Structural Patterns of Research Article Titles:
An Exploratory Study in Five Branches of Linguistics

Naruthai Surapongraktrakool
Department of Foreign Languages, Faculty of Humanities
Kasetsart University
Email: naruthai.s@ku.th

Abstract

This study explores the constructions of research article titles in five branches of linguistics by examining two key aspects: title length and syntactic structures. Data from 15 specialized journals covering language teaching, computational linguistics, psycholinguistics, sociolinguistics, and stylistics were analyzed. The findings reveal that, on average, computational linguistics titles are the shortest at nine words while titles in other branches tend to be 12–13 words long. Total of 12 types of title structures were identified, with single-unit titles, especially single phrases, being most common in computational linguistics and multi-unit titles, specifically phrase/phrase structures, being more frequent in other branches. Independent clauses are found most often in psycholinguistics titles. This study provides useful guidelines for researchers seeking to craft titles in these main branches of linguistics and encourages instructors in English for Academic Purposes to incorporate the structural patterns of titles into their lessons.

Titles may be the smallest and shortest part among all components of a research article, but they bear a crucial responsibility to inform readers about the study as well as to create a positive impression on potential audiences. Therefore, Swales (1990) referred to them as "serious stuff" (p. 224). The title is typically the first part readers encounter, whether through traditional library searches or the modern technology of online searches. It serves as a crucial initial point of reference for them to determine whether the article aligns with their interests or needs (Hyland & Zou, 2022), which may subsequently lead them to the paper's abstract and its content. Creating article titles is challenging (Kumar, 2013; Moore, 2020), particularly for novice researchers and students

who must carefully select words and structural patterns to convey the essence of their work, capture readers' interest, and invite them to read further.

The grammatical structure and length of a title is important. Studies have shown that the title's syntactic structure plays a vital role in attracting readers, with interrogative titles (Jamali & Nikzad, 2011) and shorter titles (Jamali & Nikzad, 2011; Paiva et al., 2012) proving more appealing to diverse audiences. Additionally, research has linked the title structure consisting of two parts with a balanced length, separated by a colon, to increased citation rates (Nair & Gibbert, 2016), highlighting the impact of title choices on the recognition of research.

When crafting research titles, key factors to be considered include word choice, title structure, and overall length. Interestingly, Rath (2011) found that most manuals of style tended to focus on content and keywords in the title rather than providing suggestions about title structure or title length. While writers may have a solid grasp of the essence and critical keywords of their work, effectively assembling them into a title can be a distinct challenge. Questions about length, structural patterns, and overall composition may leave them uncertain, as there is often no definitive right or wrong approach. This uncertainty can lead to difficulty in crafting a title that not only sells but also provides the precise information that readers are seeking. Choices of construction and length play a pivotal role in how information is presented and received by the audience. Experienced researchers might not perceive this as an issue, given their exposure to numerous titles, which grants them a broad perspective on title structures and common practices. However, new members of the academic community may require more support at the beginning of their career. Thus, a general idea of prevailing title length and constructions can serve as a helpful guide for struggling students and novice researchers.

Within the field of linguistics, research has reported various findings in terms of title length and common patterns. Different findings among studies on the same discipline may not be surprising, though, since trends in research article titles have undergone a dramatic change over time (Jiang & Hyland, 2023; Xiang & Li, 2020). This, however, seems an inadequate explanation for previous research conclusions in linguistics title constructions, since the findings from the same time period exhibited divergent results. An example of this was reflected in Pearson's (2020) study, which specifically examined titles within the subdiscipline of English as a Second Language (ESL) writing. Pearson's (2020) report on common title structures, along with those of others like Hyland and Zou (2022), and Xiang and Li (2020) that used data from the same time frame, revealed contrasting outcomes. This discrepancy may be attributed to the unique subdisciplinary characteristics of the data.

Linguistics is continuously expanding its scope, covering a number of main branches, for example, applied linguistics, computational linguistics,

psycholinguistics, sociolinguistics, and stylistics, which emerge from the core areas (phonetics, phonology, syntax, semantics, and pragmatics) that serve as the foundational knowledge for these branches (Finch, 2008). The integration of various fields of knowledge creates a distinct academic environment with unique characteristics, and "sometimes a sub-discipline forms its own conventions" (Nagano, 2015, p. 134). Therefore, to thoroughly explore the titles in the field of linguistics, the selection of journals for the study is a crucial consideration. It is important to account for the variation in subdiscipline natures to ensure a valid empirical generalization of title characteristics that are possibly common in the field.

Given that "there are marked disciplinary preferences when it comes to titles" (Swales & Feak, 2012, p. 379), and each discipline has its own unique characteristics that can impact title patterns (Nair & Gibbert, 2016), the interdisciplinary features of various branches of linguistics may also affect title creation, resulting in structural differences among the branches. However, variations in title patterns across subdisciplines of linguistics have not been examined in the existing studies, and Pearson's (2020) work centered on titles within the subdiscipline of ESL, which was not a comparative study of different branches. Therefore, this exploratory study seeks to provide initial insights into the patterns and characteristics of research article titles in the field of linguistics and investigate subdisciplinary differences, serving as a starting point for more extensive investigations in the future. By employing five main branches of linguistics proposed by Finch (2008), who clearly defined scopes that effectively presented their interdisciplinary aspects, this paper extracted 750 titles from 15 specialized journals publishing articles in applied linguistics, computational linguistics, psycholinguistics, sociolinguistics, and stylistics. The study focuses on two key features—length and syntactic structure—exploring (a) the length of titles, (b) the structural organization of titles as single-unit or multi-unit, (c) the syntactic structure of both single-unit and multi-unit titles, and (d) noticeable differences across branches. The findings of this study offer general guidelines, examples, and statistics on title construction variation for writing titles in different subdisciplines of linguistics. The study serves as a supplement to style manuals, presenting a range of possibly preferred title lengths and alternative structures that go beyond the most frequently found constructions, to encourage titling creativity.

Research Article Titles

Since the 1970s, several studies have examined the structure and content of research article titles, including works by Diener (1984), Diodato (1982), Kuch (1978), and Peritz (1984). However, it was not until Swales (1990) highlighted the crucial role of titles in academic works that this topic began to receive greater attention. Over the past two decades, research has expanded to

include a range of publications investigating not only syntactic constructions and semantic content but also the relation to various factors, such as the number of authors (e.g., Hudson, 2016; Lewison & Hartley, 2005; Yitzhaki, 1994), article length (e.g., Yitzhaki, 2002), downloads and citation rates (e.g., Jamali & Nikzad, 2011; Nair & Gibbert, 2016), and the relationship between keywords in titles and citations (Yang, 2019). The research has examined titles in a variety of works, including textbooks, review papers, research articles, proceedings papers, theses, and dissertations. More recent studies (e.g., Hyland & Zou, 2022; Jiang & Hyland, 2023; Pearson, 2020; Xiang & Li, 2020; Xie, 2020) have concentrated on the titles of research articles themselves, indicating a growing interest in this area.

In the studies of research article titles, two important areas have received attention: title length and syntactic structure. Their relation to readers' interest was studied by Jamali and Nikzad (2011), and Paiva et al. (2012). Jamali and Nikzad (2011) reported that readers tended to download interrogative titles more than other types, while in terms of length, longer titles received slightly less attention than shorter ones. Paiva et al. (2012) agreed with these findings, reporting that articles with short titles had higher view counts. These studies emphasize the role of title length and syntactic structure in attracting readers.

Investigation Across Disciplines

Numerous authors have suggested that title length varies significantly across disciplines (Milojević, 2017; Nagano, 2015; Soler, 2007). Although some authors have found that titles in the sciences were longer than those in the social sciences (Nagano, 2015) and linguistics (Busch-Lauer, 2000; Haggan, 2004; Soler, 2007; Xie, 2020), others, such as Hyland and Zou (2022), reported opposite findings and argued that titles in social sciences and humanities, which included linguistics in their categorization, were longer than those in the sciences. Jiang and Hyland (2023) concluded that the length between 11 and 15 words was the most common for recent titles found in both soft and hard sciences due to the similar need to make the title more descriptive to attract readers' attention. Title lengths in linguistics varied from 7.98 (Soler, 2007) to 8.8 (Haggan, 2004), 9.08 (Busch-Lauer, 2000), 10.8 (Gesuato, 2008), 11.2 (Xiang & Li, 2020), 12.2 (Pearson, 2020), and 13.4 (Hyland & Zou, 2022) words per title, and titles seem to be increasing in length over time (Xiang & Li, 2020).

In terms of syntactic structures, nominal group titles (those having nouns as head words) were found to be more prevalent in the hard sciences (Méndez et al., 2014; Wang & Bai, 2007) and in both the hard and soft sciences (Soler, 2007). Also, more single titles (those consisting of one construction ending with one or no punctuation marks) were observed in the hard sciences, while compound titles (those comprised of two parts separated by a punctuation mark

such as a colon or question mark) were more frequent in the soft sciences (Busch-Lauer, 2000; Hyland & Zou, 2022; Jiang & Hyland, 2023; Nagano, 2015; Xie, 2020). In linguistics, different findings were reported on the prevalence of compound titles compared to single constructions. Although some studies, such as Gesuato (2008), Haggan (2004), and Soler (2007), reported that compound titles were less common than single constructions, Busch-Lauer (2000), Cheng et al. (2012), Hyland and Zou (2022), and Xiang and Li (2020) found the opposite. Moreover, Xiang and Li (2020) emphasized that longer titles have become more popular among linguistic researchers over time. Recent studies seem to agree with those findings, except Pearson (2020), who focused on titles in written feedback in ESL writing, a subdiscipline of linguistics. Pearson (2020) found that nominal structures had the highest frequency, followed by compound titles. This highlights the possibility of different findings in subdiscipline-level studies.

Considering the wide range of findings on titles in the field of linguistics and the various approaches used to present research data, novice researchers may find it difficult to apply them effectively as guidelines. For example, some frameworks have presented data that may have been intended for a different purpose and not as guidelines, such as total word counts of all titles in the corpus (e.g., Gesuato, 2008; Haggan, 2004; Soler, 2007; Xiang & Li, 2020; Xie, 2020). Also, mean values of title length have been presented in most studies, but a more extensive presentation of the range of possibly preferred title lengths would be more illuminating. Additionally, although some studies have provided broader categorizations of trends as a general or an alternative view (e.g., Busch-Lauer, 2000; Hyland & Zou, 2022), a more detailed categorization of all patterns found may be more useful for beginners, because it would not only suggest the most-preferred constructions but also clearly show alternative possibilities that are slightly different from the norms.

Additionally, previous studies have reported varied results on title length and syntactic structures. This may be attributed to the imbalanced representation of subdisciplines in the source journals selected by the studies. Although some studies focused on "highest rated" (Hyland & Zou, 2022, p. 3) journals with high impact factors (Soler, 2007; Xie, 2020) and international prestige (Cheng et al., 2012), others targeted journals that mainly publish research on language and not applied studies (Xiang & Li, 2020) or retrieved journals from "on-line queries" (Gesuato, 2008, p. 134) and specific academic holdings (Haggan, 2004). Consequently, the existing studies have yielded mixed findings regarding title structure in linguistics and have not examined variations in title patterns across subdisciplines. The only subdisciplinary study found in the literature was one centered on titles in written feedback, which was not a comparative study of different branches.

Previous research prompts further investigation into subdisciplines and their potential differences. Therefore, an in-depth study focusing on the subdisciplinary level is needed, and such a study could reveal evidence of the predominant forms of titles in these branches.

Main Branches of Linguistics

Linguistics, as the scientific study of language, intersects with various academic disciplines, blending insights from cognitive science, anthropology, psychology, computer science, and more. It not only deepens our understanding of language but also fosters a rich interdisciplinary exploration of human cognition, communication, and culture through its branches. Some main branches of linguistics proposed by Finch (2008) are as described below.

Applied linguistics involves the application of linguistic knowledge to real-life situations (Cook, 2003). Finch (2008), and Schmitt and Celce-Murcia (2020) included language teaching and learning in applied linguistics and considered it the most prominent area of this branch, as other areas such as second language acquisition (SLA) are applied to develop language teaching (Allan et al., 2010).

Computational linguistics aims to develop algorithms and computational models that can process and analyze human language. Experts from various fields of knowledge beyond linguistics, including computer science, mathematics, and statistics, must collaborate in order to create a model of natural language that can understand and interact with humans such as popular artificial intelligence (AI) systems like ChatGPT.

Psycholinguistics mainly studies the relationship between psychological processes and language use. In essence, it is the study of language and the mind (Aitchison, 2008; Finch, 2008). It examines the processes involved in how humans acquire, interpret, and produce language. As an interdisciplinary field of linguistics and psychology, it includes cognitive science and neuroscience (Fernández & Cairns, 2018).

Sociolinguistics explores the connections between language use in different social contexts and social factors, such as age, gender, identity, class, race, and region. Combining knowledge of sociology and linguistics, sociolinguistics seeks to explain language variety and the changes in language use that occur within a society.

Stylistics uses linguistic methodology and analysis to study style in both literary and nonliterary texts (Jeffries & McIntyre, 2010) to gain insight into how speakers select language use according to the situation, their relationship with other speakers, and the way they communicate. It can be applied to understand language use in various contexts, such as sociology, politics, and media studies, to describe the language use of powerful people (Finch, 2008).

The interdisciplinary nature of linguistics suggests that the branches may have distinct characteristics, potentially resulting in different frequencies of syntactic structures in the research article titles.

Syntactic Structure

Biber et al.'s *Longman Grammar of Spoken and Written English* (1999) categorized grammatical constructions found in English as phrases and clauses.

Phrase

Phrases can be single-word or multi-word structures. Each type of word (noun, verb, adjective, adverb, or preposition) serves as the head in a specific phrase category: *noun phrase* (NP), *verb phrase* (VP), *adjective phrase* (AP), *adverb phrase* (AdvP), and *prepositional phrase* (PP). The following are examples with the head word in bold type:

- a beautiful **flower**   (NP)
- have been **studying**   (VP)
- incredibly **delicious**   (AP)
- quite **suddenly**   (AdvP)
- **in** the park   (PP)

Clause

Clauses are organized and built primarily around a central component known as the "verb phrase," which plays a crucial role in expressing the action, state, or event that the clause is conveying. They are categorized into independent and dependent clauses.

Independent clauses are self-contained grammatical units which are "not part of any larger structure" (Biber et al., 1999, p. 202). They can contain sub-elements like embedded clauses or be connected to other independent clauses through coordination. Based on communicative purposes, there are four structural types:

- *Declarative clauses* make statements, e.g., "She is reading a book."
- *Interrogative clauses* ask questions, e.g., "Is she reading a book?"
- *Imperative clauses* issue commands or requests, e.g., "Come here."
- *Exclamative clauses* express exclamations, e.g., "What a beautiful day!"

Dependent clauses are those embedded within larger clauses, and the larger clauses they are part of are independent clauses referred to as main clauses. They can be categorized into finite or non-finite dependent clauses based on their verb phrase structure.

- *Finite dependent clauses* have specific verb forms indicating tense or modality such as nominal clauses, adverbial clauses, and relative clauses (in bold), e.g., "I believe [**that he is honest**]," "[**After the rain**

stopped], we went out," and "The book [**that she's reading**] is interesting."
- *Non-finite dependent clauses* do not have verb phrases marked for tense or modality and often lack an explicit subject and subordinator such as infinitive clauses, *ing*-clauses, and *ed*-clauses (in bold), e.g., "She traveled [**to explore different cultures**]," "She enjoys [**swimming in the ocean**]," and "The document, [**signed by the CEO**], is now official."

In some cases, such as a chapter title in an academic text, dependent clauses can be used separately without being part of a larger clause; these are referred to as "unembedded dependent clauses," for example, "Renorming super-reflexive Banach spaces" (Biber et al., 1999, p. 223).

Thematic Roles

Thematic roles are labels used to identify the underlying relationships between verbs and the arguments (noun phrases or pronouns) that they take in a sentence. The following are major and minor thematic roles according to Timyam's *An Introduction to English Linguistics* (2010).
- *Agent*: the entity that performs an action
- *Theme*: the entity that undergoes the effect of an action; the entity to which a property is attributed
- *Source*: the starting point for a movement
- *Goal*: the end point for a movement
- *Location*: the place where an action occurs
- *Possessor*: the entity possessing an entity
- *Experiencer*: the entity experiencing a psychological state
- *Stimulus*: the entity triggering a psychological state
- *Benefactive*: the entity benefiting from an action
- *Instrument*: the entity used by the agent in bringing about an event
- *Causer*: the agent who brings about an event involving other participants
- *Force*: the inanimate agent
- *Comitative*: the entity accompanying some other entity
- *Direction*: a type of goal, with no implication that the end point is reached
- *Path*: the route along which movement occurs
- *Temporal*: expressions of time and duration
- *Purpose*: expressions of purpose and reason

In the sentence "She analyzed data with a computer program," the subject "she" serves as the agent, signifying the entity performing the action of

analysis. "Data" takes on the role of the theme, representing the entity that is the object of the analysis, undergoing the action. Finally, "a computer program" functions as the instrument, indicating the tool or means employed by the agent (she) to carry out the analysis of the data.

Semantic Domains of Verbs

Verbs were classified into seven semantic domains by Biber et al.'s *Longman Grammar of Spoken and Written English* (1999) as follows.

- *Activity verbs* usually refer to a volitional activity (e.g., *buy, open, work, go, try, use, make, take, get*). They are also sometimes used to express events that occur without the volition of an agent (e.g., *move, give*).
- *Communication verbs* involve communication activities, particularly verbs describing speech and writing (e.g., *ask, shout, describe, claim, offer, thank*).
- *Mental verbs* refer to mental states and activities. They do not involve physical action. Some convey volition; others do not. They express mental states or processes (e.g., *think, know*), emotions expressing attitudes or desires (e.g., *love, want*), perception (e.g., *see, taste*) and receipt of communication (e.g., *read, hear*). They describe mental activities that are relatively dynamic (e.g., *calculate, consider, decide, discover, examine, solve, read*) and some are more stative (e.g., *believe, remember, understand, feel, hate, prefer, suspect*)
- *Causative verbs* (e.g., *allow, cause, force, enable, require, let*) indicate that some person or thing helps to bring about a new state of affairs.
- *Occurrence verbs* report events that occur without an actor. Often the subjects of these verbs are affected by the event described by the verb (e.g., *become, change, develop, die, grow, happen, occur, increase*).
- *Existence/relationship verbs* report a state that exists between entities. Copular verbs (e.g., *be, seem, appear*) are some of the most common. Other verbs are not copular verbs, but report a particular state of existence (e.g., *exist, live, stay*) or a particular relationship between entities (e.g., *contain, include, involve, represent, has*).
- *Aspectual verbs* (e.g., *begin, continue, keep, start, stop, finish*) characterize the stage of progress of an event or activity.

Methodology
Framework

The data categorization framework was developed from the previous research conducted by Afful and Ankomah (2020), and Gesuato (2008). Their categorization highlights clear structural patterns and enhances accessibility, making it suitable for a guideline. Afful and Ankomah (2020), and Gesuato

(2008) both calculated title length by word count and presented average numbers of words. In terms of title structures, Gesuato (2008) categorized titles into single-unit and multi-unit structures based on the presence of punctuation marks (colons, semicolons, periods, question marks, or dashes) and then analyzed them further as NP, VP, AP, PP, clauses, and combinations of these phrases. Similarly, Afful and Ankomah (2020) identified three categories: single-unit, compound-unit, and complex-unit titles, separated by the number of punctuation marks, and identified the structures in detail as NP, PP, finite clauses, non-finite clauses, and sentences. The categorization framework, derived from these two established research studies, provides readers with a clear, user-friendly tool for structural guidance. By combining these sources, the framework simplifies complex structural patterns, distinguishes between single-unit and multiple-unit structures, and offers detailed combinations. This clarity is invaluable to readers seeking accessible title construction guidelines.

For the structural analysis, *Longman Grammar of Spoken and Written English* (1999) by Biber et al., as described under the "Syntactic Structure" section earlier in this article, was used as the grammatical framework to categorize the title structures, as it offers a contemporary and comprehensive view of English grammar with a straightforward approach. Since it is known for its reliance on large linguistic corpora, it can provide accurate and up-to-date insights into how language is used in various contexts, including titles.

From these frameworks, this study analyzed (a) the length of titles by counting the number of words, (b) the structural organization of titles as single-unit or multi-unit, (c) the syntactic structure of both single-unit and multi-unit titles, and (d) noticeable differences across branches.

Research Sample

This study explores 750 titles of research articles from 15 specialized journals publishing articles relevant to five main branches of linguistics categorized by Finch (2008): applied linguistics, computational linguistics, psycholinguistics, sociolinguistics, and stylistics. The area of language teaching was selected to represent the branch of applied linguistics, since Finch (2008) and Schmitt and Celce-Murcia (2020) included it in applied linguistics and considered it the most prominent area of this branch. Therefore, the term "language teaching" was used in place of "applied linguistics."

Data Collection

Data were collected from journals indexed in Scopus using the SCImago Journal Rankings (SJR) for 2021, as listed in Appendix. The search results were organized based on the SJR indicator, which "expresses the average number of weighted citations received in the selected year by the documents published in the selected journal in the three previous years" (SCImago, n.d.).

The journals were chosen according to the following conditions:
- They are ranked in the top 100 journals in the subject category of "Linguistics and Language" on the SCImago Journal Rankings (SJR) for 2021.
- Their area of interest and expertise corresponds with a specific branch of linguistics. The journals' aims and scopes were thoroughly reviewed. Those that cover a wide range of topics in linguistics were excluded.

The criteria for selecting titles from each journal were as follows:
- They are from articles published between 2017 and 2021.
- The articles must be pertinent to the field of linguistics. The keywords and abstracts were reviewed to confirm their relevance.
- They are research article titles. Other types of articles, such as editorials, commentaries, monographs, and book reviews, were excluded as they are not the focus of the study.

Steps in data collection were as follows:
1. Three journals were chosen for each branch from the top position.
2. From each selected journal, 50 titles were collected, starting from the most recent issues, totaling 150 titles per branch.
3. The article titles were obtained by accessing the online tables of contents on the selected journals' websites and manually inputting each title into an Excel file.
4. The data were organized into separate Excel sheets for each journal and labeled with its corresponding branch of linguistics. In cases of uncertainty about the word category in the title, the abstract was consulted to determine its meaning and appropriate categorization.
5. Since the number of published articles varied among journals selected for the same branch, in cases where a particular journal had fewer published articles within the selected time frame, additional titles were included from the other two journals to attain a total of 150 titles per branch. The same sample size across all branches ensure that the data collection approach is consistent, enhancing the reliability of the findings.

The details of the corpus can be found in Table 1.

Table 1
Corpus Details

Branches	Journals	Titles
Computational linguistics	3	150
Language teaching	3	150
Psycholinguistics	3	150
Sociolinguistics	3	150
Stylistics	3	150
Total	15	750

A collection of 750 research article titles was compiled, with 150 titles from three journals selected for each branch, to facilitate diversity and effective comparative analysis across the branches.

Validity and Reliability

The validity and reliability of the measurement procedures applied in this title analysis were examined. To establish validity, two experts in English linguistics were consulted, and their valuable feedback led to adjustments and refinements of the methodology. Additionally, an intercoder reliability assessment was conducted by two experts in English structure, who independently evaluated 20% of the randomly selected coded data on syntactic structures, thematic roles, and semantic domains of verbs. Agreement rates of 99.33% for syntactic structures, 100% for thematic roles, and 100% for semantic domains of verbs were attained. In the case of disagreement, constructive discussions were held to reach a final consensus.

Data Analysis

1. Title Length

For title length, an Excel formula was used to generate the number of words for each title by counting the words separated by a space, and the AVERAGE function was used to determine the average length of titles in the corpus. Hyphenated words, such as "data-to-text" and "task-oriented," as well as abbreviations such as "BERT" and "ERP," were each counted as one word. Additionally, infinitive verbs, such as "to understand" and "to inhale," were each manually counted as one word.

2. Structural Organization

In the first step of coding, the single-unit titles were separated from the multi-unit titles. A single-unit title was defined as a title that consisted of one syntactic construction ending with one or no punctuation marks. In contrast, a multi-unit title comprises two parts separated by a punctuation mark such as a colon or question mark.

3. Syntactic Structure

       Syntactic structures were manually coded according to Biber et al.'s *Longman Grammar of Spoken and Written English* (1999) as described under the "Syntactic Structure" section earlier in this article. Frequency counts and percentages were computed to analyze syntactic constructions, with the data being calculated by branch and type of pattern.

Table 2

Examples of Data Analysis

Title	Length (words)	Structural Organization	Syntactic Structure
Dependency-Based Syntax-Aware Word Representations	4	Single-unit	PH (NP)
How Speakers Continue with Talk After a Lapse in Conversation	10	Single-unit	DC (FC)
Gaze Direction Signals Response Preference in Conversation	7	Single-unit	IC (Declarative)
Waiting to Inhale: On Sniffing in Conversation	6	Multi-unit	DC/PH (NFC/PP)
How Effective Are Intentional Vocabulary-Learning Activities? A Meta-Analysis	8	Multi-unit	IC/PH (Interrogative/NP)

Note. PH = phrase, DC = dependent clause, IC = independent clause, NP = noun clause, PP = prepositional phrase, NFC = non-finite clause, FC = finite clause.

       Table 2 above illustrates how data was analyzed in terms of length, structural organization, and syntactic structure where the structure in parentheses represents sub-categories of the syntactic structure.

Findings

1. Title Length

       Although the length of titles varied across different branches, the average title length for four of the branches was quite similar (12–13 words), while computational linguistics had a distinctly lower value (9 words), as shown in Figure 1 below.

Figure 1
Title Length by Branch
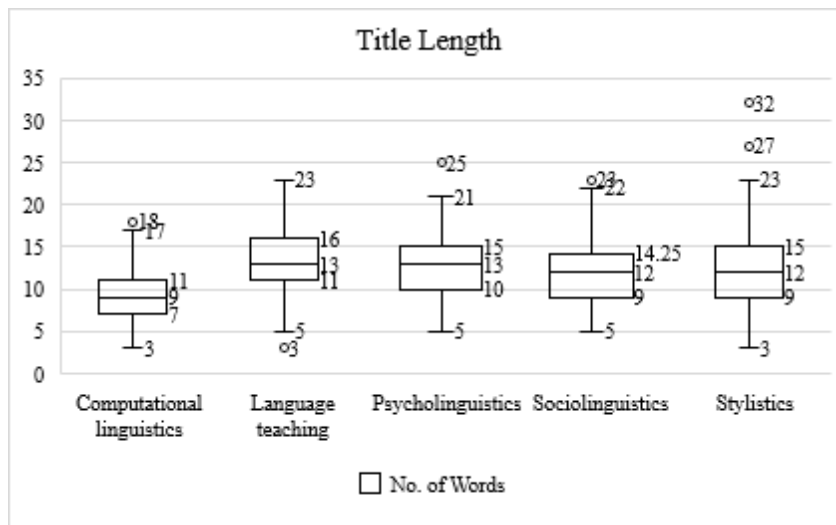
In Figure 1, the boxes contain the middle 50% of the data, and the median lines within them represent the central tendency of title length, which precisely matched the mean length in this study (i.e., 9, 12, and 13 words). The whiskers extend to the shortest and longest non-outlier titles (e.g., 5 and 21 words, respectively, in psycholinguistics). Finally, outliers, titles significantly longer or shorter than the others, are plotted as individual dots beyond the whiskers (e.g., 3 words in language teaching and 27 words in stylistics).

The box and whisker plot provides a clear representation of data density, showing where most of the data is located. The box indicates that the title lengths of half of the collected data fall within this range. A narrow box, such as that seen in computational linguistics, indicates that the data were more closely clustered. Fifty percent of the titles were in the range of 7 to 11 words in length, which suggests a high degree of consistency, or a low level of variability in overall title length within computational linguistics. The following are examples of titles with average lengths:

- Sketch-Driven Regular Expression Generation From Natural Language and Examples (9 words) (computational linguistics)
- On the Relevance and Accountability of Dialect: Conversation Analysis and Dialect Contact (12 words) (sociolinguistics)
- The Role of Language Teacher Metacognition and Executive Function in Exemplary Classroom Practice (13 words) (language teaching)

The similar trend in title length is also reflected in Table 3, which categorizes titles into single-unit and multi-unit structures.

Table 3

Average Title Length of Single-Unit and Multi-Unit Titles by Branch

Branch	Single-Unit (words)	Multi-Unit (words)
Computational linguistics	8.3	11.1
Language teaching	11.7	14.6
Psycholinguistics	11.1	14.5
Sociolinguistics	9.0	13.2
Stylistics	9.4	13.6

Table 3 shows that the average title length in computational linguistics was the lowest in both categories.

2. Structural Organization

The titles were also classified into single-unit and multi-unit titles according to their structural organization. The proportions of both single-unit and multi-unit titles are shown in Figure 2 below.

Figure 2

Percentage of Frequency of Single-Unit and Multi-Unit Titles by Branch
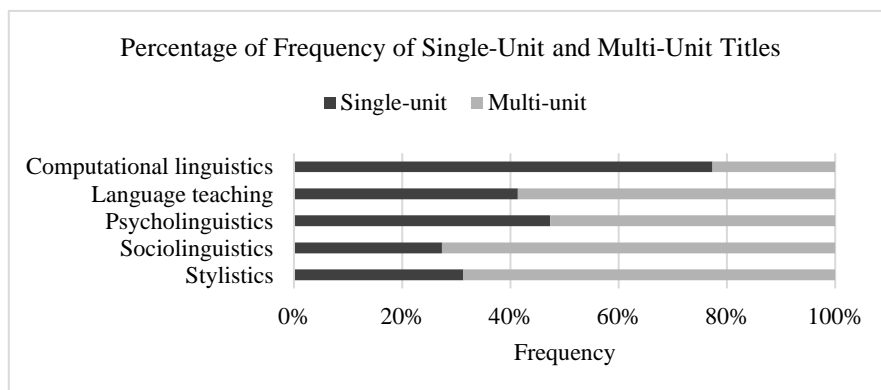
Computational linguistics titles again stood out with a unique feature. Single-unit titles were more prevalent in computational linguistics (77.3%), while most titles in all other branches comprised multiple units. Sociolinguistics titles had the fewest single-unit titles among the branches (27.3%), followed by stylistics (31.3%), language teaching (41.3%), and psycholinguistics (47.3%).

3. Syntactic Structure of Single-Unit and Multi-Unit Titles

There are three structures of the titles: phrase (PH), dependent clause (DC), and independent clause (IC). The frequency of occurrence of the types of

constructions used in single-unit and multi-unit titles is presented in Table 4 with the top three highest frequency counts in bold type.

Table 4

Frequency of Construction Types by Branch

Unit	Type	Computational Linguistics		Language Teaching		Psycho-linguistics		Socio-linguistics		Stylistics	
		f	%	*f*	%	*f*	%	*f*	%	*f*	%
Single-unit	PH	**89**	**59.3**	**43**	**28.7**	**36**	**24**	**36**	**24**	**40**	**26.7**
	DC	**25**	**16.7**	**17**	**11.3**	7	4.7	5	3.3	5	3.3
	IC	2	1.3	2	1.3	**28**	**18.7**			2	1.3
Multi-unit	PH/PH	**20**	**13.3**	**52**	**34.7**	**40**	**26.7**	**65**	**43.3**	**49**	**32.7**
	PH/DC	5	3.3	9	6	5	3.3	8	5.3	17	11.3
	PH/IC			1	0.7	6	4			2	1.3
	DC/PH	3	2	11	7.3	7	4.7	**17**	**11.3**	**20**	**13.3**
	DC/DC			2	1.3	1	0.7	4	2.7	5	3.3
	DC/IC			1	0.7	4	2.7				
	IC/PH	4	2.7	9	6	16	10.7	13	8.7	9	6
	IC/DC	2	1.3	2	1.3			2	1.3	1	0.7
	IC/IC			1	0.7						
Totals (*f*)		150		150		150		150		150	

Note. PH = phrase, DC = dependent clause, IC = independent clause.

For single-unit titles, all branches favored PH in their titles, with computational linguistics having the highest frequency ($f = 89$). NP was the most common phrase type found in all branches (e.g., "Adaptive Semiparametric Language Models"). DC, on the other hand, were most frequent in the computational linguistics dataset ($f = 25$), with most being non-finite clauses beginning with a present participle nonfinite verb (e.g., "Revisiting Multi-Domain Machine Translation"). The occurrence of IC (e.g., "Morphological preview effects in English are restricted to suffixed words") was only notable in psycholinguistics, while other branches had fewer or no such occurrences.

For multi-unit titles, PH/PH construction was commonly used in all branches, most of which were NP/NP (e.g., "Disaster Linguicism: Linguistic Minorities in Disasters"). The frequency of PH/PH exceeded that of PH and became the most common pattern of titles in all branches except computational linguistics, where single-unit titles predominated.

Table 5 provides the summary of the top three patterns of research article titles found in each branch, along with examples.

Table 5

Summarization of Top Three Title Patterns for Each Branch

Branch	Approximate Length (words)	Syntactic Structure	Title
Computational linguistics	7–11	1. PH	A Graph-Based Framework for Structured Prediction Tasks in Sanskrit
		2. DC	Characterizing English Variation Across Social Media Communities With BERT
		3. PH/PH	ParsiNLU: A Suite of Language Understanding Challenges for Persian
Language teaching	11–16	1. PH/PH	Teacher Metacognitions About Identities: Case Studies of Four Expert Language Teachers in China
		2. PH	The Effects of Textually Enhanced Captions on Written Elicited Imitation in L2 Grammar
		3. DC	Establishing Appropriate Cut Scores of a Standardized Test for a Local Placement Context
Psycholinguistics	10–15	1. PH/PH	Lexical Constraints on the Prediction of Form: Insights From the Visual World Paradigm
		2. PH	Interplay of Morphological Configuration and Language Switching in Numerical Processing and Word Processing.
		3. IC	Linguistic Focus Guides Attention During the Encoding and Refreshing of Working Memory Content
Sociolinguistics	9–14	1. PH/PH	Incomplete Neutralization in African American English: The Case of Final Consonant Voicing
		2. PH	Copula Variation in Asturian Spanish and the Multidimensionality of Stancetaking in Interaction
		3. DC/PH	Incorporating Translation into Sociolinguistic Research: Translation Policy in an International Non-Governmental Organisation
Stylistics	9–15	1. PH/PH	Negativity and Positivity Biases in Economic News Coverage: Traditional Versus Social Media
		2. PH	Four Ways of Delivering Very Bad News in a Japanese Emergency Room
		3. DC/PH	Turning the Passer-By Into a Customer: Multi-Party Encounters at a Market Stall

Note. PH = phrase, DC = dependent clause, IC = independent clause.

4. Noticeable Differences Across Branches

Two branches stand out for their noticeable differences in syntactic structures: computational linguistics, which predominantly employs three specific constructions with a limited variety of others, and psycholinguistics, characterized by a higher frequency of independent clauses in titles. To gain a

deeper understanding of the subdisciplinary characteristics reflected in these preferred titles, an investigation of the words used in these structures is presented below.

4.1 Computational Linguistics Titles
The thematic roles of NPs in the postmodifiers of NP titles were 39.1% purpose (e.g., "Self-Supervised Regularization for Text Classification"), 27.2% possessor (e.g., "Consistent Transcription and Translation of Speech"), and 22.8% instrument (e.g., "AMR-To-Text Generation With Graph Transformer").

For non-finite clause titles, 42.9% were instrument (e.g., "Parsing Chinese Sentences With Grammatical Relations"), and purpose and location were 21.4% each (e.g., "Pretraining the Noisy Channel Model for Task-Oriented Dialogue", and "Maintaining Common Ground in Dynamic Environments").

The semantic domains of verbs in non-finite clauses were 48% activity verbs (e.g., "Supertagging the Long Tail With Tree-Structured Decoding of Complex Categories") and 40% mental verbs (e.g., "Learning an Executable Neural Semantic Parser").

In NP/NP titles, 52.6% had an algorithm name as the first NP and a description of the algorithm as the second NP (e.g., "MasakhaNER: Named Entity Recognition for African Languages" and "LINSPECTOR: Multilingual Probing Tasks for Word Representations").

4.2 Psycholinguistics Titles
Independent clauses accounted for 36.7% of all titles examined, in both single-unit and multi-unit cases. Among them, declaratives were the most common, representing 76.4%, while interrogatives accounted for 23.6%. The average length of single-unit independent clause titles was 12 words.

Semantic domains of verbs were 40.4% causative (e.g., "Magnitude Sound Symbolism Influences Vowel Production") and 40.4% existence/ relationship (e.g., "Speech Spoken by Familiar People Is More Resistant to Interference by Linguistically Similar Speech").

Discussion
By comparing the five branches of linguistics in this study, certain similarities and differences were identified. Notably, the titles in computational linguistics exhibited the most distinct differences both in title length and structure. Given its association with certain scientific disciplines such as computer science, mathematics, and engineering (Jurafsky & Martin, 2009), this branch may produce results that closely align with these fields. The average title length in computational linguistics (9 words) falls within the range of Anthony's (2001) findings for average computer science title lengths, which varied from 8

to 9.9 words. The average lengths of titles in mathematics and engineering found by Hyland and Zou (2022) were also similar at 9 and 9.7 words, respectively. The average title length was found to be longer (12–13 words) in other branches which primarily draw upon social sciences and humanities, focusing on various aspects of language and language use. In terms of the most preferred range of title length, three of these branches closely align with earlier studies of the related disciplines: language teaching (11–16 words) and education (11–15 words) as reported by Hyland & Zou (2022), psycholinguistics (10–15 words) and psychology (11–15 words) as observed by Jiang & Hyland (2023), and sociolinguistics (9–14 words) and sociology (9–15 words) based on the research by Nagano (2015). Since no relevant studies on title lengths in a related field to stylistics, such as communication, have been found, their relations cannot be addressed.

The findings that single-unit titles were the most prevalent in computational linguistics, which agree with Hyland and Zou (2022), who classified mathematics and engineering as hard sciences and noted that the use of single-unit titles was much more common in the hard sciences. Multi-unit titles were more common in language teaching (58.7%), psycholinguistics (52.7%), and sociolinguistics (72.7%), which are consistent with earlier studies in related disciplines. In Hyland & Zou's study (2022), 70.7% of titles in education featured multi-unit structures, while Jiang & Hyland (2023) reported 50.5% in psychology, and Nagano (2015) found 75% in sociology. Although there were substantial differences in the percentages between language teaching and education, the percentages in psycholinguistics and psychology, and sociolinguistics and sociology were relatively close. This may suggest a potential relationship between the characteristics of the disciplines and the prevalence of multi-unit titles in the respective branches.

The dominance of NP/NP structure in multi-unit titles may be attributed to the common recommendation in guidelines to prioritize the use of keywords as reported in Rath's (2011) study, and keywords are predominantly nouns. The *Publication Manual of the American Psychological Association* (2020) also advises authors to create concise and focused titles, and avoid unnecessary words which can lengthen titles and potentially confuse indexers. NPs can effectively help avoid the complex verb structures often found in clauses, which tend to increase title length. However, while single-unit NPs are generally shorter than multi-unit NP/NPs, there is a clear preference for using multi-unit NP/NP constructions over single-unit NPs. This preference may be due to the idea that multi-unit titles can include more information by either providing more detail on the topic or by offering a more specific and precise description (Hyland & Zou, 2022). Multi-unit NP/NP titles may be necessary to effectively convey the key aspects of the research. Titles with two parts may be used to provide a

more specific perspective on a topic, setting them apart from similar works (Jiang & Hyland, 2023).

The following sections discuss the major subdisciplinary differences in patterns used, which may be influenced by the nature of the branch.

Computational Linguistics Titles: Short and Simple

Computational linguistics involves utilizing computational tools and techniques to analyze and model natural language data. Research in this field tends to develop and evaluate new methods and algorithms for processing natural language. This aspect of research is reflected in postmodification of the NP and non-finite clause titles, where instrument and purpose are commonly found. This suggests they focused on either proposing a method for something or identifying the algorithm or application. Moreover, non-finite clause titles focused on action or activity expressed by the semantic domains of the non-finite verbs. Within these clauses, a notable presence of activity verbs and mental verbs can be identified, representing both physical and cognitive actions.

Although it could be argued that methods can be found in titles of research in any field, methods are particularly distinctive in computational linguistics. This emphasis on methods and algorithms in the titles may cause them to be shorter and simpler than those in other branches. Multi-unit NP/NP, which was the most common title pattern in other branches, was only the third most common pattern in computational linguistics. However, despite being classified as multi-unit titles, the information conveyed in these titles was often as simple as single-unit titles. This is demonstrated by an algorithm name and a description of the algorithm presented in the first and second NP of the title. This pattern corresponds with Anthony's (2001) finding that the "Name: Description" structure is common in hanging titles (two-part titles separated by a punctuation mark) in computer science. The focus of these titles was on the algorithms used, which is similar to the focus of a single-unit NP.

In computational linguistics, the preference for short and simple titles may also be driven by the fact that the research is primarily targeted towards specialists in the field, resulting in relatively more technical and straightforward titles as compared to other fields. Concise titles that are direct and straightforward are more likely to attract the attention of specialists who are looking for specific methods or techniques. Descriptive titles are unnecessary as long as the methods are clear to the authors' peers. The titles may be difficult to understand for a general audience because they assume a level of technical expertise and familiarity with the field's concepts and terminology. As Crystal (2017) noted, "what scientists write in a technical journal will be clear to their colleagues, but unclear to non-specialists" (p. 114). For example, in the titles "Tree Structured Dirichlet Processes for Hierarchical Morphological Segmentation" and "Using Semantics for Granularities of Tokenization," the

terms "Dirichlet," "granularities," and "tokenization" are likely not easily understood by general audiences.

Psycholinguistics Titles: Independent Clauses

Another notable characteristic of titles can be found in psycholinguistics. Unlike other branches, psycholinguistics typically used independent clauses more frequently in both single-unit and multi-unit titles.

From the data, psycholinguistics research titles tended to include independent clauses more commonly than those in other branches, reflecting the focus on the main findings of the research. Experimental methods are central to psycholinguistics research (Allan et al., 2010; Vorwerg, 2012), with a strong emphasis on empirical results. Independent clauses titles tend to be "more assertive about the outcome of the study" (Kumar, 2013, p. 362), which is evident from the semantic domain of the verbs used in the titles (e.g., Sentence Context Guides Phonetic Retuning to Speaker Idiosyncrasies). Causative and existence/relationship verbs were mainly used to present the research results to the audience, explaining causal relationships or depicting states and relationships between entities. In terms of titles with interrogative clauses, asking a question can be more engaging and thought-provoking, which is apparent in Jamali and Nikzad's (2011) study where interrogative titles were downloaded more than declaratives. Rather than presenting the results directly, the titles encouraged readers to actively consider the research question and find the answers within the article.

Since psycholinguistics combines knowledge from psychology and linguistics, the use of independent clauses found in this study may also be influenced by changing conventions in the field of psychology. Jiang and Hyland (2023) found that, in terms of highlighted aspects of the paper presented in titles, the frequency of psychology titles mentioning results rose between 1960 and 2020.

Accessibility to a general audience is also likely a contributing factor to the prevalence of independent clause titles in psycholinguistics research articles. Such titles can communicate the main findings more widely in a way that is easily understood by a broad audience, including nonexperts. Unlike computational linguistics, which has a narrower focus, psycholinguistics "covers a very large territory" (Finch, 2008, p. 193), and its research on language processing and comprehension can be of broad interest to readers both within and outside of the field. Practical applications of psycholinguistic research can include, for example, language teaching and learning (Siyanova-Chanturia & Martinez, 2015), clinical settings for designing therapy to treat patients (Carragher et al., 2015), assisting survey designers in enhancing survey questions (Lenzner et al., 2010), and improving spam and phishing detection technologies (Xu & Rajivan, 2023). It has potential implications for a range of

fields beyond language processing, so its audience may seek out the research results that they can apply rather than the methods used. Consequently, clear titles can ensure that the article is easily understood by those interested in the specific outcome.

However, the fact that independent clause titles tend to be longer than other structures may raise concerns about accessibility. Longer titles can be more challenging to read and may reduce accessibility for some audiences. Nonetheless, a well-crafted, independent clause title can be concise and effective in providing a clear focus for the article. In this study, the average length of single-unit independent clause titles in psycholinguistics is comparable to the length of all single-unit titles in language teaching. Even though most of the single-unit titles in language teaching were phrases and clauses, with only two occurrences of independent clauses, the length of independent clause titles in psycholinguistics was not significantly longer.

These findings suggest that some branches of linguistics exhibit distinctive characteristics that are reflected in the creation of titles, presenting different aspects of research to their respective audiences and forming conventions within the specific academic community. Two branches displayed distinctive characteristics in their title constructions. Computational linguistics titles were notably shorter than titles in other branches, often utilizing a single NP to propose an algorithm and a non-finite clause to describe the method. On the other hand, psycholinguistics titles frequently employed independent clauses to present study results, which was more common than in other branches. Among the remaining branches, the NP/NP structure was the most preferred title format. This study thus contributes to a better understanding of how subdisciplinary differences influence the formation of research article titles within the field of linguistics.

Conclusion

This study explored the variations in title length and structural patterns of research article titles across different branches of linguistics. Noticeable pattern differences found in this study were mainly attributed to variations in the selected sections of the research article: methods aimed at a specific group of readers or outcomes targeting a broader audience, which serves as empirical evidence to support that branches of linguistics have distinctive natures and characteristics that are reflected in their title creation. This study thus has answered calls for more research into the subdisciplinary level of titles, an area that has not received much attention in the existing literature (Pearson, 2020).

By raising awareness of the distinctive natures of different branches of linguistics, this study suggests that English for Academic Purposes (EAP) instructors take this aspect into consideration and incorporate them into their lessons to develop title creation skills among students in linguistics. Learning to

craft titles in accordance with the preferences of specific linguistic subdisciplines may help ensure that research articles are more accessible to the expected audience in the academic community. This can help such students "be successful in international communication" (Kanoksilapatham, 2012, p. 306), since the title is the introduction to the research article that, according to Kanoksilapatham, scholars use as a primary means to communicate academically and professionally. Moreover, as students gain insights into title patterns in different linguistic branches, they may be inspired to design research that addresses interdisciplinary questions or explores linguistic phenomena across subfields.

This study's findings can serve as a reference for beginners and novice researchers. However, the appropriate title length and style may vary depending on the content of the research article and intended audience. Writers can consider various statistics from this study in their title creation. They can decide whether to use a certain pattern that may provide an opportunity to stand out from the crowd if they so wish. Because there is a great deal of variation in language use, novice writers should not limit themselves to commonly used patterns; instead, they should be encouraged and feel free to exercise creativity and expand their expression in their writing (Nagano, 2015). Additionally, they should consult the author guidelines of their target journal for further guidance on title formatting and style, including any specific character limits for the running head, which may affect the title designs.

This study has limitations that should be noted. First, the sample size in this exploratory study is small and further research is encouraged to confirm and expand upon these findings. Second, this study examined the word types used in the distinctive structures of titles within two specific branches. It would be valuable to explore the characteristics of words in other constructions and branches as well. However, accurate analysis of technical terms may require the expertise of specialists in those fields. Future research could benefit from collaborative efforts between linguists and researchers specializing in specific disciplines to investigate word use and content in other branches of linguistics.

Acknowledgements

References

Afful, J. B. A., & Ankomah, C. (2020). Deconstructing applied linguistics conference paper titles: A syntactic analysis. *Journal of Language and Linguistic Studies*, *16*(3), 1219–1236. https://doi.org/10.17263/jlls.803669

Aitchison, J. (2008). *The articulate mammal: An introduction to psycholinguistics* (5th ed.). Routledge.

Allan, K., Bradshaw, J., Finch, G., Burridge, K., & Heydon, G. (2010). *The English language and linguistics companion*. Palgrave Macmillan.

American Psychological Association. (2020). *Publication manual of the American Psychological Association* (7th ed.). https://doi.org/10.1037/0000165-000

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

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

Busch-Lauer, I.-A. (2000). Titles of English and German research papers in medicine and linguistics theses and research articles. In A. Trosborg (Ed.), *Analysing professional genres* (pp. 77–94). John Benjamins Publishing Company. https://doi.org/10.1075/pbns.74.08bus

Carragher, M., Sage, K., & Conroy, P. (2015). Outcomes of treatment targeting syntax production in people with Broca's-type aphasia: Evidence from psycholinguistic assessment tasks and everyday conversation. *International Journal of Language & Communication Disorders*, *50*(3), 322–336. https://doi.org/10.1111/1460-6984.12135

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. https://journal.aall.org.au/index.php/jall/article/view/178

Cook, G. (2003). *Applied linguistics*. Oxford University Press.

Crystal, D. (2017). *Making sense: The glamorous story of English grammar*. Oxford University Press.

Diener, R. A. V. (1984). Informational dynamics of journal article titles. *Journal of the American Society for Information Science*, *35*(4), 222–227. https://doi.org/10.1002/asi.4630350405

Diodato, V. (1982). The occurrence of title words in parts of research papers: Variations among disciplines. *Journal of Documentation*, *38*(3), 192–206. https://doi.org/10.1108/eb026728

Fernández, E. M., & Cairns, H. S. (2018). Prologue. In E. M. Fernández & H. S. Cairns (Eds.), *The handbook of psycholinguistics* (pp. xxiii–xxv). Wiley Blackwell.

Finch, G. (2008). *Key concepts in language and linguistics* (2nd ed.). Palgrave Macmillan.

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.

Haggan, M. (2004). Research paper titles in literature, linguistics and science: Dimensions of attraction. *Journal of Pragmatics*, *36*(2), 293–317. https://doi.org/10.1016/S0378-2166(03)00090-0

Hudson, J. (2016). An analysis of the titles of papers submitted to the UK REF in 2014: Authors, disciplines, and stylistic details. *Scientometrics*, *109*(2), 871–889. https://doi.org/10.1007/s11192-016-2081-4

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

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

Jeffries, L., & McIntyre, D. (2010). *Stylistics*. Cambridge University Press.

Jiang, F. K., & Hyland, K. (2023). Titles in research articles: Changes across time and discipline. *Learned Publishing*, *36*(2), 239–248. https://doi.org/10.1002/leap.1498

Jurafsky, D., & Martin, J. H. (2009). *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition* (2nd ed.). Prentice Hall.

Kanoksilapatham, B. (2012). Structure of research article introductions in three engineering subdisciplines. *IEEE Transactions on Professional Communication*, *55*(4), 294–309. https://doi.org/10.1109/TPC.2012.2223252

Kuch, T. D. C. (1978). Relation of title length to number of authors in journal articles. *Journal of the American Society for Information Science*, *29*(4), 200–202. https://doi.org/10.1002/asi.4630290408

Kumar, M. J. (2013). Making your research paper discoverable: Title plays the winning trick. *IETE Technical Review*, *30*(5), 361–363. https://www.tandfonline.com/doi/abs/10.4103/0256-4602.123113

Lenzner, T., Kaczmirek, L., & Lenzner, A. (2010). Cognitive burden of survey questions and response times: A psycholinguistic experiment. *Applied Cognitive Psychology*, *24*(7), 1003–1020. https://doi.org/10.1002/acp.1602

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

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

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

Moore, T. (2020). Academic clickbait: The arcane art of research article titling. *The Australian Universities' Review*, *62*(1), 54–56. https://search.informit.org/doi/10.3316/informit.006130266170400

Nagano, R. L. (2015). Research article titles and disciplinary conventions: A corpus study of eight disciplines. *Journal of Academic Writing*, 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

Paiva, C. E., Lima, J. P. d. S. N., & Paiva, B. S. R. (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

Pearson, W. S. (2020). Research article titles in written feedback on English as a second language writing. *Scientometrics*, *123*(2), 997–1019. https://doi.org/10.1007/s11192-020-03388-7

Peritz, B. C. (1984). On the informativeness of titles. *International Classification*, *11*(2), 87–89. https://doi.org/10.5771/0943-7444-1984-2-87

Rath, A. (2011). A discourse analysis of manuals of style: The specification of function and structure for composition of research article titles. *Journal of Research in Applied Linguistic Studies*, *2*(2), 3–17.

Schmitt, N., & Celce-Murcia, M. (2020). An overview of applied linguistics. In N. Schmitt & M. P. H. Rodgers (Eds.), *An introduction to applied linguistics* (3rd ed., pp. 1–16). Routledge.

SCImago. (n.d.). *Scimago journal & country rank*. Scimago Lab. Retrieved July 25, 2022, from https://www.scimagojr.com/journalrank.php

Siyanova-Chanturia, A., & Martinez, R. (2015). The idiom principle revisited. *Applied Linguistics 36*(5), 549–569. https://doi.org/10.1093/applin/amt054

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

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

Swales, J. M., & Feak, C. B. (2012). *Academic writing for graduate students: Essential tasks and skills* (3rd ed.). University of Michigan Press ELT. https://doi.org/10.3998/mpub.2173936

Timyam, N. (2010). *An introduction to English linguistics*. Thailand Research Fund.

Vorwerg, C. (2012). Experimental methods in psycholinguistics. In A. Ender, A. Leemann, & B. Wälchli (Eds.), *Methods in contemporary linguistics* (pp. 363–386). De Gruyter Mouton. https://doi.org/10.1515/9783110275681.363

Wang, Y., & Bai, Y. (2007). A corpus-based syntactic study of medical research article titles. *System*, *35*(3), 388–399. https://doi.org/10.1016/j.system.2007.01.005

Xiang, X., & Li, J. (2020). A diachronic comparative study of research article titles in linguistics and literature journals. *Scientometrics*, *122*(2), 847–866. https://doi.org/10.1007/s11192-019-03329-z

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

Xu, T., & Rajivan, P. (2023). Determining psycholinguistic features of deception in phishing messages. *Information and Computer Security*, *31*(2), 199–220. https://doi.org/10.1108/ICS-11-2021-0185

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. (1994). Relation of title length of journal articles to number of authors. *Scientometrics*, *30*(1), 321–332. https://doi.org/10.1007/BF02017231

Yitzhaki, M. (2002). Relation of the title length of a journal article to the length of the article. *Scientometrics*, *54*(3), 435–447. https://doi.org/10.1023/A:1016038617639

Appendix

List of Journals

Computational linguistics
- *Transactions of the Association for Computational Linguistics*
- *Artificial Intelligence*
- *Computational Linguistics*

Language teaching
- *Modern Language Journal*
- *TESOL Quarterly*
- *System*

Psycholinguistics
- *Journal of Memory and Language*
- *Bilingualism: Language and Cognition*
- *Journal of Experimental Psychology: Learning Memory and Cognition*

Sociolinguistics
- *Journal of Sociolinguistics*
- *Language in Society*
- *Language Variation and Change*

Stylistics
- *Research on Language and Social Interaction*
- *Communication Research*
- *European Journal of Communication*