Embedding in research article titles: The case of science and history

Alvin Ping Leong
Nanyang Technological University

Abstract
As the first few words to greet the reader, titles of research articles play a crucial role in influencing reader decisions on whether to access the articles. Over the years, various features of titles have been investigated, but the extent of their grammatical complexity has not received adequate research attention. This diachronic study addresses this gap by focusing on grammatical embedding in research article titles. One thousand titles from *Science* and *English Historical Review* (from 1886 to 2024), representing the disciplines of science and history, were compared. The analysis revealed that in science titles, the use of embedded noun phrases (NPs) in postmodifiers decreased steadily after peaking in 1950. The use of embedded NPs in history titles lagged behind ranking NPs, with the gap widening after 1950; by 2024, the gap between ranking and embedded NPs was at its widest. Collectively, both science and history titles displayed a tendency to use less embedding, particularly after 1950. This study highlights the changing titling preferences in science and history, and provides some guidance to help writers make their titles less complex. Several areas for further study are also recommended.

Keywords: titles, research articles, science, history, embedding, grammatical complexity

1 Introduction

Titles of academic papers are a fascinating area of study. They are the first lines of text to greet the reader, who must then decide whether to download the manuscript and (hopefully) read it. Diao (2021) observes aptly that the "[t]itle of a research article is an abstract of the abstract" (ibid. 6041), implying that the primary focus of titles in academic disciplines is not simply to attract the reader's attention, but to provide sufficient information to help the reader better understand the core details of the research article.

How this information is conveyed in titles has been investigated widely in a number of studies. These include corpus-based work on the use of humor (Heard et al. 2023; Sagi and Yechiam 2008), quoted speech (Pearson 2021), and even punctuation marks (Diao 2021; Hao 2024). A range of disciplines have also been investigated in both synchronic and diachronic studies, including psychology (Hallock and Bennett 2021; Sagi and Yechiam 2008), pragmatics (Li and Xu 2019), medicine (Kerans et al. 2020), and a mix of social science and STEM (i.e., science, technology, engineering, and mathematics) disciplines (Hyland and Zou 2022; Jiang and Hyland 2023), among others.

Taken together, these studies reveal a trend toward lengthier titles, an increasing use of sentences, and a sustained preference for noun phrases. This has led some writers to offer advice, particularly to early-career scholars, on the crafting of titles to boost readership and, hopefully, citations of their research efforts. Kenny (2023), for instance,

VERTAISARVIOITU
KOLLEGIALT GRANSKAD
PEER-REVIEWED
www.tsv.fi/tunnus

Finnish Journal of Linguistics 38 (2025), 89–108
DOI: 10.61197/fjl.157572

advocates the use of clauses and recommends that titles be as descriptive as possible. Other pieces of advice, though, may need to be evaluated more carefully, given the current research findings. Moore (2020) promotes brevity, arguing that "articles with longer titles were found to be downloaded less often and were less cited" (ibid. 56). This advice, however, runs against findings that show a move toward longer titles, which, according to Jiang and Hyland (2023: 242), help writers to "distinguish [their papers] from the herd and appeal to wide readership", a clear departure from Moore's view.

Indeed, on this count, pragmatism, rather than adherence to often de-contextualized recommendations, may better serve the needs of writers. For instance, in an editorial for *Langmuir*, a chemistry journal, Heinz and Walker (2023) advocate that titles should typically not exceed 15 words. Interestingly, barely a year later in the journal's first issue of 2024, 33 out of the 106 titles in that issue exceeded the word limit. So, while it is technically true that the titles on the whole were less than 15 words long, it is also not atypical of them to exceed this limit.

We therefore need a closer look at how the norms of titling have shifted over time, reflecting the changing concerns of writers, and what these changes, if any, imply about the future direction of these norms.

Current diachronic work on article titles has already addressed numerous issues, such as lexis, syntax, and informativeness. However, a more careful investigation about grammatical complexity—specifically in relation to embedding and condensed phrasing—appears lacking. This is particularly needed in the case of academic writing, which has often been characterized as being structurally compact (Biber et al. 2022). The work of Jiang and Hyland (2023) also reminds us that attention should be paid to disciplinary influences on titling norms. The two broad disciplines that are often viewed to be markedly different from each other are the sciences and humanities, as each has its own distinct writing style (Grech 2019). How grammatical complexity is manifested in the titles in these two disciplines remains to be explored.

This paper addressed these gaps by investigating the grammatical complexity of 1,000 research article titles taken from two journals representing science and history. The latter is often seen as a core discipline of the humanities; as Denecke (2021) notes, the only discipline representing the humanities in the *Oxford Handbook of Global Studies* (Juergensmeyer et al. 2019) is history. The science and history titles were sampled from five time periods, from 1886 to 2024. The notion of grammatical complexity, with a focus on noun phrases and embedding, was based on the rank scale of the Hallidayan framework (Halliday and Matthiessen 2014). It is hoped that the study will shed some light on the disciplinary norms of titling, and what this implies for future trends.

This paper is organized as follows. Section 2 offers a review of relevant studies on research article titles. Section 3 elaborates on the concepts of the rank scale and grammatical complexity. The details of the corpus and the method of analysis are described in Section 4. The findings are presented and discussed in Sections 5 and 6. The concluding section summarizes the major findings and recommends areas for further work.

2 Literature review

Studies on various aspects of research article titles generally follow two broad themes—how these aspects have changed over time, and how they presently compare between different fields of study. We begin with diachronic studies. Consistently, available studies

have shown that titles have increased in word length over the years (Jiang and Hyland 2023; Jiang and Jiang 2023; Li and Xu 2019). This remains true even when the comparative time periods are not very far apart (e.g., 1980–2018) (Xiang and Li 2020). The current average word length of titles is about 15 to 16 words (Jiang and Jiang 2023; Kerans et al. 2020), although differences across disciplines and journals are only to be expected.

This increase in word length naturally brings about changes in other aspects as well. In the discipline of pragmatics, Li and Xu (2019) found that the increase in word length corresponded with an increase in lexical density and a preference for compound titles, where distinct parts of the title are separated by a punctuation mark, usually the colon. Li and Xu (2019) argue that these indicators together suggest a "change in informativeness and complexity of titles" (ibid. 1626). This is echoed in the work of Xiang and Li (2020), who note that titles of linguistics articles had also become more diversified, containing information pertaining to not merely the research topic, but also the results, datasets, and methodologies.

In a major study comprising 36,000 titles spanning four decades (1960–2020) and involving six disciplines, Jiang and Hyland (2023) found an increasing use of indicatives (rather than merely noun phrases) and discipline-specific differences in the use of single or compound titles; whereas the hard sciences preferred single titles for their directness, the softer disciplines favored compounds in view of the interpretive nature of the research. As Pearson (2020) notes, "the predilection for compound patterns reflects [the writer's] ability to present the relationship between the constitutive elements as complex or ambiguous" (ibid. 1010).

Apart from differences in the use of single or compound titles, synchronous comparative studies have also revealed other discipline-specific differences. One of them, interestingly, is the word length of titles; although titles have progressively lengthened over time, as seen above, certain disciplines may exert greater pressure on writers to make titles more informative, and so lengthen them. Hyland and Zou (2022), for instance, found that titles in the social sciences were comparatively longer than STEM titles, explaining that "social scientists work in areas which often span more disparate audiences which may lack the background needed to see the importance and coverage of the work" (ibid. 5). Another difference concerns the use of questions. While they are rarely used across disciplines, they are even more rare in the STEM disciplines (Cook and Plourde 2016; Hyland and Zou 2022). This is perhaps unsurprising, given the focus of the hard sciences on objectivity, precision, and clarity. Questions, on the other hand, incorporate an element of non-assertiveness and open-endedness in titles, and are thus normally avoided in such disciplines.

Yet, issues concerning word length, and the use of questions and simple/compound titles relate to only the more visible and broader components of titles. While they do influence reader decisions to access or bypass the manuscript (Hallock and Bennett 2021), a subtler aspect of complexity has not always been addressed adequately in the literature. This relates not merely to the nomenclature adopted in the respective disciplines. Technical terms, after all, are unavoidable in academic writing, contributing in part to what Biber and Gray (2016: 3) term "a form of deliberate obscurity". Rather, the issue of concern is a more fundamental, *grammatical* one; it relates to the extent to which titles contain compressed and embedded elements, since these can make the reading process unnecessarily difficult.

A case in point is the ubiquitous noun phrase. These phrases are widely acknowledged to be pervasively used in titles (Hao 2024; Xiang and Li 2020). This should come as little

surprise since nouns are naming words, and it would be impossible to craft a title without using at least one nominal element. Notwithstanding this, the use of noun phrases in the titles of research articles remains under-studied. This is compounded by the use of the general label "NP" in some studies for all types of noun phrases (Hao 2024; Xie 2020). As a welcome contrast, Li and Xu (2019) established a more detailed classification of noun phrases and identified four sub-categories—(a) noun phrase + prepositional phrase, (b) noun phrase + noun phrase, (c) present participle phrase, and (d) prepositional phrase. Unfortunately, for some reason, the researchers did not follow these labels consistently. For instance, in (1–2) below, only (1) was categorized as a noun phrase + prepositional phrase, but (2) was regarded as containing merely noun phrases, even though each noun phrase has a prepositional phrase postmodifier (examples from Li and Xu 2019: 1624):

(1) Pragmatic markers in Chinese

(2) Pragmatic markers in contrast: The case of 'well'

Li and Xu's (2019) analysis also appears to be broad-based, as it excluded embedded elements in several instances. In (3), the underlined segment was labeled simply as a noun phrase, although it has a prepositional phrase postmodifier with an embedded noun clause containing a second prepositional phrase.

(3) 'I just found your blog'. <u>The pragmatics of initiating comments on blog posts</u>

Ignoring such embedded elements runs the risk of under-reporting the true extent of their use, and the broader characteristic(s) of the discipline under study. To understand more fully the concept of embedding, we turn next to discuss the issue of grammatical complexity.

3 Grammatical complexity

The lay understanding of complexity is often viewed in terms of difficulty; something that is harder to do or understand is typically described as being complex. But this is only partly true in writing. The complexity of a text also needs to be understood in terms of its having different but related parts. In Michael Halliday's systemic functional grammar, for instance, a sentence comprising a single clause is termed a simplex (Halliday and Matthiessen 2014: 430), but if the sentence contains two or more clauses, the term "clause complex" is used instead. When applied to grammatical units in general, then, we may understand complexity in the following way: "grammatical complexity is conceptualized as essentially equivalent to structural elaboration, with longer linguistic units and units having more structural embedding being considered more complex" (Biber et al. 2022: 2).

As we see here, it is not merely the adding of component parts that makes a linguistic unit complex. The *embedding* of these parts also plays a role. What counts as embedding is helpfully explained using the Hallidayan rank scale, which hierarchically arranges grammatical units as depicted in Figure 1 (Halliday and Matthiessen 2014: 115).

Clause
|
Word phrase
|
Word
|
Morpheme

Figure 1. Hallidayan rank scale for grammatical units (adapted)

According to this scale, any unit that occurs at a rank beneath its own rank is considered embedded. For example, noun clauses are embedded because their grammatical function (as the subject, object, or complement of the larger clause) operates at the rank of the word phrase. Similarly, prepositional phrases that function as noun phrase postmodifiers operate at the rank of the word, and are therefore considered embedded.

Examples of embedding are given in (4–5). Following the Hallidayan convention, embedded clauses are enclosed within double square brackets [[…]], and embedded phrases within single square brackets […]. All examples from hereon are taken from this study's corpus.

(4) Experimental hypertension [[produced by a plastic capsule [[applied to the kidney]]]] *(Science, 1950)*

(5) Richard Champion and the Rockingham Whigs: The Aristocratic politics [of a Bristolian Quaker merchant [in the age [of the American revolution]]] *(History, 2023)*

In (4), the sheer amount of information packed into 12 words increases the cognitive load on the reader. The focus here is on the outcome (hypertension); the method (involving the plastic capsule and the placement of the capsule) is described in two embedded clauses, implying that these pieces of information are subsidiary to the outcome. In (5), we see a similar cognitive demand on the reader. The title describes not merely Aristocratic politics but specifically that involving a particular merchant and a particular era. This is achieved through the use of three nested prepositional phrases, embedded in a Chinese-box fashion.

Embedding is not an uncommon phenomenon in academic writing; the literature, in fact, suggests that it is one of the hallmarks of scholarly writing. As Biber et al. (2022) note, "what we commonly find in research writing are linguistic units with phrases embedded in phrases" (ibid. 3). In the case of titles, which favor succinctness, it is only reasonable to expect that embedding is also pervasively employed. Yet, such complexity issues have not been widely addressed in research work on titles.

This paper is an attempt to address this gap by investigating the extent of embedding in the titles of science and history research articles over the last 138 years. This in no way suggests that embedding is the only marker of complexity; other syntactic features, such as premodification, do also contribute to grammatical complexity. Embedding in titles, though, remains an under-explored feature, and focusing on it complements what we already know about the broader, more obvious characteristics of titles.

4 Data and methodology

4.1 Selection of corpus

The corpus comprised 1,000 titles of research articles published in the journals *Science* and *English Historical Review*, representing science and history, respectively. The reason why these two major disciplines were selected is that they are often taken to be on two ends of the continuum, due chiefly to their markedly different research areas and methodologies (Courtney and Courtney 2008). Whereas science investigates physical entities governed by natural law, history explores past events based on artefacts and documentary evidence. In science, theories and hypotheses can be tested in reproducible experiments, but these are not quite possible in the case of history. Investigating these two disciplines can therefore offer us a glimpse as to whether grammatical complexity is a feature that is discipline-specific, or generalizable across two very different disciplines.

The two journals were selected because of their heritage. *Science* is a prestigious journal with a long history. Established in 1880, it publishes articles in all fields of science. According to Scimago (2023), it was ranked first in the history and philosophy of science, and second in multidisciplinary science for the year 2023. *English Historical Review* was established in 1886, only six years after the inception of *Science*, and is the oldest English-language scholarly journal in the discipline of history. Although it is presently a Q2 journal, and is therefore not as highly ranked as *Science*, it nevertheless offers a glimpse of the continuous development of history titles across the years.

Research article titles were selected over five time periods—1886, 1900, 1950, 2000, and 2024. In the first four time periods, the first 100 titles of each year were recorded. In the final time period (2024), the 100 most recent titles at the time of analysis (June 2024) were recorded. This resulted in a total of 1,000 titles, as shown in Table 1.

Table 1. Number of research article titles in each time period

Journal	Time period					Total
	1886	1900	1950	2000	2024	
Science	100	100	100	100	100	500
English Historical Review	100	100	100	100	100	500
Total	200	200	200	200	200	1,000

It should also be pointed out that the indicated year of each time period is merely a marker. This is because it was not always possible to obtain 100 titles in a single year, particularly in the case of history titles. For instance, the first 100 titles from *English Historical Review* spanned a few years, 1886–1888.

4.2 Method of analysis

Preliminary analyses of the corpus were conducted to compute the word length and lexical diversity of each title. Lexical diversity is a measure of the range of words used in a text segment, and can provide a further characterization of the range of words used in the title. Lexical diversity was measured using Guiraud's Index (GI) (Guiraud 1960) instead of the traditional type-token ratio (TTR). TTR is commonly used as an approximate measure of the lexical variation of a text segment; the formula is the simple division of the total

number of distinct words (types) by the total number of words (tokens) in the segment, i.e., $\frac{types}{tokens}$. However, an increase in the word length of the segment results in a larger denominator, which in turn lowers the TTR unless there is a corresponding increase in the number of types. Since the focus is more on types than tokens, GI compensates for this by taking the square root of the denominator. This revision, also called root TTR for obvious reasons, takes the following form: $\frac{types}{\sqrt{tokens}}$.

On a note of caution, though, there is presently no adequate measure to compute the lexical diversity of short texts such as titles (Hess et al. 1986; Liimatta 2024). Even workarounds such as the moving-average TTR (Covington and McFall 2010) are hampered by the brevity of titles in the corpus, the shortest of which comprised only a single word. In spite of this word-length constraint, a measure was needed to provide at least an approximation of the lexical diversity of titles, and it was for this reason that GI was selected for its straightforward computation and general usefulness (see Daller and Xue 2010). The GI scores reported here, therefore, are not intended to be definitive, but indicative.

In the actual analysis, all clauses and word phrases, including their embedded versions, were counted. In the analysis, the usual abbreviations for linguistic elements were used, as follows: NP (noun phrase), PP (prepositional phrase), and RC (relative clause). The categorization adhered to the guidelines in Carter and McCarthy (2006). As a word phrase is defined by its head word, coordinated heads (e.g., "Queen Elizabeth and the Valois Princes") were regarded as belonging to separate word phrases. A sample of the analysis of the first three titles from the first issue of *English Historical Review* is provided in Figure 2.

	HISTORY 1886	Words (Token)	Type	Guiraud's Index	Declarative (clause)	Question (clause)	Question (non-clause)	NP	PP	NP (embedded)	PP (embedded)	RC (embedded)
1	German schools [of history]	4	4	2				1		1	1	
2	Homer and the early history [of Greece]	7	7	2.64575				2		1	1	
3	The tyrants [of Britain, Gaul, and Spain A.D. 406-411]	9	9	3				1		3	1	

Figure 2. Sample analysis of the first three titles in *English Historical Review*

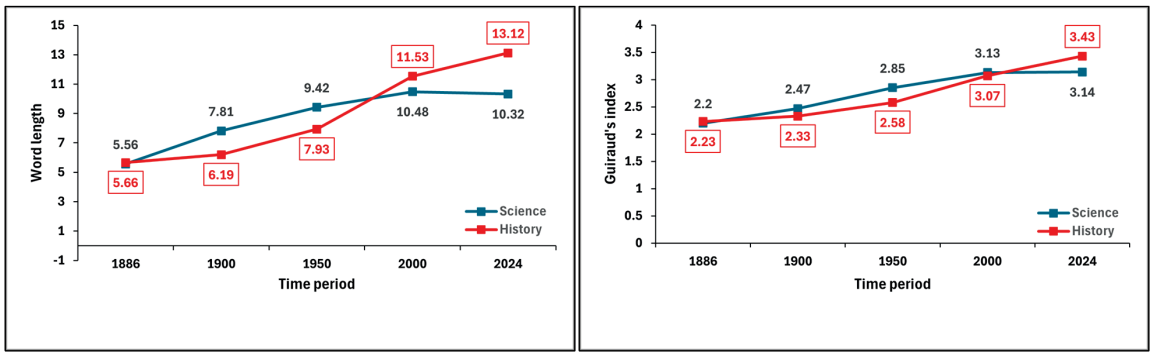
4.3 Statistical analysis

All statistical tests were conducted using *Real Statistics Resource Pack* (Zaiontz 2022), a Microsoft Excel add-in. Word length and GI differences for each time period were tested using the omnibus one-way analysis of variance (ANOVA) test, and the Tukey honestly significant difference post-hoc test for all significant ANOVA results. Statistical differences in word length and GI between science and history titles for each time period were determined using the independent-samples *t*-test. The data on the various categories of (embedded) grammatical units were analyzed using the χ^2 test for independence and, where needed, the two-sample test of proportions. The significance level for all statistical tests was α=.05. In this paper, significant differences are marked with a single asterisk for $p<.05$, and double asterisks for $p<.01$.

5 Findings

5.1 Word length and Guiraud's Index

The word lengths of and GI scores for both science and history titles increased over the years (Figure 3). ANOVA results were significant for both science [$F(4, 495)=32.45$, $p<.001$**] and history [$F(4, 495)=91.07$, $p<.001$**] titles. Significant ANOVA results were also returned for the GI scores for both disciplines [science: $F(4, 495)=69.28$, $p<.001$**; history: $F(4, 495)=108.28$, $p<.001$**].



(a) Mean word lengths (b) Mean GI scores

Figure 3. Mean word lengths and lexical diversity scores

These word-length results mirror those reported in other studies (Jiang and Hyland 2023; Li and Xu 2019; Xiang and Li 2020), but they also highlight distinct patterns in both disciplines (Table 2; only significant results are reported). In science titles, Tukey post-hoc comparisons showed that the differences in word lengths were statistically significant only from 1886 to 1950. In the case of history titles, on the other hand, pairwise comparisons for all time periods, with the exception of 1886 vs. 1900, were statistically significant. The upward trend involving lexical diversity was similarly clear. Pairwise comparisons of the GI scores for all time periods were statistically significant, with the exception of 2000 vs. 2024 (science) and 1886 vs. 1900 (history).

Table 2. Multiple comparisons of the word length and GI scores across time periods in science and history titles (MD = mean difference; SE = standard error)

Period		Word length				Guiraud's Index			
		MD	SE	*p* value	Cohen's *d*	MD	SE	*p* value	Cohen's *d*
Science									
1886	1900	2.25	0.36	<.001**	0.62	0.27	0.05	.002**	0.53
	1950	3.86	0.36	<.001**	1.07	0.65	0.05	<.001**	1.30
	2000	4.92	0.36	<.001**	1.36	0.93	0.05	<.001**	1.86
	2024	4.76	0.36	<.001**	1.32	0.94	0.05	<.001**	1.89

1900	1950	1.61	0.36	.015*	0.45	0.38	0.05	<.001**	0.77
	2000	2.67	0.36	<.001**	0.74	0.66	0.05	<.001**	1.33
	2024	2.51	0.36	<.001**	0.69	0.67	0.05	<.001**	1.36
1950	2000	1.06	0.36			0.28	0.05	<.001**	0.56
	2024	0.90	0.36			0.29	0.05	<.001**	0.58
2000	2024	0.16	0.36			0.01	0.05		
History									
1886	1900	0.53	0.35			0.10	0.05		
	1950	2.27	0.35	<.001**	0.66	0.35	0.05	<.001**	0.76
	2000	5.87	0.35	<.001**	1.70	0.84	0.05	<.001**	1.82
	2024	7.46	0.35	<.001**	2.16	1.11	0.05	<.001**	2.41
1900	1950	1.74	0.35	.004**	0.50	0.25	0.05	.001**	0.54
	2000	5.34	0.35	<.001**	1.55	0.74	0.05	<.001**	1.60
	2024	6.93	0.35	<.001**	2.01	1.01	0.05	<.001**	2.20
1950	2000	3.60	0.35	<.001**	1.04	0.49	0.05	<.001**	1.06
	2024	5.19	0.35	<.001**	1.50	0.76	0.05	<.001**	1.65
2000	2024	1.59	0.35	.011*	0.46	0.27	0.05	<.001**	0.59

How science and history titles compared with each other is summarized in Table 3. The word lengths of science titles were higher than those of history titles in the earlier time periods (1900, 1950) but appeared to stabilize at about 10 words per title from 1950 onwards. The word lengths of history titles, on the other hand, increased steadily from 1900, surpassing those of science titles in 2000 and 2024. This same trend was observed for the GI scores (Figure 3), although the extent is less marked. As compared to history titles, science titles were lexically richer in 1900 and 1950, but less so in 2024.

Table 3. Descriptive statistics of the word length and GI scores in science and history titles (SD = standard deviation)

Period	Science		History		*t*(198)	*p* value	Cohen's *d*
	Mean	SD	Mean	SD			
Word length							
1886	5.56	2.97	5.66	2.68	0.25		
1900	7.81	4.52	6.19	2.36	3.18	.002**	0.45
1950	9.42	4.72	7.93	2.94	2.68	.008**	0.38
2000	10.48	2.80	11.53	4.41	2.01	.045*	0.28
2024	10.32	2.44	13.12	4.34	5.63	<.001**	0.80
GI							
1886	2.20	0.50	2.23	0.48	0.40		
1900	2.47	0.53	2.33	0.39	2.07	.039*	0.29
1950	2.85	0.66	2.58	0.44	3.41	<.001**	0.48
2000	3.13	0.38	3.07	0.52	0.95		
2024	3.14	0.36	3.34	0.47	3.38	<.001**	0.48

Taken together, both disciplines displayed a general increase in mean word lengths and GI scores over the years, although these appear to have plateaued in science titles in the later time periods. The rising trend in history titles since 1900, on the other hand, was an uninterrupted one. In particular, the higher mean GI score for history titles in 2024 implies not merely greater lexical diversity, but possibly greater complexity. This is because this diversity could be manifested in terms of new or embedded terms, adding to a higher level of information density (Zhou et al. 2023). As Baese-Berk et al. (2021) note (emphasis mine), "samples with greater diversity may also include less repetition, more switches among topics, and use of multiple lexical items to refer to the same concept. Each of these factors could make it *more* challenging for a listener to understand what is being said" (ibid. 3).

The preliminary indications at this stage, then, suggest a move toward complexity in history titles, but away from complexity in science titles. Whether this is indeed borne out when considering embedded elements and other indicators is discussed next.

5.2 Postmodifiers

Since research article titles convey key details about the topic of investigation, NP titles were dominantly used in the corpus. These ranged from the simpler, sometimes single-word, NPs in 1886 (6–7) to the more complex constructions in more recent years (8).

(6) Longevity *(Science, 1886)*

(7) Mather and Randolph *(History, 1886)*

(8) Massively parallel characterization [of regulatory elements [in the developing human cortex]] *(Science, 2024)*

As illustrated in (8) above, the longer constructions often contained embedded elements, the vast majority of which were PPs. These postmodifiers were found in about four in 10 of all NPs. In general, the use of PP postmodifiers was roughly the same between science and history titles, with the exception of 1950 and 2024, when more PP postmodifiers were found in science titles (Table 4). With the marked decrease in the use of PP postmodifiers in science titles between 2000 and 2024 (z=2.66, p=.007**), though, indications suggest a gradual stabilization in their use to about 35% in both disciplines.

Table 4. Proportions of NPs with PP postmodifiers to all NPs

Period	Science			History			z stat	p value
	NP all	NP[PP]	%	NP all	NP[PP]	%		
1880	224	96	42.86	228	88	38.60	0.92	
1900	274	125	45.62	245	116	47.35	0.39	
1950	332	177	53.31	312	121	38.78	3.70	<.001**
2000	341	167	48.97	418	163	39.00	2.76	
2024	301	116	38.54	491	170	34.62	1.11	.006**

By contrast, clausal postmodifiers, as exemplified in (9–10) below, were exceedingly rare in the corpus, appearing in just 44 science NPs (2.99%) and 12 history NPs (0.71%).

(9) A three-dimensionally architected electronic skin [[mimicking human mechanosensation]] *(Science, 2024)*

(10) The term ealdorman [in the translations [[prepared at the time [of King Alfred]]] *(History, 1953)*

The avoidance of clausal postmodifiers is due perhaps to the brevity constraint of titles, which could also account for why, when clauses need to be used, non-finite versions were preferred. Finite clausal postmodifiers, which tend to be lengthier, were found in only five science titles (11); none was found in history titles.

(11) A subclass [of Ras proteins [[that regulate the degradation [of IκB]]]] *(Science, 2000)*

The load of embedding thus appears to be borne largely by PPs, not clauses. On the face of it, this preference for word-phrase postmodifiers (over clausal ones) implies a move toward simplicity. A more careful consideration, though, is needed. This is because PPs are *compressed* structures. The Hallidayan framework, in fact, regards PPs as "a contraction of a clause" (Halliday and Matthiessen 2014: 362–363), arguing that the preposition can be viewed as a minor verb with the NP as its complement. This is particularly evident in the case of participial prepositions such as "regarding" and "including", among others, where the grammaticalization of such prepositions from participle verbs is well documented (e.g., Skiba 2021). As Halliday and Matthiessen (2014) explain, using various examples:

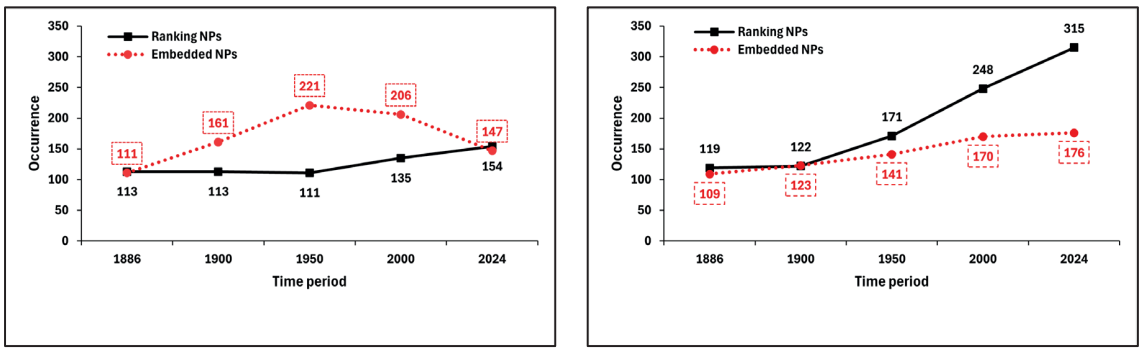
> "Thus the internal structure of *across the lake* is like that of *crossing the lake*, with a non-finite verb as Predicator. In some instances there is a non-finite verb that is more or less interchangeable with the preposition, e.g. *near/adjoining* (*the house*), *without/not wearing* (*a hat*), *about/concerning* (*the trial*). There is, in fact, an area of overlap between prepositional phrases and non-finite clauses; some instances can be interpreted as either, and some non-finite verb forms can be classified as prepositions, e.g. *regarding*, *considering*, *including*." (ibid. 424–425)

The fact that PP postmodifiers occurred more in science than history titles in general therefore points to a greater tendency toward structural contraction in the former. This feature of writing—as being compact and dense—has long been observed about science writing (e.g., Biber and Gray 2016), and we see traces of it here in science titles as well. Notwithstanding this, we should also be mindful of the falling trend in the use of PP postmodifiers in science titles from 1950/2000 to 2024, as this may indicate a gradual move *away* from the inherent complexity of such postmodifiers. For history titles, apart from the spike in 1900 (47.35%), the proportions of PP postmodifiers were relatively flat, ranging from 34–39%. To examine further the extent of complexity between the two disciplines, we turn next to embedded noun phrases.

5.3 Embedded noun phrases

For clarity, in this paper, NPs that operate at their rightful rank on Halliday's rank scale (Figure 1) are termed *ranking* NPs, to distinguish them from embedded NPs. As discussed earlier, embedded NPs occurred largely in PP postmodifiers and, to a much lesser extent,

clausal postmodifiers. The occurrences of ranking and embedded NPs in science and history titles are shown in Figure 4.



(a) NPs in science titles (b) NPs in history titles

Figure 4. Occurrences of ranking and embedded NPs in science and history titles

The use of both types of NPs in science and history titles was comparatively the same in 1886, a time when titles were much shorter and simpler. Beyond 1950, however, the paths between the two disciplines diverged (see Table 5). The use of ranking NPs in history titles outpaced those in science titles; by 2024, there were twice as many ranking NPs in history titles than in science titles (315 ~ 154). Indeed, the trend in history titles appears to be a conspicuous move away from the use of embedded NPs, relying instead on ranking NPs to carry the information load. As shown in Figure 4(b), the use of embedded NPs lagged behind ranking NPs, with the gap between them widening after 1950.

Table 5. Comparisons of ranking and embedded NPs in science and history titles

Period	Science	History	χ^2 stat	*p* value
Ranking NPs				
1886	113	119	0.16	
1900	113	122	0.35	
1950	111	171	12.77	<.001**
2000	135	248	33.34	<.001**
2024	154	315	55.27	<.001**
Embedded NPs				
1886	111	109	0.02	
1900	161	123	5.09	.024*
1950	221	141	17.68	<.001**
2000	206	170	3.45	
2024	147	176	2.60	

Where science titles are concerned, the bell shape of the embedded-NP line in Figure 4(a) is an interesting development, suggesting a possible trend reversal in future years. On the one hand, in three of the five time periods (1900, 1950, 2000), more embedded NPs were used in science titles than ranking NPs, implying that science titles were generally more complex than history titles, at least during the period 1900 to 1950. However, given

the progressively declining use of embedded NPs in science titles from 1950 onwards, it is unclear whether this trend will be reversed in the future. The latest statistics for 2024 indicate a near 1:1 match for ranking and embedded NPs in science titles.

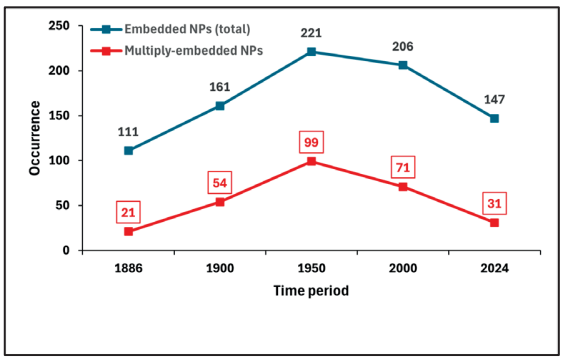
To find out whether the trends observed in Figure 4 for science and history titles differed when considering multiply-embedded NPs, a deeper level of analysis was conducted. Multiply-embedded NPs are those that are nested within another embedded NP. We have already seen examples of such multiply-embedded NPs in (5) and (8). Here is another to illustrate the complexity involved when processing such multiply-embedded elements:

(12) Working-class women and the buying and selling [of stolen goods [in urban communities [in the north west [of England] and Belfast, 1918–1960]]] *(History, 2022)*

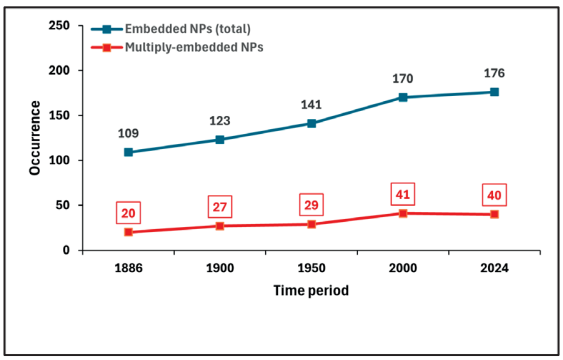
As segmented above, there are four multiply-embedded NPs in (12). For the reader, details regarding the transactions of goods need to be unpacked in a stepwise fashion. The goods in question are not ordinary, but stolen goods, and the transactions involve working-class women in urban communities rather than rural ones. These urban communities, more specifically, are restricted to only one geographical part of England, and the whole of Belfast. As can be seen, the processing of these varied pieces of information can be fairly demanding. Given the potential of these multiply-embedded NPs to add to the complexity of titles, the question that naturally arises is whether the use of such grammatical units mirror or depart from the shape of the dotted red lines in Figure 4.

As seen in Figure 5, the lines representing embedded and multiply-embedded NPs take roughly the same shape. In science titles, the proportion of multiply-embedded NPs to total embedded NPs was highest in the 1950 time period (44.80%), but this dipped by more than half in the 2024 time period (21.09%; z=5.86, p<.001**). Thus, the use of embedded NPs in science titles had not just fallen; the proportion of multiply-embedded NPs had decreased as well, a further indication of the trend away from complexity.

In history titles, by contrast, the use of multiply-embedded NPs was relatively stable across the time periods, occurring in about one in five embedded NPs. Given the increasing use of ranking NPs, as shown earlier in Figure 4(b), the overall trend for history titles also appears to be moving away from complexity.



(a) Embedded and multiply-embedded NPs in science titles



(b) Embedded and multiply-embedded NPs in history titles

Figure 5. Occurrences of embedded and multiply-embedded NPs in science and history titles.

Lastly, comparing the use of multiply-embedded NPs between disciplines revealed that the wide gap between science and history titles from 1900 to 2000 all but vanished in the 2024 time period (Table 6). This suggests a convergence of the two disciplines toward a simpler style of titling in recent years, to make titles more accessible in an increasingly competitive environment that places a high premium on scholarly publications (McGrail et al. 2006).

Table 6. Comparisons of multiply-embedded NPs in science and history titles

Period	Science	History	χ^2 stat	*p* value
1886	21	20	0.02	
1900	54	27	9.00	.003**
1950	99	29	38.28	<.001**
2000	71	41	8.04	.005**
2024	31	40	1.14	

5.4. Use of sentences

Apart from NPs, the analysis also revealed another indicator, the use of full-sentence declaratives, in support of the general trend toward non-complexity, particularly in science titles. The inclusion of sentences in titles is a recent development, and differences in how they are used in science and history titles suggest a disciplinary influence.

Full-sentence declaratives in both disciplines were found in only the 2024 time period, although it is acknowledged that some could have appeared in earlier titles that were not included in the corpus. Such declaratives occurred far more frequently in science than history titles. From nil entries in the earlier time periods, 41 science titles in 2024 were framed as declaratives. What is noteworthy about such titles is that in their expanded form, embedding was somehow minimized, as seen in (13–14).

(13) Vitamin D regulates microbiome-dependent cancer immunity *(Science, 2024)*

(14) Large-scale chemoproteomics expedites ligand discovery and predicts ligand behavior [in cells] *(Science, 2024)*

For instance, there is no embedding in (13), and in (14), only one of the four NPs (i.e., "cells") is embedded. Although declaratives are generally lengthier than word phrases, and longer segments do carry the potential to contain more embedded elements, the need to keep titles brief appears to have curtailed the occurrence of these elements. In science titles, declaratives also make the information much clearer, as compared to the more compact structure of word phrases. In (13–14), the declaratives make explicit the outcomes of the respective studies, which can be highly useful to scholars scanning titles to make quick decisions about whether to read the papers. This is very different in the case of phrasal titles (15), which generally mention only broader details.

(15) Single-cell genomics and regulatory networks [for 388 human brains] *(Science, 2024)*

In history titles, full sentences were used infrequently, appearing in only five titles in the 2024 time period. Like the sentences used in science titles, they generally did not contain embedded elements. However, unlike science titles, the history sentences functioned differently. All five sentences appeared as quoted texts in the initial segments of compound titles (16–17). There was also some variability—imperatives, not just declaratives, were used, as in (17).

(16) 'The true physicians here are the padres': British Christian army chaplains and the liberation [of Bergen-Belsen] *(History, 2023)*

(17) 'Take power—vote Liberal': Jeremy Thorpe, the 1974 Liberal revival, and the politics [of 1970s Britain] *(History, 2022)*

These quoted elements, what Chen and Liu (2023) refer to as the rhetorical part of titles, are "informationally optional in that knowledge about the subject […] can be retrieved from the second part even if the rhetorical part is deleted" (ibid. 3). The sentences, in other words, provide merely a teaser to contextualize the study, mirroring the research approach that is characteristic of this discipline (Bentrovato 2017). For this reason, perhaps, a compound format is needed for such titles. The quoted segments cannot stand alone; they need to be elaborated in the second part, all of which take the familiar form of NPs. Hence, unlike the sentences in science titles, the key information load is still largely borne by NPs in history titles.

6 General discussion

Collectively, the results suggest a move away from complexity in both science and history titles, notwithstanding the increase in word length and GI scores in both disciplines. This is seen not only in the use of direct markers of complexity—i.e., postmodifiers, (multiply-) embedded NPs—but also in the recent predisposition to incorporate sentences into titles.

This may appear to be an unexpected result, given the usual complaints that scholarly writing is difficult to read. Biber and Gray (2016), for instance, aptly summarize the somewhat negative view of academic writing in this way: "a common stereotype of academic prose is that it is deliberately complex, and more concerned with impressing readers than communicating ideas—all making it needlessly difficult to understand" (ibid. 1). It is therefore tempting to assume that titles of research articles would be similarly complex. Also, since titles are brief texts, the packing of information into a few words may inadvertently result in a complex construction. This may be true, but the current corpus has shown that we may need to revisit this assumption, and the larger stereotype of academic writing. I raise two points in this regard.

First, what we understand by complexity may be influenced more by the *lexical sophistication* of the text that we read, rather than its grammatical complexity. The use of technical terms and unfamiliar phrases can make it difficult for a non-specialist to fully understand the text. In the case of titles, which are much shorter, this can seem daunting. Consider, for instance, the examples in (18–19), both taken from the 2024 time period.

(18) The odd-number cyclo[13]carbon and its dimer, cyclo[26]carbon *(Science, 2024)*

(19) Chartist studies and Malcolm Chase: A re-appreciation *(History, 2023)*

If one has no prior knowledge of what cyclocarbons are, what the numbers in square brackets refer to, or what "dimer" means, the title in (18) can appear complex. Indeed, science writing, characterized by its use of specialized vocabulary and acronyms, can appear impenetrable to many people (Barnett and Doubleday 2020; Plavén-Sigray et al. 2017). This is true of writing in the humanities as well. If the reader is unfamiliar with chartism or Malcolm Chase, (19) can also seem inaccessible. Such issues, though, are lexical in nature, not grammatical. In fact, both (18) and (19) do not contain any embedded elements; in the context of this study, they are considered grammatically simple.

As alluded to in Section 2, the use of specialized vocabulary and technical terms is unavoidable. They are needed to prevent ambiguity or vagueness—a particular chemical compound (cyclocarbon) or a nineteenth-century working-class movement (chartism) can only be known by a certain name. The use of embedded elements, by contrast, is a matter of choice. The tentative indications from this study are that scientists and historians are choosing to avoid embedding in a seeming effort to make titles less complex.

Is this reflective of a larger trend in academic writing? This leads me to the second point. Several studies investigating grammatical aspects of scholarly articles have observed a falling trend in the use of the passive voice (Banks 2017; Leong 2021a; Seoane 2006), and a simpler clause structure in science writing as compared to humanities writing (Leong 2021b). Where embedding is concerned, a recent study found fewer instances of embedded clauses in biology articles as compared to classics articles (Leong 2023). These studies suggest that science writing, on the whole, has become grammatically less complex, motivated perhaps by the larger goal to make it more accessible to a wider readership, particularly in the present era of multidisciplinary studies. Humanities writing, by all accounts, appears to be more grammatically complex than science writing, but as related work on the former is much less extensive as compared to the latter, further research is needed for a clearer characterization of this genre.

Hence, the trend toward grammatical non-complexity in titles, at least where embedding is concerned, thus seems to be aligned with the larger trend in science writing. Whether this is also reflective of the trend in history writing is uncertain at this stage.

7 Conclusion

This study sought to investigate the extent of grammatical embedding in science and history research article titles sampled from two journals. It was found that the word lengths of science titles increased from 1886 to 1950, but plateaued at about 10 words per title thereafter. By comparison, the word lengths of history titles saw a steady increase from 1900 to 2024; the final mean length in the 2024 time period was 13.12 words. The lexical diversity of science and history titles also increased over the years, although the GI scores for science titles remained steady in 2000 and 2024 (3.13~3.14). For history titles, the growth was a progressive one from 1900 (2.47) to 2024 (3.43).

The majority of titles in both disciplines were found to be phrasal elements, with PPs as the main postmodifiers. By contrast, full-sentence titles were confined to only the 2024 time period, and involved fewer instances of embedding. PP postmodifiers were found in about 40% of all NPs across the time periods. Between 2000 and 2024, the proportions of PP postmodifiers to all NPs decreased markedly in science titles; by comparison, the proportions in history titles were generally even. In the 2024 time period, the proportions in both disciplines hovered at about 35%.

Where the use of embedded NPs was concerned, the trend line for science titles was bell-shaped, suggesting a possible trend reversal toward using fewer embedded NPs in the future. By 2024, the numbers of ranking NPs and embedded NPs in science titles were about equal (154 ranking NPs to 147 embedded NPs). In history titles, a gradual increase in the use of embedded NPs was observed, but these lagged behind ranking NPs, with the gap widening after the 1950 time period. By 2024, the gap was at its widest (315 ranking NPs to 176 embedded NPs). The use of multiply-embedded NPs in science titles also followed a bell shape. The proportion of multiply-embedded NPs to total embedded NPs from 1950 to 2024 fell by more than half in science titles (44.80% ~ 21.09%). In history titles, the proportions across time periods were relatively stable at about 20%.

The diachronic angle adopted in this study allows future trends to be inferred. It also allows us to better evaluate the various pieces of advice offered on titling. All too common is the tip for titles to be clear and concise. For instance, "the title should reflect, as accurately, completely, and concisely as possible, the central message of your study" (Cook and Bordage 2016: 1101), and "[t]itles that are accurate and descriptive but concise are important" (Schaak 2022: 250). Such tips do have merit, but they lack detail. Further, making titles concise and yet informative is challenging since these two attributes generally move in opposite directions. The major findings of this study provide some guidance from a grammatical perspective. Conciseness basically means keeping titles to the current norm of between 10 and 15 words. NP titles are still preferred, but the information load is increasingly borne by ranking NPs, rather than embedded ones. Even in sentence titles, which have recently become popular in science, embedding is minimized.

Apart from inferences regarding future trends, investigating titles this way also helps to heighten awareness of a less-obvious feature of titles. Grammatical embedding is not always as obvious to the eye as the use of specialized terms, which doubtless give disciplines like science and history a sense of complexity. The appropriate use of embedded elements can go some way to make the writing less complex; grammar also has a role to play.

The limitations of this study are that only two journals were involved. Other journals may have specific requirements, which could affect the results to some extent (Kerans et al. 2020). Given the growing international, diverse range of authors (Wang et al. 2024), the results could also be affected by the varying language abilities of the authors, as not all of them might be native users of English. A broader, comparative look involving more journals, disciplines, and authors are therefore crucially needed. Work involving interdisciplinary research articles (e.g., in medical humanities) would also be interesting to explore their titling norms. For instance, would the norms lean more toward those of the sciences or the humanities?

Lastly, some consideration also needs to be made about how titles are perceived by the people who really matter, the readers. It is a rather curious state of affairs that investigative work on titles, including this present one, does not always include the views of readers. A notable exception is the survey by Hallock and Bennett (2021) involving psychology undergraduates. Among other things, they found that their respondents preferred longer titles, and titles framed as questions. Such reader-based studies, though, are a rarity. More work is certainly needed to find out what readers, both specialists and non-specialists, prefer about the form and content of titles, and what they require from them. This will help writers to craft titles that truly meet the needs of their readers and, hopefully, lead them to read the articles.

References

Baese-Berk, Melissa M. & Drake, Shiloh & Foster, Kurtis & Lee, Dae-Yong & Staggs, Cecelia & Wright, Jonathan M. 2021. Lexical Diversity, Lexical Sophistication, and Predictability for Speech in Multiple Listening Conditions. *Frontiers in Psychology* 12. 661415. https://doi.org/10.3389/fpsyg.2021.661415

Banks, David. 2017. The extent to which the passive voice is used in the scientific journal article, 1985–2015. *Functional Linguistics* 4(1). 12. https://doi.org/10.1186/s40554-017-0045-5

Barnett, Adrian & Doubleday, Zoe. 2020. The growth of acronyms in the scientific literature. *eLIFE* 9. e60080. https://doi.org/10.7554/eLife.60080

Bentrovato, Denise. 2017. History textbook writing in post-conflict societies: From battlefield to site and means of conflict transformation. In Psaltis, Charis & Carretero, Mario & Čehajić-Clancy, Sabina (eds.), *History education and conflict transformation: Social psychological theories, history teaching and reconciliation*, 37–76. London: Palgrave Macmillan. https://doi.org/10.1007/978-3-319-54681-0_2

Biber, Douglas & Gray, Bethany. 2016. *Grammatical complexity in academic English: Linguistic change in writing*. Cambridge: Cambridge University Press.

Biber, Douglas & Gray, Bethany & Staples, Shelley & Egbert, Jesse. 2022. *The register-functional approach to grammatical complexity: Theoretical foundation, descriptive research findings, application*. New York: Routledge.

Carter, Ronald & McCarthy, Michael. 2006. *Cambridge grammar of English*. Cambridge: Cambridge University Press.

Chen, Xinren & Liu, Hao. 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

Cook, David A. & Bordage, Georges. 2016. Twelve tips on writing abstracts and titles: How to get people to use and cite your work. *Medical Teacher* 38(11). 1100–1104. https://doi.org/10.1080/0142159X.2016.1181732

Cook, James M. & Plourde, Dawn. 2016. Do scholars follow Betteridge's law? The use of questions in journal article titles. *Scientometrics* 108. 1119–1128. https://doi.org/10.1007/s11192-016-2030-2

Courtney, Michael & Courtney, Amy. 2008. Epistemological distinctions between science and history. *arXiv*: 0803.4245. https://doi.org/10.48550/arXiv.0803.4245

Covington, Michael A. & McFall, Joe D. 2010. Cutting the Gordian knot: The moving-average type–token ratio (MATTR). *Journal of Quantitative Linguistics* 17(2). 94–100. https://doi.org/10.1080/09296171003643098

Daller, Helmut & Xue, Huijuan. 2010. Lexical richness and the oral proficiency of Chinese EFL students: A comparison of different measures. In Daller, Helmut & Milton, James & Treffers-Daller, Jeanine (eds.), *Modelling and assessing vocabulary knowledge*, 150–164. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511667268.011

Denecke, Wiebke. 2021. Comparative global humanities now. *Journal of Word Literature* 6(4). 479–508. https://doi.org/10.1163/24056480-20210004

Diao, Junli. 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

Grech, Victor. 2019. Writing a scientific paper (WASP): An overview of differences in styles between the sciences and the humanities. *Early Human Development* 129. 96–99. https://doi.org/10.1016/j.earlhumdev.2018.12.008

Guiraud, Pierre. 1960. *Problèmes et méthodes de la statistique linguistique*. Paris: Presses Universitaires de France.

Halliday, Michael A. K. & Matthiessen, Christian M. I. M. 2014. *Halliday's introduction to functional grammar*. London: Routledge.

Hallock, Robert M. & Bennett, Tara N. 2021. I'll read that!: What title elements attract readers to an article? *Teaching of Psychology* 48(1). 26–31. https://doi.org/10.1177/0098628320959948

Hao, Jialiang. 2024. Titles in research articles and doctoral dissertations: Cross-disciplinary and cross-generic perspectives. *Scientometrics* 129. 2285–2307. https://doi.org/10.1007/s11192-024-04941-4

Heard, Stephen B. & Cull, Chloe A. & White, Easton R. 2023. If this title is funny, will you cite me? Citation impacts of humour and other features of article titles in ecology and evolution. *Facets* 8. 1–15. https://doi.org/10.1139/facets-2022-0079

Heinz, Ozge & Walker, Gilbert. 2023. How to write an insightful article I: Impactful trio—title, abstract and TOC graphic. *Langmuir* 39(6). 2089−2091. https://doi.org/10.1021/acs.langmuir.3c00191
Hess, Carla W. & Sefton, Karen M. & Landry, Richard G. 1986. Sample size and type-token ratios for oral language of preschool children. *Journal of Speech, Language, and Hearing Research* 29(1). 129–134. https://doi.org/10.1044/jshr.2901.129
Hyland, Ken & Zou, Joanna Hang. 2022. Titles in research articles. *Journal of English for Academic Purposes* 56. 101094. https://doi.org/10.1016/j.jeap.2022.101094
Jiang, Feng Kevin & Hyland, Ken. 2023. Titles in research articles: Changes across time and discipline. *Learned Publishing* 36. 239–248. https://doi.org/10.1002/leap.1498
Jiang, Gianna Kexin & Jiang, Yajun. 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
Juergensmeyer, Mark & Sassen, Saskia & Steger, Manfred B. (eds.). 2019. *Oxford handbook of global studies*. Oxford: Oxford University Press.
Kenny, Amanda. 2023. The importance of article titles. *Nurse Education Today* 131. 105983. https://doi.org/10.1016/j.nedt.2023.105983
Kerans, Mary Ellen & Marshall, Jane & Murray, Anne & Sabaté, Sergi. 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
Leong, Ping Alvin. 2021a. The passive voice in scholarly writing: A diachronic look at science and history. *Finnish Journal of Linguistics* 34. 77–102. https://journal.fi/finjol/article/view/103168
Leong, Ping Alvin. 2021b. Writing in the sciences and humanities: A clause-complex perspective. *Word* 67(2). 137–158. https://doi.org/10.1080/00437956.2021.1909866
Leong, Ping Alvin. 2023. Discipline-specific writing and embedded clauses: The case of cell biology and classics. *Word* 69(4). 362–379. https://doi.org/10.1080/00437956.2023.2270875
Li, Zhijun & Xu, Jinfen. 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
Liimatta, Aatu. 2024. Text length and short texts: An overview of the problem. In Kaunisto, Mark & Schilk, Marco (eds.), *Challenges in corpus linguistics: Rethinking corpus compilation and analysis*, 106–125. Amsterdam: John Benjamins.
McGrail, Matthew R. & Rickard, Claire M. & Jones, Rebecca. 2006. Publish or perish: A systematic review of interventions to increase academic publication rates. *Higher Education Research and Development* 25(1). 19–35. https://doi.org/10.1080/07294360500453053
Moore, Tim. 2020. Academic clickbait: The arcane art of research article titling. *Australian Universities' Review* 62(1). 54–56. https://files.eric.ed.gov/fulltext/EJ1245084.pdf
Pearson, William 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, William 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
Plavén-Sigray, Pontus & Matheson, Granville James & Schiffler, Björn Christian & Thompson, William Hedley. 2017. Research: The readability of scientific texts is decreasing over time. *eLIFE* 6. e27725. https://doi.org/10.7554/eLife.27725
Sagi, Itay & Yechiam, Eldad. 2008. Amusing titles in scientific journals and article citation. *Journal of Information Science* 34(5). 680–687. https://doi.org/10.1177/0165551507086261
Schaak, Raymond E. 2022. Increase your impact: Writing tips to reach a broader audience. *ACS Nanoscience* 2. 250–252. https://doi.org/10.1021/acsnanoscienceau.2c00030
Scimago. 2023. *Scimago journal and country rank*. http://www.scimagojr.com/
Seoane, Elena. 2006. Changing styles: On the recent evolution of scientific British and American English. In Dalton-Puffer, Christiane & Kastovsky, Dieter & Ritt, Nikolaus & Schendl, Herbert (eds.), *Syntax, style and grammatical norms: English from 1500–2000*, 191–211. Bern: Peter Lang.
Skiba, Michael. 2021. *Participial prepositions and conjunctions in the history of English*. Munich: utzverlag GmbH.
Wang, Jue & Frietsch, Rainer & Neuhäusler, Peter & Hooi, Rosalie. 2024. International collaboration leading to high citations: Global impact or home country effect? *Journal of Informetrics* 18(4). 101565. https://doi.org/10.1016/j.joi.2024.101565

Xiang, Xuechun & Li, Jing. 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
Xie, Shaoliang. 2020. English research article titles: Cultural and disciplinary perspectives. *Sage Open* 10(2). 1–12. https://doi.org/10.1177/2158244020933614
Zaiontz, Charles. 2022. *Real statistics using Excel*. http://www.real-statistics.com/
Zhou, Xinye & Gao, Yuan & Lu, Xiaofei. 2023. Lexical complexity changes in 100 years' academic writing: Evidence from Nature Biology Letters. *Journal of English for Academic Purposes* 64. 101262. https://doi.org/10.1016/j.jeap.2023.101262

Contact information

Alvin Ping Leong
Nanyang Technological University
Language & Communication Centre
e-mail: alvin.leong@ntu.edu.sg