Françoise Salager-Meyer*, María Ángeles Alcaraz Ariza**

Titles are "serious stuff"\(^1\): a historical study of academic titles\(^2\)

ABSTRACT

In this paper we carried out a diachronic analysis (1840-2009) of a corpus of 180 medical case report titles drawn from the British Medical Journal. We analyzed a series of quantitative variables (number of authors and their institutional affiliation, title length, and punctuation/grammatical data) and qualitative variables (authors' collaboration and types of titles). The results of our research show various shifts over the period studied that could be attributed to the following factors: 1) the progressive professionalization of medicine; 2) the need of disciplinary teams to conduct an ever-increasing complex research; and 3) the increased specialization and the growing complexity of medical science. The only variable that has remained constant over the years is the nominal nature of case report titles. It could then be stated that case report titles would distinguish themselves from research article titles, which are being characterized by a certain tendency towards verbalization.

Key words: English, medicine, case reports, titles, diachronic analysis

\(^1\) Swales (1990: 224)

\(^2\) This paper forms part of a wider research that is supported by Grant M-976-09-06A from the Scientific, Technological, Humanistic and Artistic Research Center (CDCHT) from the University of The Andes, Mérida, Venezuela.

* Correspondence address: Françoise Salager-Meyer, Graduate School of Medicine, Universidad de Los Andes, Mérida, Venezuela, e-mail address: francoise.sm@gmail.com

** Correspondence address: María Ángeles Alcaraz Ariza, Department of English Studies, University of Alicante, Alicante (Spain), e-mail address: ariza@ua.es
1. INTRODUCTION

1.1. On the importance of titles in scientific research

In our paper published in 1991 (Salager-Meyer 1991), we argued that, because of the tremendous growth in the number of journals, hence of papers, published and because of the interdisciplinary nature of research, scientists would have to rely more and more on abstracts as a short, concise, complete and accurate source of information. Scientific paper abstracts are indeed a time-saving device that helps readers to decide whether the whole article is worth reading or not. Today, because of the hyperproduction of professional literature which is estimated to double every 12 years (Stix 1994), most scientists are content with reading the title only of the papers they deem interesting for their research purposes.

As a matter of fact, there is evidence that doctors sometimes make clinical decisions from the titles of journal articles (Haynes et al. 1990; Goodman 2000). This is why titles should convey effectively the topic of the report, and, if possible, the design of the reported investigation, while attracting the attention of and informing the primary target audience, editors and reviewers. Titles should therefore be clear, accurate and precise (Swales and Feak 1994; Day 1998; Hartley 2008). The more precise and accurate indeed the title is, the easier it will be for bibliographers to compile data for indexing, abstracting and other documentation purposes. However, it is not long ago, as Goodman et al. (2001) report, that monographs about writing scientific papers (Zeiger 2000) have begun to stress the importance and pivotal role of titles.

One of the reasons why, despite their succinctness, titles are ”serious stuff” (Swales 1990: 224) is because, as Yitzhaki (1994) rightfully posits, the function of titles is to draw readers’ attention to a paper and to indicate its content from a short glimpse, thus allowing readers to decide whether the paper deserves further reading. According to Swales and Feak (1994), a title should indicate the topic and scope of the study, and be self-explanatory to readers in a given discipline. We can thus acknowledge, along with many other researchers (e.g., Bird and Knight 1975; Diener 1984; Nahl-Jakobovits and Jakobovits 1987; Alley 1996; Yitzhaki 1994; Whissel 1999; Anthony 2001; Gross et al. 2002; Haggan 2004), that titles –these small but key front rhetorical devices– play a capital role in scientific research in the sense that they are the gateway that represents the reader’s first encounter with a document, whether it is a research article, a thesis, a conference paper, a review paper, etc.

Other scholars have stressed that titles should be as informative as possible in order to facilitate the process of storing, searching and retrieving the information (Black

The aforesaid clearly underscores that the field called "titleology" (Biacchi 2003, cited in Soler 2011: 124) has grown quite substantially since Swales claimed in 1990 that titles were an issue in academic genres that had not been fully addressed. As Soler (2011) remarks, since its creation, the field has diversified itself through a heterogeneous range of topics addressed by applied linguists, information scientists and psychologists.

Because the vast and rich literature on the subject has examined the issue from a range of various perspectives, we will now present the results and conclusions of the most significant papers that have been published on the topic. We will classify that research into four groups, although, as we will see, in some cases there is an overlapping of approaches: 1) the research conducted from a discipline-specific perspective; 2) the research that tackled the issue of academic titles from a cross-disciplinary angle; 3) the research that dealt with the cross-generic aspect of the topic; and finally, 4) the research that considered academic titles from a cross-cultural angle.

1.2. Review of literature on academic titles

1.2.1. Discipline-specific studies on academic titles

Some research on academic titles (by far, the minority) has been conducted in mono-disciplinary contexts. Gesuato (2009), the most recent discipline-specific study, examined the differences and similarities of titles among four genres in the field of Linguistics, viz books, research articles (RAs), dissertations, and proceedings papers. She specifically examined the following variables: length (calculated as the number of running words), lexical density, syntactic encoding, structural organization, subphrasal syntax, content analysis and information sequencing. The findings of her study showed that the similarities among the titles of these four genres outweigh the differences. The author attributes this phenomenon to the commonality of the genre, to the linguistic encoding and the communicative goals of titles, and to the fact that the genres she studied belonged to the same discipline.

In the field of Psychology, Whissell (1999) found that, on average, titles of RAs were 12-word long, while Wang and Bai (2007) observed that in Medical RAs, the average length of the titles was 10.9 words, with 99% of them being realized as nominal groups and 75% characterized by the presence of single heads. Ninety-eight percent did not have any subtitles, and 68% were accompanied by post-modifying prepositional phrases.
Forray and Woodilla (2005) explored the ways in which temporality is invoked and represented in Management RA titles. The authors observed that temporality manifested itself in punctuation and word choice, among others. Goodman et al. (2001), for their part, examined the content of the titles of 420 peer-reviewed medical RAs and queried editors of the journals on editorial policies and practices in relation to titles. The authors noted that editors occasionally modified titles to increase their clarity and informativity, and the few journals having a policy on titles addressed the issue of title length only.

1.2.2. Cross-disciplinary research

The vast majority of the research that has been conducted on academic titles has adopted a cross-disciplinary approach. To our knowledge, the first study on titles of academic research that adopted such a perspective is that of Buxton and Meadows (1977) on RAs from the natural and the social sciences. The results of their study indicate that titles in the natural sciences have a higher information content than those in the social sciences, and that the overall length of the titles in both the natural and the social sciences increased over time, which corresponds to an informativity increase. Another cross-disciplinary research is that of Yitzhaki (1994) who examined RA titles in the hard sciences, social sciences and humanities, measuring title informativity and its possible correlation with the number of authors. In the scientific fields, Yitzhaki found that there was a moderate positive correlation between the number of authors and the number of content words in the titles. By contrast, in the social sciences, the correlation was found to be rather low and relevant to a few titles only. As for the humanities, no correlation was found between the number of authors and the number of content words. Yitzhaki argues that the correlation he detected in the scientific disciplines can be accounted for by the high frequency of multiple-authored papers in science journals.

Other cross-disciplinary studies on academic titles were those conducted by Fortanet et al. (1997, 1998) and Haggan (2004) who reported differences as well as similarities in the syntax of titles across disciplines. Fortanet et al. (1997, 1998) analyzed 200 titles of RAs in Computer Science, Applied Linguistics, Business and Economics, and Chemistry, and reported that the Chemistry titles were the longest, while those in Linguistics were the shortest. They also found that the most common syntactic structure of titles was made up of a ‘premodifier + a head + a postmodifier’. Head combinations were found to be more frequent in Linguistics and Business and Economics titles, while combinations of pre- and post-modifiers were more frequent in Chemistry and Computer science. Linguistics and Business and Economics titles displayed a majority of -ing forms, while their Chemistry and Computer Science
counterparts exhibited a balanced distribution between -ing and -ed forms. Finally, the Linguistics, Economics and Business titles favored the use of definite articles, while the Chemistry and Computer Science titles showed an even distribution of definite and indefinite articles. Haggan (2004), for his part, compared over 700 RA titles in Literature, Linguistics and Education and found similar syntactic and structural choices in the three disciplines.

Other studies examined the structure and wording of titles. Anthony (2001) studied the length, word frequency, prepositions, and punctuation marks in the titles of various Computer Science sub-disciplines. The average title length was 9 words, with most titles ranging from 6 to 12 words. On average, 2-unit titles, with a colon separating them, made up about 13% of the data. The two most frequent semantic relationships holding between the two parts of a title were ‘name: description’ and ‘topic: scope’, but with considerable variation across RAs. There were statistically significant differences among the RA titles in relation to the frequency of specific words showing the sub-disciplinary content specificity. For his part, Afful (2004) explored the variations of dissertation titles in English Studies and Engineering in a corpus of 798 titles. The variables studied included text length, text structure and the use of prepositions. He recorded both differences and similarities between the titles of the two disciplines.

The studies of Dillon (1982), Lewinson and Hartley (2005), and Hartley (2007) focused on the use of colons in titles. These punctuation marks are used to mark two information units indicating either the general framework of the article and the specific topic of the document, or the topic and the method. Dillon (1982) explored the titles of 1,150 RAs in Education, Psychology and Literary Criticism published between 1880 and 1980, and noticed a steady increase in the use of colons across these three disciplines. In their 1997 study, Fortanet et al. also reported that the colon, semicolon, and full stop were the most frequent punctuation marks in Business and Economics titles, and the least common in Computer Science titles. For their part, Lewinson and Hartley (2005) also reported that titles with colons were longer and more informative than those without colons. In a similar vein, Hartley (2007) found that disciplinary differences exist in the use of colons, with a greater use in the Arts than in the Sciences.

1.2.3. Cross-generic studies on academic titles

In her study on 480 journal review papers and 90 RA titles in the Biological and Social Sciences, Soler (2007) identified four main structural constructions, viz., nominal groups, compounds, full sentences and question titles. The most common construction across disciplines and genres was the nominal group. The full-sentence
construction was found to be a generic and disciplinary peculiarity of Biology RAs, whereas the question construction was infrequently used, most commonly though in review papers. Soler moreover analyzed the length of the titles across disciplines, and found that the average number of words in the titles she analyzed was 15.48 words in Medicine, 14.98 in the Natural Sciences, 10.89 in the Social Sciences, and 7.98 in Linguistics.

1.2.4. Cross-cultural research on academic titles

Nord (1995) and Busch-Lauer (2000) investigated the influence of language/culture (here taken as language variations) in academic titles from different perspectives. Nord (1995) examined the functional value of 12,000 titles across various academic publications in English, French, German and Spanish. She recorded the same frequency hierarchy of what she called ‘the optional functions’ of titles and a lack of culture-specific variations in genres like poems and scholarly articles. Busch-Lauer (2000), for her part, evaluated a corpus of 150 German and English titles in Linguistics and Medicine collected from RAs and conference papers, and 25 English titles written by German researchers. She observed that Linguistics titles were shorter than Medicine titles, and that German titles were shorter than English ones. Moreover, Medicine titles written in English preferred a mono-structure format, while those written in German favored a title-subtitle structure. In general, the Medicine titles were long, precise, and informative, while those from the Linguistics field were short, vague, abstract, catchy and reflected individual stylistic preferences.

To sum up, the above review of the literature related to titles of academic writings shows that such titles vary and, at the same time, display similarities across a number of factors and in several dimensions, such as structure, syntactic encoding, length, wording, use of punctuation marks, informativity, functions, and information sequencing. Our review of the literature also shows that there are several gaps in the research on academic titles. We could, for example, mention that, by far, the academic genre that has been "over-investigated" is, as in many other linguistic and rhetorical studies, the RA. By contrast, case report (CR) titles, including, of course, medical CRs, have not been researched at all. The same remark applies as well to diachronic studies on academic titles. Of all the studies we mentioned above, only one (Dillon (1982) dealt with academic titles from a diachronic standpoint.

It is thus our intention here to fill that conceptual gap by presenting the results of a diachronic analysis of a corpus of CR titles from 1840 to the present (see ‘Corpus’ below) and compare them with the results obtained by previous research on titles in other scientific genres, such as the research paper and the review article. More precisely, the present study aims at answering questions related to the evolution of the
type of CR titles, their length, their grammatical and syntactic complexity, and their authorship practices. By examining authorship data, this study seeks to develop, inter alia, a sense of the collaborative practices of medical CR writers.

2. CORPUS and METHODS

We analyzed a corpus of 180 randomly selected CR titles divided into three blocks comprising 60 CR titles each: Block A from 1840 to 1850; Block B from 1920 to 1930, and Block C that covers the year 2009. Titles from Blocks A and B were drawn from one single journal, the *British Medical Journal (BMJ)*. Since the *BMJ* stopped publishing case reports in the late 1990’s, Block C titles were drawn from the *BMJ Case Reports*, which was launched by the end of 2008 and whose 2008 and 2009 issues are freely accessible on line. This explains why we chose the year 2009 as our Block C.

Neither the *BMJ* nor the *BMJ Case Reports* has a stated policy regarding the writing of CR titles. The only policy the *BMJ* has addresses the length of titles and the (non) use of abbreviations.

Twenty-nine variables were recorded in each of the 180 titles. These were divided into two categories: 1) numerical or quantitative variables (those that can be counted), and 2) categorical or qualitative variables that cannot be counted but answer a yes/no question.

The following variables belong to the numerical/quantitative group:

1. Number of authors and their institutional affiliation (from the United Kingdom, from Europe but outside the United Kingdom and from outside Europe);
2. Title length. All the words included in the title were counted. The concept of "word" was defined as the unit occurring between spaces. Each word making up abbreviations was counted as one word: for example, "SAIM [(segmental absence of intestinal musculature] (C)³] was counted as four different words. Compound words and hyphenated words were counted according to the number of their semantic components. For example, "gall-bladder" (B) was counted as two words;
3. Number of titles that start with the expression "(A) case of ....";
4. Punctuation data. The number of commas, colons, semi-colons and full stops;

3 The letter at the end of each example refers to the Block from which the example was drawn.
5. Grammatical data. The frequency of present and past participles, of compound nouns [e.g. "blood serum" (B)] and compound adjectives [e.g. "high-tension current" (B)], of prepositions, coordinating and subordinating conjunctions and relative pronouns.

To the categorical/qualitative group, belong:

1. The absence of author’s affiliation;
2. The different types of authors’ collaboration (local, national or international);
3. The different types of titles:
   – Verbal vs. nominal. A verbal title –also called "informative", "declarative" (Smith 2000: 915) or "assertive sentence title" (Rosner 1990: 108)– contains an active verb with a full sentence that usually states the findings or the conclusion of the research being reported (e.g. "Fibrates help lower LDL"). By contrast, a nominal title, also called "indicative", does not contain any conjugated verb (e.g. "Help of Fibrates in LDL lowering");
   – General subject or "topic" titles, such as "Papilliferous Carcinoma of the Thyroid Gland" (B);
   – Attention-bidding titles, such as "Wired bladder in a cordless era" (C);
   – Question titles, e.g. "Seizure or syncope? A channelopathy with cardiac and cerebral manifestation" (C);
   – Research procedure titles, i.e. those that contain a statement of purpose, method and/or outcome, such as "Case of recovery from taking a large quantity of tincture of opium" (A), which mentions both the method (taking a large quantity of tincture of opium) and the outcome (recovery).

The above-mentioned variables were recorded in each CR title according to the interpretative skills of the authors of this paper. Ambiguous and doubtful cases were measured against the interpretation provided by an English-speaking medical doctor.

3. RESULTS and DISCUSSION

3.1. Title types: Indicative/Nominal Group titles

As we said in the previous section, all the titles we analyzed consist of more or less expanded nominal phrases, also called "indicative titles", which give a straightforward presentation of the object of the study. Here are three examples, one from each Block:
1. Chloroform in catalepsy (A)
2. Tumours of the frontal lobe of the brain (B)
3. Remodelling of coronary arteries (C)

This result clearly corroborates those of previous research on scholarly paper titles that also found a marked preponderance of nominal/indicative titles over verbal/informative/full sentence titles. It is when referring to the _evolution_ of scientific titles that our results do contrast with those of previous research. Indeed most research on the topic has underlined a shift over time towards sentence (informative) titles. Almost twenty years ago, Berkenkotter and Huckin (1995), for instance, already reported that titles of _research articles_ had become more informative over time. The findings of their research indeed showed that in the 1970’s, full sentence titles were very rare, and that in the mid-1990s, they constituted more than 20% of all journal articles and were especially common in biology.

Since we did not find a single case of verbal/informative title in our CR corpus, it would seem that the use of full sentences in scientific title writing is a _generic_ question. Indeed, the research we mentioned above deals with _research and review articles_, whereas ours exclusively focused on CRs. We can thus assert that CR titles –at least those published in the _British Medical Journal_– have always been, and still are, written as nominal phrases.

### 3.2. Title length

The length variable was clearly associated with Block C, which means that the information load and semantic richness of CR titles have increased over time. In this respect, our findings corroborate the results of previous studies. Lewinson and Hartley (2005), for example, reported a 1.25-fold increase in _research paper_ titles between 1970-1974 and 2005-2009, and Goodman (2011) found an approximate doubling in the number of words in _research article_ titles since the 1970s.

The _coordinating conjunction_ variable was found to be intimately related to title length and to be clearly identified with Block C. Indeed, the most numerous the coordinating conjunctions in a title, the longer the title. The two most frequent coordinating conjunctions found in Block C were "and" and "or".

The _colon_ variable was also found to characterize Block C, which means that its use has also increased over time. This too has a direct bearing on title length. It has been shown indeed that titles with colons (also called "compound titles", Hartley 2007) are longer on average and contain more information than titles without them.
In Blocks A and B, colons were mainly used to introduce the findings of a surgical procedure (ex. 4 below) or the surgical procedure itself (example 5):

4. Case of lithotomy: the calculus weighing seven ounces (A)
5. Strangulated umbilical hernia: Resection of gangrenous ileum at the age of 69 (B)

Characteristic of these two Blocks as well, but more frequent in Block A than in Block B, was the use of two colons in the same title, where the first colon introduces the consequence of the event described in the first part of the title (example 6) or a surgical procedure (example 7), and the second precedes the treatment outcome, either death or recovery.

6. Poisoning by Fowler’s solution: Abortion: Mortal fainting (A)
7. Fibroids complicating pregnancy: Hysterectomy: Recovery (B)

By contrast, in Block C, colons are mostly used to underline the rarity of the CR (examples 8 and 9 below):

8. Atypical uterine leiomyoma: a rare variant of a common problem (C)
9. Papulonecrotic tuberculids: a rare cutaneous manifestation of tuberculosis in pregnancy (C)

All in all, our findings lead us to put forward the hypothesis that the longer titles from Block C are explained by the fact that today’s titles require more detailed information about the type of disease and its consequences, the uniqueness of the CR, its educational value and its originality. In short, today more bottom-line information is being loaded into the most highly foregrounded part of any article, i.e. the title.

3.3. Syntactic complexity

Block A titles were generally understandable to the layman. The great majority of Block A titles started with the expression "Case of", as the following example illustrates:

10. Case of traumatic tetanus (A)
11. Case of valvular disease of the heart (A)

Such titles were usually very short and syntactically and semantically rather simple. But CR titles became more and more complex, both semantically and syntactically.

The increasing syntactic complexity and semantic richness of CR titles are not only related to increasing length (see above), but also to the increasing number of compound nouns and adjectives in Block C as a way to condense information (Salager-Meyer 1984). What in Block A or B would have been expressed as "Case of short-
sightedness cured by operation" (A) would in Block C be rendered as "Operation-cured shortsightedness". Here-below are three examples of titles with several compound nouns and adjectives:

12. Eruptive xanthomas with Koebner phenomenon, type 1 diabetes mellitus, hypertriglyceridaemia and hypertension in a 41-year-old man (C)
13. Treatment of chronic bleeding of the small intestine in Rendu-Osler-Weber disease with argon plasma coagulation under double-balloon enteroscopy (C)
14. Thyroid storm induced by trauma due to spear-fishing gun trident impaction in the neck (C)

The higher frequency of compound nouns and adjectives in Block C is directly related to the low frequency of prepositions recorded in that Block. Prepositions, especially of, by, in, and with, were indeed found to be a distinctive feature of Block A titles, as in the following examples:

15. Case of acute laryngitis, with remarks on Dr. Wardele’s cases of spasm glottis (A)
16. Case of emphysema occurring in child-birth (A)
17. Case of varicose aneurism, cured by ligature of the brachial artery (A)

What is interesting to observe, too, is the fact that not only are compound nouns and adjectives more numerous in Block C than they are in Blocks A and B, but they are also longer, as examples 12, 13 and 14 above and examples 18 and 19 below illustrate:

18. Diagnostic difficulty of pulmonary embolus in a bariatric patient and complication of therapeutic dose low-molecular weight heparin to the surgical anastomosis (C)
19. Secondary bronchiolitis obliterans organising pneumonia in a patient with carbamazepine-induced hypogammaglobulinaemia (C)

3.4. Commas and past participles (Block A)

The numerical variables ‘commas’ and ‘past participle’ and the categorical variables ‘mention of methods/treatment’ and ‘mention of outcome’ were all clearly associated with Block A. This is explained by the fact that in the mid-19th century, all past participles expressed either a therapeutic procedure (example 20 below) and/or a surgical outcome (example 21 below) and were preceded by a comma.

20. Case of congenital obliteration of the os uteri, cured by operation (A)
21. Case of compound fracture of the femur, the limb saved by sawing off the end of the bone (A)
These two examples show that at that time much emphasis was put on the treatment administered and/or the surgical procedure performed and their final outcome (see use of colons for introducing results/outcomes in Block A titles, section 3.2 above).

3.5. Title type diversity (Block C)

General subject titles, also called "topic titles", such as:

22. Pneumonic haemorrhagic effusion into pleura (B)
23. The therapeutic value of oxygen in pulmonary lesions (B)

were clearly characteristic of Block B. These titles rather look like editorial titles or titles of oral communications.

Conversely, question titles were found to be more frequent in Block C than in the remaining two Blocks. Here are two examples of question titles:

24. Giant cutaneous melanomas: evidence for primary tumour induced dormancy in metastatic sites? (C)
25. Mesodiverticular band simulating acute appendicitis? (C)

Attention-bidding titles were found to be also clearly more frequent in Block C than in the remaining two Blocks. Apart from the example given in Section 2 above, here is an additional one:

26. "Metallic taste": search for the needle in a haystack (exemplary diagnostic measures and successful minimal invasive endoscopic treatment of a needle-like copper-containing foreign body in the gastric wall) (C)

We can then see that both question and attention-bidding titles, although not very frequent, are more characteristic of today's CR titles than they were in the mid-19th and mid-20th century.

3.6. Authorship and collaboration practices

The institutional affiliation of 9 out of 60 (15%) of CR authors in the mid-19th century was not identified probably because it was "obvious" that they worked at a British institution, the British Medical Journal being a British journal. This practice had totally disappeared by the mid-20th century where all authors' institutional affiliation was mentioned in the CR bylines.

Our findings also disclosed that the total number of authors recorded in Block C was much greater than that recorded in either Block A or B, i.e. it has been increasing over time. Not only has the overall number of authors increased over time, but
so have collaboration practices. There was indeed no collaboration whatsoever in Blocks A and B. Indeed, the absence of the three variables that refer to some kind of collaboration—whether local, national or international—were found to be associated with Blocks A and B. By contrast, their presence was closely related to Block C titles, thus underlining the fact that collaboration is an increasingly important factor in today’s CR writing.

It is finally interesting to observe that the local collaboration variable characterizes Block C more than the national and international variables do. This clearly corroborates the results of very recent research findings that show that physical location seems to influence to an appreciable extent those with whom one will work.

4. CONCLUSIONS

The following factors could account for the various shifts observed: 1) the progressive professionalization of medicine; 2) the need of multidisciplinary teams to conduct an ever-increasing complex research; and 3) the increased specialization and the growing complexity of medical science.

The only variable that has remained constant over the years is the nominal nature of case report titles. In that sense, CRs would distinguish themselves from research article titles.

Despite its appeal, our research is not without its shortcomings. To identify the general trend in CR title formulation, a larger corpus that would cover a wider range of titles from writers from different linguistic backgrounds would be needed. Nationality and mother tongue probably influences title realization. This could be an avenue of research in its own right. Our choice of medical journals also makes us less confident to make broad generalizations on title formation in the discipline and genre analyzed here.

REFERENCES


Françoise Salager-Meyer, María Ángeles Alcaraz Ariza: Titles are "serious stuff": a historical...