

MAKING SENSE OF TITLES: HOW PARTS OF RESEARCH ARTICLE TITLES RELATE

ROBIN LEE NAGANO

This paper reflects the search for an appropriate system of categorization to apply to the content of research article titles. The title conveys more than a series of key words describing research; it also suggests how information relates and may use rhetorical devices. Here a corpus of titles of journal articles in the multidisciplinary area of financial literacy is used for a pilot study on categorizing the information content multiple-unit titles. A system previously applied to applied linguistics titles is used and potential adaptations to better suit titles in financial literacy are discussed.

Introduction

Swales & Feak discuss titles in their academic writing textbook, and suggest that common relations between units of multiple-unit titles are: Problem: Solution; General: Specific; Topic: Method; Major: Minor.¹ However, they were unable to back up this list by data, as title studies began to be published only later.

There seem to be two approaches in the title literature that classify the information or rhetorical function of elements. One focuses on the information given in the title, usually considering the title as a whole. The other approach focuses on multiple-unit titles, identifying the function of each unit and looking for patterns in the relation between the units.

The majority of researchers looking at information and rhetorical elements of titles have focused on one particular discipline and worked out a system for titles in that particular discipline. For example, Anthony proposed five components for the titles of research articles in the discipline of computer science:

- Name of Approach/Algorithm/Application etc.
- Description of Approach/Algorithm/Application etc.
- Topic of Research Article
- Scope of Research Article
- Method of Research.²

Anthony found that a *Name: Description* format was quite common in some of the computer science journals he investigated, primarily used for articles introducing a

¹ SWALES – FEAK 1994, 209.

² ANTHONY 2001, 189.

new system, approach, or algorithm, such as “KLAIM: A Kernel Language for Agents Interaction and Mobility” or “Veinerization: A new shape description for flexible skeletonization”.³

However, researchers in other fields have rarely if ever identified “Name” units. Moattarian & Alibabae found no examples in a three-discipline corpus of 420 titles, dropping it from their classification system for titles in applied linguistics, civil engineering, and dentistry.⁴ This is a good example of how the discourse of a discipline affects titles of articles within that discipline.

Cheng et al. first attempted to apply Anthony’s categories to a corpus of applied linguistics titles and decided that modification was needed.⁵ They worked out a set of eleven categories made up of different combinations of the elements: to Anthony’s categories of topic, scope, method, and description, they added source, question, and metaphor in order to better describe the titles in their corpus. These categories were adopted by Slougui in a study on dissertation titles in the field of teaching English as a second/foreign language. She contrasted titles written by UK-based dissertation writers (not necessarily native speakers of English) and those in a non-EFL context (Algerian universities), identifying eight difference combinations used by the UK-based writers but only five combinations by the other group, which overwhelmingly chose a Topic: Source arrangement.⁶ Slougui slightly modified the categories of Cheng et al., re-naming “metaphor” as “rhetorical device”, a broader and more descriptive label.

Gesuato analyzed titles in the discipline of linguistics, using the classifications of topic (general topic and specific topic, if needed), method, and context, and paying attention to the sequencing of the elements in multiple-unit titles. For research articles, she found the arrangement *General topic: Specific topic* made up 57% of the titles.⁷ No definitions were provided, which may be one reason that her categories appear not to have been adapted by other researchers. The category ‘context’ is an intriguing way to group together information on groups of people, places, situations, and other information that gives the reader a better idea of what to expect in the paper.

Another influential system has been that of Goodman et al., who chose the categories of methods, dataset, results, conclusions, or topic only for a title corpus in medicine.⁸ In their study, ‘dataset’ referred to the name or acronym of a particular study (such as “the PRIME-MD 1000 Study”). Note that this study did not examine the relations between units, but focused on the information available in the whole title, whether it consisted of one or more units. Sahragard & Meihami adopted the labels of Goodman et al. but added definitions more suitable for the discipline

³ ANTHONY 2001, 191.

⁴ MOATTARIAN – ALIBABAE 2015, 37.

⁵ CHENG – KUO – KUO 2012, A8.

⁶ SLOUGUI 2018, 11.

⁷ GESUATO 2008.

⁸ GOODMAN – THACKER – SIEGEL 2001, 76.

of applied linguistics.⁹ The two sets of definitions are compared in *Table 1*, casting some light on what is relevant to one discipline but not the other.

Table 1. Definitions for content categories

Category	GOODMAN et al. (2001, 76)	SAHRAGARD & MEIHAMI (2016, 1320)
Topic only	Title noted a subject but did not include information on other categories listed below, or it seemed ambiguous.	Title indicates a subject but does not include information on other categories listed below, or it seems ambiguous.
Methods/design	Title specified the topic and an approach to study design, data management, or analysis (such as case-control, cohort, effectiveness, efficacy, frequency, incidence, mortality, prevalence, surveillance, survival, trend, or validity study; meta-analysis; randomized clinical trial; or sensitivity-specificity or cost-effectiveness analysis); or title indicated an investigation of an epidemic or outbreak; or title provided an incomplete description of a method (such as assessment, evaluation, population sample, or comparison).	Title specifies the topic and an approach to study design, data management, or analysis (such as case-control, cohort, effectiveness, efficacy, frequency, incidence, prevalence, trend, or validity study; meta-analysis; randomized), or provides a vague description of a method (such as assessment, evaluation, or comparison) or refers to the statistical tests used in a study (ANOVA, ANCOVA, MANOVA). It may contain information about the method of the study but not its design (for example randomization), or vice versa.
Dataset	Title indicated the topic and name or acronym of a specific study (such as “the PRIME-MD 1000 Study”).	Title indicates the topic and name or acronym of a specific study, which can reveal the origin of the dataset considered or the population sample consulted.
Results	Title contained the topic and quantitative information (a specific value), semiquantitative or ordinal information (such as increased, decreased, high, or low), or some other specification of a relation (such as association, change, correlation, determinants, effect, evidence, impact, influence, outcomes, predictors, relation, remission, risk,	Title includes the topic and quantitative information (a specific value), semiquantitative or ordinal information (such as increased, decreased, high, or low), or some other specification of a relation (such as association, change, correlation, determinants, effect, evidence, impact, influence, outcomes, predictors, relation, remission, risk, variabil-

⁹ SAHRAGARD – MEIHAMI 2016.

	variability, or variation) regarding what the authors found.	ity, or variation) about the findings. Other titles may contain terms such as role, effectiveness, relationship, etc.
Conclusion	Title included the topic and an unequivocal statement based on the analysis of the reported evidence.	Title includes the topic and an unequivocal statement based on the analysis of the reported evidence such as the implications of a study or the overall conclusion.

The five categories of Goodman et al. were also used in a diachronic study by Xiang & Li comparing titles in two fields in the humanities: linguistics and literature.¹⁰ In this case, dataset information was present in over 30% of the titles in literature, which may sound surprising, but the authors stated that they interpreted the name of the literary work as a dataset. They found that linguistics articles gave information on the dataset (no details given) in only 4-6% of the titles investigated.

Researchers working with classification of information in other fields have come up with different labels to describe information relevant to their investigations. Méndez et al. narrowed down the categories to statement of purpose (similar to topic?), method and outcome in an analysis of titles in astrophysics.¹¹ In contrast, Kerans et al., studying research article titles in clinical medicine journals, added more detail than the Goodman categories, searching for information about research design/methods, results, patient population, treatment, clinical context, and geographic location.¹²

Genre studies in the academic discourse of multidisciplinary/interdisciplinary fields are still not very common. Samraj noted that discourse can be influenced by the interdisciplinary nature of a topic in which researchers are from different backgrounds, perhaps creating a greater need for explicit statements, as the amount of shared knowledge may be difficult to judge in a topic with fuzzy boundaries. She also suggests that a field with an applied nature may involve more persuasive discourse justifying practical applications.¹³

In this pilot study, a set of research article titles from the multidisciplinary topic of financial literacy is analyzed using the categories of Cheng et al. for identifying relations between units of multiple-unit titles. The aim is to find whether the categories are suitable to a different discipline and to consider points that may require modification in order to move towards the creation of a widely applicable set of categories to investigate the information and rhetorical content of titles.

¹⁰ XIANG – LI 2020, 857–858, 861.

¹¹ MÉNDEZ – ALCARAZ – SALAGER-MEYER 2014, 2336.

¹² KERANS – MARSHALL – MURRAY – SABATÉ 2020, 134.

¹³ SAMRAJ 2005, 152.

Corpus and methods

The corpus used for this pilot study consists of titles of research articles dealing with financial literacy. The Scopus database was searched for research articles published in English in 2021 whose titles contain the term “financial literacy”. Of the 181 results, several were omitted due to insufficient information on the indexing status of the journal; only those articles published in journal with quartile ratings for 2021 were included. The articles were published in journals classified in business management and accounting, finance, psychology, education, and numerous other fields and appear in a wide variety of journals, as is typical of papers in financial literacy (Goyal & Kumar¹⁴). The final title corpus contains 168 titles. This paper concentrates on the relation between units of multiple unit titles: the corpus contains 87 2-unit titles and one 3-unit title, so $n = 88$. Coding was carried out by the researcher by hand.

Results and discussion

The 88 multiple-unit titles in my financial literacy corpus were coded using the categories of Cheng et al. (topic, scope, method, description, source, question, and metaphor) and the order of the units was indicated. A total of twenty combinations were found; the most frequent are shown in Figure 1 in comparison with Cheng et al.’s results.

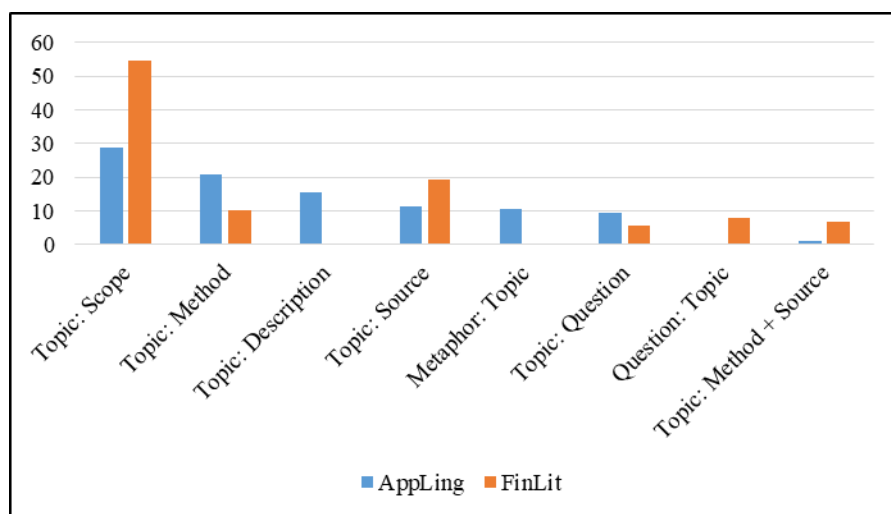


Figure 1. Rhetorical combinations of multiple-unit titles with at least 5% representation in one corpus: CHENG et al. (2012) for applied linguistics (AppLing) and financial literacy (FinLit) results

¹⁴ GOYAL – KUMAR 2021.

The results for the two corpora shown in Figure 1 reveal that three unit relation types that appear in both corpora are Topic: Scope, Topic: Method, and Topic: Source. As financial literacy has a strong social-sciences component, it is not surprising that similarities are found. Another shared feature is questions, supporting results of other cross-disciplinary studies (see Nagano¹⁵). However, the frequency of some combinations differs quite widely between the two corpora. It is difficult to determine whether these differences are due to disciplinary differences or simply to differences in coding. While Cheng et al. reported eleven different combinations,¹⁶ my preliminary results showed twenty combinations, with the most complex being *Topic: Method + Scope + Source* for the title *Financial Literacy Level: An Empirical Study on Savings, Credit and Budget Management Habits in High School Students*.¹⁷ However, I considered other classifications for this title, notably *Description: Method + Topic + Source*.

While attempting to use the system reported in Cheng et al., I encountered several difficulties in classification. The vital category of ‘topic’ is itself a problem. Let us take the imaginary title “The effect of financial literacy on investment decisions of retired people in California”. Is the entire title the topic? Or perhaps ‘the effect of financial literacy on investment decisions’? Or ‘financial literacy’? ‘Investment decisions’? What level of specificity is stipulated? “Scope” seems to refer to a narrower version of the topic, introducing particular factors or parameters investigated, but figuring out where topic ends and scope begins can be a challenge. This is one instance where there may be a benefit to splitting information into two units, allowing both authors and readers to move from a more general description to specifics.

Another category that caused difficulty for me was “description”. There was only one instance that fit the sense introduced by Anthony used in combination with a name: *ASSET: A new measure of economic and financial literacy*. I finally considered it as a comment or wider interpretation of the study or its application. The term “metaphor”, in my judgement, fit only one title in the FinLit corpus: “*If You Fall, Stand Up Again’: The Moral Nature of Financial Literacy in the Global South.*”

Besides the lack of guidelines in applying the category labels, the most important issue I identified is the use of “question” as a category. The problem with this widespread practice is that the syntactic form of an interrogative (including question fragments) is being conflated with its informative or rhetorical content, thus masking part of its contribution to the information gained by the reader. On the other hand, it is important to register that a non-standard method has been used; the author has chosen to design a title that does not follow the typical choice of a noun phrase. For instance, the original title *Whom to educate? Financial literacy and investor awareness* could have been designed using the very common pattern

¹⁵ NAGANO 2010, 104.

¹⁶ CHENG – KUO – KUO 2012, a-4, Table 3.

¹⁷ Titles from the FinLit corpus are given in italics.

of coordinated noun phrases, something like ‘Financial literacy, investor awareness, and targets of education’. If the question is not considered, the information on the target population for education gets lost. Similarly, the original title *Women and financial literacy in Spain. Does marital status matter?* introduces the study’s important variable of marital status. It could perhaps have been worded ‘The effect of marital status on the financial literacy of women in Spain’, with a rather different effect (and affect). Thus, both the information value of the question and its rhetorical message need to be recognized in order to describe the role of the question in the title. For this reason, as I work further with analyzing the informative and rhetorical patterns of titles, I will experiment with expanding the term “rhetorical device” to mark questions (as well as metaphors, quotations, and so on) and use it in tandem with other categories, such as *Rhetorical device* + *Scope*. It is important to note that questions occur quite regularly in some fields and rarely or never in others.¹⁸

Similar issues arise with any non-nominal title, such as title units beginning with –ing forms. In this example, both units use this format, but the underlying information differs: *Taking Financial Literacy Downtown: Leveraging Community Partnerships to Increase Communication of Essential Financial Information in an Urban Area*. The first unit is probably intended mainly for rhetorical effect, as something to catch the eye of the reader, as well as active associations connected with the idea of ‘taking something downtown’. The second unit inserts the verb “leveraging”, rather than a nominal version such as ‘Communication of essential financial information in an urban area through community partnerships’, for a much more action-based impression (along with the use of ‘increase’ to indicate positive change). Both units include information on the study, mediated through the rhetorical choice of the –ing form. Several researchers working with content analysis have commented on the ambiguity of this form, and also the need for human analysis to classify such constructions. Goodman et al. pointed out the example of “Increasing prescription of drugs for secondary prevention after myocardial infarction”: is it a description of a trend in medical practice or a suggestion for treatment?¹⁹ Li & Xu suggest referring to abstracts to clear up ambiguity,²⁰ but we should bear in mind that the reader will often encounter titles with no supporting information available (in the Table of Contents or a list of database search results).

There are several limitations to the current pilot study. Naturally, the limited size of the corpus is one issue. A more serious one is the subjective and rather intuitive categorization that I have employed. Because of potential ambiguity in titles themselves, as well as in categorizing them, title studies focusing on content often have more than one coder and note that reasonable familiarity with the discipline is

¹⁸ SOLER 2007, 97: questions in 5% of linguistics titles; NAGANO 2010, 104: questions in 2% or less in natural sciences and engineering, 5–13% for humanities and social science disciplines.

¹⁹ GOODMAN – THACKER – SIEGEL 2001, 87.

²⁰ LI – XU 2019, 1625.

needed by coders²¹, while Anthony used four people for categorizing, three of whom were specialist informants²². A more formal study would require a proper taxonomy, multiple coders, analysis of agreement rates, and a protocol for resolving differing judgements. It would also be worth exploring associations between informative and rhetorical content and syntactic style, both in multiple-unit and single-unit titles.

Conclusions

This paper has reported on an effort to apply a particular set of content analysis categories to multiple-unit titles of research articles in the multidisciplinary area of financial literacy. The categories of Anthony designed with computer science titles in mind were modified by Cheng et al. for use in analyzing titles in applied linguistics. While the categories captured much of the information type, further adjustment would seem to be needed, along with a standardized set of definitions for labels. The category ‘question’ is particularly problematic, as it reveals only the form and nothing about the content. With further modification, this could become a content-analysis tool that would allow comparison among studies of title features.

If a reliable system for analyzing the content of titles and title units can be worked out, then it could prove useful in comparing titles in different disciplines or genres. Currently, the lack of a standardized method makes it difficult to compare data from studies of different disciplines. Knowing more about title features can help writers to design titles that are appropriate to their discipline and also to become aware of the variety of options and explore new types of titles. In a multidisciplinary area such as financial literacy, a content analysis can be one tool to explore aspects of different influences and communities of contributors.

References

ANTHONY 2001

Laurence ANTHONY: Characteristic features of research article titles in computer science. *IEEE Transactions on Professional Communication* 44. 2001, 187–194.

CHENG – KUO – KUO 2012

Stephanie W. CHENG – Chih-Wei KUO – Chih-Hua KUO: Research article titles in applied linguistics. *Journal of Academic Language & Learning* 6/1. 2012, 1–14.

²¹ KERANS – MARSHALL – MURRAY – SABATÉ 2020, 130.

²² ANTHONY 2001, 189.

GESUATO 2008

Sara GESUATO: Encoding of information in titles: Academic Practices across four genres in linguistics. In: *The Role of E-corpora in Translation and Language Learning*. Ed. C. TAYLOR. Trieste, Edizioni Università di Trieste, 2008, 127–157.

GOODMAN – THACKER – SIEGEL 2001

Richard A. GOODMAN – Stephen B. THACKER – Paul Z. SIEGEL: What's in a title? A descriptive study of article titles in peer-reviewed medical journals. *Science Editor* 24. 2001, 75–78.

GOYAL – KUMAR 2021

Kirti GOYAL – Satish KUMAR: Financial literacy: A systematic review and bibliometric analysis. *International Journal of Consumer Studies* 45/1. 2021, 80–105.

KERANS – MARSHALL – MURRAY – SABATÉ 2020

Mary Ellen KERANS – Jane MARSHALL – Anne MURRAY – Sergi SABATÉ: Research article title content and form in high-ranked international clinical medicine journals. *English for Specific Purposes* 60. 2020, 127–139.

LI – XU 2019

Zhijun LI – Jinfen XU: The evolution of research article titles: the case of Journal of Pragmatics 1978–2018. *Scientometrics* 121. 2019, 1619–1634.

MÉNDEZ – ALCARAZ – SALAGER-MEYER 2014

David I. MÉNDEZ – M. Ángeles ALCARAZ – Françoise SALAGER-MEYER: Titles in English-medium Astrophysics research articles. *Scientometrics* 98. 2014, 2331–2351.

MOATTARIAN – ALIBABAEI 2015

Aasa MOATTARIAN – Ahmad ALIBABAEI: Syntactic Structures in Research Article Titles from Three Different Disciplines: Applied Linguistics, Civil Engineering, and Dentistry. *The Journal of Teaching Language Skills* 7/1. 2015, 27–50

NAGANO 2010

Robin Lee NAGANO: Question forms in journal article titles. In: *IV. Alkalmazott Nyelvészeti Doktorandusz Konferencia*. Budapest, MTA Nyelvtudományi Intézet, 100–110.

SAHRAGARD – MEIHAMI 2016

Rahman SAHRAGARD – Hussein MEIHAMI: A diachronic study on the information provided by the research titles of applied linguistics journals. *Scientometrics* 108. 2016, 1315–1331.

SAMRAJ 2005

Betty SAMRAJ: An exploration of a genre set: Research article abstracts and introductions in two disciplines. *English for Specific Purposes* 24/2. 2005, 141–156.

SLOUGUI 2018

Doudja SLOUGUI: Dissertation titles in EFL and UK-based contexts: How much do they differ? *ASp* 74. 2018, 20.

SOLER 2007

Viviana SOLER: Writing titles in science: An exploratory study. *English for Specific Purposes* 26/1. 2007, 90–102.

SWALES – FEAKE 1994

John M. SWALES – Christine B. FEAKE: *Academic Writing for Graduate Students: A course for nonnative speakers of English*. Ann Arbor, The University of Michigan Press, 1994.

XIANG – LI 2020

Xuechun XIANG – Jing LI: A diachronic comparative study of research article titles in linguistics and literature journals. *Scientometrics* 122. 2020, 847–866.