. The qualities of good titles are listed in some studies. For example, authors must keep titles concise, specific, descriptive, and neutral; avoid interrogative or declarative components; and avoid acronyms (Dewan & Gupta, 2016). Swales and Feak (2012, p. 379) state that titles should be "self-explanatory," and Bahadoran et al. (2019) propose that a well-written title is informative, accurate, specific, concise, and unambiguous. Conciseness and informativeness are frequently reported as vital for generating good titles. These two characteristics are indicated by an essential feature of titles: length.

Thus, title length is frequently studied as a significant factor influencing the scholarly impact of RAs. Contradictory conclusions have been reached concerning title length and RA citation numbers. Positive correlations have been observed between title length and citation number (e.g., Braticevic et al., 2020; Habibzadeh & Yadollahie, 2010; Jacques & Sebire, 2009). Longer titles are more informative because they contain more keywords, making articles more accessible for readers to retrieve from databases. However, opposing findings have been reported in which shorter titles, which are succinct and straightforward, were shown to attract more views, downloads, and citations (Gnewuch & Wohlrabe, 2017; Hallock & Dillner, 2016; Jamali & Nikzad, 2011; Paiva et al., 2012). Therefore, writers often face the dilemma of balancing title conciseness and informativeness.

Heterogeneity has been observed in the length of titles within the soft and hard sciences. Some studies have demonstrated that titles in hard disciplines are longer than those in soft disciplines (Nagano, 2015; Soler, 2007), while others have revealed the opposite (Hyland & Zou, 2022; Lewison & Hartley, 2005). One explanation for this is the discrepancy between disciplines and journals, as the conventions of particular journals influence the results (Hyland & Zou, 2022). Another reason is that title patterns might change; thus, studies in different years could yield different results.

Accordingly, diachronic studies of titles emerged over the last decade. Whissell (2013) analyzed titles in psychology over 45 years and found an increasingly frequent use of longer titles. This trend has also been observed in other disciplines, such as applied linguistics (Jiang & Jiang, 2023), pragmatics (Li & Xu, 2019), linguistics and literature (Xiang & Li, 2020), and economics (Guo et al., 2015). Jiang and Hyland (2023) confirmed that titles in the six disciplines selected have been increasing in length, and the number of compound titles has increased in softer sciences over the past 20 years. Furthermore, the titles of articles in applied linguistics journals increasingly provided more information about the

research method/design from 1975 to 2015 (Sahragard & Meihami, 2016). Although Méndez et al. (2014) noted the same title length findings, they also observed an opposite trend regarding title structure and content information in astrophysics between 1998 and 2012; namely, more nominal phrase titles instead of compound titles were used, and more information about the results was included instead of information pertaining to methods. The aforementioned studies show that RA titles appear to have increased in length and informative nature over recent decades but with disciplinary variations.

In addition, disciplinary variations in the structure and content of titles have been discussed. NP structures are claimed to be prevalent in both soft (e.g., Nagano, 2015; Soler, 2011; Xie, 2020) and hard sciences (e.g., Diao, 2021; Morales et al., 2020; Wang & Bai, 2007). Moreover, researchers in soft disciplines have frequently used compound titles with colons as the major linking devices (e.g., Hartley, 2007a, 2007b; Jalilifar, 2010). Haggan (2004), for example, showed that science papers have the lowest percentage (21.5%) of compound titles, while this type appears three times more often in literature papers (60.8%). Furthermore, the information presented in titles varies across disciplines. Goodman et al. (2001) noted that the contents of RA titles of medical journals include "Topic only" and "Methods/design," "Dataset," "Results," and "Conclusion," and the first two types of information are most common. Kerans et al. (2020) explored titles in top-ranked general and specialty clinical medicine journals and reported massive variations regarding whether methods and results were mentioned. In contrast, in astrophysics RA titles, with minor variations, "Methods" appear less frequently than "Purpose" and "Results" (Méndez et al., 2014).

The above studies indicate a wealth of exploration of RA titles within one discipline or across several disciplines synchronically or diachronically, contributing to our knowledge of title characteristics. Cross-generic studies comparing RA titles with other academic titles have been conducted recently; for example, comparative studies on the titles of RAs and review articles (Kerans et al., 2016; Soler, 2007, 2011), conference papers (Hartley, 2007a), case reports (Salager-Meyer et al., 2013), and theses and dissertations (Jalilifar, 2010). The findings of these studies confirm the cross-generic variation in titles. Soler's (2007, 2011) surveys on research and review paper titles are noteworthy. She examined the structural construction of titles in the biological sciences and social sciences, reporting that structure is the distinctive feature between the titles of these two genres (Soler, 2007). Continuing the previous study, she investigated crosslinguistic and cross-generic differences between English and Spanish title constructions. The results revealed variations in title length depending on discipline and language, indicating that the predominance of nominal-group titles is not confined to discipline, genre, or language (Soler, 2011).

A DD is regarded as "the longest and most challenging piece of assessed writing" (Thompson, 2013, p. 284), with few graduate students ever encountering a comparable task. Consequently, DDs hold significant academic value and have garnered increasing amounts of attention in recent decades (Qiu & Ma, 2019; Zhou & Jiang, 2023). As the most advanced educational genre (Kawase, 2015), the DD is an "under-theorized, under-studied, and under-taught text" (Paré et al., 2009, p. 179) that embodies a complex and multifaceted nature (Paré, 2019) characterized by distinctive requirements and traditions that diverge from those of RAs. Swales (2004) noted that DD introductions present a general framework that resembles RAs. Accordingly, comparative analyses of certain sections of these two genres have been performed. For example, Kawase (2015) investigated the utilization of metadiscourse in the introductions of DDs and subsequently published RAs, revealing that a substantial number of authors exhibit a heightened frequency of metadiscourse use in their article introductions. As asserted by the author, these differences can be

attributed to the characteristics of DDs as an educational genre and RAs as a professional genre where authors must navigate intense competition to have their manuscripts accepted for publication. From a genre perspective, El-Dakhs (2018) compared 200 abstracts of DDs from renowned American and British universities with 200 abstracts of RAs in journals with high rankings and concluded that abstracts of DDs and RAs exhibit dissimilarities stemming from their representation of distinct genres, thereby exerting a significant influence on the composition of all sections and underscoring the significance of investigating disparities among corresponding segments of DDs/RAs in relation to their contextual factors. Previous studies on DD titles have been cross-disciplinary or cross-generic, focusing on the differences and similarities between the predominant title forms (Jalilifar, 2010; Soler, 2018). Jalilifar et al. (2010) performed a comparative study on MA and Ph. D theses titles and RA titles in linguistics, reporting that the thesis titles seemed to have a higher level of informativeness, whereas RA titles exhibited a wider range of structural diversity. Although these studies have revealed specific titular characteristics of DDs, it is noteworthy that no prior comparative studies on DD and RA titles across various academic fields have been conducted. Hence, we aim to address the following research questions:

(1) What are the word lengths of RA and DD titles by discipline?
(2) Does the use of punctuation marks highlight disciplinary and generic differences?
(3) Are there disciplinary and generic similarities or differences in the structure and content information of these titles?

Data and methods

Data collection

A bespoke corpus of 1600 titles was constructed and included two subcorpora: RA titles and DD titles (800 per group). Two hundred titles were chosen for each of the following selected disciplines under the categories recognized by SCImago[2] (Hyland & Zou, 2022): linguistics (Ling), economics (Econ), medicine (Med) and computer science (CS). These were representative of soft (e.g., Gesuato, 2008; Haggan, 2004; Nagano, 2015; Soler, 2007; Xie, 2020) or hard (e.g., Anthony, 2001; Soler, 2011; Wang & Bai, 2007; Yitzhaki, 2002) disciplines, and research was conducted on RA titles in these disciplines, facilitating a comparative analysis of the present study with existing research findings. The titles were chosen from original research articles and doctoral dissertations published between 1 January 2017 and 31 December 2021. Four contemporary high-performing journals (see Table 1) in each discipline and DDs from prestigious universities in the US and UK were selected considering the following parameters:
RA titles:

(1) RAs from top-tier journals with high impact factors[3] ranging from 14.4 to 168.9 in hard disciplines and 2.3 to 13.7 in soft disciplines (Web of Science) were included.

Table 1 Selected journals for the RA title subcorpus

Soft Disciplines	Journals	Hard disciplines	Journals
Linguistics	*Applied Linguistics* (IF 3.6) *Modern Language Journal* (IF 4.9) *System* (IF 6.0) *TESOL Quarterly* (IF 3.2)	Medicine	*British Medical Journal* (IF 107.7) *Journal of the American Medical Association (JAMA)* (IF 120.7) *Lancet* (IF 168.9) *The New England Journal of Medicine* (IF 158.5)
Economics	*Journal of Accounting & Economics* (IF 5.9) *Journal of Economic Growth* (IF 2.3) *Journal of Economic Perspectives* (IF 8.4) *Quarterly Journal of Economics* (IF 13.7)	Computer Science	*Artificial Intelligence* (IF 14.4) *IEEE Transactions on Pattern Analysis and Machine Intelligence* (IF 23.6) *Nature Machine Intelligence* (IF 23.8) *Science Robotics* (IF 25.0)

The hard-discipline journals are all SCIE-indexed (Science Citation Index—Expanded), while the soft-discipline journals are SSCI-indexed (Social Sciences Citation Index).
(2) Titles from research articles were included, and reviews, perspectives, editorials, and discussions were excluded.

DD titles:

(1) DDs were included from US and UK universities among the top 200 in "QS World University Rankings 2022".[4]
(2) The main subject of the DDs was limited to the terms of the included disciplines ("linguistics," "economics," "medicine," and "computer science"); interdisciplinary DDs involving other disciplines were excluded.

We applied stratified random sampling to select and download all the titles from the Web of Science (WoS) and ProQuest databases and cross-checked the data manually to ensure that they conformed to the abovementioned parameters. RAs and DDs from prestigious journals and universities were selected to guarantee that the chosen works effectively exemplified high standards of writing norms and conventions in their respective fields while also receiving approval from esteemed authorities within the academic community, such as journal editors, reviewers and university professors, whether they serve as readers or examiners (El-Dakhs, 2018). The quality of RAs is typically judged by the journals in which they are published, which provides a more representative measure of good RAs according to discipline than does including titles from random journals of varying quality and uncertain value in the field (Jiang & Hyland, 2023). The writers' identities, whether they were native or nonnative speakers, were considered inconsequential because their writings garnered approval from esteemed gatekeepers within the academic community, indicating adherence to established quality standards in the field. Thus, differentiating the cultural backgrounds of the writers was not prioritized in this study, as all accepted RAs and DDs met the necessary criteria for publication or examination (El-Dakhs, 2018). This approach is applicable to the present research, which focuses on the established norms of RAs and DDs in accordance with the consensus among members of the academic community.

Methods

All the RA titles were downloaded from WoS, and DD titles were collected from ProQuest and stored in a Microsoft Excel spreadsheet for coding and further analysis. The title word count, coding of punctuation marks, structure, and content information types were all assessed using Excel. The corpus was coded manually and searched or calculated automatically using formulas in the software. The researcher coded, classified and examined the data twice with a time interval of two weeks. Only eight discrepancies appeared, and these were resolved by discussion with two professors in applied linguistics. We used SPSS 22.0 to determine the statistical significance of the results. Specifically, chi-square tests were used.

---

[4] https://www.topuniversities.com/university-rankings/world-university-rankings/2022

The data analysis procedure followed the sequence of four title features: length, punctuation usage, structure, and content information type. The length of the titles was determined by the "word count formula" in Excel. Abbreviations, acronyms, and hyphenated words were counted as one word (Li & Xu, 2019; Milojevic, 2017; etc.), while words with an apostrophe were automatically counted by Excel as two words in some cases (e.g., "Teachers' Perceptions") and one word in other cases (e.g., "Student's"). The counting results were checked manually by the authors. A few examples are as follows:

> *Consumption of ultra-processed foods and cancer risk: results from NutriNet-Sante prospective cohort* (Med RA) (12 words)
> *Investigating Syntactic Complexity in EFL Learners' Writing across Common European Framework of Reference Levels A1, A2, and B1* (Ling RA) (18 words)
> *A three-fold approach to the imperative's usage in English and Dutch* (Ling RA) (11 words)

The punctuation marks were searched and counted in spreadsheets, and ten punctuation mark types were identified. The number of each type is the number of title(s) containing the particular punctuation, as in the following examples:

> *Reconfigurable magnetic microrobot swarm: Multimode transformation, locomotion, and manipulation* (CS RA) (colon; comma)
> *Modernizing Retailers in Emerging Markets: Investigating Externally-Focused and Internally-Focused Approaches* (Econ DD) (colon; hyphen)

Title structure was classified based on preliminary research and previous studies (e.g., Diao, 2021; Haggan, 2004; Li & Xu, 2019; Soler, 2007; Wang & Bai, 2007; Xie, 2020). The following five types were identified: nominal phrase (NP), compound (CP), full sentence (FS), verbal phrase (VP), and prepositional phrase (PP) titles. See Table 2.

The category of content information shows the information type provided by the titles. Based on the framework of Goodman et al. (2001) (Li & Xu, 2019; Sahragard & Meihami, 2016; Xiang & Li, 2020, etc.), the titles were divided into Topic Only, Method/design, Dataset, Results, and Conclusion titles. Table 3 shows the descriptions of each category. One title can include several types of information listed in the framework above; such titles were counted separately, as in the following example:

> *Association between active commuting and incident cardiovascular disease, cancer, and mortality: prospective cohort study* (Med RA) (Methods/design + Results)

This title contains information on Methods/design ("prospective cohort study") and Results ("association").

Results and discussion

Title length

Title length is considered crucial for RA retrieval and subsequent citation and for DDs to pass the examiners' review, thereby enhancing the academic significance of these works and aligning with the requirements of journals or dissertation committees. Longer titles provide a more comprehensive description of studies, increasing their discoverability in

Table 2 Classification of title types by structure

Title type	Description	Examples
NP title	Only including one part and commencing with (a) nominal group(s)	*Rate-Amplifying Demand And The Excess Sensitivity Of Long-Term Rates* (Econ RA) *Quantum Algorithms for Learning and Learning Algorithms for Quantum* (CS DD)
CP title	Containing two parts that are juxtaposed on either side of a punctuation mark (e.g., colon, question mark, period, etc.)	*Deep Learning Systems for Pneumothorax Detection on Chest Radiographs: A Multicenter External Validation Study* (CS RA) *Processing Context-Sensitive Expressions: The Case of Gradable Adjectives and Numerals* (Ling DD)
FS title	In the form of a full sentence	*Deep learning decodes the principles of differential gene expression* (CS RA) *Do Economists Swing for the Fences after Tenure?* (Econ RA)
VP title	Beginning with a verb, whether in gerund or to-infinitive	*Developing Data Efficient Algorithms in Artificial Intelligence* (CS DD) *Handling time varying confounding in observational research* (Med RA)
PP title	Initiated by a preposition	*On Non-culminating Accomplishments in Mandarin* (Ling DD) *On Adopting Additive Manufacturing for the Mass Customization of Orthopedic Implants* (Med DD)

Table 3 Framework for examining titles according to content information (adapted from Goodman et al., 2001 and Sahragard & Meihami, 2016)

Type of information	Description
Topic only	Title notes a subject but does not include information on other categories listed below, or it seems ambiguous
Methods/design	Title specifies the topic and an approach to study design, data management, or analysis (such as case−control, cohort, effectiveness, efficacy, frequency, incidence, prevalence, trend, or validity study; meta-analysis; randomized; or sensitivity–specificity or cost-effectiveness analysis); or it provides an incomplete description of a method (such as assessment, evaluation, or comparison)
Dataset	Title indicates the topic and name or acronym of a specific study, which can reveal the origin of the dataset considered or the population sample consulted
Result	Title contains the topic and quantitative information (a specific value), semiquantitative or ordinal information (such as increased, decreased, high, or low), or some other specification of a relation (such as association, change, correlation, determinants, effect, evidence, impact, influence, outcomes, predictors, relation, relationship, remission, risk, role, variability, or variation) about the results
Conclusion	Title includes the topic and an unequivocal statement based on the analysis of the reported evidence, such as the implications of a study or the overall conclusion

Table 4 Length of titles across disciplines and genres

Soft Disciplines				Hard Disciplines			
Discipline	No. of words	No. of title	Title length (words/title)	Discipline	No. of words	No. of title	Title length (words/title)
Econ RA	1799	200	9.00	Med RA	3287	200	16.44
Ling RA	2529	200	12.65	CS RA	1925	200	9.63
Econ DD	1412	200	7.06	Med DD	2547	200	12.74
Ling DD	2025	200	10.13	CS DD	1689	200	8.45
Total	7765	800	9.71	Total	9448	800	11.81

online literature searches and enhancing their potential impact (Habibzadeh & Yadollahie, 2010; Jacques & Sebire, 2009); however, other studies have shown that shorter titles may also imply a more precise focus and have the potential to attract a greater number of views and citations (Jamali & Nikzad, 2011; Paiva et al., 2012). As the aforementioned studies focused on RA titles across various disciplines, it became apparent that the average length of titles is heavily influenced by disciplinary conventions. Our study confirms disciplinary variations in title length for both RAs and DDs across four disciplines. Additionally, generic variations were detected in this study.

Overall, the average length of titles in Med and Ling is longer than that of titles in CS and Econ within the two respective genres (shown in Table 4), which is partly in line with the observations of Haggan (2004) and Soler (2007, 2011) that the average length of RA titles is longer in natural science than in social science. Specifically, the titles of Med RAs (average length: 16.44 words) were the longest among the four disciplines, while the titles of Econ DDs (average length: 7.06 words) were the shortest.

The rationale behind the variations in preferences for title length likely resides within disciplinary community conventions and the requirements set by academic journals. For example, Med RAs, which have the longest titles on average, tend to have an analogical structure, as shown below:

> *Neoadjuvant atezolizumab in combination with sequential nab-paclitaxel and anthracycline-based chemotherapy versus placebo and chemotherapy in patients with early-stage triple-negative breast cancer (IMpassion031): a randomised, double-blind, phase 3 trial* (28 words) (Med RA)
>
> *Pembrolizumab versus paclitaxel for previously treated, advanced gastric or gastro-oesophageal junction cancer (KEYNOTE-061): a randomised, open-label, controlled, phase 3 trial* (20 words) (Med RA)
>
> *Filgotinib as induction and maintenance therapy for ulcerative colitis (SELECTION): a phase 2b/3 double-blind, randomised, placebo-controlled trial* (17 words) (Med RA)

This structure contains dense information and a more specialized vocabulary, distinguishing one study from other articles and attracting a significant number of authors working in the field (Hyland & Zou, 2022), representing the journal's tradition.

From a cross-generic perspective, the average length of RA titles was longer than that of DD titles in all four selected disciplines. Genre-specific features and requirements may explain these findings. For RAs, given the growing competition to captivate readers and garner citations, authors and editors need to be aware of the significance of informative titles (Jiang & Jiang, 2023). In contrast, DDs, which function as "a final examination in a long student career" (Johns & Swales, 2002, p. 16), should comply with the guidelines of their respective universities; for example, the "Guidelines for the PhD Dissertation" of Graduate School of Arts and Sciences in Harvard University states that "the title should be as concise as possible, consistent with giving an accurate description of the dissertation.[5]" Thus, DD titles could be more concise and shorter. Taking the titles of Econ DDs, which have the shortest average length, as an example, we frequently observe the following structures:

> *Essays on Economics of Immigration* (5 words)s (Econ DD)
> *Essays on Finance and Macroeconomics* (5 words) (Econ DD)
> *Essays on Healthcare Economics* (4 words) (Econ DD)

An analysis of the abstracts and contents of these DDs revealed that they were "topic-based" DDs, which "typically commence with an introductory chapter which is then followed by a series of chapters which have titles based on sub-topics of the topic under investigation" (Paltridge, 2002, p.132). Within the community of economics, titles tend to follow the structure of "essays on + (a relatively broad) topic".

Punctuation marks

As shown in Table 5, the titles in soft disciplines included 463 (57.9%) with punctuation marks and 337 (42.1%) without; in the hard disciplines, 478 titles (59.8%) included

---

[5] https://www.hsph.harvard.edu/wp-content/uploads/sites/74/2014/09/guidelines-for-the-PhD-dissertation.pdf

Table 5 Titles with/without punctuation marks

| | Soft Disciplines n (%) | | | | | Hard Disciplines n (%) | | | | |
| | Econ | | Ling | | Total | Med | | CS | | Total |
	RA	DD	RA	DD		RA	DD	RA	DD	
Titles with punctuation	139 (69.5)	55 (27.5)	149 (74.5)	120 (60.0)	463 (57.9)	166 (83.0)	115 (57.5)	115 (57.5)	82 (41.0)	478 (59.8)
Titles without punctuation	61 (30.5)	145 (72.5)	51 (25.5)	80 (40.0)	337 (42.1)	34 (17.0)	85 (42.5)	85 (42.5)	118 (59.0)	322 (40.3)
Total	200 (100)	200 (100)	200 (100)	200 (100)	800 (100)	200 (100)	200 (100)	200 (100)	200 (100)	800 (100)

Table 6 Usage of punctuation marks

| Punctuation | Soft Disciplines | | | | Hard Disciplines | | | |
| | Econ | | Ling | | Med | | CS | |
	RA	DD	RA	DD	RA	DD	RA	DD
Colon	83	30	101	75	119	48	40	30
Hyphen	34	22	67	46	128	69	87	52
Comma	26	18	20	31	66	28	13	16
Parentheses	1	3	5	8	26	6	1	0
Apostrophe	12	1	30	5	2	6	1	0
Quotation marks	0	0	5	4	0	0	0	0
Question mark	28	1	8	3	0	3	3	0
Slash	0	0	3	3	1	3	0	0
Period	1	1	5	1	2	2	1	0
Exclamation point	0	0	0	1	0	0	0	0

punctuation, and 322 (40.3%) did not. The chi-square test was used to check for statistical difference, and the results demonstrated that overall, there was no significant difference (Pearson chi-square $= 0.581$, $p = 0.446$, $p > 0.05$) between soft and hard disciplines in terms of punctuation marks. Generally, from the cross-disciplinary perspective, titles share more similarities than differences in punctuation mark use since more titles have punctuation. Nevertheless, Econ and CS DDs are exceptions: the number of titles without punctuation is greater than that of those with punctuation.

Punctuation in titles functions to "coordinate structures, negotiate text space, and express authors' intention and emotions" (Diao, 2021, p. 6051). Thus, punctuation is an indicator of title complexity: the higher the percentage of punctuation usage is, the more complex the title. Table 5 also shows distinct differences between the two genres, as the percentage of titles with punctuation marks among RAs is greater than that among DDs across all four disciplines. The statistical test revealed a significant difference (Pearson chi-square $= 100.133$, $p = 0.000$, $p < 0.05$) between RAs and DDs regarding punctuation use in titles; overall, in this context, RA titles are more complex than DD titles.

Ten punctuation marks were identified from the corpus: colon, hyphen, comma, parenthesis, apostrophe, quotation mark, question mark, slash, period, and exclamation point. The apparent disciplinary differences between the soft and hard disciplines are shown in Table 6. Punctuation marks were more varied in soft disciplines, especially linguistics; nine and ten punctuation types were observed in the Ling RA and Ling DD titles, respectively. Titles in both genres of Econ and Med included a similar number of punctuation types, while titles of CS DDs had the fewest punctuation types (three out of the ten). From the cross-generic perspective, punctuation usage is similar across RAs and DDs in all four disciplines; for example, the same seven punctuation types appear in Econ RAs and Econ DDs.

In addition to punctuation type, the frequency of each type varied between soft and hard disciplines. The most recurrent punctuation in the soft disciplines was the colon, and the second most frequent was the hyphen. In contrast, in all four hard disciplines, more RA titles than DD titles had colons or hyphens.

Colons connect components in compound titles (Morales et al., 2020) and are frequently used in titles (Haggan, 2004; Hartley, 2007a, 2007b; Soler, 2011). The popularity of colons in

titles is probably because titles with colons are more informative than those without (Lewinson & Hartley, 2005). Authors in both genres across all disciplines prefer to use colons to compile compound titles to achieve informativeness.

Hyphens account for a relatively high percentage of punctuation marks in titles across all disciplines and genres. Hyphens combine two or more words to create new words, for example, "Data-Driven," "Alpha-T-Catenin," and "Quasi-experiment." The use of hyphens can be interpreted as increasing lexical density and complexity to achieve informativeness and conciseness. Hyphens also appear as non-alphanumeric characters in compound titles to relate the two components, serving the same function as colons, increasing the use of this type of punctuation in titles.

The comma is the third most frequent punctuation mark in titles in most disciplines and genres except in Econ and Ling RAs. The comma functions like a conjunction, joining a series of words to show the parallel components in titles. The relatively high-frequency use of commas indicates the multiple objects of study, as the following examples show.

> *Essays on Asset Pricing, Portfolio Choice, and International Finance* (Econ DD)
> *Abstraction, Generalization, and Embodiment in Neural Program Synthesis* (CS DD)
> *Assessing Epidemiology, Treatment Selection, and Outcomes of Head and Neck Merkel Cell Carcinomait* (Med DD)

The apostrophe in Ling RAs, parenthesis in Med RAs, and question mark in Econ RAs demonstrate the highest frequency among the remaining punctuation types. Apostrophes transform different words into plurals, contractions, and possessives, making the titles more logical. Among Ling RAs, the possessive apostrophe is most commonly used; for example, "students'," "teachers'" and "children's" represent the experimental group or participants of the studies. Parentheses are used to convey technical information in titles, most often to introduce acronyms:

> *Accuracy of the Patient Health Questionnaire-9 (PHQ-9) for screening to detect major depression: individual participant data meta-analysis* (Med RA)
> *Strengthening the reporting of observational studies in epidemiology using mendelian randomisation (STROBE-MR): explanation and elaboration* (Med RA)

Parentheses frequently appear in Med RA titles, indicating that medicine scholars tend to use abbreviations to avoid redundant and lengthy repetition of the technical terms. A question mark indicates that a sentence is a question and is found in CP or FS titles. The use of question marks in titles has three motivations: "to formulate the central issue of the study," "to stimulate interest," and "to express the dubious nature of results" (Ball, 2009, p. 676–677). This finding shows that using question marks in Econ RA titles is traditionally preferred, corresponding with the high percentage of FS titles (interrogative sentences) in this category.

In summary, regarding punctuation usage, RA titles are more complex than DD titles are, revealing a generic difference between the two genres. Regarding discipline, titles in hard and soft disciplines demonstrate similar complexity in terms of punctuation use; however, the latter show greater diversity. The colon, hyphen, and comma outweigh other punctuation in frequency, which could be explained by the title structure variable.

Structure

Figure 1 shows the occurrence of the five categories of structures found in the titles analyzed, namely, NP, CP, FS, VP, and PP. The preference for title types is clear: overall,
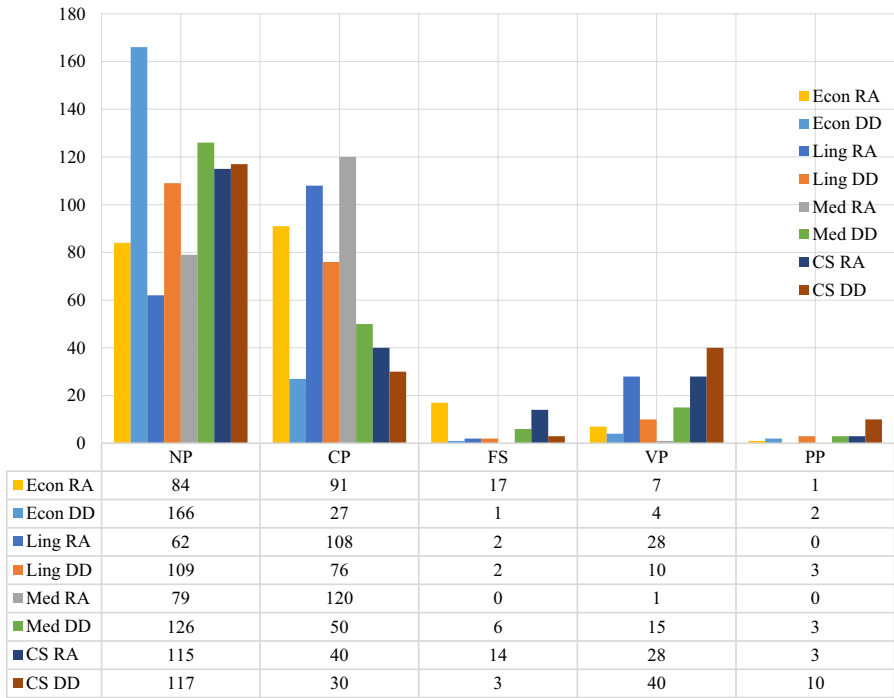
	NP	CP	FS	VP	PP
Econ RA	84	91	17	7	1
Econ DD	166	27	1	4	2
Ling RA	62	108	2	28	0
Ling DD	109	76	2	10	3
Med RA	79	120	0	1	0
Med DD	126	50	6	15	3
CS RA	115	40	14	28	3
CS DD	117	30	3	40	10

Fig. 1 Title structures across disciplines and genres

NP and CP are the two most preferred structures in all disciplines and genres analyzed, while PP is the least common type. The Pearson chi-square test results showed that there was no significant difference in the use of NP titles between soft and hard disciplines (Pearson chi-square $= 0.643$, $p = 0.422$, $p > 0.05$). In contrast, a significant difference existed in the use of CP titles (Pearson chi-square $= 10.726$, $p = 0.001$, $p < 0.05$). Regarding genre, we detected significant differences between RA and DD titles in the use of CP (Pearson chi-square $= 86.429$, $p = 0.000$, $p < 0.05$) and NP (Pearson chi-square $= 79.629$, $p = 0.000$, $p < 0.05$).

In line with the findings of previous studies (e.g., Soler, 2007; Wang & Bai, 2007; Xie, 2020), our results show a stronger preference for NP titles in the two genres across the four disciplines, except in Ling and Med RAs. NP titles, as single-format titles, are unadorned and explicit, as they directly and straightforwardly present the information of the study to the community (Hyland & Zou, 2022). They have the "powerful ability to compact information in an economical way through various pre- and post-modifiers" (Wang & Bai, 2007, p. 395). The prevalence of NP titles observed in this study suggests that the authors of both RAs and DDs favor the linear construction of NP titles that is widely accepted in each community.

CP titles account for a large proportion of Med, Ling, and Econ RA titles. CP titles are divided by punctuation marks into more than two parts with specific or logical relationships to allow readers to obtain more information (Li & Xu, 2019). Our observation from Ling and Econ RA titles underpins the previous statement that CP titles are "extremely commonplace across the soft knowledge disciplines" (Hyland & Zou, 2022,

p. 7). However, Med RA authors also show this preference for NP titles, which contradicts Wang and Bai's (2007) finding that 99.0% of Med RAs have NP titles. One explanation for this difference is that the titles of the aforementioned groups were collected only from the *New England Journal of Medicine*, and the journal's requirements for RA titles are distinct from those of other journals. Another interesting finding is that most Med RA titles were CP (60.0%) or NP (39.5%), and only one title was VP; thus, Med RA titles demonstrated the lowest structural variation. Relating this finding to the previous discussion of title length and punctuation in Med RAs, it seems justifiable to conclude that although only representing the fewest title types, Med RA titles have the highest complexity and informativeness in the corpus since CP titles are common and the titles include more punctuation and are longest on average.

Overall, a relatively small proportion of titles are FS; however, according to our counts, such titles demonstrate a relatively greater percentage of occurrence in Econ and CS RAs than in other disciplines and genres. Since economics is a social science field, this observation contradicts the findings of previous studies to a certain degree because it has been claimed that the FS structure is specific to titles in the natural sciences (Berkenkotter & Huckin, 1995; Li & Xu, 2019; Soler, 2007). FS titles accounted for 8.5% of the Econ RA titles, ranking third among all title types, just behind CP and NP, and all in the corpus were interrogative. For example,

> *Are CEOs' purchases more profitable than they appear?* (Econ RA)
> *Was the post-1870 fertility transition a key contributor to growth in the West in the twentieth century?* (Econ RA)

This finding also explains why the question mark usage frequency was highest in the titles of Econ RAs among all the topics studied. Hyland (2004) stated that social sciences, which encompass the study of human subjects, rely on qualitative analyses or statistical probabilities to construct and represent knowledge. Consequently, comprehensive exposition and a cautious approach in articulating assertions need to be used.

In summary, DDs in all four disciplines show a surprising frequency of NP titles, while for RA titles, authors switch their choices between NP and CP. These findings are somewhat consistent with those of previous studies on RA titles (Haggan, 2004; Soler, 2007, 2011; Wang & Bai, 2007; Xie, 2020), which indicated that NP is the most common title type across disciplines, genres, and cultures. The exception observed in this study, based on our analysis, is that the presence of CP titles can be attributed to disciplinary and generic characteristics, as they are predominantly found in Ling and Med RAs, thereby confirming the adherence to disciplinary conventions in constructing titles for RAs.

Content information

The five categories of content information in the titles are unevenly distributed across all four disciplines and two genres (see Table 7), revealing disciplinary and generic variations. Noticeably, Topic Only titles are the most frequent in both Econ and CS. This consistency is understandable in that the specific norms of the two disciplines lead authors to prefer titles that do not mention detailed research information but instead provide descriptive information about the topic, yielding concise and implicit titles. Topic-only titles are less informative (Sahragard & Meihami, 2016), are typically shorter and represent single structures (NP, VP, PP, and FS). Now, we can speculate on the interrelationship among the three parameters, namely, length, structure, and

Table 7 Content information by discipline and genre

Information Categories	Soft Disciplines n (%)				Hard Disciplines n (%)			
	Econ		Ling		Med		CS	
	RA	DD	RA	DD	RA	DD	RA	DD
Topic only	117 (58.5)	159 (79.5)	45 (22.5)	43 (21.5)	30 (15.0)	64 (32.0)	115 (57.5)	108 (54.0)
Methods/design	13 (6.5)	15 (7.5)	57 (28.5)	26 (13.0)	109 (54.5)	56 (28.0)	62 (31.0)	62 (31.0)
Dataset	38 (19.0)	18 (9.0)	110 (55.0)	138 (69.0)	79 (39.5)	68 (34.0)	6 (3.0)	21 (10.5)
Result	34 (17.0)	11 (5.5)	38 (19.0)	28 (14.0)	72 (36.0)	45 (22.5)	2 (1.0)	7 (3.5)
Conclusion	9 (4.5)	3 (1.5)	14 (7.0)	4 (2.0)	6 (3.0)	14 (7.0)	15 (7.5)	2 (1.0)

We observe that some titles could include more than one type of information, so the sum of the counts of title information does not equal the number of research titles investigated

content information: Econ and CS titles with more Topic Only types are shorter on average, and the incidence of CP structures is lower (except for among Econ RAs).

The Methods/design category is most prevalent among Med RAs, accounting for 54.5% of the total. This category is less frequently used in other disciplines and genres. This finding is inconsistent with previous findings regarding Med RA titles. Goodman et al. (2001) found Topic Only titles to be the most common (40%) among Med RAs and indicated that Methods/design titles were less common (30%). These two studies collected titles from the same four medical journals but during different periods, thus presenting conflicting results. This lack of conformity could be explained by changes in journal policies and editorial boards (Sahragard & Meihami, 2016; Whissell, 2013). Regarding Med DD titles, among which content information types are distributed more evenly, a lower proportion mentioned Methods/design. Thus, Med RA titles indicate a clear difference between the two genres. One reason for the heterogeneity between genres is that authors in the medical field write their titles strictly in conformity with specific journal policies to make their papers publishable and allow readers to quickly discover their work after publication (Whissell, 2013).

The Dataset category shows uniqueness and importance among Ling RAs and DDs, as 68.5% and 69.0%, respectively, of these titles include data, indicating a sharp difference from other disciplines. Mentioning research data could provide more detailed information, e.g., in linguistics, the language data investigated and/or the research participants (Sahragard & Meihami, 2016), are included in this category. Titles containing data information are prevalent among Ling RAs and DDs; the following are examples:

Exploring Social and Linguistic Diversity across African Americans from Rochester, New York (Ling DD)
Referential cohesion in Swedish preschool children's narratives (Ling RA)

The use of the Results category varies among the four disciplines. In soft disciplines, the numerical results are not contrastive, while discipline disparity exists within hard disciplines. More titles containing information about results are found in Med than in CS, especially in Med RAs. Notably, among the Med RAs, 90.3% of the titles with results information also contained other types of information, supporting the previous statement regarding the preference for more informative and complicated titles among Med RA authors. Examples are shown below:

Effectiveness of therapeutic heparin versus prophylactic heparin on death, mechanical ventilation, or intensive care unit admission in moderately ill patients with covid-19 admitted to hospital: RAPID randomised clinical trial (RA Med) (Result, Dataset, and Methods/design)
Association of Gestational Diabetes With Maternal Disorders of Glucose Metabolism and Childhood Adiposity (RA Med) (Result and Dataset)

Conclusion information is included in the titles across genres and disciplines least frequently, which is consistent with the findings of previous diachronic studies on applied linguistics journal titles (Li & Xu, 2019; Sahragard & Meihami, 2016), which showed that titles with conclusion information have always been uncommon. Titles containing conclusion information are more implicit than those containing Method/design and Results information (Haggan, 2004; Jalilifar, 2010). Hence, conclusion information is not frequently included in titles.

Conclusion

We examined four key features of RA and DD titles in soft and hard disciplines, namely, length, punctuation marks, structure, and content information, and the findings confirmed the disciplinary and generic variations. The title length results partially contradict previous findings that titles in hard disciplines are usually longer than those in soft disciplines (Haggan, 2004; Soler, 2007, 2011). The average length of Econ titles is the shortest in both genres, and that of Med titles is the longest. From the cross-generic perspective, the average length of RA titles was longer than that of DD titles across all disciplines selected. The punctuation mark analysis revealed that, in general, more titles in hard disciplines than in soft disciplines contain punctuation, with the exception of CS DDs. More RA titles than DD titles employ punctuation in all disciplines, suggesting the greater complexity of RA titles in punctuation use. Regarding punctuation variety, there was no noticeable difference across the four disciplines and two genres, except for CS DDs. Moreover, the examination of title structure is partly congruent with the findings of previous research concluding that NP titles and CP titles are the two most common types and that PP titles are the least common (Soler, 2007; Wang & Bai, 2007; Xie, 2020). The title content information analysis revealed variations across disciplines. The authors of both genres, with the exception of Med, prefer titles containing the same type of content information.

It is challenging to provide an explanation for these variances across disciplines. Hyland and Zou (2022) contend that it is crucial for scientists in the hard disciplines to differentiate their research from that of others. They operate within a highly competitive research landscape with substantial investments that necessitates rapid publication to establish early precedence. Furthermore, their focus lies within a distinct and easily identifiable field of study. However, social scientists' studies often encompass diverse and interdisciplinary aspects, thereby posing difficulties in identifying a cohesive target audience. Therefore, RA titles in the soft disciplines need to be more captivating and precise, aiming to capture readership for the article rather than assuming it. Nieuwenhuis (2023) discovered a greater prevalence of the same creative components in the titles of soft science RAs than in those of hard science RAs, which is consistent with the assertion made by Hyland and Zou (2022). Our findings support this assertion, as the use of question marks in Econ RA titles appears to be more frequent compared to other disciplines. Questions employed in titles can be perceived as attention-grabbing rhetorical elements (Chen & Liu, 2023), serving as a marketing tool to pique readers' interest (Ball, 2009; Nieuwenhuis, 2023).

The present research findings provide empirical evidence supporting the notion that RAs and DDs are distinct genres with their own genre-specific features (El-Dakhs, 2018; Kawase, 2015; Koutsantoni, 2006), even within the smallest subgenre, the title. This can inform both pedagogical and academic implications. First, the results of this study can be applied to research writing or academic writing classes or guidebooks to help teachers educate students on composing titles for RAs or DDs by instructing them on various classifications of titles (Sahragard & Meihami, 2016). Specifically, titles are generated not from an infinite range of options but rather from a relatively limited subset that reflects authors' understanding of their respective communities (Hyland & Zou, 2022). ESP instructors should also encourage students in these four disciplines to employ the appropriate NP or CP titles that are commonly utilized within their respective fields while incorporating frequently observed information into the titles. Furthermore, the findings of the current research could offer valuable insights for authors aspiring to publish internationally across the four disciplines under investigation, particularly within the sixteen selected journals

examined in this study (Xiang & Li, 2020). Additionally, this study also offers ESP/EAP researchers and practitioners some inspiration for title research, contributes to the literature on titleology, and supports the statement that titles are "serious stuff" (Swales, 1990, p. 224).

Finally, the current study is not without its limitations. The syntactic structures and information types of the titles were manually analyzed. Despite the author's consistent adherence to established procedures, the classification may still entail a certain degree of subjectivity. Moreover, we recognize that this was a relatively small-scale study, and much more must be done in this area. Further studies may consider academic titles from additional disciplines, e.g., creative disciplines and emerging disciplines, and titles of under-researched academic genres, such as academic blog posts and short videos.

Declarations

References

American Psychological Association. (2020). *Publication manual of the american psychological association* (7th ed.). American Psychological Association.

Anthony, L. (2001). Characteristic features of research article titles in computer science. *IEEE Transactions on Professional Communication, 44*(3), 187–194. https://doi.org/10.1109/47.946464

Bahadoran, Z., Mirmiran, P., Kashfi, K., & Ghasemi, A. (2019). The principles of biomedical scientific writing: Title. *International Journal of Endocrinology and Metabolism, 17*, e98326. https://doi.org/10.5812/ijem.98326

Ball, R. (2009). Scholarly communication in transition: The use of question marks in the titles of scientific articles in medicine, life sciences and physics 1966–2005. *Scientometrics, 79*(3), 667–679. https://doi.org/10.1007/s11192-007-1984-5

Berkenkotter, C., & Huckin, T. N. (1995). News value in scientific journal articles. In C. Berkenkotter & T. N. Huckin (Eds.), *Genre knowledge in disciplinary communication: Cognition, culture, power* (pp. 27–44). Routledge.

Bramoullé, Y., & Ductor, L. (2018). Title length. *Journal of Economic Behavior & Organization, 150*, 311–324. https://doi.org/10.1016/j.jebo.2018.01.014

Braticevic, M. N., Babic, I., Abramovic, I., Jokic, A., & Horvat, M. (2020). Title does matter: A cross-sectional study of 30 journals in the medical laboratory technology category. *Biochemia Medica, 30*(1), 128–133. https://doi.org/10.11613/bm.2020.010708

Bunton, D. (2002). Generic moves in PhD theses introductions. In J. Flowerdew (Ed.), *Academic discourse* (pp. 57–75). Routledge. https://doi.org/10.4324/9781315838069

Chen, X., & Liu, H. (2023). Academic "click bait": A diachronic investigation into the use of rhetorical part in pragmatics research article titles. *Journal of English for Academic Purposes, 66*, 101306. https://doi.org/10.1016/j.jeap.2023.101306

Cheng, S. W., Kuo, C.-W., & Kuo, C.-H. (2012). Research article titles in applied linguistics. *Journal of Academic Language and Learning, 6*(1), A1–A14.

Dewan, P., & Gupta, P. (2016). Writing the title, abstract and introduction: looks matter! *Indian Pediatrics, 53*(3), 235–241. https://doi.org/10.1007/s13312-016-0827-y

Diao, J. (2021). A lexical and syntactic study of research article titles in library science and scientometrics. *Scientometrics, 126*, 6041–6058. https://doi.org/10.1007/s11192-021-04018-6

El-Dakhs, D. A. S. (2018). Why are abstracts in PhD theses and research articles different? A genre-specific perspective. *Journal of English for Academic Purposes, 36*, 48–60. https://doi.org/10.1016/j.jeap.2018.09.005

Finlay, C. S., Sugimoto, C. R., Li, D., & Russell, T. G. (2012). LIS dissertation titles and abstracts (1930–2009): Where have all the librar* gone? *The Library Quarterly, 82*(1), 29–46. https://doi.org/10.1086/662945

Fox, C. W., & Burns, C. S. (2015). The relationship between manuscript title structure and success: Editorial decisions and citation performance for an ecological journal. *Ecology and Evolution, 5*, 1970–1980. https://doi.org/10.1002/ece3.1480

Gesuato, S. (2008). Encoding of information in titles: Academic practices across four genres in linguistics. In C. Taylor (Ed.), *Ecolingua. The role of E-corpora in translation and language learning* (pp. 127–157). EUT Edizioni Università di Trieste.

Gnewuch, M., & Wohlrabe, K. (2017). Title characteristics and citations in economics. *Scientometrics, 110*(3), 1573–1578. https://doi.org/10.1007/s11192-016-2216-7

Goodman, R. A., Thacker, S. B., & Siegel, P. Z. (2001). What's in a title? A descriptive study of article titles in peer-reviewed medical journals. *Science Editor, 24*(3), 75–78.

Guo, S., Zhang, G., Ju, Q., Chen, Y., Chen, Q., & Li, L. (2015). The evolution of conceptual diversity in economics titles from 1890 to 2012. *Scientometrics, 102*, 2073–2088. https://doi.org/10.1007/s11192-014-1501-6

Habibzadeh, F., & Yadollahie, M. (2010). Are shorter article titles more attractive for citations? Cross-sectional study of 22 scientific journals. *Croatian Medical Journal, 51*(2), 165–170. https://doi.org/10.3325/cmj.2010.51.165

Haggan, M. (2004). Research paper titles in literature, linguistics and science: Dimensions of attractions. *Journal of Pragmatics, 36*, 293–317. https://doi.org/10.1016/S0378-2166(03)00090-0

Hallock, R. M., & Dillner, K. M. (2016). Should title lengths really adhere to the American Psychological Association's twelve word limit? *American Psychologist, 71*(3), 240–242. https://doi.org/10.1037/a0040226

Hartley, J. (2005). To attract or to inform: What are titles for? *Journal Technical Writing & Communication, 35*, 203–213. https://doi.org/10.2190/NV6E-FN3N-7NGN-TWQT

Hartley, J. (2007a). Planning that title: Practices and preferences for titles with colons in academic articles. *Library & Information Science Research, 29*(4), 553–568. https://doi.org/10.1016/j.lisr.2007.05.002

Hartley, J. (2007b). Colonic titles! *The Write Stuff, 16*(4), 147–149.

Hyland, K. (2004). Disciplinary interactions: Metadiscourse in L2 postgraduate writing. *Journal of Second Language Writing, 13*, 133–151. https://doi.org/10.1016/j.jslw.2004.02.001

Hyland, K., & Tse, P. (2005). Hooking the reader: A corpus study of evaluative that in abstracts. *English for Specific Purposes, 24*(2), 123–139. https://doi.org/10.1016/j.esp.2004.02.002

Hyland, K., & Zou, H. (2022). Titles in research articles. *Journal of English for Academic Purposes, 56*, 101094. https://doi.org/10.1016/j.jeap.2022.101094

Jacques, T. S., & Sebire, N. J. (2009). The impact of article titles on citation hits: An analysis of general and specialist medical journals. *Journal of the Royal Society of Medicine Short Reports, 1*(2), 1–5. https://doi.org/10.1258/shorts.2009.100020

Jalilifar, A. (2010). Writing titles in applied linguistics: A comparative study of theses and research articles. *Taiwan International ESP Journal, 2*(1), 27–52.

Jalilifar, A., Hayati, A., & Mayahi, N. (2010). An exploration of generic tendencies in Applied Linguistics titles. *Journal of Faculty of Letters and Humanities, 5*(16), 35–57.

Jamali, H. R., & Nikzad, M. (2011). Article title type and its relation with the number of downloads and citations. *Scientometrics, 88*(2), 653–661. https://doi.org/10.1007/s11192-011-0412-z

Jiang, F. K., & Hyland, K. (2023). Titles in research articles: Changes across time and discipline. *Learned Publishing, 36*, 239–248. https://doi.org/10.1002/leap.1498

Jiang, G. K., & Jiang, Y. (2023). More diversity, more complexity, but more flexibility: Research article titles in *TESOL Quarterly*, 1967–2022. *Scientometrics, 128*, 3959–3980. https://doi.org/10.1007/s11192-023-04738-x

Kawase, T. (2015). Metadiscourse in the introductions of PhD theses and research articles. *Journal of English for Academic Purposes, 20*, 114–124. https://doi.org/10.1016/j.jeap.2015.08.006

Kawase, T. (2018). Rhetorical structure of the introductions of applied linguistics PhD theses. *Journal of English for Academic Purposes, 31*, 18–27. https://doi.org/10.1016/j.jeap.2017.12.005

Kerans, M. E., Marshall, J., Murray, A., & Sabaté, S. (2020). Research article title content and form in high-ranked international clinical medicine journals. *English for Specific Purposes, 60*, 127–139. https://doi.org/10.1016/j.esp.2020.06.001

Kerans, M. E., Murray, A., & Sabaté, S. (2016). Content and phrasing in titles of original research and review articles in 2015: Range of practice in four clinical journals. *Publications, 4*(2), 11. https://doi.org/10.3390/publications4020011

Koutsantoni, D. (2006). Rhetorical strategies in engineering research articles and research theses: Advanced academic literacy and relations of power. *Journal of English for Academic Purposes, 5*(1), 19–36. https://doi.org/10.1016/j.jeap.2005.11.002

Lewison, G., & Hartley, J. (2005). What's in a title? Numbers of words and the presence of colons. *Scientometrics, 63*(2), 341–356. https://doi.org/10.1007/s11192-005-0216-0

Li, Z., & Xu, J. (2019). The evolution of research article titles: The case of Journal of Pragmatics 1978–2018. *Scientometrics, 121*, 1619–1634. https://doi.org/10.1007/s11192-019-03244-3

Méndez, D. I., ÁngelesAlcaraz, M., & Salager-Meyer, F. (2014). Titles in English-medium Astrophysics research articles. *Scientometrics, 98*(3), 2331–2351. https://doi.org/10.1007/s11192-013-1174-6

Milojevic, S. (2017). The length and semantic structure of article titles—evolving disciplinary practices and correlations with impact. *Frontiers in Research Metrics and Analysis, 2*, 2. https://doi.org/10.3389/frma.2017.00002

Morales, O. A., Perdomo, B., Cassany, D., Tovar, R. M., & Izarra, É. (2020). Linguistic structures and functions of thesis and dissertation titles in dentistry. *Lebende Sprachen, 65*(1), 49–73. https://doi.org/10.1515/les-2020-0003

Moslehi, S., & Kafipour, R. (2022). Syntactic structure and rhetorical combinations of Iranian English research article titles in medicine and applied linguistics: A cross-disciplinary study. *Frontiers in Education, 7*, 935274. https://doi.org/10.3389/feduc.2022.935274

Nagano, R. L. (2015). Research article titles and disciplinary conventions: A corpus study of eight disciplines. *Journal of Academic Writing, 5*, 133–144. https://doi.org/10.18552/joaw.v5i1.168

Nair, L. B., & Gibbert, M. (2016). What makes a "good" title and (how) does it matter for citations? A review and general model of article title attributes in management science. *Scientometrics, 107*(3), 1331–1359. https://doi.org/10.1007/s11192-016-1937-y

Nieuwenhuis, J. (2023). Another article titled "Should I Stay or Should I Go?" or, the mass production of academic research titles. *The Information Society, 39*(2), 123–129. https://doi.org/10.1080/01972243.2022.2152916

Paiva, C., Lima, J., & Paiva, B. (2012). Articles with short titles describing the results are cited more often. *Clinics, 67*(5), 509–513. https://doi.org/10.6061/clinics/2012(05)17

Paltridge, B. (2002). Thesis and dissertation writing: An examination of published advice and actual practice. *English for Specific Purposes, 21*(2), 125–143. https://doi.org/10.1016/S0889-4906(00)00025-9

Paré, A. (2019). Re-writing the doctorate: New contexts, identities, and genres. *Journal of Second Language Writing, 43*, 80–84. https://doi.org/10.1016/j.jslw.2018.08.004

Paré, A., Starke-Meyerring, D., & McAlpine, L. (2009). The dissertation as multi-genre: Many readers, many readings. In C. Bazerman, A. Bonini, & D. Figueiredo (Eds.), *Genre in a changing world* (pp. 179–193). The WAC Clearinghouse and Parlor Press.

Pearson, W. S. (2020). Research article titles in written feedback on English as a second language writing. *Scientometrics, 123*, 997–1019. https://doi.org/10.1007/s11192-020-03388-7

Pearson, W. S. (2021). Quoted speech in linguistics research article titles: Patterns of use and effects on citations. *Scientometrics, 126*, 3421–3442. https://doi.org/10.1007/s11192-020-03827-5

Qiu, X., & Ma, X. (2019). Disciplinary enculturation and authorial stance: comparison of stance features among master's dissertations, doctoral theses, and research articles. *Ibérica, 38*, 327–348.

Sahragard, R., & Meihami, H. (2016). A diachronic study on the information provided by the research titles of applied linguistics journals. *Scientometrics, 108*, 1315–1331. https://doi.org/10.1007/s11192-016-2049-4

Salager-Meyer, F., Alcaraz-Ariza, M. A., & Briceño, M. L. (2013). Titling and authorship practices in medical case reports: A diachronic study (1840–2009). *Communication & Medicine, 10*(1), 63–80. https://doi.org/10.1558/cam.v10i1.63

Soler, V. (2007). Writing titles in science: An exploratory study. *English for Specific Purposes, 26*(1), 90–102. https://doi.org/10.1016/j.esp.2006.08.001

Soler, V. (2011). Comparative and contrastive observations on scientific titles written in English and Spanish. *English for Specific Purposes, 30*(2), 124–137. https://doi.org/10.1016/j.esp.2010.09.002

Soler, V. (2018). Estudio exploratorio de títulos de tesis doctorales redactados en lengua Española [Exploratory study of Ph.D. thesis titles written in Spanish]. *Lebende Sprachen, 63*(2), 374–392. https://doi.org/10.1515/les-2018-0022

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

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

Swales, J., & Feak, C. B. (2012). *Academic writing for graduate students*. Ann Arbor, MI, The University of Michigan Press.

Taş, E. E. I. (2008). A corpus-based analysis of genre-specific discourse of research: The PhD thesis and the research article in ELT. *Doctorate thesis at Middle East Technical University (Turkey)*. http://citeseerx. ist.psu.edu/viewdoc/download?doi=10.1.1.633.4476&rep=rep1&type=pdf

Thompson, P. (2005). Points of focus and position: Intertextual reference in PhD theses. *Journal of English for Academic Purposes, 4*(4), 307–323. https://doi.org/10.1016/j.jeap.2005.07.006

Thompson, P. (2013). Thesis and dissertation writing. In B. Paltridge & S. Starfield (Eds.), *The handbook of English for specific purposes* (pp. 283–299). West Essex.

Wang, Y., & Bai, Y. (2007). A corpus-based syntactic study of medical research article titles. *System, 35*, 388–399. https://doi.org/10.1016/j.system.2007.01.005

Whissell, C. (2012). The trend toward more attractive and informative titles: American psychologist 1946–2010. *Psychological Reports, 110*(2), 427–444. https://doi.org/10.2466/17.28.pr0.110.2.427-444

Whissell, C. (2013). Titles in highly ranked multidisciplinary psychology journals 1966–2011: More words and punctuation marks allow for the communication of more information. *Psychological Reports, 113*(3), 969–986. https://doi.org/10.2466/28.17.PR0.113x30z5

Xiang, X., & Li, J. (2020). A diachronic comparative study of research article titles in linguistics and literature journals. *Scientometrics, 122*, 847–866. https://doi.org/10.1007/s11192-019-03329-z

Xiao, W., & Sun, S. (2020). Dynamic lexical features of PhD theses across disciplines: A text mining approach. *Journal of Quantitative Linguistics, 27*(2), 114–133. https://doi.org/10.1080/09296174.2018.1531618

Xie, S. (2020). English research article titles: Cultural and disciplinary perspectives. *SAGE Open, 10*(2), 1–12. https://doi.org/10.1177/2158244020933614

Yang, W. (2019). A diachronic keyword analysis in research article titles and cited article titles in applied linguistics from 1990 to 2016. *English Text Construction, 12*(1), 84–102. https://doi.org/10.1075/etc.00019.yan

Yitzhaki, M. (2002). Relation of the title length of a journal article to the length of the article. *Scientometrics, 54*, 435–447. https://doi.org/10.1023/A:1016038617639

Zhou, H., & Jiang, F. K. (2023). "The study has clear limitations": Presentation of limitations in conclusion sections of PhD dissertations and research articles in applied linguistics. *English for Specific Purposes, 71*, 34–47. https://doi.org/10.1016/j.esp.2023.02.001