1

Academic writing: Challenging the stereotypes

1.1 Academese: Obtuse or informational?

Many speakers of English share the view that the language of academic writing is peculiar, not only different from everyday speech, but also different from most other registers of English. These perceived differences are not neutral. Rather, a common stereotype of academic prose is that it is deliberately complex, and more concerned with impressing readers than communicating ideas – all making it needlessly difficult to understand.

These attitudes are reflected in the label *academese*, which is usually defined with negative connotations. For example, the *Tameri Guide for Writers* (www.tameri.com) defines *academese* as 'an artificial form of communication commonly used in institutes of higher education designed to make small, irrelevant ideas appear important and original. Proficiency in academese is achieved when you begin inventing your own words and no one can understand what you are writing.'

In fact, the suffix *–ese* seems to have taken on this broader meaning when attached to other roots. Thus, according to the on-line site www.wiktionary.org, *officialese* is 'the typical language of [written] official documents, legalistic and pompous'; *legalese* is 'wording that resembles how a lawyer writes, especially such that is confusing to the layperson'; and *bureaucratese* is 'any language containing many non-essential words intended to imply more importance or intelligence than actually present'.

The primary focus of these negative attitudes is on the use of rare and obscure words and phrases. The perception is that these words are often chosen to impress readers rather than efficiently convey new information. The portrayal of academic prose in similar terms has been around for some time, as evidenced by Orwell's famous critique in 1946:

> The writer either has a meaning and cannot express it, or he inadvertently says something else, or he is almost indifferent as to whether

> his words mean anything or not. This mixture of vagueness and sheer incompetence is the most marked characteristic of modern English prose. . .
>
> In certain kinds of writing, particularly in art criticism and literary criticism, it is normal to come across long passages which are almost completely lacking in meaning.
>
> Orwell, George. 1946.
> 'Politics and the English Language'
> Horizon, 13(76): 252–265

Here we see Orwell being especially critical of academic writing in the humanities (specifically art criticism and literary criticism). The enterprise of research in the humanities is focused on our everyday experience. Unlike scientists, humanities researchers are not discovering and documenting new natural phenomena and processes. Rather, their work is more interpretive, attempting to describe and understand the human experience. Like the author of Ecclesiastes, readers might believe that 'there is nothing new under the sun' when it comes to discussions of the human experience – and as a result, expect that there would be little need for new technical vocabulary and complex grammatical constructions to discuss these familiar ideas, relationships, and experiences. This seems to be Orwell's underlying complaint: humanities researchers mostly discuss phenomena that are already familiar to us all, but they do it in a style that is indecipherable to non-specialist readers.

These stereotypes about academese in humanities writing continue right up to the present. For example, Pinker (2014: para 4) discusses the question of 'Why academics stink at writing?', and notes that:

> The most popular answer outside the academy is the cynical one: [. . .] Scholars in the softer fields spout obscure verbiage to hide the fact that they have nothing to say. They dress up the trivial and obvious with the trappings of scientific sophistication, hoping to bamboozle their audiences with highfalutin gobbledygook.

Texts that seem to illustrate this prose style are unfortunately not hard to find. For example, the journal *Philosophy and Literature* sponsored a 'Bad Writing Contest' from 1995–1998, which celebrated 'the most stylistically lamentable passages found in scholarly books and articles published in the last few years' (http://denisdutton.com/bad_writing.htm). The contest lampooned academese simply by quoting the academic prose of famous literary scholars. For example, the winner of the 1998 contest was Judith Butler, 'a Guggenheim Fellowship-winning professor of rhetoric and comparative literature at the University of California at Berkeley'. The first-prize sentence, singled out for its 'anxiety-inducing obscurity', reads as follows:

Text Sample 1.1
Literary criticism article

> The move from a structuralist account in which capital is understood to structure social relations in relatively homologous ways to a view of hegemony in which power relations are subject to repetition, convergence, and rearticulation brought the question of temporality into the thinking of structure, and marked a shift from a form of Althusserian theory that takes structural totalities as theoretical objects to one in which the insights into the contingent possibility of structure inaugurate a renewed conception of hegemony as bound up with the contingent sites and strategies of the rearticulation of power.

Judith Butler. 1997.
'Further Reflections on the Conversations of Our Time'
Diacritics, 27(1): 13–15.

As noted earlier, the popular explanation for academese in humanities writing is that it represents a form of deliberate obscurity. Pinker (2014), however, proposes a slightly different explanation, based on the underlying goals of humanities versus science research writing. Pinker describes the primary goal of humanities writing as 'self-presentation' rather than the communication of information. This goal results in a 'self-conscious style', where the author's primary concern is 'to escape being convicted of philosophical naïveté about his own enterprise'. In contrast, science research writing is associated with a 'classic style' focused on the communication of information: 'The writer can see something that the reader has not yet noticed, and he orients the reader so she can see for herself. The purpose of writing is presentation, and its motive is disinterested truth. [. . .] The writer knows the truth before putting it into words; he is not using the occasion of writing to sort out what he thinks' (Pinker 2014: para 11).

The linguistic characteristics of the 'classic' style associated with science writing are dramatically different from the 'self-conscious' style of humanities writing. Science prose can be boring, relying on simple syntactic constructions, as in:

Text Sample 1.2
Biology research article

> The neurites are black on a yellow-brown background in the original preparation (see Figure 5). One neurite can be traced coursing through the basement membrane of the epidermis (arrow). The neurites appear to penetrate the cytoplasm of the epidermal cells (see also Fig. 11).

Bryce L. Munger. 1965.
'The Intraepidermal Innervation of the Snout Skin of the Opossum'
Journal of Cell Biology, 26(1): 79–97.

Although there are often technical terms in this style, those terms usually refer to physical entities rather than abstract concepts. The overall stylistic effect is to emphasize the direct communication of information (even if that information is still unclear to non-specialist readers because of the technical vocabulary).

Our goals in the present book are not to enter into the debate about 'good' and 'bad' academic writing, or the motivations for the 'academese' found in many academic texts. But we *are* interested in the linguistic characteristics of academic writing, including the differences between humanities writing and science writing. And we are especially interested in the ways in which academic writing has changed historically in its grammatical style.

Given the striking linguistic differences between present-day humanities writing versus science writing, readers probably will not be surprised to learn that academic writing styles have undergone major linguistic change over the past few centuries. In particular, it would be easy to assume that academic writing has become increasingly complex over time, and that humanities writing has been especially influenced by these changes, resulting in the grammatically-elaborated styles of academese illustrated in Text Sample 1.1. At the same time, it might be assumed that science prose has somehow resisted these historical changes, and thus retained the grammatically-simpler styles of earlier centuries.

It turns out that those expectations are wrong in nearly every way. In the following chapters, we show that:

- There are different types of grammatical complexity.
- Complexity in humanities writing is associated with structural elaboration.
- Humanities academic writing has changed little over the last few centuries in its structurally elaborated grammatical style.
- Complexity in science writing is associated with structural compression, not structural elaboration.
- Science writing has undergone extreme historical change related to the use of these structural compression features.

In the next section, we provide more detailed discussion of the grammatical features associated with academic writing, and the striking grammatical differences between humanities and science writing. This is followed by sections that set the stage for our analyses: previous research on grammatical change, and the discourse factors that influence change. In contrast to most previous research – which focuses on spoken interaction as the breeding ground for linguistic innovations – we argue here that historical

change can occur in writing as well as speech. The following chapters present the results of empirical corpus analyses, showing that this theoretical possibility has indeed occurred. In fact, the corpus analyses in following chapters show that academic science writing has been the locus of some of the most dramatic grammatical changes that have occurred in English over the past three centuries.

1.2 Describing the grammar of academic writing

Researchers in the humanities would argue that they do not simply document human experiences that are familiar to us all; rather, they are constantly offering new interpretations of those experiences, and as a result, coining new terms that reflect those new interpretations. Thus, although these researchers discuss familiar situations and experiences, they do so in highly technical ways that are not easily understood by the non-expert. For example, consider the following excerpt from a literary criticism article:

Text Sample 1.3
Modern literary criticism article

THE MANUSCRIPT OF CHARLOTTE BRONTE'S "Villette" (1853)

Published in the wake of the Great Exhibition of 1851, it maps out the contours of interiority in a world newly captivated by the peculiar resonance of things. Though Bronte liked to think that her novel "touche[d] on no matter of public interest," its conception of the psychological interior was significantly inflected by its setting in mid century Thing City (Letters 3: 75). Villette places interiority in an intimate connection with object-filled interiors even as it hopes for an inner life that eludes the varied fetishisms of Thing City. This nostalgia for a more pristine and private form of psychological depth is, in turn, articulated in terms that reveal how entrenched persons are in the public empire of things. Villette constitutes an attempt to negotiate between a critique of commodity fetishism and a paradoxically fetishistic preoccupation with the traces and tokens of inner life. The novel suggests that the bourgeois subject, though it comes into being through its relations with things, is defined by the nostalgic notion that its true interiority has been lost under the pressure of things.

Eva Badowska. 2005.
'Choseville: Brontë's "Villette" and the Art of Bourgeois Interiority'
PMLA, 120: 1509–1523.

One of the most salient characteristics of this text is the highly specialized vocabulary. Abstract technical terms are common in this passage, such as *interiority* and *fetishisms*. In addition, relatively common words like *things*,

objects, *contours*, *interior, depth,* and *life* are used together with technical terms to produce complex phrases referring to highly abstract concepts, such as *the peculiar resonance of things*, *object-filled interiors*, *the contours of interiority, nostalgia for a more pristine and private form of psychological depth, the public empire of things,* and *a paradoxically fetishistic preoccupation with the traces and tokens of inner life*. Taken together, these abstract terms and complex phrases make it difficult for the non-specialist reader to understand the content of the literary analysis presented in the article.

Research writing in the sciences is probably even more difficult for the non-expert to understand, but its linguistic style is less often singled out for criticism. In part, this is because there is less of a mismatch between our expectations and the discourse style. The general goal of science research is to discover new information about the natural world, identifying new phenomena and explaining previously identified phenomena and patterns. Scientists require new words to refer to these previously undocumented phenomena and processes, resulting in prose that can be almost unintelligible to the non-expert. However, since the non-specialist reader has no prior experience with or existing knowledge of those phenomena and processes, we have little expectation that we should fully understand these texts. The following passage from a biochemistry research article illustrates:

Text Sample 1.4
Modern biochemistry article

> Several kinases phosphorylate vimentin, the most common intermediate filament protein, in mitosis. Aurora-B and Rho-kinase regulate vimentin filament separation through the cleavage furrow-specific vimentin phosphorylation. Cdkl also phosphorylates vimentin from prometaphase to metaphase, but its significance has remained unknown. Here we demonstrated a direct interaction between Plk1 and vimentin-Ser$_{55}$ phosphorylated by Cdkl, an event that led to Plk1 activation and further vimentin phosphorylation. Plk1 induced the phosphorylation of vimentin-Ser82, which was elevated from metaphase and maintained until the end of mitosis. This elevation followed the Cdkl-induced vimentin-Ser$_{55}$ phosphorylation, and was impaired by Plk1 depletion.
>
> T. Yamaguchi et al. 2005.
> 'Phosphorylation by Cdk1 induces Plk1-mediated
> vimentin phosphorylation during mitosis'
> *Journal of Cell Biology*, 171(3): 431–436.

Similar to Text Sample 1.3, the most obvious characteristic of this biochemistry article is its technical vocabulary. Most readers of the present book have never encountered the verb *to phosphorylate*, or nouns like *kinases, vimentin*, *prometaphase, metaphase,* and *phosphorylation*. Since we

have no idea what these terms refer to, the entire passage has little meaning for us. Thus, the vocabulary of science research writing is similarly technical to the vocabulary of humanities prose, supporting the stereotype that all academic writing is complex and hard to understand.

It is much more difficult to notice the typical grammatical structures used in these academic texts. However, when considered at the grammatical level, we discover that there are important linguistic differences among the various disciplines of academic prose; and in particular, the language of science research writing is quite different from the language of humanities prose. Such considerations lead in turn to other surprising findings that contradict previous claims and assumptions about academic writing. In particular, the corpus-based analyses presented in the following chapters challenge four major stereotypes about academic prose:

1. all kinds of academic prose are essentially the same
2. academic prose employs complex and elaborated grammar
3. academic prose is maximally explicit in meaning
4. academic prose is conservative and resistant to linguistic change

Along the way, we challenge two basic theoretical assumptions that have been widely adopted in previous linguistic research:

1. grammatical complexity is equivalent to structural elaboration, realized especially through the increased use of dependent clauses
2. grammatical changes are initiated in speech; grammatical innovations do not occur in writing

In the following sections, we briefly introduce each of these widely held beliefs and begin to challenge these assumptions.

1.2.1 *Academic written texts: All basically the same?*

Although writing researchers and students talk about 'academic prose' as if it were a well-defined construct, there are actually many ways in which academic texts differ from one another. For example, we discussed in preceding sections how humanities research writing differs in its goals and topics from science research writing. Despite those differences, though, all academic written texts can seem similar linguistically, especially in contrast to the discourse styles that we are familiar with in conversation or popular written registers like fiction.

Characteristics of this general academic-prose style are evident in both Text Samples 1.3 and 1.4 above. For example, both texts rely on specialized technical vocabulary, including many nominalizations (nouns that are formed from verbs or adjectives by the addition of a suffix):

Literary criticism text:

> *exhibition, interiority, resonance, conception, connection, fetishisms, preoccupation, relations*

Biochemistry text:

> *separation, cleavage, phosphorylation, significance, interaction, activation, ability, elevation, depletion*

The two texts also share a reliance on some specialized grammatical features, which contribute to the perception that the texts are somehow peculiar and difficult to understand. For example, passive voice – a grammatical feature often associated with academic prose – is commonly used in both of these texts:

Literary criticism text:

> *a world newly captivated by . . . things*
>
> *its conception . . . was significantly inflected by its setting*
>
> *this nostalgia . . . is, in turn, articulated . . .*
>
> *the bourgeois subject . . . is defined by the nostalgic notion that its true interiority has been lost*

Biochemistry text:

> *vimentin-Ser$_{55}$ phosphorylated by Cdkl*
>
> *the phosphorylation of vimentin-Ser$_{82}$ . . . was elevated from metaphase and maintained until the end of mitosis*
>
> *This elevation . . . was impaired by Plk$_1$ depletion*

However, a more careful analysis uncovers ways in which these two texts differ in their grammatical characteristics, reflecting the different norms of their academic disciplines. For example, attributive adjectives (i.e., adjectives that pre-modify a head noun) are very common in humanities academic writing, and Text Sample 1.3 above illustrates this characteristic:

Literary criticism text:

> *Great Exhibition*
>
> *peculiar resonance*
>
> *public interest*

psychological interior

intimate connection

inner life

varied fetishisms

a more pristine and private form of psychological depth

public empire

a paradoxically fetishistic preoccupation

nostalgic notion

Attributive adjectives are generally less common in science research writing, and Text Sample 1.4 is typical in this regard, with only two examples: *direct interaction* and *further vimentin phosphorylation*.

However, science prose tends to employ a more specialized grammatical device to modify noun phrases: nouns that occur as pre-modifiers of a head noun. Here again, Text Sample 1.4 illustrates the typical pattern:

Biochemistry text:

filament protein

vimentin filament separation

the cleavage furrow-specific vimentin phosphorylation

Plk1 activation

vimentin phosphorylation

Plk1 depletion

These pre-modifying nouns can sometimes be compounded with participles, as in:

***Cdk1-induced** vimentin-Ser55 phosphorylation*

In contrast, pre-modifying nouns are much less common in humanities writing, as is illustrated by the literary criticism text, which employs only two occurrences:

Literary criticism text:

Thing City

commodity fetishism

Further consideration of Text Sample 1.4 illustrates an additional grammatical device that is much more common in science writing than humanities writing: appositive noun phrases. These are noun phrases that modify the immediately preceding head noun, with no overt grammatical connector. For example:

Biochemistry text:

> ***vimentin***, *the most common intermediate filament protein*
>
> *a direct* ***interaction*** *between Plk1 and vimentin-Ser55 phosphorylated by Cdkl, an event that led to Plk1 activation and further vimentin phosphorylation*

Table 1.1 summarizes these differences in the preferred grammatical features of the humanities versus science writing, illustrated from Text Samples 1.3 and 1.4.

Comparing the grammatical structure of these two texts from a quantitative perspective uncovers additional differences. For example, the two passages have important differences in their preferred sentence structures. Texts Samples 1.3a and 1.4a below are copies of the two text samples, highlighting the verbs in both passages. Although the literary criticism text is considerably longer than the biochemistry text (179 words versus 94 words), both passages consist of six sentences. Thus, the sentences are much longer in the literary criticism text than the biochemistry text. A related difference is the density of verbs in the two passages: the literary criticism text utilizes three to four verbs in each sentence, while the biochemistry text uses only one to two verbs per sentence.

Text Sample 1.3a
Literary criticism text (Badowska 2005)

Verbs marked in **bold underline**

> **Published** in the wake of the Great Exhibition of 1851, it **maps** out the contours of interiority in a world newly **captivated** by the peculiar resonance of things. Though Bronte **liked** to **think** that her novel "**touche[d]** on no matter of public interest," its conception of the psychological interior **was** significantly **inflected** by its setting in mid century Thing City (Letters 3: 75). Villette **places** interiority in an intimate connection with object-filled interiors even as it **hopes** for an inner life that **eludes** the varied fetishisms of Thing City. This nostalgia for a more pristine and private form of psychological depth **is**, in turn, **articulated** in terms that **reveal** how **entrenched** persons **are** in

Table 1.1 *Preferred grammatical features of Text Sample 1.3 (literary criticism) versus Text Sample 1.4 (science)*

Linguistic Characteristic	Pattern	Literary Criticism Text (Text Sample 1.3)	Biochemistry Text (Text Sample 1.4)
Technical vocabulary and nominalizations	+ literary criticism + biochemistry	*exhibition, interiority, resonance, conception, connection, fetishisms, preoccupation, relations*	*separation, cleavage, phosphorylation, significance, interaction, activation, ability, elevation, depletion*
Passive voice	+ literary criticism + biochemistry	*a world newly captivated by . . . things* *its conception . . . was significantly inflected by its setting* *this nostalgia . . . is, in turn, articulated . . .* *the bourgeois subject . . . is defined by the nostalgic notion that its true interiority has been lost*	*vimentin-Ser55 phosphorylated by Cdkl* *the phosphorylation of vimentin-Ser82 . . . was elevated from metaphase and maintained until the end of mitosis* *This elevation . . . was impaired by Plk1 depletion*
Attributive Adjectives	+ literary criticism - biochemistry	*Great Exhibition, peculiar resonance, public interest, psychological interior, intimate connection, inner life, varied fetishisms, a more pristine and private form of psychological depth, public empire, a paradoxically fetishistic preoccupation, nostalgic notion*	*direct interaction, further vimentin phosphorylation*
Nouns as noun pre-modifiers	- literary criticism + biochemistry	*Thing City, commodity fetishism*	*filament protein, vimentin filament separation, the cleavage furrow-specific vimentin phosphorylation, Plk1 activation, vimentin phosphorylation, Plk1 depletion*
Noun + participle as noun pre-modifier	- literary criticism + biochemistry	– no examples –	*Cdkl-induced vimentin-Ser55 phosphorylation*
Appositive noun phrases	- literary criticism + biochemistry	– no examples –	*vimentin, the most common intermediate filament protein* *a direct interaction between Plk1 and vimentin-Ser55 phosphorylated by Cdkl, an event that led to Plk1 activation and further vimentin phosphorylation*

the public empire of things. Villette **constitutes** an attempt to **negotiate** between a critique of commodity fetishism and a paradoxically fetishistic preoccupation with the traces and tokens of inner life. The novel **suggests** that the bourgeois subject, though it **comes** into being through its relations with things, **is defined** by the nostalgic notion that its true interiority **has been lost** under the pressure of things.

Text Sample 1.4a
Biochemistry text (Yamaguchi et al. 2005)

Verbs marked in **bold underline**

Several kinases **phosphorylate** vimentin, the most common intermediate filament protein, in mitosis. Aurora-B and Rho-kinase **regulate** vimentin filament separation through the cleavage furrow-specific vimentin phosphorylation. Cdkl also **phosphorylates** vimentin from prometaphase to metaphase, but its significance **has remained** unknown. Here we **demonstrated** a direct interaction between Plk1 and vimentin-Ser$_{55}$ phosphorylated by Cdkl, an event that **led** to Plk1 activation and further vimentin phosphorylation. Plk1 **induced** the phosphorylation of vimentin-Ser$_{82}$, which **was elevated** from metaphase and **maintained** until the end of mitosis. This elevation **followed** the Cdkl-induced vimentin-Ser$_{55}$ phosphorylation, and **was impaired** by Plk1 depletion.

It turns out that this difference in the density of verbs reflects an even more important difference in the typical syntactic structures of the two texts. At the clause level, the biochemistry text employs very simple syntactic structures, with few embedded dependent clauses. Some sentences consist of only a single main clause; for example:

*kinases **phosphorylate** vimentin. . .*

*Aurora-B and Rho-kinase **regulate** vimentin filament separation. . .*

Other sentences consist of two coordinated main clauses:

*Cdkl also **phosphorylates** vimentin [. . .], but its significance **has remained** unknown*

*This elevation **followed** [. . .] phosphorylation, and **was impaired** by Plk1 depletion*

In fact, there are only two dependent clauses in the entire text passage from the science article. These are both relative clause structures modifying a head noun:

*an event that **led** to Plk1 activation*

*the phosphorylation of vimentin-Ser82, which **was elevated** from metaphase [. . .]*

In contrast, the literary criticism text is characterized by a high density of dependent clauses. In fact, every sentence in this text employs multiple dependent clauses, often with multiple levels of embedding. These include many different types of dependent clauses:

Complement clauses:

***liked** to **think** [. . .]*

***think** that her novel "**touche[d]** [. . .]*

***reveal** how **entrenched** [. . .]*

***suggests** that the bourgeois subject [. . .] **is defined** [. . .]*

Adverbial clauses:

***published** in the wake of the Great Exhibition*

*Though Bronte **liked** [. . .]*

*even as it **hopes** [. . .]*

*though it **comes** [. . .]*

Relative clauses and noun complement clauses:

*a world newly **captivated** by the peculiar resonance of things*

*an inner life that **eludes** [. . .]*

*terms that **reveal** [. . .]*

*an attempt to **negotiate** [. . .]*

*the nostalgic notion that its true interiority **has been lost** [. . .]*

These two texts, which are typical of their respective disciplines, illustrate the fallacy inherent in the stereotype that all academic writing is the same. Rather than being homogeneous, consideration of actual text examples shows that there are systematic grammatical differences in the research

writing from different academic disciplines. When compared to conversation or popular written registers (like fiction or newspaper reportage), it makes sense to treat academic writing as a general register with distinctive grammatical characteristics. At the same time, though, there are systematic patterns of linguistic variation that distinguish the different types of academic writing. These differences are related to two other stereotypes about academic writing: the beliefs that academic prose employs complex and elaborated grammar, and that it is maximally explicit in meaning. We turn to those stereotypes in the following section.

1.2.2 Academic writing: Complex grammar and explicit meanings?

Grammatical features like the ones described in the last section contribute to the general perception that academic writing is more complex, structurally elaborated, and explicit in meaning than most other spoken and written registers. In fact, two of the most strongly held beliefs about academic writing are the stereotypes that it is extremely complex and elaborated in terms of its characteristic grammatical structures, and that it is maximally explicit in its presentation of information.

Hughes (1996: 33–34) notes these characteristics, writing that spoken grammar employs 'simple and short clauses, with little elaborate embedding', in contrast to written discourse, which employs 'longer and more complex clauses with embedded phrases and clauses', 'explicit and varied marking of clause relations', 'explicit presentation of ideas', and 'explicit indication of text organization'.

The stereotypes that academic writing is elaborated and explicit are pervasive among writing teachers and researchers. For example, a simple database search of ERIC identified 114 published research articles where 'writing' and 'elaborate/elaborated/elaboration' appeared together, and 367 articles where 'writing' and 'explicit' occurred together. Although there is no consensus on the specific meanings of the terms 'elaboration' and 'explicitness' in these articles, there is clearly general consensus that these qualities can be attributed to academic writing as a whole, including both student writing and professional research writing. These views are often mentioned in passing, treated as background information rather than as a claim that requires evidence. Thus, consider the following quotes presented in Table 1.2 [with **emphasis added**].

These characterizations vary in the specific target of their descriptions (e.g., professional research writing or student academic writing). Further,

Table 1.2 *Characterizations of academic writing as elaborated and/or explicit*

Professional academic research writing is elaborated and/or explicit[…]	'[. . .] in academic writing [. . .] **elaborated** structures are generally preferred as they facilitate the readers' understanding of the text.' (Hyland & Tse 2005: 127)
. . .including across disciplines	'I have sought to make two broad claims in this paper: that **elaboration is a complex and important rhetorical function in academic writing**, and that its use varies according to discipline.' (Hyand 2007: 284)
	'Here then we see the emphasis of the soft knowledge fields on the discursive exploration of possibilities and limiting conditions, identifying and **elaborating** relationships in argument.' (Hyland 2008: 11)
	'This reflects the more discursive and evaluative patterns of argument in the soft knowledge fields, where persuasion is more **explicitly** interpretative and less empiricist [. . .] The presentation of research is therefore altogether more discursively **elaborate**, [. . .]' (Hyland 2008: 16)
. . .increasingly so over time	'Moreover, the [historical] change to a more **elaborate** presentation of the data-analysis procedures may serve to strengthen the dependability (e.g. consistent, faithful, stable, unbiased.), accuracy (e.g. **clearer, precise**) and aptness (e.g., well-suited, appropriate, suited) of the findings to be reported subsequently in the Results section [. . .]' (Li & Ge 2009: 98)
Student academic writing that is high-quality and/or advanced is elaborated and/or explicit	'Students [writing chemistry lab reports] engage in **elaborated** discourse with a **high degree of specificity** [. . .] Once they have focused on salient data and evidence, **elaborated** forms of discourse arrange information into **more complex and explicit representations** reflective of canonical scientific ideas.' (Wright 2008: 292)
	Myhill (1999) identifies **elaboration and use of subordination as features which tend to characterise high quality Grade A writing** [. . .]' (Keen 2004: 95) 'The redrafting process facilitates "**progressively more extended clause planning and greater elaboration**.' (Keen 2004: 96)

some of the descriptions are more nuanced and specific, identifying particular ways in which academic writing is elaborated and/or explicit. However, these descriptions all reflect the dominant and pervasive assumption that academic writing can be described as both elaborated and explicit.

For the researcher or teacher of academic writing, it would be nearly impossible to read current writing research and not form the general belief that academic writing is 'complex', 'elaborated', and 'explicit'.

In descriptive linguistics, grammatical complexity and structural elaboration are strongly associated with the use of dependent clauses. By definition, a 'simple' clause has only a subject, verb, and object or complement. A 'simple' noun phrase has a determiner and head noun. Additions or modifications to these patterns result in 'complex' and 'elaborated' grammar. In particular, linguists have traditionally singled out dependent clauses as the most important type of grammatical complexity and structural elaboration (often contrasted with simple clauses or coordinated clauses; see, e.g., Huddleston 1984: 378; Willis 2003: 192; Purpura 2004: 91; Carter & McCarthy 2006: 489).

Building on this view, researchers in applied linguistics and second language writing development have operationalized the measurement of grammatical complexity in terms of the *T-unit*: a main clause and all associated dependent clauses. Wolfe-Quintero, Inagaki, and Kim (1998: 118–119) recommend variables like the number of dependent clauses per T-unit as the 'best [. . .] complexity measures so far', and numerous researchers in applied linguistics have applied similar measures in their research (see, e.g., Ellis & Yuan 2004; Brown et al. 2005; Larsen-Freeman 2006; Nelson & Van Meter 2007; Li 2000; Brown et al. 2005; Norrby & Håkansson 2007).

The preoccupation with dependent clauses as the best reflection of grammatical complexity leads to some bizarre conclusions when we consider actual texts from different registers. For example, conversational interactions regularly employ many embedded dependent clauses. Thus, consider the following two typical utterances from a casual dinner conversation:

Jill mentioned twice today
[that I need
[to find something
[to put my trophy on]]]

It was just one of those things
[where I think
[Paul's gotten to the point
[where he won't just go on and accept
[what she says]]]]

It turns out that finite dependent clauses are actually more common in conversation than in academic writing (see discussion in Chapter 3). Thus, according to the stereotypical view of grammatical complexity, we

would be forced to conclude that conversation is more complex than academic writing – a conclusion completely at odds with the perceptions of any university student!

Of course, multiple embedded clauses do result in discourse that is more complex than alternative styles that rely on simple clauses, all other things being equal. And some written academic texts illustrate this type of grammatical complexity. For example, the literary criticism text discussed in the last section provides a good example of academic writing that employs extensive use of dependent clauses. In fact, all sentences in that passage incorporate multiple dependent clauses, often with multiple levels of embedding. Thus, consider the structure of the last sentence from Text Sample 1.3, with three embedded dependent clauses (a verb complement clause, an adverbial clause, and a noun complement clause):

*The novel **suggests***
[that the bourgeois subject,
*[though it **comes** into being through its relations with things],*
***is defined** by the nostalgic notion*
*[that its true interiority **has been lost** under the pressure of things]]*

However, the stereotype that there is only one type of grammatical complexity – associated with clausal embedding – fails to capture the differences between conversational discourse and many sub-registers of informational written discourse. And the stereotype that all written academic texts are grammatically complex because they employ frequent dependent clauses is far from correct. For example, Text Sample 1.4a illustrated a passage from science research writing with surprisingly few occurrences of dependent clause structures. If we applied the traditional measures of grammatical complexity that focus on clausal embedding, we would conclude that this science text was not complex, because it employs little structural elaboration in the form of embedded dependent clauses.

So why do we have the perception that science writing is complex? Consider the first two sentences from Text Sample 1.4. Both sentences have a simple clause structure with only a single main verb and no embedded dependent clauses:

1. *Several kinases **phosphorylate** vimentin, the most common intermediate filament protein, in mitosis.*
2. *Aurora-B and Rho-kinase **regulate** vimentin filament separation through cleavage furrow-specific vimentin phosphorylation.*

However, even though there are no dependent clauses, these sentences convey much embedded information. That information is expressed in embedded phrases rather than dependent clauses. However, the same information could be conveyed more explicitly through the use of multiple clauses. Thus compare 1b and 2b to the original sentences 1 and 2:

1b. *Vimentin **is** the most common intermediate filament protein; these proteins **are structured** to **form** intermediate filaments; kinases **phosphorylate** vimentin; this process **occurs** during the process of mitosis.*

2b. *Vimentin filament **can be separated**; Aurora-B and Rho-kinase **regulate** that separation; the separation **occurs** when something **phosphorylates** vimentin specifically where a furrow **begins** the process of **creating** cleavage*

The important point for our purposes here is to challenge the stereotype that all academic writing is grammatically complex because it employs frequent dependent clauses. Rather, Text Sample 1.4 above illustrates a grammatical discourse style where information is conveyed through phrasal devices rather than through the use of dependent clauses. These phrasal features include nouns as a pre-modifier of another noun (e.g., *filament protein, vimentin filament separation, cleavage furrow-specific vimentin phosphorylation*), appositive noun phrases (e.g., *vimentin, the most common intermediate filament protein*), and prepositional phrases (e.g., *in mitosis, through [. . .] phosphorylation*).

These devices constitute a second major type of grammatical complexity, *not* associated with structural elaboration. In fact, they have exactly the opposite function: maximally compressing structure rather than elaborating structure. One of our primary goals in the following chapters is to document this type of grammatical complexity, describing the use of structurally compressed grammatical devices in academic prose (especially in science research writing). In the process, we directly challenge the stereotype that the complexities of academic writing involve extensive elaboration with frequent dependent clauses.

A related stereotype is the belief that academic writing is maximally explicit in the expression of meaning. These two stereotypes go together because it is commonly assumed that academic writing must be structurally elaborated in order to explicitly convey precise information. However, it turns out that the complexity devices actually preferred in science research writing (i.e., compressed phrasal devices) result in a major *reduction* in explicitness.

Clausal forms of expression are considerably more explicit than phrasal features, because they grammatically specify the meaning relationships among elements. For example, there are five meaning elements in the

noun phrase *cleavage furrow-specific vimentin phosphorylation*, but there are absolutely no overt grammatical signals to tell us the meaning relationships among those elements. Because we are not specialists in genetic biology, we are not certain about the intended meaning relationships among these five elements; our best guess is captured by the following clausal representation:

> *something* ***phosphorylates vimentin specifically*** *where a* ***furrow*** *begins the process of creating* ***cleavage***

Of course, to be certain about the intended meaning relationships here, we would need advanced expert knowledge in biochemistry. Our point here, though, is to illustrate the compressed grammatical style of science research writing. There is an incredible amount of information compressed into the noun phrase *cleavage furrow-specific vimentin phosphorylation*, using only phrasal grammatical devices. And there are no explicit indications of the intended meaning relationships among the elements in that noun phrase. Thus, rather than being maximally explicit in the expression of meaning, it might be more accurate to portray certain types of academic prose (especially science research writing) as being maximally *inexplicit* in the expression of meaning. In Chapter 6, we return to the lack of explicitness resulting from the use of phrasal grammatical devices, challenging yet another of the widely-held stereotypes about academic research writing.

1.2.3 Academic writing: Resistant to change or dynamic?

Another stereotype about academic writing is that it is conservative and therefore resistant to historical change, in contrast to spoken discourse, which is constantly in flux. In part, this stereotype is tied to beliefs about grammatical correctness as embodied in standard English. Academic prose is singled out as the register most likely to conform to the rules of standard English, in part because those rules are overtly taught in educational contexts. The norms of grammatical correctness are in turn assumed to be fixed and resistant to change, especially as they are codified in a written record:

> Writing, however, is an artificial, conscious activity, and thus it is easy to resist language change in writing. We are taught to do just this, and therefore most written language is an artificial representation, omitting the signs of change which the real language, the spoken one, is full of. (McWhorter 2001: 17)

As a result, there is a widespread perception that written academic registers resist historical change and thus employ essentially the same grammatical style today as academic texts from a few centuries ago.

Hundt and Mair (1999) focus on academic writing as being especially resistant to change, in contrast to popular written registers that are 'agile' and receptive to change:

> journalistic prose and academic writing seem to be two poles on a scale of written genres differing in the degree to which they are a) open to innovations or b) are prone to retain conservative forms. What we are proposing is a cline of openness to innovation ranging from "agile" to "uptight" genres', with academic writing fitting on the 'uptight' end of the continuum [. . .] Because academic prose is less affected by economic pressures to win bigger audiences, it can afford to be more "old-fashioned". (Hundt & Mair 1999: 236)

In contrast, spoken discourse is constantly in flux, and thus it has been the focus for research on historical change. Language change in the spoken mode is assumed to eventually bring about change in the written mode:

> Indeed, writing slows language change down somewhat even on the spoken level, as writing reinforces our sense of "language" as a dismebodied blueprint to be followed or flouted .... No matter what the authority of the written form, or how tenaciously it holds on to the past, or how absurd the gulf between the written and the spoken form becomes, the spoken form always, always keeps on changing – and ultimately drags the written form reluctantly with it. (McWhorter 2001: 17)

This latter process can be documented in modern written academic texts that have begun to use innovative grammatical features adopted from speech. For example, the following sentences from modern academic research articles might appear to be unremarkable, but they each contain a grammatical feature that would have been virtually unattested in academic writing from earlier centuries:

1. *Orwell, whether he knew it or not, was actually advocating critical support for Churchill's war government. In truth, Macdonald was being a bit disingenuous.*
2. *The exact meaning of "fixation" on the molecular level will have to be settled, however, by the study of well defined molecular species.*
3. *Then the nucleic acid of the particle becomes released and reaches the cellular site where its replication is going to take place.*

These sentences illustrate grammatical innovations that emerged in speech in recent centuries: the more widespread use of progressive verbs (especially progressive stative verbs, like *being*) in #1; the semi-modal *have to* in #2; and the semi-modal *be going to* in #3. Such features reflect the general historical trend towards more 'colloquial' styles over the last

century, where grammatical innovations from spoken discourse are becoming increasingly prevalent in written discourse (see, e.g., Hundt & Mair 1999; Mair 2006; Leech et al. 2009). This historical drift has especially affected popular written registers like fiction and newspaper writing (see the corpus analyses in Chapters 3 and 4). Academic research writing has for the most part resisted the adoption of these colloquial features. While features like stative progressive verbs and semi-modals can be traced back to the eighteenth and early nineteenth centuries in popular written registers, they have only recently begun to appear in academic writing (and they are still extremely rare in that register). Thus, although it is possible to find instances of innovative colloquial features in academic writing, the evidence in this regard overwhelmingly supports the stereotype that academic writing is conservative and resistant to historical change.

In other respects, though, academic writing has been at the forefront of historical change. These grammatical innovations do not involve the use of stigmatized colloquial or non-standard forms, and thus they have generally not been noticed by usage commentators. And historical linguists have assumed that linguistic change originates in speech (and then spreads to popular written registers), and thus they have generally disregarded the possibility of historical change in the use of other grammatical devices that originates in academic writing. However, empirical corpus-based research shows that the grammatical discourse style of academic writing – especially science research writing – has changed radically over the past 300 years. One of our primary goals in the present book is to document those patterns of change.

To illustrate these historical developments, consider the following nineteenth century academic texts, comparing their grammatical characteristics to the modern humanities and science texts discussed in Sections 1.2.1 and 1.2.2. Text Sample 1.5 is taken from an 1852 book on the early history of Otsego County in New York, while Text Sample 1.6 is from a science article published in the *Philosophical Transactions of the Royal Society of London* in 1800.

Text Sample 1.5
Nineteenth century history

> Discarding all claims to literary merit, I hope to contribute something that shall instruct and amuse the rising generation; the tendency of which, shall be to make them better citizens, and fit them to discharge the various duties of life, with credit to themselves and benefit to our common country. Individual incidents unimportant as they appear, go to make up when combined, what we call history; and though generally over-looked, except as connected with the great men of the earth,

yet as little rills and streamlets feed the mighty rivers, that sweep along with irresistible force; so individual events help to make up the great historical mass, that floats down the current of time.

Levi Beardsley. 1852.
'*Reminiscences: Personal and other incidents; Early settlement of Otsego County; Notices and anecdotes of public men; . . .*'
Published by Charles Vinten.

Text Sample 1.6
Nineteenth century science

The discoveries which have been made with respect to light, as it proceeds immediately from the sun, are many and important; but the observations on that species of light which is spontaneously emitted from various bodies, are not only few in number, but in general very imperfect. The author is therefore desirous of drawing the future attention of the philosopher more particularly to this subject, and of communicating his own experiments and observations upon it, to this learned Society.

By the spontaneous emission of this light, the author wishes to distinguish it from all kinds of artificial phosphorus; which, as he apprehends, differ essentially, in some of their properties from that light of which he means to treat. And, by its adhesion to bodies with some degree of permanency, he distinguishes it from that transient sort of light which is observable in electricity, in meteors, and in other lucid emanations. The light which is the subject of this paper, he shall therefore beg leave to discriminate by the name of spontaneous light.

Nathaniel Hulme. 1800.
'Experiments and Observations on the Light which is spontaneously emitted, with some Degree of Permanency, from various Bodies'
Philosophical Transactions of the Royal Society of London, 90: 161–187.

These nineteenth century academic texts are similar grammatically to present-day humanities texts. For example, the nineteenth century history text illustrates a dense use of attributive adjectives; in fact, almost every noun phrase is modified by an attributive adjective in this passage:

Attributive adjectives:

literary merit

rising generation

better citizens

various duties

common country

individual incidents

great men

little rills and streamlets

mighty rivers

irresistible force

individual events

great historical mass

The nineteenth century science text is also similar to modern humanities writing in its use of attributive adjectives as well as frequent nominalizations:

Attributive adjectives:

various bodies

learned Society

artificial phosphorus

spontaneous light

Nominalizations:

discoveries

observations

adhesion

permanency

Attributive adjective + nominalization:

future attention

spontaneous emission

lucid emanations

At a more specific level, the preferred syntactic structures employed in the nineteenth century history article (Text Sample 1.5) are in many

ways similar to those used in the modern literary criticism article (Text Sample 1.3), especially with respect to the use of long sentences incorporating multiple dependent clauses; these include all major types of dependent clauses:

Complement clauses:

I hope to contribute something . . .

shall be to make them better citizens. . .

make up [. . .] what we call history. . .

help to make up the great historical mass . . .

Adverbial clauses:

Discarding all claims. . .

to discharge the various duties. . .

to make up. . .

when combined. . .

though generally over-looked

except as connected with the great men . . .

as little rills and streamlets feed the mighty rivers . . .

Relative clauses:

something that shall instruct and amuse. . .

the tendency of which, shall be. . .

the mighty rivers, that sweep along . . .

the great historical mass, that floats down the current . . .

The main difference between the nineteenth century history text and the modern humanities text is the extent to which they rely on dependent clauses: the nineteenth century humanities text is marked by this characteristic to an even greater extent than modern humanities prose. The result is often incredibly long sentences that are extremely complex in terms of their clausal embedding. Thus, consider the clausal structure of the last sentence from Text Sample 1.5:

Individual incidents, go
[to make up
[when combined]
[what we call history]]
and
[though generally over-looked
[except as connected with the great men of the earth]]
[yet as little rills and streamlets feed the mighty rivers,
[that sweep along with irresistible force]]
so individual events help
[to make up the great historical mass,
[that floats down the current of time]]

This comparison of nineteenth century and twentieth century humanities texts shows that there has been historical change in the *extent* to which elaborated structures (with multiple dependent clauses) are used, but there is little evidence of change in the *types* of grammatical structures that are employed. In contrast, a comparison of the nineteenth century science article in Text Sample 1.6 and the modern science article in Text Sample 1.4 illustrates historical change at a much more fundamental level. The nineteenth century science text is similar to humanities texts (both nineteenth and twentieth centuries) in utilizing extensive clausal embedding. Text Sample 1.6 particularly relies on relative clauses, but adverbial clauses are also used:

Relative clauses:

The discoveries which have been made...

that species of light which is spontaneously emitted...

all kinds of artificial phosphorus; which, as he apprehends, differ essentially...

that light of which he means to treat

that transient sort of light which is observable in electricity...

the light which is the subject of this paper...

Adverbial clauses:

as it proceeds immediately from the sun

as he apprehends

Table 1.3 *Summary of major grammatical differences between four text excerpts*

	Nineteenth Century	Present Day
Humanities	*Text Sample 1.5: history* + attributive adjectives ++ clausal embedding	*Text Sample 1.3: literary criticism* + attributive adjectives + clausal embedding
Science	*Text Sample 1.6: physics* +nominalizations + clausal embedding (relative clauses) - nouns as noun pre-modifiers - appositive noun phrases	*Text Sample 1.4: biochemistry* + nominalizations - clausal embedding ++ nouns as noun pre-modifiers + appositive noun phrases

The density of these dependent clause structures is not nearly as great as in the nineteenth century history text, but the general tendency is similar: to structure discourse through the use of embedded dependent clauses. This tendency is especially noteworthy when we contrast the grammatical style of this nineteenth century science text with the preferred grammatical structures in Text Sample 1.4 (the modern science research article). On the one hand, the nineteenth century science text relies on dependent clause structures to a much greater extent than the modern science text. But the more interesting difference concerns the grammatical features that are not found in the nineteenth century text: nouns as pre-modifiers of a head noun, or appositive noun phrases post-modifying a head noun. These features represent important grammatical innovations that are found in modern science research writing, but are generally much less common in other written registers (including modern humanities academic writing).

Table 1.3 summarizes the trends that we have identified so far, comparing two parameters of variation: discipline (humanities vs. science) and historical period (nineteenth century vs. modern day). While some linguistic features are consistently used in academic writing across disciplines or across historical periods, others show marked differences along these same parameters. In particular, the table highlights the peculiar nature of modern science research writing, strikingly different from present-day humanities writing as well as nineteenth century science writing. This register is the locus of the types of grammatical change documented in the present book. As summarized in Table 1.3, grammatical features like pre-modifying nouns and appositive noun phrases were rare in all registers in the nineteenth century. In contrast, these phrasal devices are so prevalent in present-day academic science writing that they usually escape notice.

The present book describes the distinctive grammatical characteristics of modern science research writing, and documents the dramatic patterns of historical change towards that grammatical style. Those changes include a strong decrease in the use of dependent clauses, accompanied by a strong increase in the use of phrasal modifiers. As a result, modern science writing regularly employs phrasal complexity features that are minimally used in earlier historical periods (e.g., pre-modifying nouns and appositive noun phrases). These historical developments are of particular interest because they contradict generally accepted claims about historical change in English. In particular, we show that

1. grammatical innovation and change occurs in writing as well as in spoken interaction; and
2. academic written registers are *not* conservative and resistant to change.

Thus, the analyses reported in the following chapters show how certain kinds of grammatical features have expanded historically in their frequency of use, their grammatical productivity, and their meanings and functions – and how all of these changes have occurred primarily in academic writing (especially science research writing), reflecting the special communicative characteristics of that register.

1.3 Grammatical change in English: Below the radar?

Historical change in English is an important topic treated in almost all introductory textbooks on linguistics and sociolinguistics. For the most part, though, these treatments focus on changes in pronunciation/phonology or lexical change. For example, several books document phonological changes in English like the loss of initial [h] in 'which', 'whether', and 'whine', or lexical changes like *chesterfield* > *couch* and *serviette* > *napkin* (see, e.g., Holmes 1992; Chambers 2002). In many cases, these changes result in present-day patterns of regional or social dialect variation. Examples of this type include the loss or retention of post-vocalic [r], the merger of [t, d] with [Θ, ð], and the vowel shifts found in American northern cities (see, e.g., Wolfram and Schilling-Estes 2005).

Grammatical change is generally disregarded in introductory linguistics textbooks. The few books that deal with this topic at all tend to focus on morphological change, such as the regularization of verb paradigms for irregular verbs (e.g., *dreamt* → *dreamed*), or vice versa (*sneaked*→ *snuck*). Some books also discuss patterns of morphological change (or resistance to change) that have resulted in present-day dialect differences (see, e.g.,

Wolfram and Schilling-Estes 2005: 85–92; Finegan 2008). Such changes include the loss of third person singular verb inflection (*he runs → he run*) and the development of second person plural pronouns in many regional dialects of English (e.g., *y'all, youse, yoins*).

In this regard, the coverage in linguistics textbooks is typical of linguistics research generally, which places 'much more emphasis on phonological and lexical phenomena than on grammatical ones' (Leech et al. 2009: 7). One explanation for this fact is that we just fail to notice grammatical change: that 'we are considerably less perceptive of it than of other kinds of linguistic change' (Strang 1970: 60, cited in Leech et al. 2009: 7). Leech et al. go on to discuss the difficulty in noticing grammatical changes. One obvious factor is the nature of the phenomena. When the phonetics of a word begins to change, we notice that some people pronounce the word differently from others. When a new word comes into the language, we notice that some people are using the word and others not. From a synchronic perspective, such changes usually manifest themselves as social patterns of dialect variation; consequently, we readily notice some groups of speakers using pronunciations and words that are not used by other groups of speakers. Such patterns of variation can evolve rapidly, so that there are often dramatic phonological and lexical differences from one generation of speakers to the next (see, e.g., Labov 1994, 2001, 2010).

In contrast, grammatical change is usually more abstract, involving the use of different phrase types and different clause types, and changes in the ways in which those structures are configured relative to one another. Probably an even more important consideration is that grammatical change proceeds much more slowly than lexical/phonological change, and it is less clearly associated with dialect variation.

The grammatical changes that have been most often noticed are typically lexico-grammatical: changes in the grammatical function of specific words. Prescriptive grammarians often target specific lexico-grammatical features associated with the general 'decay' of written English. Most of these pundits focus on present-day use and perceived misuse, but the targeted features are often specific words undergoing grammatical change. Examples include variation in the past participle form for irregular verbs (e.g., *dove* vs. *dived*, *learnt* vs. *learned*, *lit* vs. *lighted*), preposition choice (e.g., *different from* vs. *different than*), *less* versus *fewer* with count nouns, *like* functioning as a subordinator or a preposition, the conversion of adjectives to adverbs (e.g., *real good; go fast/slow*), and the choice between *that, who,* and *whom* as relative pronouns with an animate head noun.

Similarly, sociolinguistic studies of grammatical change usually focus on developments that affect specific words. For example, numerous studies

have documented the recent grammatical change associated with features like the new quotative verbs (e.g., *go, be like, be all*), the discourse marker *like*, and *so* used as an intensifier modifying an adjective (see the survey of research in Tagliamonte 2006). Changes of this type affect only a few specific words, they occur in a relatively short time span, and they are usually associated with strong patterns of social dialect variation; as a result, such changes are often noticed and described by professional linguists as well as the wider public.

In contrast, other grammatical changes are general structural shifts, influencing an entire construction type or grammatical system. For example, over the last few centuries there have been general changes in the use of modal verbs (e.g., Leech 2003, 2011; Millar 2009), verb phrases with marked aspect (e.g., Smith & Rayson 2007; Nesselhauf 2007; Mair and Hundt 1995; Smith 2002; Hundt 2004; Mair 2006), and complementation patterns (e.g., Mair 1995, 2002; Rudanko 2010). Such changes have a lexical aspect, in that they gradually work their way through the lexicon. However, these are not innovations restricted to a few individual words. Further, in contrast to the restricted lexico-grammatical changes discussed in the last paragraph, these general structural changes occur gradually over the course of several centuries, and they typically do not serve to distinguish among social dialects during their period of transition. As a result, such historical developments have attracted much less attention than short-term, socially-stigmatized grammatical changes in the use of particular words.

For example, progressive aspect has increased strongly in use over the past three centuries, developing new meanings and functions which have gradually occurred with an ever increasing set of lexical verbs (see Hundt 2004; Mair 2006; Leech et al. 2009). If we compare the end points of this historical change, the differences are as dramatic as the recent development of new quotative verbs like *go* and *BE all*. However, historical change in the use of progressive aspect has been much more gradual, occurring over a period of centuries rather than a period of a few years. And there is no particular verb that we can point to as the target of this change; rather, the use of progressive aspect has been generally increasing for an ever expanding set of verbs.

These two characteristics – long-term change (vs. rapid change) and change influencing a grammatical system (vs. change that targets a few specific words) – are typical of many grammatical developments in English. Such patterns of change are not salient to speakers of a language at any given point in time, and they are not obviously associated with synchronic patterns of dialect variation. As a result, linguistic change of

this type has generally been disregarded in comparison to lexical and phonological changes. However, several recent books on the history of the English language include descriptions of such grammatical change. In the following section, we survey some of the major features discussed in those treatments.

1.4 Two types of grammatical change

A long-term perspective on the history of English makes it clear that there have been numerous important grammatical changes in English over the course of the last 1,000 years. These include the change to a fixed SVO word order, the loss of most inflectional morphology (especially case suffixes), the increase in the range of function words (including prepositions, auxiliary verbs, and infinitive marker *to*), and more recently the introduction of the dummy auxiliary DO (see Rissanen 1999; van Gelderen 2006).

However, grammatical changes over the last 300 years – the period of Modern English – have been much less dramatic. There have been numerous interesting developments, but they differ in kind from the earlier dramatic changes. Recent historical developments include many of the features that have been discussed by prescriptivists, such as changing preferences for irregular versus regular verb forms, the choice of *fewer* versus *less*, and the decline in use for the relative pronoun *whom*. However, there have also been changing patterns of use for numerous other grammatical features, documented in survey articles by Barber (1964: 130–144), van Gelderen (2006: 171–174, 214–220), and Fennell (2001: 144–145, 173–174), as well as in the more extended treatments by Denison (1998), Mair (2006), and Leech et al. (2009). Some of the more important changes are:

- development of auxiliary-like uses for *want to* → *wanna*
- increase in the use of the semi-modals be *going to* (*gonna*) and *have got to* (*gotta*)
- decrease in use of the core modal verbs
- increasing use of analytical rather than synthetic comparison
- increasing use of the *s*-genitive, especially with non-human nouns
- decreasing use of the *of*-genitive
- increasing use of the 'mandative' subjunctive
- decrease in the use of *shall* as a future marker
- increasing use of progressive constructions

– development of the progressive-passive
– development of the *get* passive
– overall decreasing use of the *be* passive
– increasing tendency to omit the relative pronoun *that*
– increase in the number and type of multi-word verbs
– changing complementation patterns with non-finite clauses

These linguistic developments illustrate two major types of grammatical change:

a) the emergence and development of new grammatical features (including new grammatical uses of an individual word); and
b) shifts in the typical use and functions of core grammatical features

The first type of grammatical change has been studied mostly under the rubric of 'grammaticalization', which focuses on the way in which content words evolve over time to be used as grammatical function words. Examples from the list mentioned earlier include the use of *have to* and *got to* as semi-modals, *wanna* with modal auxiliary functions, and *get* as an auxiliary verb in passive constructions. Other examples include the use of *well* as a discourse marker, *pretty* as a hedge or intensifier, and sequences like *in spite of, with regard to,* and *because of* used as complex prepositions (see, e.g., Krug 2000; Tagliamonte 2004; Hopper & Traugott 2003; Lindquist and Mair 2004; Nevalainen 2004; Hoffmann 2005).

The new quotative verbs illustrate a related kind of grammatical change. Some of these forms – like *go* – have evolved from a lexical verb to be used in the more grammaticalized function of marking reported speech and thought. The quotative *be all* also seems to have evolved from a lexical (adjectival) use to this more grammaticalized function. However, it could be argued that *be like* evolved in the opposite direction, from a function word (a preposition or subordinator) to a more lexicalized grammatical function as a quotative marker. Regardless, these are all examples of grammatical innovation that result in particular words changing to be used with new grammatical functions.

In contrast, the second type of grammatical change is completely different in kind: it focuses on the extent to which a grammatical feature is used rather than the development of new grammatical classes. This type of grammatical change affects the overall frequency of a feature and the co-occurrence of the grammatical feature with an ever increasing set of words; these changes are also often associated with an increase in the meanings and functions expressed by the grammatical feature. Examples

of this type from the above list include the increasing use of the *s*-genitive, progressive verbs, and multi-word verbs. At the same time, other features have declined in use and functionality over the past few centuries (e.g., modals and passive voice verbs).

It turns out that most grammatical change in English over the past 300 years has been quantitative rather than categorical. Even grammatical innovations, like the development of semi-modals and the *get*-passive, have gradually continued to increase in frequency and functionality over this period. Thus, a quantitative perspective on grammatical change has become increasingly important in recent research:

> . . .changes in the realm of syntax are often a function of quantity, rather than quality; that is, certain structures have expanded in number and frequency of occurrence during the PDE period. (Fennell 2001: 173)

> Since relatively few categorical losses or innovations have occurred in the last two centuries, syntactic change has more often been statistical in nature, with a given construction occurring throughout the period and either becoming more or less common generally or in particular registers. The overall, rather elusive effect can seem more a matter of stylistic than of syntactic change. . . (Denison 1998: 93)

Most of the diachronic changes that we document in this book are of this type, relating to the frequency of occurrence and the functionality of grammatical features, rather than the development of completely new grammatical categories. It turns out that grammatical change of this type has been especially prevalent in academic writing, unlike the lexico-grammatical innovations typical of speech.

1.5 The locus of historical change in English: Speech or writing?

It will be clear from the preceding sections that we are interested primarily in grammatical innovations that have occurred in the discourse of academic writing. For many historical linguists, this focus is a non-starter. That is, most previous research on grammatical change has focused on spoken discourse, and it has often overtly excluded the possibility of grammatical change in writing.

The view that language change occurs in speech – and not in writing – can be attributed to several sources. At the most basic level, this view reflects the widespread belief in linguistics that spoken language is 'real', while written language is a derivative artifact (see discussion in Biber 1988: 5–7). In the early American linguistic tradition, Sapir and Bloomfield both held this position:

> writing is 'visual speech symolism' (Sapir 1921: 19–20)
>
> writing is 'not language, but merely a way of recording language by visible marks' (Bloomfield 1933: 21)

And numerous other linguists in more recent decades have reiterated this position:

> 'speech is fundamental and writing . . . only a secondary derivative' (Hall 1964: 8–9)
>
> Written communication is 'derivative of the face-to-face conversational norm' (Fillmore 1981: 153)
>
> 'spoken language is "true" language, while written language is an artifact' (Aronoff 1985: 28)
>
> 'Language, Saussure taught us, is first and foremost a spoken system – writing is a secondary coding, but speech is primary' (Schneider 2002: 67 – citing Saussure 1916)
>
> 'Writing, however, is an artificial, conscious activity, and thus it is easy to resist language change in writing. We are taught to do just this, and therefore most written language is an artificial representation, omitting the signs of change which the real language, the spoken one, is full of.' (McWhorter 2001: 17)

Nearly all introductory textbooks on linguistics continue to reflect this bias in their reference to 'speakers' and 'speech communities', with little or no acknowledgement of 'writers' or 'writing communities'.

Historical linguists who study recent change in English have similarly been preoccupied with the characteristics of speech, describing the ways in which English has become increasingly colloquial and informal in recent centuries. The linguistic features associated with these changes originated in speech, and they have gradually been adopted in writing. Thus, to the extent that historical researchers have considered change in writing, they have focused mostly on the increasing use of colloquial features borrowed from speech, such as first person pronouns, contractions, and semi-modals. This historical trend towards a more colloquial style has been documented by numerous discourse analysts and corpus linguists, and is referred to as the 'drift' of written registers towards more 'oral' styles (Biber and Finegan 1989a), 'informalization' (Fairclough 1992), and 'colloquialization' (Hundt and Mair 1999; Mair 2006; Leech et al. 2009). For example, Mair (2006) describes this process as follows:

> For some time now, for example, and increasingly so since the 1960s and 1970s, an egalitarian and informal communicative culture has

> been promoted in the public domain which has brought the norms of writing closer to the norms of spoken usage. In grammatical terms, this has favored the rapid disappearance of archaisms such as *upon* for *on* or the subjunctive in all but its mandative use, and led to a decrease in the popularity of typical markers of formal and written style such as the passive voice. On the other hand, it has facilitated the spread of informal grammatical options such as contractions, the *going to*-future, or certain types of progressives into domains in which they used to be rare. (Mair 2006: 88)

Earlier investigations of these patterns suggested that they represent a general historical trend in English. However, subsequent investigations show that there are important differences among written registers. For example, Biber and Finegan (1997/2001) show that written registers like fictional novels and personal letters have been strongly influenced by the drift towards more colloquial linguistic styles, but written academic registers (especially science research articles and medical research articles) have not participated in that trend. Hundt and Mair (1999) also note this difference, distinguishing between 'agile' written registers (e.g., newspaper prose) that are receptive to these changes, and 'up-tight' written registers (e.g., academic prose) that resist historical change towards the increased use of colloquial features.

A complementary approach to historical change has focused on the ways in which grammatical innovations emerge in language use, carried out under the rubrics of 'grammaticalization', 'usage-based approaches', 'emergent grammar', and the general study of 'frequency effects' in language use. These approaches share a focus on grammatical use in natural discourse, describing how new grammatical constructions and/or functions emerge from natural communicative situations.

Most usage-based studies of grammatical change focus on spoken interaction, using the term 'speaker' as a cover term for the addressor or producer of discourse, and often explicitly framing the discussion in terms of 'utterances' and conversational 'interaction'.[1] Written discourse has generally received little attention, with most studies simply disregarding the possibility that grammatical innovations might develop in natural written communication.

For example, an exclusive focus on spoken interaction is front-and-center in the edited book on *Interaction and Grammar* by Ochs, Schegloff,

[1] A few usage-based studies acknowledge the possibility of historical change in writing as well as speech. For example, Traugott (2003: 125) mentions the possible role of the 'writer' in her definition of 'subjectification': the tendency of meanings 'to become increasingly based in the SP[eaker]/W[riter]'s subjective belief state or attitude to what is being said and how it is being said'.

and Thompson (1996), which approaches grammar as part of the social practices associated with conversational interaction. The studies in this book analyze transcripts of conversations, showing how grammar emerges with special functions in spoken interactional contexts. A similar implicit emphasis on spoken interaction can be observed in so-called usage-based models of grammar (see, e.g., Langacker 1987; Kemmer & Barlow 2000). While these descriptions tend to be based on intuitive notions of spoken interaction, rather than direct analysis of conversational texts, they still implicitly frame the discussion relative to spoken discourse; for example, the 'usage events' that form the foundation of the usage-based model are 'utterances' produced by 'speakers' (Langacker 2000: 9).

Similarly, in describing the motivations and enabling factors of grammaticalization, Hopper and Traugott (2003: 71) note that previous research has focused on 'the role of speakers and hearers negotiating meaning in communicative situations'. Bybee and Hopper (2001) are also typical in their focus on spoken interaction, noting that:

> The notion of emergence . . . relativizes structure to speakers' actual experience with language, and sees structure as an on-going response to the pressure of discourse. . .
>
> The distribution and frequency of the units of language are governed by the content of people's interactions. . .
>
> Patterns of use . . . deal with patterns of occurrence of morphosyntactic structures in natural conversation. . . (Bybee & Hopper 2001: 3)

Croft (2000) is especially emphatic in arguing that language change occurs in utterances produced by speakers in conversational interaction; for example:

> language use is essentially a joint act between speaker and addressee . . . Language is a fundamentally social interactional phenomenon. So is language change. (Croft 2000: 87)

Given this background, it is not surprising that many studies of grammaticalization have focused on grammatical features that are common in conversation but rarely used in writing, such as the English semi-modals (e.g., *have to, got to*), discourse markers (e.g., *well, like*), intensifiers (e.g., *so*), and quotative verbs (e.g., *go, be all*).

As noted above, most scholars simply disregard the possibility that grammar might also emerge in written use. One of the few exceptions is Croft (2000), who directly considers the possibility but then strongly argues against it:

> One might speculate that the advent of the written medium led to directed evolution in the development of these construction types [e.g., nominalizations, participles, attributive adjectives] . . . Typological research indicates, however, that all of these construction types are present in most if not all unwritten languages. (Croft 2000: 83)
>
> It is thus possible that there is directed change in the advent of the written register. However, expansion into the new linguistic niche results in the evolution at most of new degrees of syntactic complexity – multiple iterations and embeddings of structures – rather than in developing completely new grammatical structures. (Croft 2000: 83–84)

In this passage, Croft is specifically discussing grammatical innovation from a typological perspective: can a new type of grammatical construction emerge in a written language in comparison to the types of constructions that already exist in the world's spoken languages? However, his basic position is in line with the generally accepted view of most historical linguists: that grammatical innovations occur in speech, not writing; and that any grammatical developments in writing are derivatives from speech.

Of course, historical linguists do not have access to speech samples from earlier historical periods, and thus they are forced to rely on the analysis of written texts. It might be supposed that this methodological requirement would have forced linguists to consider the possibility of grammatical change originating in writing. But that has not occurred. In fact, the methodological focus of historical research has been on ways to filter out the assumed 'obscuring' effect of writing, in the attempt to uncover the 'true' patterns of language use that occurred in speech from earlier periods.

Schneider (2002) has written one of the most thorough treatments of this methodological issue, under the promising title 'Investigating variation and change in written documents'. This title might lead the reader to expect discussion of the importance of grammatical change in writing. However, the very first sentence of the article immediately quashes that possibility, asserting that speech is primary and writing merely 'a secondary coding' (Schneider 2002: 67). Schneider goes on to discuss the implications of this view for historical research:

> Normally, as variationist linguists we are not directly interested in the written record as such [. . .] – its function is predominantly to serve as a clue, a pathway to the variation and change of the language system in itself. More directly, most written records of interest in this context represent a speech act [. . .]. In such cases, the written record functions as a filter: [. . .] the rendering of the speech event is only indirect and imperfect. (Schneider 2002: 67)

> A written record of a speech event stands like a filter between the words as spoken and the analyst. As the linguist is interested in the speech event itself [. . .], a primary task will be to "remove the filter" as far as possible. (Schneider 2002: 68)

In sum, Schneider argues that historical linguists have no choice but to analyze written documents. However, they should have no expectation of uncovering grammatical changes that emerge in writing. Rather, the focus is exclusively on speech (representing the real 'language system'), and the methodological challenge is therefore to uncover the patterns of use that existed in speech, based (unfortunately) on the analysis of written texts.

Culpeper and Kytö (2010) directly address this methodological challenge by carrying out detailed linguistic analyses of spoken dialogue as represented in speech-based registers from Early Modern English (e.g., legal trial proceedings, witness depositions, drama). Other researchers (e.g., Nevalainen and Raumolin-Brunberg 1996; Denison 1998) have collected texts from informal/colloquial registers, like personal letters or diaries, in the hope that these will approximate the underlying linguistic system found in speech from earlier historical periods. These studies have been highly important for the wealth of detailed information that they provide about the typical linguistic patterns of use in earlier periods. However, they also reflect the dominant view that the primary inquiry of interest in historical linguistics concerns the patterns of change that occurred in speech. Written records allow us to infer what those changes have been, but they are assumed to be of no interest in themselves. And in particular, the language of informational written prose is considered to be far removed from the kinds of discourse that historical linguists should be analyzing to uncover genuine historical change.[2]

In the present book, we take a fundamentally different position, arguing that the communicative demands of any register – whether spoken or written – have the potential to facilitate the emergence of grammatical uses (and possibly new constructions) associated with those communicative needs. The first author has been advocating this position for the last 25 years; for example:

[2] Our characterization of the state-of-the-art here disregards the long tradition of research by historical corpus linguists such as Rissanen, Nevalainen, and Crespo, who have carried out numerous historical studies of corpora consisting of written documents (specifically written academic documents in the case of many of the corpora constructed at the University of Helsinki and the University of A Coruña).

> . . .the two modes of communication [speech and writing] have quite different strengths and weaknesses, and they therefore tend to be used in complementary situations. From this perspective, neither can be said to be primary: they are simply different. The linguistic characteristics of each mode deserve careful attention, and the relationship between the two modes must be investigated empirically rather than assumed on an a priori basis. (Biber 1988: 9)

Biber and Conrad (2009: 6) further develop this argument, claiming that 'linguistic features are always functional when considered from a register perspective'; that 'linguistic features tend to occur in a register because they are particularly well suited to the purposes and situational context of the register'. These functionally-motivated patterns of linguistic variation characterize present-day spoken registers (Chapter 4) and written registers (Chapter 5), as well as historical change in writing (Chapter 6).

Fox (2007: 299) has also advocated investigation of this possibility, stating that: 'The relationship of written language to language-in-interaction is complex and worthy of independent study'. Thus, Fox argues that written discourse is produced under completely different circumstances from spoken discourse, and that these characteristics might have direct grammatical consequences:

> Writers and readers typically have no time constraints placed on their production and comprehension, a fact which presumably allows more complex syntactic structures to arise . . . On the other hand, speakers and recipients in real-time conversation have immense time pressures on them . . . [creating] a tendency in conversation towards shorter and syntactically simpler utterances. . . (Fox 2007: 314)

> It is clear . . . that writing alone, at leisure . . . is a different grammatical enterprise than is designing an utterance, in real time . . . the grammar of written language thus needs to be taken up as a separate investigation. . . (Fox 2007: 315)

The present book explores this possibility, based on the underlying assumption that the communicative circumstances of any register will provide the context for functionally-motivated grammatical change. In particular, we focus on academic research writing as a register that differs in almost every way from face-to-face conversation:

- written rather than spoken
- monologic rather than interactive and co-constructed
- requiring specialized, professional background knowledge, but no assumption of personal background knowledge
- slowly produced and carefully revised and edited

If grammatical change is influenced by the pressures of the communicative situation, there is every reason to expect that the grammatical devices and functions emerging in written academic writing will be strikingly different from those that have emerged from conversational interactions.

The analyses in the following chapters show convincingly that historical written texts should not be treated as a 'filter' that obscures true historical change. Rather, extensive grammatical change has occurred in writing. These patterns of change have affected different grammatical characteristics from the colloquial features described in most previous research. However, the grammatical innovations of written discourse cannot be discounted as mere 'iterations and embeddings'. In particular, focusing on noun phrase constructions in English, the following chapters document extensions in the range of grammatical variants, the range of lexical associations, and the range of grammatical/semantic functions – in addition to dramatic quantitative changes in the frequency of use.

1.6 Overview of the present book

The present book challenges previous assumptions about grammatical complexity, historical linguistic change, and academic writing, documenting the ways in which academic writing has been linguistically innovative over the past two centuries. The primary focus of the book is the description of phrasal complexity features and the associated phrasal discourse style that is typical of present-day science research writing. From a synchronic perspective, the book documents how academic written registers are unlike other registers – spoken or written – in their reliance on phrasal complexity features rather than a clausal style of discourse. From a diachronic perspective, the book documents linguistic innovations that have emerged in writing (especially science research writing) but have generally not been adopted in speech.

We base our descriptions on large-scale corpus analysis that allows us to track long-term change in the grammatical structure and discourse style of academic writing, as compared to a range of other written and spoken registers. Chapter 2 introduces the corpus-based approach to linguistic analysis, as well as the specific corpora and analytical methods used for our study. It begins with a discussion of recent trends and issues in diachronic corpus-based research. The chapter then introduces the corpora that serve as the basis for our analyses; these span 300 years and represent a range of registers (e.g., everyday conversation, newspaper writing, fiction, academic writing). The academic writing corpus is designed to represent several sub-registers, including humanities prose, social science articles, and natural science articles.

Chapter 2 also outlines the research methods for our analyses, including corpus annotation, interactive and automatic analysis techniques, and quantitative research designs. Finally, Chapter 2 introduces and exemplifies the grammatical features discussed in the book, including the full set of clausal and phrasal complexity features.

Chapters 3 through 6 present the major findings of our research. Chapters 3 and 4 present quantitative corpus findings, while Chapters 5 and 6 are more interpretive, presenting the results of detailed linguistic analyses. These chapters are sequenced as follows: Chapter 3 describes synchronic patterns of register variation; Chapter 4 presents the results of similar analyses carried out from a diachronic perspective; Chapter 5 presents detailed linguistic analyses to support the claim that these historical developments are not 'merely' stylistic changes in frequency; and finally, in Chapter 6, we turn to the question of explicitness, showing how phrasal complexity features are less explicit in meaning than alternative clausal features.

Chapter 3 provides a synchronic description of the grammatical discourse style of academic writing, focusing especially on features associated with grammatical complexity and elaboration, but also describing the use of common, core grammatical features. A major goal of this chapter is to show that traditional operational definitions of 'complexity' – in terms of clausal embedding – fail to capture the actual complexities of academic writing, which are phrasal in nature. Through a series of case studies, we undertake synchronic comparisons between academic writing and other spoken and written registers to demonstrate that this phrasal discourse style is highly distinctive – occurring only in written registers with highly informational purposes and specialized audiences. Thus, this chapter documents the distinctive grammatical characteristics of modern academic research writing by comparing it to conversation, textbooks, classroom teaching, fiction, and newspapers. We conclude Chapter 3 with a case study that considers variation *within* modern day academic research writing. Chapter 3 demonstrates that academic writing – and especially science research writing – can be regarded as an outlier: a register unlike nearly all other spoken and written registers.

Chapter 4 then turns to a diachronic analysis of this same range of registers, to document the evolutionary path that academic research writing has taken over the past 300 years. We investigate the use of colloquial features as well as complexity features, and we consider both clausal as well as phrasal complexity features. The results show that academic writing has been resistant to change in the use of colloquial

Table 1.4 *Challenging basic assumptions about grammatical complexity, linguistic change, and academic writing through corpus-based analyses*

	There is a single construct of grammatical complexity: structural elaboration realized through dependent clauses	Grammatical changes are initiated in speech; grammatical innovations do not occur in writing	All kinds of academic prose are essentially the same	Academic prose employs complex and elaborated grammar	Academic prose is maximally explicit in meaning	Academic prose is conservative and resistant to linguistic change
Chapter 3: Phrasal vs. clausal discourse styles (synchronic)	✓		✓	✓		
Chapter 4: The historical evolution of phrasal discourse styles	✓	✓	✓	✓		✓
Chapter 5: The functional extension of phrasal devices	✓	✓		✓	✓	✓
Chapter 6: The loss of explicitness	✓			✓	✓	

features, but it has actually been the locus of grammatical change in the use of phrasal complexity features. Comparison of sub-registers within academic writing show that specialist science research articles have been especially important for these grammatical innovations.

Chapter 5 presents detailed linguistic analyses to support the claim that these historical developments are not 'merely' stylistic changes in frequency. Rather, this chapter documents major extensions in the lexical associations, meanings, and discourse functions of phrasal complexity features. Chapter 6 continues our functional analysis of the phrasal discourse style of academic writing, turning to an unintended consequence of these grammatical innovations: the loss of explicitness. Through detailed analyses of particular phrasal devices, we discuss how phrasal noun modifiers compress information into dense constructions that are considerably less explicit in meaning than alternative clausal features.

Taken together, the analyses presented in Chapters 3 through 6 directly challenge many stereotypes and assumptions about grammatical complexity, historical change, and academic writing. We summarize the overall progression of the book in Table 1.4.

The final chapter, Chapter 7, summarizes the main themes of the book, leading into a discussion of applied implications. First, the chapter revisits these basic assumptions about complexity, linguistic change, and academic writing, highlighting how the analyses presented in Chapters 3–6 counter previous stereotypes. We then turn to the implications that result from our corpus-based findings. We discuss implications for theories of grammatical complexity, arguing that phrasal modifiers and structural compression are equally important to, but fundamentally different from, clausal modifiers and structural elaboration. We also consider implications for linguistic change, arguing that grammatical change *can* emerge in writing, and that in fact certain types of linguistic change must be considered within the context of written language. Finally, we explore applied implications related to the study of language development and the teaching of academic reading and writing (English for Academic Purposes).