

Nominalizing the verb phrase in academic science writing

Douglas Biber and Bethany Gray
Northern Arizona University

(Please note the tables and figures for this paper can be found in a separate Excel file.)

1. Introduction

One of the most distinctive linguistic characteristics of modern academic writing is its reliance on nominalized structures. These include nouns that have been morphologically derived from verbs (e.g., *development*, *progression*) as well as verbs that have been ‘converted’ to nouns (e.g., *increase*, *use*). Almost any sentence taken from an academic research article will illustrate the use of such structures. For example, consider the opening sentences from three education research articles; derived nominalizations are underlined and converted nouns given in *italics*:¹

Now that programed instruction has emerged from the laboratories of experimental psychologists and become a bona fide teaching *aid*, consideration is being given to the expansion and utilization of the media by which *programs* are presented. [1965, Acad-NS]

With the *strain* that increased enrollments are placing on the resources of many colleges and universities, these institutions are becoming more concerned with minimizing the inefficiency arising from student *transfers*. [1965, Acad-NS]

This paper reports an analysis of Tucker's central prediction system model and an empirical comparison of it with two competing models. [1965, Acad-NS]

In most cases, it is easy to restate the information using verbs and clauses. For example:

someone is considering the way in which the media has expanded and is being utilized

students are transferring from colleges and enrolling in other colleges

we analyzed a model that predicts systems, and compared it to other models

However, in academic writing it is more common to use nominalized rather than verbal structures to express information of this type.

Halliday (2004: xvi-xvii) refers to this characteristic of academic writing as a heavy reliance on ‘grammatical metaphor’, where nouns are used with meanings beyond their normal ‘category meanings’. Thus, according to Halliday, the category meaning of nouns is to refer to ‘entities’, but nominalizations refer to ‘processes’ and ‘qualities’, taking on instead the category meanings of typical verbs and adjectives.

Halliday provides several examples to illustrate how discourse in written scientific prose exploits grammatical metaphor. For example, he contrasts alternative styles of expression, as in the following examples from Halliday (2004: 34):

1. *Glass cracks more quickly the harder you press on it.*
2. *Cracks in glass grow faster the more pressure is put on.*
3. *Glass crack growth is faster if greater stress is applied.*
4. *The rate of glass crack growth depends on the magnitude of the applied stress.*
5. *Glass crack growth rate is associated with applied stress magnitude.*

The first of these examples is the most ‘congruent’, where the meanings of words correspond to the expected meanings of the grammatical categories used. For example, the verbs *cracks* and *press* are used to refer to those processes. In contrast, examples 4 and 5 illustrate a dense use of grammatical metaphor, with qualities and processes being expressed by nouns rather than adjectives and verbs.

Halliday (1988; reprinted in Halliday 2004, Chapter 5) has further claimed that science prose followed a historical progression similar to that illustrated in #1-#5 above. Thus, illustrated from scientific texts written by Chaucer (1390), Newton (1704), Priestley (1760’s), and modern texts, Halliday argues that science prose has shifted historically from congruent styles of expression to a dense reliance on grammatical metaphor. This is represented schematically as the following historical development:

A happens; so *x* happens
 Because *a* happens, *x* happens
 That *a* happens causes *x* to happen
 Happening *a* causes happening *x*
 Happening *a* is the cause of happening *x*
 (Halliday 2004: 155)

Although they have been influential, Halliday’s contributions in this area have been primarily theoretical rather than descriptive. Thus, they have been fairly limited in the scope of the grammatical features considered and are not supported by empirical analysis. To address some of these concerns, Banks (2008) adopts essentially the same theoretical framework but undertakes an empirical/quantitative analysis of the historical development of scientific writing. Banks considers three major linguistic features in his analysis: 1st person pronouns, passive verbs, and nominalizations. Regarding the last, Banks’ findings support the conclusion that nominalizations have increased in use historically, in both physical and biological science (see Banks 2008: 124). Most of these nominalizations are used to refer to ‘material processes’ (e.g., *separation*, *emergence*), but ‘mental processes’ (e.g., *calculation*) and ‘verbal [i.e., communicative] processes’ (e.g., *description*) are also important (see Banks 2008: 125-130).²

As noted above, Halliday and Banks focus on the role of nominalizations in academic writing. However, there is actually a suite of other linguistic features that are similar in that they present information through phrasal structures rather than through the use of verbs and clauses (cf. Wells 1960, who described the opposition of ‘nominal’ versus ‘verbal’ styles). For the most part, these features are noun phrase modifiers, including

attributive adjectives (*basic opposition*), nouns as pre-modifiers of a head noun (*government officials, air flow limitations*), prepositional phrases as post-modifiers of a head noun (*a rapid increase in the size of the egg*), and appositive noun phrases (*The root, the only efficient part, is perennial*). Several studies have documented strong historical increases in the use of this collection of features in academic writing (see, e.g., Biber and Clark 2002; Biber 2009, Biber and Conrad 2009, Biber and Gray 2010, 2011; Leech, et al. 2009, Mair 2006).

Multi-Dimensional (MD) studies of register variation (e.g., Biber 1988) have shown that the discourse style of written informational registers is fundamentally different from spoken registers with respect to many of these same linguistic features.³ Thus, 'Dimension 1' from the 1988 study shows that written informational registers rely heavily on nouns, attributive adjectives, and prepositional phrases; in contrast, spoken registers use many verbs (especially present tense and mental verbs) and employ extensive clausal embedding (with finite adverbial clauses and complement clauses) together with a dense use of pronouns, stance features, and reduced structures. Biber and Finegan (1997/2001) use the 1988 MD analysis to document the patterns of historical register change over the past three centuries. With respect to Dimension 1, academic writing has steadily evolved towards increasingly large 'negative' scores, associated with increased use of noun phrase structures; registers like drama and fiction shift historically towards larger 'positive' scores, associated with increased use of verbs, clauses, and other 'involved' features.

However, none of these previous historical studies have focused specifically on the use of verbs. Rather, most of these studies have documented the increased use of nominal structures, with the assumption that there must have been a corresponding decline in use for verbal structures. Thus, processes are referred to with nominalizations rather than verbs, and information is packaged into phrases as nominal modifiers rather than as clauses that incorporate verbs.⁴ The general assumption has been that the historical shift towards a nominal style entails a shift away from a verbal style. However, it is possible in theory for the two to be independent developments. Thus, the first goal of the present paper is to investigate this assumption empirically: Does an increased use of nominal structures correspond to a decreased use of verbs?

The findings presented in Sections 3 and 4 below are based on an analysis of a large-scale corpus, and these investigations generally confirm prior expectations: nominal structures have increased in written informational registers, while verbs and clausal structures have decreased. However, the two trends are not equal in scope: the increase in nominal structures is considerably greater than the decrease in verbal structures. In addition, the findings show the centrality of register influences. That is, the shifts in nominal and verbal styles have occurred to differing extents in various registers. Three situational factors seem to be involved here: the written mode (offering the possibility of extensive planning, revision, and editing), informational purpose, and specialist audience. Spoken registers and written registers that do not have a primary informational purpose (e.g., drama and fiction) have generally not participated in these historical changes. But even written informational registers differ depending on more specific situational factors.

In our study here, we consider four informational written registers that participate in these historical changes to differing extents. At one extreme, science research articles have a highly restricted purpose of presenting new information (the research findings) to

a narrow group of highly specialized readers. In many cases, this information is so ‘new’ that it is completely outside the realm of normal experience for non-specialist readers. Non-science research articles (e.g., in sociology, political science, education, history, etc.) present new information about more familiar topics, like the difficulty level of school reading texts, or the role of churches in World War II. These articles can also be less informational in purpose, including descriptions and argumentative discussion which offer new perspectives on previously known information (e.g., debating the historical importance of the scientific revolution during the seventeenth century). Because the content is somewhat less technical, the readership is also somewhat less specialized. (Thus, educated adults can comprehend most texts across a wide range of disciplines in the humanities and social sciences, in contrast to the impenetrable content of many texts in the hard sciences.) Popular science articles also deal with technical information, but are packaged in a way that is accessible to a wider readership. And finally, at the other extreme, newspaper articles have an informational purpose but are packaged in a way that is accessible to a very wide readership. As we show below, these fine-grained differences in purpose and audience correspond to systematic differences in the extent to which historical linguistic changes have occurred.

One aspect of our findings is especially surprising: the increase in nominal structures is considerably greater in magnitude than the corresponding decrease in verbs and clausal structures. We document these patterns (Section 3), and then focus on a more detailed analysis of verbs, exploring in Section 4 historical changes in the specific sets of verbs used across historical periods. In that section we show that many verbs have decreased in use, including the copula *BE* and other existence verbs, while only a few verbs increase in use over time. First, however, we turn to a description of the corpora and methods for the present study.

2. Corpora and methods

In order to investigate historical change in the use of nominal versus verbal/clausal structures, we analyze diachronic corpora of four registers, described in Table 1. Although our focus in this chapter is on written academic prose, we document overall trends in the use of nominal linguistic features across three other registers by way of comparison: newspaper reportage, fiction, and drama. Academic prose and newspapers represent informational writing, while fiction and drama represent non-informational registers, with drama providing some indication of the characteristics of speech across historical periods.

Table 1. Overall corpus composition

Number of texts				
Register	18th Century	19th Century	20th Century	Total Texts
Academic Prose	76	151	600	827
Newspapers	30	80	180	290
Fiction	47	82	86	215
Drama	23	30	32	85

Size in words				
Register	18th Century	19th Century	20th Century	Total Words
Academic Prose	270,908	924,560	4,011,020	c. 5.2 million
Newspapers	29,420	75,243	141,375	c. 250,000
Fiction	1,481,377	6,164,848	4,185,107	c. 12 million
Drama	21,007	28,510	29,954	c. 80,000

The newspaper corpus is composed of articles from two sources. From 1700-1850, the newspaper texts come from the ARCHER corpus (see Biber, Finegan and Atkinson, 1994). From 1850 on, the corpus includes both texts from ARCHER and texts sampled from a variety of topic areas within *The New York Times* (NYT). The fiction corpus is composed of fictional novels published in English, including text excerpts from ARCHER and full texts collected for the purposes of this study. The drama corpus comes from ARCHER.

The academic prose corpus is composed of texts from a variety of sources. For the eighteenth and nineteenth century sub-corpora, we began with the science/medical research articles included in ARCHER and the *Corpus of English Texts on Astronomy* (CETA, see Crespo and Moskowich 2010), and then supplemented that collection with a sample of texts taken from the *Philosophical Transactions of the Royal Society* (PT, for 1800, 1850, and 1875) and *Science* (for 1875).

In order to more fully investigate linguistic change in the twentieth century, we distinguish among three sub-categories of academic prose: technical science research writing, technical non-science research writing, and popular science writing. These three sub-corpora are detailed in Table 2 below.

Table 2. Sub-corpora within twentieth century academic prose

Number of texts						
Register	1900	1925	1965	1985	2005	Total Texts
Science research	50	50	81	64	52	297
Non-Science research	-	-	88	88	87	263
Popular science	-	-	-	40	-	40

Size in words						
Register	1900	1925	1965	1985	2005	Total Words
Science research	374,692	324,611	467,077	282,503	363,481	c. 1.8 million
Non-Science research	-	-	477,764	581,050	861,111	c. 1.9 million
Popular science	-	-	-	278,731	-	c. 280,000

Twentieth century science writing is represented by texts from ARCHER, *PT*, and *Science* for 1900 and 1925. For 1965-2005, science and non-science writing are represented by the *Twentieth Century Research Article* corpus, which is composed of professional research articles from 11 academic journals. The science writing sub-corpus includes texts in biology, medicine, ecology, and physiology. The non-science sub-corpus includes texts in education, psychology, and history. The final sub-corpus, popular science writing, is represented by texts from the journals *Philosophical Transactions of the Royal Society* and *Science*. While texts from these two journals are used to represent science research writing in earlier periods, we consider the modern (1985) samples from these journals to be ‘popular’ science due to changes in their primary communicative purposes and audience: surveying previous research findings (rather than presenting the results of current research studies), and written for a more generalist audience (rather than specialists in a particular subdiscipline). Thus, for our more detailed analyses of linguistic change in the twentieth century, we consider the three sub-registers within academic prose separately in relation to differences in their audience and informational purpose.

In sum, the four informational registers studied in this chapter (including the three sub-registers of academic writing) vary according to the degree to which their readerships are specialized and the degree to which their purposes are informational and technical (described in Table 3). The comparison corpora, fiction and drama, on the other hand, have primarily wide audiences and aesthetic/entertainment purposes. The results from the corpus investigations will be analyzed in relation to these situational characteristics.

Table 3. Audience and Purpose of Written Registers

Register	Audience	Purpose
Science	highly specialized	highly informational, technical
Non-Science	specialized	highly informational, less technical
Popular Science	wide readership	highly informational, less technical
Newspaper	wide readership	informational, not technical
Fiction	wide readership	aesthetic/entertainment
Drama	wide audience (spoken)	aesthetic/entertainment

All corpora were ‘tagged’ using an automatic grammatical tagger developed by the first author (see Biber 1988). Additional computer programs were written to calculate rates of occurrence for the linguistic features associated with nominal and verbal styles. Rates of occurrence were calculated for each text in the corpus by ‘norming’ the raw counts for each linguistic feature to a standard rate per 1,000 words (see Biber, Conrad, and Reppen 1998: 263-264). To measure nominal style, the computer programs counted the occurrences of nominalizations, nouns, noun + noun sequences, and relative clauses based on information contained in the grammatical tags. A second computer program counted features associated with verbal style. This program identified verbs based on tag information, and counted the number of simple verbs, (no markers of perfect/progressive aspect, passive voice, or modality), perfect aspect verb phrases, passive voice verb phrases, and verb phrases containing modals. In addition, more detailed rates of occurrence were computed for three structures with the verb *be*: *BE* + adjective, *it + is* + adjective + *that/to*, and *there + is*.

Lastly, a third computer program was written to identify the specific verbs that have increased or decreased in use. For this analysis, the program created a list of all finite verbs (based on tag information) occurring in the corpus and calculated frequency information for the use of each verb in each time period. This analysis was restricted to science writing in two time periods, 1900 and 2005.

3. Overall change in the use of nominal and verbal features

As noted in Section 1 above, nominalizations have been studied as one of the most distinctive linguistic characteristic of scientific writing (e.g., Halliday 2004, Banks 2008). However, a historical increase in the use of nominalizations is shared generally across informational written registers, rather than being uniquely characteristic of science. Thus, Figure 1 shows that nominalizations have increased strongly in both academic prose and newspaper writing, while they have actually decreased in use in novels and drama. In fact, as Figure 2 shows, nominalizations have actually increased more in non-scientific academic writing (research articles in education, psychology, history, etc.) than in science research articles. Interestingly, nominalizations have also increased more in popular science articles than in specialist science research articles.

[Figures 1, 2 about here]

It turns out, though, that nominalizations are just one feature among a suite of inter-related grammatical devices that are especially characteristic of present-day informational writing, and that most of these other devices are especially characteristic of specialist science research writing. The most obvious of these other features is simply the overall use of nouns in written registers. As Figure 3 shows, nouns have increased in use in academic research writing and in newspaper prose over the past three centuries, while their use has remained relatively constant in drama and fiction. But unlike nominalizations, the use of common nouns has increased much more strongly in scientific research writing than in non-scientific writing (see Figure 4).

A similar historical trend is found for the use of nouns as nominal pre-modifiers, although the extent of change in the twentieth century is even more dramatic (see also Biber and Clark 2002; Biber and Gray 2009; Leech et al. 2009: 212-218, 303-306). Figure 5 shows that the use of these modifiers is generally rare up until the twentieth century. However, at that point there was a large increase in both academic prose and newspaper writing, and a smaller increase in fiction writing. In contrast, the use of this device in drama remains rare up to the present day. As Figure 6 shows, this increase has been especially pronounced in science research articles (compared to academic writing generally). The historical shift begins at the turn of the twentieth century, and then rapidly takes off over the period 1925-1965, continuing to increase in use right up to the present time.

[Figures 3, 4, 5, and 6 about here]

Other features that function as phrasal modifiers of nouns have followed similar historical developments (see, e.g., Biber and Gray 2011). For example, attributive adjectives have steadily increased in academic prose over the past three centuries, while their use has remained relatively constant in newspaper prose and decreased slightly in drama and fiction. Prepositional phrases as noun modifiers and appositive noun phrases have also increased dramatically over the last century. Prepositional phrases are the most common type of post-nominal phrasal modifier: c. 15 times more frequent than relative clauses in present-day academic prose. Similar to the pattern for nouns as pre-modifiers, the twentieth century witnessed a dramatic increase in the use of PPs as post-modifiers (especially for prepositions other than *of*-genitives), restricted mostly to informational written discourse. As a result, sentences like the following are common in present-day academic prose:

Specifically, we were interested in the qualitative ecological difference in emphasis between changes in composition vs. changes in relative abundance. [2005, Acad-SC]

In contrast to the dramatic increases in the use of nouns and phrasal noun modifiers, there have been only modest decreases in the use of verbs. Figure 7 shows that the use of simple verb phrases (including present tense and past tense, but with no marked aspect, voice, or modality) has remained relatively constant across time in non-science and popular science writing. Only specialist science research writing shows a decrease in use. Perfect aspect verb phrases (Figure 8) are much less common overall and show some fluctuation across periods, but there is a general decrease in use in all academic registers in the twentieth century. Although finite passive voice verb phrases are stereotypically associated with academic writing, Figure 9 shows that they have also undergone a notable decrease in use, especially in the last half of the twentieth century. Finally, the use of verb phrases with modal verbs (Figure 10) follows a similar trend, with an especially notable decrease in science research writing in the last half of the twentieth century.

The general decrease in the use of verbs has also affected the use of dependent clauses, even when they are noun modifiers. Thus, despite the large increase in the use of nouns, the use of finite relative clauses has declined steadily in academic prose, with a more pronounced twentieth century decline in science research articles (see Figure 11).

[Figures 8, 9, 10, and 11 about here]

Overall, the increase in nominal style and decrease in verbal style is mostly a twentieth century development. Nouns began to increase in use slowly in the nineteenth century, and then dramatically increased in use in the twentieth century. Likewise, verbs remained relatively constant in use throughout the nineteenth century, and it has only been in the twentieth century that we see a decrease in use. Thus, there are two general complementary trends in academic writing: an increase in the use of nouns and phrasal noun modifiers, and a decrease in the use of verbs and clausal structures. Those trends can be observed in all written academic registers, but they are considerably stronger in specialist science research writing than in other academic sub-registers.

However, as Figure 12 shows, the two trends are not at all of equal importance. That is, from the nineteenth to the twentieth centuries, nouns have increased in use by nearly 125 per 1,000 words, while verbs have decreased in use by only c. 15 per 1,000 words. Details of this disparity can be seen by comparing Figures 1-6 (for nominal structures) to Figures 7-11 (for verbs), paying attention to the differences in scale. So, for example, nouns as pre-modifiers have undergone a large increase in science research writing: from c. 20 per 1,000 words in 1900 to almost 80 per 1,000 words in 2005 science research writing (Figure 6); in contrast, simple verbs have decreased only from c. 50 per 1,000 words to c. 38 per 1,000 words (Figure 7).

[Figure 12 about here]

Taken together, these findings lead to several additional research questions, including:

1. Specifically which verbs are used in different historical periods? Do we find the same verbs used at the beginning of the twentieth century in science writing as in late twentieth century science writing? In that case, the major historical shift would be just a decline in use for those verbs. Or do we find different sets of verbs being used, accompanying the overall decline in frequency?
2. What are the additional nouns being used in late twentieth century science writing when compared to nineteenth century science writing? To what extent has the range of noun types expanded? Are these noun types newly-coined technical terms, common nouns converted from verbs or adjectives, or other common nouns used with new technical meanings?
3. From a textual perspective, what are the other word classes that have declined in use to offset the dramatic increase in nouns? Thus, Figure 12 can be interpreted in a very mechanical way: In a typical nineteenth century science text, 25% of the words were nouns (250 per 1,000 words), while in a typical twentieth century science text, 33% of the words were nouns (c. 330 per 1,000 words). Thus, an additional 8% of a typical science text is realized as nouns. In contrast, verbs have decreased in use by only 2.5% (i.e. from 9% of the text in the nineteenth century to 7.5% of a typical text in the twentieth century). So the question here is what else has been eliminated from science texts to make room for the large number of additional nouns?

Research questions 2 and 3 are clearly important, but they are outside the scope of the present volume. While we plan to investigate these questions in future work, in the remainder of the present paper we focus on Research Question 1. Thus, we document specifically how the use of verbs has changed historically over the course of the twentieth century. We restrict the analysis to science research writing, since it is the register that has undergone the most extreme shift in use.

4. A comparison of verb use in science research articles: 1900 versus 2005

As Table 4 shows, there has been little historical change in the total number of different verbs used in science research articles from 1900 versus 2005; both historical periods use around 2,500 verb types. Rather, the major historical change has occurred for the number of high frequency verbs. Thus, in the 1900 sub-corpus, 302 different verb types occur with frequencies over 50 times per million words, while only 238 different verb types occur with this frequency in 2005. The difference is even stronger for extremely high-frequency verbs, with over twice as many verb types in 1900 than in 2005.

Table 4. Number of verb types in science research articles in 1900 and 2005

Number of verb types	1900	2005
Total	2,500	2,300
Occurring > 50 times per million words	302	238
Occurring > 500 times per million words	19	10
Occurring > 1,000 times per million words	10	4

To investigate these high-frequency verbs in more detail, we calculated difference scores for the rate of occurrence in 1900 versus 2005. Figure 13 shows the verbs that have decreased the most in use (with a decrease of more than 800 times per million words). Table 5 provides a fuller list, showing all verbs that have decreased by at least 100 times per million words.

Table 5. Verbs that have decreased in use from 1900 to 2005

Decrease > 800 times per million words	are, be, is, found, made, seen, was
Decrease > 250 times per million words	absorbed, become, becomes, been, described, find, formed, give, given, have, obtained, placed, produced, show, shows, taken
Decrease > 100 times per million words	allowed, appears, applied, became, brought, carried, come, concerned, connected, contain, contained, cut, determined, developed, drawn, employed, excited, filled, follows, gave, get, gives, has, kept, known, marked, mentioned, noted, noticed, occurs, pass, passed, pointed, produce, proved, published, reached, referred, represented, said, seem, seems, shown, stated, supplied, take, takes, tied

[Figure 13 about here]

In the following subsections, we explore the discourse factors underlying the reductions in use for the most important of these verbs.

4.1 Copula *BE*

The copula *BE* is by far the most important verb that has decreased in use over the course of the twentieth century. The 3rd person singular form, *is*, has decreased in use by 5,150 per million words. Thus, this verb form alone accounts for over 30 percent of the total decline in verbs from 1900 to 2005. Other forms of *BE* have also declined in use, including *are*, *be*, *was*, and *been*. Taken together, these forms of *BE* have declined by about 8,500 occurrences per million words, accounting for over 50 percent of the total decline in verbs. (Surprisingly, *were* has actually increased in use – see below.)

There seem to be two major trends that have resulted in this large decrease for the copula *BE*. First, meanings that were expressed as simple predicating clauses in 1900 science prose are instead often expressed with phrasal structures in 2005. And second, some of the meanings that were commonly expressed with *BE*-clauses in 1900 are much less commonly expressed in 2005.

Two specific developments can be observed in relation to the general shift from *BE* clausal structures to verbless phrasal structures. First, in 1900, adjectival characteristics are often attributed to a noun using a clausal predicative structure: NP *BE* Adj. These structures have become less common by 2005 (see Figure 14), while the use of attributive adjectives has increased. Second, in 1900 *BE*-clauses with a predicative noun phrase are often used to identify or characterize a subject noun (NP *BE* NP), while

in 2005 science prose, equivalent meanings are often expressed using appositive noun phrases. Both patterns are illustrated below.

[Figure 14 about here]

One of the most common uses of *BE* in the 1900 sub-corpus is to express simple predicative relationships between the subject noun phrase and an adjective phrase following the copula:

BE-clauses with a predicative adjective:

The first term dujdt is small compared with dpjdx, since the **resistance** to the motion **is so great** [1900, Acad-SC]

This reduction in conductivity is nearly entirely due to recombination, since the **tubing is so wide** that the **loss** due to diffusion to the sides **is inappreciable**. [1900, Acad-SC]

it must be recollected that the colour is usually contaminated with the white light which illuminates the prism or grating, and that such **illumination may be very appreciable** at a part of the spectrum where the **luminosity is very small** [1900, Acad-SC]

Predicative adjectives have decreased in use in the 2005 science articles, but similar characteristics are often attributed to nouns using adjectives as pre-modifiers, as in:

Attributive adjectives as pre-modifiers:

One might expect that the increased capacity of the coercive apparatus to take on the workers movement in such a great percentage of its mobilizations, along with its first-hand knowledge of labour's many organizational imperfections, would have led to an attenuation of police brutality. [2005, Acad-NS]

Our Girnock data show that appreciable growth can take place in late winter and early spring. [2005, Acad-SC]

Student SES had a small yet statistically significant effect on student learning (ES = 0.207); a large, significant effect on dropout (ES = -0.610); and a negligible and insignificant effect on student transfer (ES = 0.039). [2005, Acad-NS]

Predicative noun phrases following *BE* are usually more complex than adjectives, and thus they cannot be easily rephrased as simple nominal pre-modifiers. For example,

the velocity of the gases is the same at every point [1900, Acad-SC]
 the solution obtained by the stream-line method is a very close approximation. [1900, Acad-SC]

The living larva of this stage is a very fine object. [1900, Acad-SC]

In many cases, though, these structures can easily be rephrased with a phrasal post-modifier, usually an appositive noun phrase. Thus compare:

Structure with a *BE*-clause

sodium chloride is a plasma constituent
The Tonic is a central tone or klang
The first stage in the process is the
addition of the phenolphthalein indicator

Structure with appositive noun phrase

sodium chloride, a plasma constituent
The Tonic (a central tone or klang)
addition of the phenolphthalein indicator
(the first stage in the process)

One major use of *BE* + predicative noun phrase in 1900 research articles is to identify or characterize elements used in the study or in the text itself (e.g., graphs, figures, photographs, variables).

Fig. 3 is a photograph of a specimen of bismuth

Curve 1 is a diagram on a small scale of the whole...

B is a constant

T is the time, in seconds

The symbols II, R0, li100, are the resistances at the observed freezing point of the alloy,

[examples from 1900, Acad-SC]

In 2005 research articles, appositive noun phrases are commonly used to express similar meanings in a more compressed manner, without the copula.⁵ For example:

Timing of peak growth coincides with the observed peak growth of our wild parr; a period when their growth effectiveness (assimilation minus BMR) was high (fig. 1d) and the period when their natural food was particularly abundant. [2005, Acad-SC]

In multivariate analyses that adjusted for age and sex, renal involvement (hazard ratio [HR] = 1.9; 95% confidence interval [CI]: 1.4 to 2.5), cardiac involvement (HR = 2.8; 95% CI: 2.1 to 3.8), pulmonary involvement (HR = 1.6; 95% CI: 1.3 to 2.2), and anti-topoisomerase I antibodies (HR = 1.3; 95% CI: 1.0 to 1.6) increased mortality risk. [2005, Acad-SC]

As these examples from 2005 research articles illustrate, it is not at all unusual to have multiple appositive noun phrases in the same clause, often modifying a single head noun, and often embedded hierarchically. For example, 'HR' and 'CI' are embedded appositive noun phrases providing acronyms for the variables 'hazard ratio' and 'confidence interval'. These appositive noun phrases all provide additional information about the head noun without the use of the copula *BE*.

The second major trend identified above – that meanings commonly expressed with *BE*-clauses in 1900 are much less commonly expressed in 2005 – relates primarily to the expression of stance. In fact, 20-30 percent of the total occurrences of *BE*-clauses in 1900

science articles are used to express stance meanings. Such meanings are much less commonly expressed in 2005 science articles.

One of the most important types of *BE*-clause used to express stance in 1900 science articles is an extraposed construction, usually with a stance or evaluation expressed as the predicative adjective. For example,

it is necessary to have definite information
 It is convenient to use tubing
 It is important that this rate of escape should be the same
 it is certain that the true curve does not exhibit irregularities
 it is clear that impregnation with lime salts has not gone on to any extent
 It is desirable to smear the edge of the tube with a trace of Vaseline
 it is essential that the time intervals should be referred to the time of origin
 It is evident that the points T and U are not very sharply defined
 it is impossible to refer to them all in detail
 It is interesting to note how like the chlorophyll bodies these cell-inclusions are
 It is obvious that such treatment would necessarily cause considerable changes to take place
 it is possible to detect any uniform changes
 it is probable that links exist
 It is easy to see that M is less than P
 it is difficult to see how they can be accounted for
 [examples from 1900, Acad-SC]

Figure 14 shows that these structures (*it is ADJ that/to*) have decreased strongly in use -- by c. 75 percent -- from 1900 to 2005.

In addition, predicative noun phrases are sometimes used in extraposed structures to express similar kinds of stance meanings:

it is a pleasant duty to acknowledge to what a very unusual extent the carrying out of my enquiry has been due to the co-operation of many friends.
 It is a mistake to suppose that the lime cells disappear from a fasting gland.
 It is a matter of some interest to determine the relation which exists between the area of an open stomate and the superficial area of the cells of the spongy parenchyma
 It is a curious and rather unaccountable feature common to all enamel development that [...] there is very little organic matter present in the finished tissue
 [examples from 1900, Acad-SC]

Even more surprising is the extent to which simple predicative structures with *BE* are used to express stance meanings in 1900 science articles. Thus consider the following typical examples:

This restriction is necessary
 The following experiment is a striking example.
 investigation is a good thing
 this matter is one of interest rather than importance...

the foregoing general indications are of much interest.
 the amounts of it recorded are of considerable significance
 " villus factors " are of doubtful value
 [examples from 1900, Acad-SC]

In many cases, these structures are used with a pronoun as subject. In these occurrences, an object or idea has been previously described, and the author is now using a *BE*-clause simply to add an evaluative comment; for example:

this is a magnificent section
 This is a deduction of some moment which has certain important applications.
 This is a most beneficent and dignified attitude
 This is the view I am inclined to take.
 some are of excellent quality
 They are of the utmost importance
 [examples from 1900, Acad-SC]

In other instances, a stance evaluation is expressed in the NP-subject rather than in the subject predicative:

this peculiarity is a function of the normal condition of the gut wall,
 A curious feature is the greater variation in c
 the crucial point is the serial number of the spinal nerves
 [examples from 1900, Acad-SC]

And in a few cases, a stance evaluation is expressed in both the NP-subject and the subject predicative:

This coincidence is a remarkable feature in the relations of gold to aluminium
 The obvious explanation of these phenomena is the assumption that the forward shifting of the fin is partially arrested in the male owing to the development of the mixipterygium
 [examples from 1900, Acad-SC]

Stance meanings are also commonly expressed in 1900 science articles when the subject predicative is a *that* complement clause. In this case, the grammatical subject of the *BE*-clause gives the stance that frames the proposition in the *that*-clause. For example,

Another case of interest is that of the nuclei produced by the discharge from a point.
Another suggestion in the paper is that the principal disturbance in an earthquake is not purely elastic,
A point to be noticed here is that the disparity obtaining between the two proximal and the two distal branches is not nearly so marked in older embryos as in the adult.
a more probable explanation in my opinion is that at the first removal the anterior parts of the visuo-sensory areas [...] were not removed

His thesis is that the rapidity of the intestinal absorption of salts varies as the rapidity of diffusion of the salts

Mr. Foerste's theory is that these deposits at one time were continuous across the present anticline

[examples from 1900, Acad-SC]

In addition to the above structures, a more specialized type of *BE*-clause was much more common in 1900 science articles than in 2005 articles: existential *there* constructions. As Figure 14 shows, these structures were over twice as common in 1900 than in 2005.

Existential *there* + *BE* constructions are usually described as a device used for focus and information packaging (see, e.g., Biber et al. 1999: 951-954.), and they do commonly serve that function in 1900 science articles; for example:

There is an almost exact correspondence between the thickness of the conjoined 1st and 2nd layers of the cortex, and the degree of amentia or dementia existing in the patient.

There is thus brought together in this Mexican Hall of the Museum the most important collection in existence for the study of the ancient civilization of Mexico and Central America.

The entrance to the lagoon is perhaps 150 feet wide, and, there is a cut through the beach rock covering the old ledge,

There are a few little strings of sericite between the calcite grains at intervals.

There are also changes in the amount of recombination of the ions and in the diffusion effect.

[examples from 1900, Acad-SC]

Surprisingly, though, an equally important use of existential *there BE*-clauses in 1900 science articles is to provide an evaluation, most often an epistemic assessment, as in:

there is no doubt but that the lagoon is rapidly filling from the sand blown in by the dunes.

there is no evidence of a convulsive centre in the pons.

there is some probability that in these tube experiments a similar relation may exist which would account for the diffusive flow...

[examples from 1900, Acad-SC]

Other examples of this type express various attitudinal stances; e.g.:

there is a very real danger in over-guidance as well as in license.

There are several interesting results in connection with molecular physics which can be obtained...

There is less substantial agreement, however, [...] as regards the question of definite retinal localization

[examples from 1900, Acad-SC]

Thus, to a large extent, the reduced use of existential *there* constructions in 2005 articles relates to the general reduction in the extent to which epistemic and attitudinal stance is expressed in those articles.

In sum, simple *BE*-clauses of all types are commonly used as a device to attribute a stance to some idea or entity in 1900 science articles. This function is much less common in 2005 articles, resulting in a lesser overall use of *BE*-clauses.

4.2. *Made, found, seen*

In addition to the copula *BE*, Table 5 above shows that three other verbs have undergone large declines in use from 1900 to 2005 science writing: *made*, *found*, *seen* (see also Figure 13). Although they have different literal meanings, all three verbs are used in 1900 science articles for the general function of describing the research process or calling attention to particular results.

Most occurrences of the verb *made* in 1900 science articles are in the passive voice. In a few cases, *made* has the literal meaning of creating or producing something, as in:

a unilateral section was made of the posterior columns
sketches of both surfaces were made for future orientation
[examples from 1900, Acad-SC]

Much more commonly, though, the verb *made* is used to identify steps in the research process, as in:

another series of experiments was made to determine the influence of concentration...
expansion experiments were made with moist air...
[examples from 1900, Acad-SC]

In most of these cases, the verb *made* has little semantic content of its own, while the processes involved are expressed through nominalizations occurring as the subject noun phrase; for example:

The **measurements** were made with a quadrant electrometer
microscopic **observations** were made with ZEISS' apochromatic lenses,
comparisons were made by adding 10 milligrams...
[examples from 1900, Acad-SC]

In many instances, the subject noun phrase refers to a mental process rather than a physical action, as in:

such **determinations** can be made
 an **attempt** was made to determine the value of k
 this **comparison** can be made by calculating the amounts...
 [examples from 1900, Acad-SC]

The same major functions are found with non-finite occurrences of *made*, used as a postnominal modifier; for example:

a series of experiments made with the diaphragms
Measurements made with a mixture of yellow and violet
observations made with a grating spectrum
 the number of readings made in each case
 [examples from 1900, Acad-SC]

The verb *found* is similarly rarely used in the literal sense of locating an object. Rather, *found* in 1900 science articles usually identifies a research result that someone has observed or discovered:

He found that the diffusivity coefficient was in all cases approximately inversely proportional to the geometrical means
 we have found that a normal solution of sodium hydroxide may be employed as an absorbing medium
 [examples from 1900, Acad-SC]

Like *made*, the verb *found* most often occurs in passive voice, in both finite and non-finite forms:

no absorption of introduced solution is found to occur,
 the excitability is found to vanish in the course of 1 minute
 It was not found possible to follow this last branch (b) far up the nerve to which it ran
 it was found that each division corresponded to a charge of '0042 electrostatic unit

the soda found in the produce had been taken up as nitrate of soda.
 The colours of the various substances found in an alloy were then fairly constant

thus sometimes small granules may be found, at other times the spheres may come to resemble closely those found in the ferment cells.
 [examples from 1900, Acad-SC]

In some cases, the verb *seen* can be almost interchangeable with *found*, usually occurring in the passive voice and referring to research results or observations:

Periods of increased respiration are seen now and again
 Many of them are seen to be in the act of disappearing
 [examples from 1900, Acad-SC]

However, *seen* usually has a more specialized function in 1900 science articles, focusing specifically on what the paper itself demonstrates, rather than on what has been observed in the research process. For example:

This hypothesis was therefore put to the test, but, as seen in Experiment VII, it is found wanting...

Slight excess of salt absorption over that of water has already been noted [...], and, as seen in Table VIII, it is a more marked feature in the lower than in the upper ileum...

Fig. 33 (Plate 20) is a field diagram, for a hollow square cylinder ; the shielding effect is seen to be very powerful.

The results are represented in fig. 6... It is seen that the velocities decrease with T, and nearly in a linear manner.

It was seen in Section 5 that the time of passage of the ions from one cylinder to the other could be varied by changing the velocity of the gas stream
 [examples from 1900, Acad-SC]

In this function, the verb *seen* often co-occurs with a modal verb:

Alloys of this composition can be seen in the photograph
 The individual grains can be seen to be distinctly flattened in a horizontal direction
 In stage 11, as can be seen in the whole egg (fig. lis), the segmentation cavity has been much diminished

the general result is the same, as may be seen from the following tabular statement
 The line of GENNARI also exists in anophthalmos, as may be seen by examination of the micro- photographs in Plate 11

In sections 8 and 9 it will be seen that the special lamination has not yet risen to the surface of the brain
 It will be seen from what follows that, when $r = 0$, N becomes infinite
 [examples from 1900, Acad-SC]

In 2005 science research articles, it is much more common to simply report research results, with few overt signals to the reader that identify a statement as research findings. The following sentences from 2005 biology/medical articles are typical, with no overt framing to tell the reader that these results have been *found* or that they can be *seen* in the associated figures and tables.

In centrobin-depleted cultures, 21% of interphase cells had four centrioles, 67% had two centrioles, 7% had one centriole, and 2% had no centriole (Fig. 6 A).

At the end of the experiment broom psyllids were present in 22 of the 55 release sites (Table 2).

Patients who underwent autopsy were younger, less likely to have had a stroke before their index MI, less likely to die of an unwitnessed clinical sudden death, and more likely to die in the hospital than were patients who died but did not undergo autopsy (Tables 1 and 2).

[examples from 2005, Acad-SC]

As a result, the verbs *made*, *found*, and *seen* have sharply declined in use from 1900 to 2005.

4.3 Other common verbs that have decreased in use in twentieth century science articles

The copula *BE* can be classified as the most common verb of ‘existence or relationship’ (see Biber et al. 1999: 364); *BE* has declined in use over the course of the twentieth century because simple relationships are instead commonly expressed through phrasal juxtaposition (see Section 4.1 above). It turns out that several other verbs that have declined in use over the twentieth century belong to this same general semantic category, including *appears*, *become/became*, *follows*, *have/has*, *represented*, and *seem/seems* (see Table 5 above). Thus consider the following typical examples from 1900 science articles:

it appears that slip-bands occur in all metals

the crystalline grains become visibly differentiated
With the continuance of the pressure the rock became plastic

...from which it follows that the series (3) is convergent.
The average of the observations works out as follows : [...]

Hence the curve must have a form similar to a continuous line
p has the value given in Equation (8)

...an amount represented by 20 on the same scale

the difference seems greater than can be attributed to instrumental errors
it seems impossible to explain how the electrode B could receive a positive charge ...
[examples from 1900, Acad-SC]

Various phrasal devices that have increased in use over the twentieth century can be used as alternatives to these verbs. For example, clauses with the copular verbs *seem* and *appear* usually express the simple existence of a state, together with an epistemic stance;

these verbs can often be rephrased with stance adverbials like *apparently* or *possibly*, or by stance nouns like *likelihood* or *possibility*; compare:

it appears that slip-bands occur in all metals [1900, Acad-SC]
slip-bands apparently occur in all metals

Clauses with *HAVE* can sometimes be rephrased using a prepositional phrase headed by *with*:

Etching a polished surface develops a multitude of facets which have the same orientation over the surface of any one grain [1900, Acad-SC]
a multitude of facets with the same orientation

Clauses with *BECOME* can be rephrased as noun phrases headed by nouns like *change*, *shift*, *transformation*, etc.:

the rock became plastic [1900, Acad-SC]
a change to a plastic state

While not all instances of these verbs can be rephrased, 2005 science articles have a strong tendency to express relationships using phrasal rather than clausal devices, resulting in the decline in use for a wide range of relationship verbs.

In addition, many of the other verbs that have declined in use are similar to *made*, *found*, and *seen* (discussed in Section 4.2 above) in that they were commonly used to describe the research process or research results in 1900 science articles. These other verbs include: *described*, *employed*, *find*, *give/gives/given/gave*, *obtained*, *occurs*, *produce*, *proved*, *show/shows/shown*, *take/takes/taken*. Thus consider the following examples from 1900:

Stanton described the relations of the fossils brought up by the well-borings to others from Florida.

nitrate of soda was employed

it was considered simpler to find the amount of air in the hydrogen after each experiment

the diagrams of Curves 2 and 5 give a record of all the freezing points observed
 This gives, for the velocity under a potential gradient of 1 volt per centimetre, less than 1/4000 centim. per second

At this moment the still liquid part of the alloy has reached the composition given by G.

The results were obtained by the use of grass nodes

supersaturation occurs in connection with thunderstorms

The results show that isotonic solutions of these two salts produce very different results. [sic]

In the case of iron it was proved definitely that under tensile stress the bands appear as soon as the yield point is passed.

All these cases show a greater amount of oxidizing enzymes
The above experiments show that there are cn_2 positive ions which pass the section of the tubes.

the results are shown in fig. 2

The time taken for the ions to be removed is somewhat longer ...
the encounters would not have taken place if the carrier were uncharged.
recombination can take place
[examples from 1900, Acad-SC]

Many of these same verbs continue to be used in 2005 science articles, but with lower frequency. As described in Sections 4.1 and 4.2 above, there are two main factors that have influenced this shift: a tendency to express information in phrases rather than clauses, and a tendency to simply not express other kinds of information. Thus compare the density of verbs in the following passages from 1900 and 2005 science research articles:

1900 science article

Before the coefficients of diffusion can be calculated from the above observations it is necessary to make a correction for the loss of conductivity due to recombination. The above experiments show that there are cn_2 positive ions which pass the section of the tubes, T1, at a distance of 1 centim from the end near the window, W14. Of these, cn_2 ions are lost in the remaining 9 centims. of the tubes. The loss is principally due to the ions coming into contact with the sides of the tubes; but the loss is also to a small extent due to collisions between positive and negative ions ; it is necessary to find how much the observed value of n_1 must be increased in order to compensate for the loss of ions arising from recombination.

2005 science article

There was no detectable effect of egg length ratio on the success of offspring under competition (Table 1). Furthermore, there was no influence of strain type on the outcome of competition in this experiment (Table 1). Relative numbers of individuals in the two broods had the strongest effect on the out-come of within-host competition (Table 1). Relative offspring size showed direct frequency dependence, where the largest offspring in a given host were predictably members of the more numerous brood inside that host (Fig. 3). There was no significant interaction between any of the main effects in the model (Table 1).

Most sentences in the 1900 excerpt are composed of a main clause plus one or more dependent clauses, resulting in the use of multiple verbs. In contrast, embedding is done with prepositional phrases or other phrasal nominal modifiers in the 2005 passage, resulting in a typical sentence structure with only one lexical verb. Even many main clause functions are expressed as phrases in the 2005 articles. For example, the specific research results are consistently referred to with an appositive noun phrase in parentheses (e.g., *Table 1, Fig. 2*), replacing clausal references like *Table 1 gives/shows...* or *As seen in Fig. 2...* Further, there is less discourse framing and less clausal expression of stance in the 2005 articles. For example, the 1900 excerpt has two occurrences of *it is necessary to...*, while the 2005 excerpt has no overt stance expressions at all. In sum, the discourse style of 1900 research articles is already notably different from most other spoken and written registers in English. However, we see a marked shift towards even more distinctive styles over the course of the twentieth century. From a grammatical perspective, there are two main correlates of that shift: an increase in the use of nominal/phrasal structures, and a decrease in the use of verbal/clausal structures.

4.4. Verbs that have increased in use in twentieth century science articles

Although the strong overall trend has been for verbs to decrease in use over the twentieth century, there have been some verbs that have undergone smaller increases in use. As Table 6 shows, the most frequent of these have increased by only c. 250 times per million words, in contrast to the much larger decreases in use seen for verbs like *is, are, be, was, made, found, and seen* (Table 5). Nevertheless, there are some interesting trends here.

Table 6. Verbs that have increased in use from 1900 to 2005

Increase > 250 times per million words	analyzed, had, performed, used, were
Increase > 100 times per million words	affect, associated, assumed, based, calculated, compared, estimated, expressed, follow, generated, included, incubated, indicate, indicates, provide, purified, recapture, regulate, regulates, reported, required, see, suggest, suggests, supported, tested, treated, use

First of all, there seems to be a general shift towards increased use of past tense forms in 2005 (versus the decrease in the use of present tense and non-finite forms for the verbs listed in Table 5 above). This shift reflects a stylistic preference for reporting results in past time, perhaps emphasizing the time-sensitive nature of research findings. In contrast, 1900 articles had a relative preference for present tense reporting of results (see above), suggesting greater generalizability of the findings.

The shift is clearly shown for the increased use of *had* (versus the decrease in use for *have* and *has*); for example:

Deletion of YCA1 had only little effect on toxin sensitivity.

Of the 60 patients, 1 had no mutation, 3 had one mutation, 7 had two mutations, and the remaining 49 had at least three mutations.

[examples from 2005, Acad-SC]

However, many of the other *-ed* forms that have increased are usually used in 2005 articles to express simple past tense rather than perfect aspect or passive voice; for example:

In this work, we analyzed skewing of X inactivation in mice

We performed time-lapse microscopy to further characterize the dynamics of RecA threads.

We used this method to analyze replication timing of the Xic locus

For our model, we assumed that by-stander CPR doubled the probability that an out-of-hospital cardiac arrest patient survived to hospital discharge

Mannino and colleagues reported rates of "low lung function" in never smokers...

[examples from 2005, Acad-SC]

The increase for *were* is more specialized. Almost all other forms of *BE* have decreased strongly in use, including present tense *is* and *are*, as well as non-finite *be*, past tense *was*, and past participle *been*. In contrast, the increased use of *were* reflects a very specific function: past time reference with a plural subject. These clauses usually express generalizations, using a plural noun with generic reference as the grammatical subject. For example:

mutant cells were significantly more sensitive against killer toxin treatment than the isogenic wild type

local cases were far more common than referred patients

[examples from 2005, Acad-SC]

The increased use of the verb *use(d)* is interesting because it is a high frequency verb with general reference. That is, it is kind of a general purpose activity verb for identifying methodological procedures. For example:

We used this method to analyze replication timing of the Xic locus.

We used the Akaike Information Criteria corrected for small sample size to rank the alternative models.

Gut contents analysis was used to characterize prey in the diet

Age-specific patterns of spots and coloration were used to identify *P. chapoda* as spiderlings.

[examples from 2005, Acad-SC]

Apparently the methods used for research procedures are more commonly identified in 2005 articles than in 1900 articles; the only comparable verb that has decreased in use is *employed* (see Table 5). However, more detailed consideration of these texts is required to confirm that possibility.

Finally, it is worth mentioning two stance verbs that have increased in use over the twentieth century: *indicate(s)* and *suggest(s)*. These are both hedging verbs, used to report research results that might need to be modified based on future research. For example,

The results indicate that coatomer-bound Cdc42 negatively regulates dynein
 These results suggest that ROS act as effectors of apoptosis in toxin-treated cells
 [examples from 2005, Acad-SC]

Interestingly, most of the epistemic stance verbs that were more common in 1900 expressed greater certainty, including *found*, *seen*, *show(s)*, *give(s)*, *known*, *proved*. Thus, the change here has been not only towards a reduced use of stance verbs, but also a shift towards a preferred use of hedging verbs rather than certainty verbs.

6. Summary and conclusion

The corpus investigations we have reported on in this chapter have shown that an increased use of nominal/phrasal structures has been a general historical change in all informational written registers, but that this change has not affected non-informational registers (spoken or written) like fictional novels and drama. In contrast, the corresponding decrease in verbal/clausal style has been much more restricted, occurring primarily in professional academic research articles, particularly science research articles. Thus, it appears that the combination of a highly specialized audience and a highly informational purpose dealing with technical information is related to the decrease in the use of verbs and clauses.

Our analysis of the particular verbs used in twentieth century science research articles showed that the overall decrease in frequency is due to a reduced use of high-frequency verbs, rather than a reduction in the number of verb types. Additionally, our analysis has shown that a reduction in the use of the copula *BE* and other existence verbs (e.g., *appears*, *become*, *follow*, *seem*, etc.) accounts for over half of the overall decline.

In Section 3, we identified three research questions arising from previous research. We have addressed the first of these in this paper. However, the second and third questions remain to be addressed. In light of our general finding that the decrease in use of verbal/clausal structures has not occurred with the same magnitude as the increase in nominal features, the third research question is particularly of interest. What other word classes have declined in use to offset the dramatic increase in nouns? What else (besides verbs) has been eliminated from science texts to make room for the large number of additional nouns? A related question is to consider how scientific discourse has changed over the course of the twentieth century, both linguistically and situationally. Both of these

questions focus on identifying how discourse has changed to compensate for the increases in nominal/phrasal structures prevalent in scientific academic writing. In on-going research, we are investigating these additional questions, with the goal of a complete account of these recent historical changes.

Notes

1. It is not always easy to determine the basic form of converted nouns or to distinguish between derivational nominalizations and converted nouns. Thus, *aid* is probably a clear example of a verb converted to a noun, but *program* might be better considered as a noun which can also be used as a verb. Regardless, the important point here is that the nominal use is typically preferred over verbal uses in academic writing.
2. Generalizations from Banks' (2008) study should be treated with caution, because the sample of texts analyzed was extremely small: only one biological science article and one physical science article for each 20-year period. As a result, the historical trends reported in the study sometimes have large fluctuations, reflecting the idiosyncratic characteristics of an individual author/text.
3. Multi-Dimensional (MD) analysis is a research approach designed to capture the patterns of linguistic variation among registers. The approach is based on analysis of a relatively comprehensive set of lexico-grammatical features in a large, multi-register corpus. Those features are reduced to underlying 'dimensions' of variation using a factor analysis, which identifies sets of linguistic features that tend to co-occur in texts. The dimensions are interpreted in functional terms, and then they can be used to compare the similarities and differences among registers. Detailed introductions to the approach are provided in Biber (1988) and Biber and Conrad (2009).
4. MD studies are a partial exception here, because they do incorporate both nominal and verbal structures. However, because the analytical focus is on clusters of co-occurring linguistic features, it is not possible to isolate the influence of any individual feature. Thus, the increasingly large 'negative' Dimension 1 scores for academic writing could be caused by one of two statistical patterns: a large historical increase in the use of noun phrase structures (the 'negative' features) with little change in the use of verbs, clauses, and 'involved' features (the 'positive' features); or a smaller historical increase in the use of noun phrase structures (the 'negative' features) with a corresponding decrease in the use of verbs, clauses, and 'involved' features (the 'positive' features).
5. These appositive noun phrases in present-day academic writing are often not strictly co-referential with the head noun, representing a major extension in function as well as frequency over earlier historical periods (see also Biber and Gray 2010).

References

- Banks, David. 2008. *The development of scientific writing: linguistic features and historical context*. London: Equinox.

- Biber, Douglas. 1988. *Variation across speech and writing*. Cambridge: Cambridge University Press.
- Biber, Douglas. 2009. Are there linguistic consequences of literacy? Comparing the potentials of language use in speech and writing. In David Olson and Nancy Torrance (eds.), *Cambridge handbook of literacy*, 75-91. Cambridge: Cambridge University Press.
- Biber, Douglas and Victoria Clark. 2002. Historical shifts in modification patterns with complex noun phrase structures: How long can you go without a verb? In Teresa Fanego, María José López-Couso and Javier Pérez-Guerra (eds.), *English historical syntax and morphology*, 43-66. Amsterdam: John Benjamins.
- Biber, Douglas and Susan Conrad. 2009. *Register, genre, and style*. Cambridge: Cambridge University Press.
- Biber, Douglas, Susan Conrad and Randi Reppen. 1998. *Corpus linguistics: Investigating language structure and use*. Cambridge University Press.
- Biber, Douglas and Edward Finegan. 1997. Diachronic relations among speech-based and written registers in English. In Terttu Nevalainen and Leena Kahlas-Tarkka (eds.), *To explain the present: Studies in changing English in honor of Matti Rissanen*, 253-76. Helsinki: Societe Neophilologique. [reprinted in Conrad and Biber 2001, 66-83].
- Biber, Douglas, Edward Finegan, and Dwight Atkinson. 1994. ARCHER and its challenges: Compiling and exploring "A Representative Corpus of of Historical English Registers". In Udo Fries, Gunnel Tottie, and Peter Schneider (eds.), *Creating and using English language corpora*. Amsterdam: Rodopi, 1-14.
- Biber, Douglas and Bethany Gray. 2011. Grammatical change in the noun phrase: The influence of written language use. *English Language & Linguistics*, 15(2), 223-259.
- Biber, Douglas and Bethany Gray. 2010. Challenging stereotypes about academic writing: Complexity, elaboration, explicitness. *Journal of English for Academic Purposes* 9, 2-20.
- Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad, and Edward Finegan. 1999. *Longman grammar of spoken and written English*. London: Longman.
- Crespo García, Begoña and Isabel Moskowich-Spiegel Fandiño. 2010. CETA in the context of the *Coruña Corpus*. *Literary and Linguistic Computing*, 25(2), 153-164.
- Halliday, M.A.K. 2004. *The language of science*. London: Continuum.
- Leech, Geoffrey, Marianne Hundt, Christian Mair, and Nicholas Smith. 2009. *Change in contemporary English: A grammatical study*. Cambridge: Cambridge University Press.
- Mair, Christian. 2006. *Twentieth-century English. History, variation and standardization*. Cambridge: Cambridge University Press
- Wells, Rulon. 1960. Nominal and verbal style. In Thomas A. Sebeok (ed.), *Style in language*, 213-220. Cambridge, MA: MIT Press.