



Phraseological patterns in reporting clauses used in citation: A corpus-based study of theses in two disciplines

Maggie Charles *

University of Oxford Language Centre, 12 Woodstock Road, Oxford OX2 6HT, United Kingdom

Abstract

This study investigates the phraseological patterning that occurs in reporting clauses used to make references to others' research. It examines finite reporting clauses with *that*-clause complement and draws upon two corpora of theses written by native speakers in contrasting disciplines: approximately 190,000 words in politics/international relations and 300,000 words in materials science. The findings show that both disciplines use significant numbers of these reporting clauses and that they most frequently occur as integral citations with a human subject. Following the work of [Francis, G., Hunston, S., & Manning, E. (1996). Collins COBUILD grammar patterns 1: Verbs. London: HarperCollins], the reporting verbs are analysed into semantic groups. Further evidence of patterning is found in both the verb groups and the tenses that occur. In both corpora, the most frequent verb group is ARGUE (e.g., *argue*, *note*, *suggest*) and the most frequent tense is present (e.g., *Skinner argues that...*). In materials science, however, there are almost as many instances of the FIND/SHOW verb group (e.g., *show*, *find*, *observe*) and these occur predominantly in past tense (e.g., *Sun (1990) showed that...*). The rhetorical functions of these patterns are discussed and explanations proposed based on genre and discipline. This research underlines the importance of phraseology in academic writing and ends by suggesting how working with patterns can be beneficial in raising students' language awareness.

© 2005 The American University. Published by Elsevier Ltd. All rights reserved.

* Tel.: +44 0 1865 283360; fax: +44 0 1865 283366.

E-mail address: maggie.charles@lang.ox.ac.uk.

Keywords: Phraseology; Disciplinary discourse; Reporting verbs; Citation; Corpus linguistics; Academic writing

1. Introduction

Citation plays a key role in academic writing. It shows how a new piece of research arises out of and is grounded in the current state of disciplinary knowledge and thus constitutes an overt manifestation of the ongoing ‘conversation of the discipline’. At the same time, it enables the writer¹ to acknowledge or take issue with the contributions of other researchers and, in displaying knowledge of the field, to establish his/her own academic authority and credibility.

Much work has therefore been devoted to the examination of citation practices. In applied linguistics, a long tradition starting with the work of Swales (1981, 1986, 1990) has seen citation as an instrument for achieving the rhetorical purposes of the writer (e.g., Berkenkotter & Huckin, 1995; Dudley-Evans, 1986; Hopkins & Dudley-Evans, 1988; Myers, 1990). Citation practices have been found to vary according to discipline (e.g., Bloch & Chi, 1995; Hyland, 1999, 2002a) and genre (Thompson & Tribble, 2001). The use of reporting clauses in citation has also been investigated, with several studies specifically focusing on the reporting verbs that occur (e.g., Hunston, 1993; Shaw, 1992; Thomas & Hawes, 1994; Thompson & Ye, 1991).

The diversity and breadth of this literature reflects the complexity of citation practices and this in turn helps to account for the difficulties that novice writers experience in learning to cite appropriately (e.g., Dong, 1996; Groom, 2000; Pecorari, 2003). Moreover, much of the material on citation offered to international students of EAP is considered inadequate (Harwood, 2004; Thompson & Tribble, 2001).

A promising means of tackling these problems is through the investigation of phraseological patterns, since such recurrent associations of words and structures can readily be used as the basis for student language awareness activities. However, to date there has been little attempt to study citation from a pattern perspective. The approach of researchers such as Hyland (2002a) and Thompson and Tribble (2001) has been to count all instances of citation and to analyse them in terms of their structure and function. One of the achievements of this work has been to underline the key importance of the reporting clause. By contrast, the present study takes the reporting clause as its starting point. In Francis, Hunston, and Manning’s (1996) terminology, it examines the verb patterns **V that** (e.g. *Skinner argues that...*) and **it be V-ed that** (e.g., *It has been reported that...*) and investigates their associated phraseology. This approach allows identification of the most frequent phraseological patterns used in citation and thus provides more detailed ‘input and support to the concept of grammatical consciousness-raising’ (Hunston & Francis, 2000, p. 262).

¹ Following the convention established by Thompson and Ye (1991), ‘writer’ is used to refer to the researchers who cite and ‘author’ for the researchers cited.

2. Data and method

The data is drawn from two corpora of theses written by native-speakers, amounting in total to just under half a million words. The corpora consist of theses from two contrasting disciplines, a social science: politics/international relations (eight MPhil theses, about 190,000 words); and a natural science: materials science (eight doctoral theses, about 300,000 words). Whole texts are used, but they exclude front matter, abstracts, bibliography and footnotes. Examples from the corpus are coded as ‘pol’ for the politics corpus and ‘mat’ for the materials corpus, followed by a number from 1 to 8 which specifies each thesis. A list of the theses with their codes appears in [Appendix 1](#).

Using WordSmith Tools ([Scott, 1999](#)), concordances were made on the search word *that* and finite reporting clauses were identified. Although omission of the *that* complementizer is possible, research has shown that its retention is ‘the norm in academic prose’ ([Biber, Johansson, Leech, Conrad, & Finegan, 1999, p. 680](#)). A back-check on three of the most frequent verbs in my corpora, *argue*, *show* and *suggest* confirms this finding. In total, only eight instances of these reporting verbs occurred without the complementizer: in politics, there were two instances of the verb *argue* and one of *suggest*; in materials, five instances of *suggest*. Thus, it is not thought that this restriction will significantly affect the frequency data.

Finally, clauses that report on the work of other researchers were distinguished. These are called here ‘research reports’ and are defined as reports with a source which is a reference to the work of others in the field. Research reports are identified by the presence of a grammatical subject which is a noun or pronoun reference to a researcher or researchers other than the thesis writer (e.g., *Copeland*; *commentators*; *he*); alternatively either a specific textual feature or the surrounding context indicates that the proposition is sourced in others’ research (e.g., *a recent study*; *Auger analysis by Ingrey*; *it is well documented that...*). Figures for research reports are given in [Table 1](#).

It is interesting to note that the frequency of research reports is very similar in the two corpora, indicating a comparable use of this pattern in reporting the literature. It is important to note, however, that this does not mean that there are similar levels of citation overall in the two corpora, but simply that politics and materials science use reporting clauses with *that* complementation to a similar extent in the performance of citation.

Table 1
Finite reporting clauses with *that*-clause complement used in research reports

Frequency measure	Politics	Materials
Number of words in corpus	190,000	300,000
Tokens	155	259
Tokens per 100,000 words ^a	81.6	86.3

^a Ideally, the frequency of the reporting clauses should be normed to a number of clauses, rather than words, but this was not possible as the corpus is not parsed for clauses.

Table 2
Frequency of three clause types per 100,000 words

Clause type	Politics	Materials
Human subject	56.3	44.7
Non-human subject	21.1	25.3
<i>It</i> subject with passive	4.2	16.3
Total	81.6	86.3

3. Research reports: analysis and findings

3.1. Clause type

In order to examine the research reports in more detail, two further types of analysis are carried out. First, three clause types are distinguished according to the grammatical subject of the reporting clause: a noun group with human reference; a noun group with non-human reference; introductory *it* followed by passive voice:

- (1) *Skinner argues* that historical texts should be read in their intellectual context. (pol6)
- (2) *A model for this system*¹⁸¹ *predicts* that 1 μ m is a reasonable order estimate for the size. . . (mat5)²
- (3) *It is thought* that this may provide some improvement (Van der Burgt 1995). . . (mat4)

Table 2 shows the frequency per 100,000 words of the three clause types in each corpus.

Disciplinary differences are seen in the frequency of these clause types: politics uses more clauses with human subjects than materials, while materials uses more clauses with non-human and *it* subjects than does politics.³ We may explain these differences by reference to differences in the ideology and epistemology of the two disciplines. As Becher and Trowler (2001) note, the construction of knowledge in the social sciences is ‘personal’, while that in the natural sciences is ‘impersonal’ and ‘objective’. However, what is more unexpected is that the relative distribution of the three clause types is the same in both corpora. Human subjects are by far the most frequent, followed by considerably lower numbers of non-human subjects and much lower frequency figures for *it* subject. It is also noticeable that, although *it* subject has almost four times as many instances in materials as in politics, the frequency figures in the two corpora for human and non-human subjects are reasonably close.

² Neither corpus is consistent in the citation convention used: some writers use author/date; others, like this writer, use a numerical convention. Further details are given in Table 4.

³ Although there is considerable variation in the clause types used by individual writers, the median figures also show some evidence of these tendencies: in politics 6.9, 2.7 and 0.3; in materials 2.9, 2.8 and 1.9 for human, non-human and *it* subject clauses, respectively.

This suggests the existence of a pattern such that when using a reporting clause with a *that*-complement to comment on the work of other researchers, writers in both disciplines tend to choose a human subject. In order to elucidate this further, it is necessary to examine the type of reference involved. Two categories of research reports are distinguished: ‘citation’ and ‘general reference’.

3.2. Citation

For the purpose of this paper, a ‘citation’ is defined as a research report which has a specific reference point that is clearly identifiable. Citations are further categorised as ‘integral’ or ‘non-integral’, following Swales (1981, 1990). However, I count only citations which use a reporting clause with *that*-complement and I use Shaw’s (1992) procedure of categorising follow-on sentences according to the form of the original citation. Thus, each occurrence of a reporting clause counts as a separate citation, although the reference point will not necessarily be found within the citation sentence itself. In the example below, my analysis counts the two reporting clauses with *that*-complement as two integral citations:

- (4) Ezz, Pope and Paider (1982) reported another intriguing feature of the YSA...
Ezz et al. found that the CRSS for [101](111) slip depended on the sense of the applied load. They also found that the asymmetry was strongly dependant on temperature and orientation. (mat6)

An integral citation is defined as a research report in which the name of the researcher appears as a sentence element with an explicit grammatical role, often, but not always, in the reporting clause. Integral citation in each of the three clause types is exemplified below:

- (5) *Copeland has argued that in effect the US ended up free-riding on the Pact...*¹⁷.
 (pol3)
 (6) *Sneddon (1965) derived...His results showed that the load displacement relationships... can conveniently be written as...* (mat3)
 (7) *...it was stated by Wells as early as 1974 that unbiased and/or unstained pn-junctions can be imaged in cleaved epitaxially grown GaAs...* (mat8)

A non-integral citation is a research report in which the name of the researcher appears in brackets, or a number refers to the name, which appears elsewhere. Non-integral citation in each of the three clause types is illustrated below:

- (8) *Commentators hope that such a remark will overcome the objection...*⁶¹ (pol4)
 (9) *...other surface analysis techniques confirm that changes do occur...*²¹ (mat2)
 (10) *It has been reported that at very low temperatures the flow stress decreases with increasing temperature up to −196 °C (Suzuki, Mishima and Miura, 1989).* (mat6)

3.3. General reference

Following Jacoby (1987), I distinguish a category of research reports in which there is no specific reference point clearly identifiable. These are here termed ‘general reference’. The category includes instances where, although there is no citational name or number, a marker indicates that the clause introduces a report on the work of other researchers. Markers in clauses with human subjects are summary references to several researchers (e.g., *writers who think that...*); those in clauses with non-human subjects are often general references to schools of thought (e.g., *the motivational approach*); markers in *it* clauses are usually adverbs (e.g., *often*). Examples of general references in the three clause types are given below:

- (11) *Rather, attitudinalists have argued* that Justices are able to pursue their ideological ends most effectively by simply voting their attitudes into law. (pol5)
- (12) *Neo-classical growth theories accept* that technological progress is central to growth. . . (pol8)
- (13) *It is well known* that uniformity along the length of the superconducting wire is one of the key factors. . . (mat4)

We should note here that some general references, particularly those in *it* clauses with an adverb marker, may apply not just to other researchers, but to all members of the field, including the thesis writer him/herself. Here, they are considered under the category of ‘research report’, since they do not refer to the work of the thesis writer alone. In the example above, we understand that all members of the field know that *uniformity along the length of the superconducting wire is one of the key factors* and that this knowledge is shared by, but not limited to, the thesis writer him/herself.

A small number of research reports cannot easily be assigned to either citation or general reference, as it is unclear whether the report refers to a specific reference point or not. Indeed in certain cases the thesis writer seems to blur this distinction, introducing indeterminacy into the text, perhaps in order to soften criticism of another researcher’s work. This can be seen in the example below, where the reader will assume that *such a position* refers to the work of Srolovitz *et al.*, despite the fact that the form of the subject does not imply the specific position of Srolovitz *et al.*, but any similar position:

- (14) Srolovitz et al. (1984) have carried out an identical simulation for $f = 0.025$. . . They suggest a volume distributed energy, such as residual strain energy. *Such a position assumes* that we can apply the results of a 2D model to the 3D case. However, it is not clear that this can be justified. (mat7)

3.4. Patterns in research reports

When we combine the analysis of reference type with that of clause type, evidence of further patterning emerges, as shown in Table 3.

In both disciplines, there is a pronounced tendency to use these clauses to make integral citations with a human subject. This use is at least three times as frequent

Table 3

Frequency of reference and clause type per 100,000 words

Reference and clause type	Politics	Materials
Integral citation human subject	45.8	42.7
Integral citation non-human subject	1.6	10.0
Integral citation <i>it</i> subject	0	3.7
Non-integral citation human subject	3.7	1.7
Non-integral citation non-human subject	5.8	12.3
Non-integral citation <i>it</i> subject	2.6	9.7
General reference human subject	6.8	0.3
General reference non-human subject	13.7	3.0
General reference <i>it</i> subject	1.6	3.0

as any other reference/clause type. Moreover, the two disciplines show similar figures: in politics 45.8 per 100,000 words; in materials, 42.7. In all of the other reference/clause types the difference between the two corpora is around a factor of two or more. The use of integral citation with a human subject leads to prominence of the cited author (Weissberg & Buker, 1990). Given the tendency in the natural sciences to downplay the role of the individual scientist, we would expect there to be far fewer instances in materials than in politics. The similarity between the two corpora in the use of this category is thus somewhat unexpected.

The predominance of integral citation is in contrast with the results of Hyland (2002a) for research articles; he found that, apart from philosophy, all the disciplines he examined showed higher figures for non-integral citation. By contrast, although the results of Thompson (2001) accorded with those of Hyland for theses in agricultural botany, the preferred choice in agricultural economics was integral citation. In comparing these findings to my own, however, we must bear in mind that my data refer only to reporting clauses, while those of Thompson and Hyland concern all citations. This suggests that it may be the **V that** pattern which tends to be used when writers construct an integral citation with a human subject. Indeed calculations for my data show that in politics 96.7% and in materials 75.7% of the integral citations examined use a human subject.

Swales (1990) suggests that citation convention (numerical or author/date) may affect the choice between integral and non-integral forms and he argues that numerical conventions predispose the writer to use non-integral citation. Table 4 shows the breakdown of citation type in the two corpora by individual writer.

In my data, six out of eight materials scientists use the author/date convention; however, it would not be accurate to attribute the relatively high incidence of integral citation solely to the citation convention used. While it is true that the two writers who use a numerical convention (mat2 and mat5) use fewer integral than non-integral citations, there are two other writers who also show low integral citation use (mat3 and mat8). It is noticeable that all four of these writers use relatively few citations of either type: their total citation frequencies are between 3.0 and 5.7. In fact, considerable individual variation is seen, with, at the other end of the scale, two writers with total citation frequencies of over 17.

Table 4

Citation convention and citation type: individual writer frequencies per 100,000 words

Thesis writer	Politics			Thesis writer	Materials		
	Citation convention	Integral	Non-integral		Citation convention	Integral	Non-Integral
pol1	Author/date	10	0.5	mat1	Author/date	8.7	3.7
pol2	Number/footnote	6.8	0	mat2	Number/endnote	0.3	2.7
pol3	Number/footnote	3.2	0	mat3	Author/date	2.3	1.0
pol4	Number/footnote	3.7	6.3	mat4	Author/date	6.0	6.0
pol5	Number/footnote	9.0	3.2	mat5	Number/endnote	1.7	4.0
pol6	Number/footnote	5.3	0.5	mat6	Author/date	16.7	2.3
pol7	Number/footnote	1.6	0.5	mat7	Author/date	16.7	1.0
pol8	Number/footnote	9.0	0	mat8	Author/date	2.7	0.7

In politics one writer follows the author/date convention and uses predominantly integral citation. However, out of the seven writers who follow the number/footnote convention, only one uses more non-integral than integral citations (pol4). In this data, then, there seems to be no correlation between the number/footnote convention and the use of non-integral forms. Indeed three of the other six number/footnote users do not employ non-integral citation at all. It would thus seem likely that the choice of integral/non-integral citation is a complex product of a number of factors including citation convention, genre, discipline and individual study type. However, this issue requires detailed investigation with much larger corpora. Suffice it to say that in the present corpora the frequency of integral citation does not seem to be due to citation convention alone. Rather I suggest that it may be explained in terms of the function of citation in the thesis.

4. Integral citation with a human subject

Since a reporting verb with *that*-complement is most frequently used in integral citation with a human subject, the remainder of this paper will concentrate on this pattern. In integral citation, the attribution is highly specific and the author by consequence highly visible. Hence, this type of citation tends to be associated with comment on an individual study rather than with generalisations about a number of studies. Given both the length of the thesis and the necessity to display knowledge, writers may choose this type of citation because it lends itself to extensive comment on the cited research. It is possible that thesis writers tend to deal with the literature by discussing individual papers at some length, rather than by subsuming a number of papers under a single comment. This explanation is also proposed by [Thompson and Tribble \(2001\)](#); however, comparative studies of theses with other genres would be necessary in order to confirm the suggestion.

The detailed discussion of specific studies clearly enables the thesis writer to take a stance as a candidate who is able to display full mastery of the literature. This necessitates situating cited authors within the field and is carried out in order to position the work of the thesis in relation to the cited research. This is illustrated below, where

the first three sentences show part of one of these extended discussions of the literature, using three integral citations with human subjects. Two further paragraphs follow, detailing problems in the area, and in the last sentence, the writer positions his research as an attempt to fill the gap in current knowledge (Hoey, 2001; Swales, 1990).

- (15) *Humphreys (1994) has shown* that extrusion conditions are important in determining texture of MA6000. In addition *he finds* that a pre-treatment before annealing can affect texture. . . *Bieler et al. (1992) report* that cube texture components can be suppressed. . .
The present work was embarked upon to investigate these issues. . . (mat7)

A different form of positioning is shown in the next example from politics, where the detailed treatment of *Holbraad's* views leads the thesis writer to pose a question, to which his next section provides the answer. The use of the verb *admit* enables the reader to predict that this answer will be negative. Thus, the subsequent question both organises the discourse and contributes to the writer's evaluation of the cited author's point (Hyland, 2002b).

- (16) *Carsten Holbraad admits* that 'subjective factors' may influence the foreign policy of individual middle powers, but *maintains* that their reactions in comparable situations should be the same. . .²⁷Does the behaviour of the states on the Canadian list of middle powers conform to that expectation?
 (pol2)

Further insight into the use of this citation form may be gained by investigating the verbs that occur. This forms the subject matter of the following section.

4.1. *Verb groups in integral citation with a human subject*

As Thompson and Ye's (1991) ground-breaking study shows, the choice of reporting verb is a key feature which enables the writer to position their work in relation to that of other members of the discipline. Thompson and Ye's research offers a threefold analysis of the evaluative potential of reporting verbs in citation (1991, pp. 372–3). First, reporting verbs show the author's stance towards the report, which may be positive, negative or neutral; second, reporting verbs can construct the writer's stance of acceptance, neutrality or rejection towards the cited research; and lastly they allow the writer's interpretation of the author's behaviour or discourse.

Thompson and Ye also distinguish three categories of reporting verbs according to the process they perform: 'textual' verbs, in which there is an obligatory element of verbal expression (e.g., *state*, *write*); 'mental' verbs, which refer to mental processes (e.g., *think*, *believe*); and 'research' verbs, which refer to processes that are part of research activity (e.g., *find*, *demonstrate*). Later studies such as those by Thomas and Hawes (1994) and by Hyland (2002a) also employ this three-way distinction,

although Hyland uses the terms ‘discourse’ and ‘cognition’ for Thompson and Ye’s textual and mental verb categories.

Since the present research takes an approach based on patterns, I make use of the groups distinguished by Francis et al. (1996) for the **V that** pattern. Four verb groups are found in my data, and are here named after the most frequent verb in the corpora: ARGUE, THINK, SHOW and FIND. Names of verb groups are printed in small capitals.

Where a verb can occur in more than one group, the context is examined in order to determine the appropriate category. Thus, following Francis et al., the verb *observe* can appear in the FIND group with the meaning of ‘becoming aware’ or ‘noticing’ and in the ARGUE group when it refers to the associated language activity. In the corpora examined here, it is interesting to note that in politics this verb is used only to refer to the language activity and thus occurs only in the ARGUE group, while in materials it refers exclusively to visual noticing and thus occurs only in the FIND group. Verb groups and explanations are adapted from Francis et al. (1996, pp. 97–101) as follows:

1. ARGUE verbs are concerned with writing and other forms of communication, e.g., *argue, suggest, assert, point out*.
2. THINK verbs are concerned with thinking, including having a belief; knowing, understanding, hoping, fearing, e.g., *think, assume, feel*.
3. SHOW verbs are concerned with indicating a fact or situation, e.g., *show, demonstrate, reveal*.
4. FIND verbs are concerned with coming to know or think something, e.g., *find, observe, discover, establish*.⁴

We may note, however, that there are clear similarities between the Thompson and Ye verb groups and those employed here: ARGUE verbs parallel the textual group; THINK verbs, the mental group; SHOW and FIND verbs, the research group. Table 5 gives the data for each verb group, showing both normed figures and percentages. Frequency figures for the individual verbs that occur in each group appear in Appendix 2 (Tables 8–11).

ARGUE verbs predominate in both corpora: in fact, 88.5% of integral citations appear with an ARGUE verb in politics and although the percentage is lower in materials, at 54.7% it is still substantial. As these verbs are concerned with communication, they primarily refer to the cited author’s text. Materials also has appreciable numbers of FIND and SHOW verbs, which deal with results and situations and are thus more likely to refer to the cited author’s research. In politics, much disciplinary activity is text-based,⁵ so the high frequency of ARGUE verbs

⁴ Tables 8–11 in Appendix 2 show the frequency figures and verb groups of all the verbs that occur in the corpora.

⁵ However, even within a single discipline there can, of course, be considerable diversity of approach. This is evident in my data, where one thesis (pol5) is a statistical study of decision-making in the US Supreme Court which relies heavily upon numerical procedures to make its case.

Table 5

Verb groups in integral citation with a human subject

Verb group	Politics		Verb group	Materials	
	Frequency per 100,000 words	% of total		Frequency per 100,000 words	% of total
ARGUE	40.5	88.5	ARGUE	23.3	54.7
THINK	3.2	6.9	FIND	11.0	25.8
SHOW	1.6	3.5	SHOW	7.3	17.2
FIND	0.5	1.2	THINK	1.0	2.3
Total	45.8	98.9	Total	42.6	100

here is not surprising. However, as materials is an experimental discipline, we might expect references to the cited author's research to be more frequent than those to their text. But even if we take all uses in both the FIND and SHOW groups to be references to the cited author's research, instances of ARGUE verbs are still in the majority, although the difference is less marked (a normed figure of 23.3 for ARGUE as against 18.3 for FIND/SHOW). In both corpora, then, there is a tendency for a human subject to co-occur with an integral citation and an ARGUE verb.

It is worth noting that if we take the ARGUE group as equivalent to 'discourse' verbs and FIND/SHOW verbs as equivalent to 'research' verbs, these figures only partially confirm those given by Hyland (2002a, p. 125). As in Hyland's data, the percentage of ARGUE verbs is higher in the social than in the natural science, while the percentage of FIND/SHOW verbs is higher in the natural than in the social science. However, the natural science data does not 'display a preference for Research-type verbs' (Hyland, 2002a, p. 125). Rather, in materials, there are similar percentages of ARGUE and FIND/SHOW verbs and indeed the ARGUE group is slightly more frequent. There are two possible reasons for this discrepancy. First Hyland's figures relate to all reporting verbs, while mine cover only reporting verbs with *that*-clause complement. Second, Hyland examines research articles, while my data relate to theses. It is possible that the difference in verb group use reflects differences between the two genres. I return to this point in discussing the function of the ARGUE verb pattern below.

So far, then, this analysis has revealed the existence of two patterns: in both corpora an integral citation with a human subject and an ARGUE verb; in materials alone, an integral citation with a human subject and a FIND/SHOW verb. However, further patterning can be seen in the tenses associated with the reporting verb groups.

4.2. Tenses in verb groups with integral citation and a human subject

When integral citation with a human subject occurs with an ARGUE verb, the reporting verb is most likely to be in the present tense. From Table 6, it is seen that in politics 88.3% of these citations use a present tense. In materials, the tendency is

Table 6

Tenses in the *ARGUE* verb group with integral citation and a human subject

Tense	Politics		Materials	
	Frequency per 100,000 words	% of total	Frequency per 100,000 words	% of total
Present	35.8	88.3	12.3	52.9
Past	2.1	5.2	8.3	35.7
Present perfect	2.6	6.5	2.7	11.4
Total	40.5	100	23.3	100

not so pronounced, but nonetheless clear: 52.9% in present as against 35.7% in past tense.

The number of instances of *FIND/SHOW* verbs is rather low for reliable generalisations to be made. Nevertheless, in contrast to the *ARGUE* group, in materials these verbs occur predominantly in the past tense. As Table 7 shows, over 70% of instances follow this pattern.

Comparison of the two verb groups shows that the frequency of the *ARGUE* group in present tense is roughly equal to that of *FIND/SHOW* verbs in past tense. Thus it would seem that, while there is one clear pattern in politics: integral citation with a human subject and a present tense *ARGUE* verb, in materials, the patterning is more complex. Two tendencies seem equally prevalent: an *ARGUE* verb in present and a *FIND/SHOW* verb in past tense, while a less pronounced tendency may also exist, to use an *ARGUE* verb in the past. Clearly, these findings need to be verified with much larger corpora. However, the use of a present tense with an *ARGUE* verb would fit in with the tendency noted above for this verb group to refer primarily to the cited author's text, while the use of a past tense with a *FIND/SHOW* verb would accord with the use of these verbs to refer to the research work itself. These figures suggest, then, that the patterning in citation is not limited to grammatical subject, citation type and verb group, but also extends to tense choice. However, given that the number of instances in this data is low, detailed discussion of this aspect of the phraseology must await further investigation using larger corpora. Hence, the following comments on the verb groups are not restricted to examples in certain tenses. I begin by discussing the *FIND/SHOW* verb group pattern in materials.

Table 7

Tenses in the *FIND/SHOW* verb group with integral citation and a human subject

Tense	Politics		Materials	
	Frequency per 100,000 words	% of total	Frequency per 100,000 words	% of total
Present	1.6	75	3.3	18.2
Past	0	0	13.0	70.9
Present perfect	0.5	25	2.0	10.9
Total	2.1	100	18.3	100

4.3. The FIND/SHOW verb group pattern

As seen in Table 5, materials makes considerably more use of the FIND and SHOW groups in comparison with very low frequencies in politics: 11.0 for FIND and 7.3 for SHOW in materials, as against 0.5 and 1.6 for the respective groups in politics. These two verb groups refer to the research activity rather than to the text and their greater use in materials is due both to the importance of experimentation and to the concern with discovery/explanation in the natural sciences (Becher & Trowler, 2001). As these factors are not so relevant to the construction of knowledge in politics, numbers are correspondingly low.

The use of an integral citation together with a SHOW or FIND verb enables the thesis writer to make the cited author prominent and thus to give personal credit to other researchers for their achievements:

- (17) *Davies and Stoloff (1965) showed* that the Young's modulus of polycrystalline Ni₃Al decreased with increasing temperature... (mat6)
- (18) *Venables and Maher (1996) actually compared* dopant contrast on identical samples in microscopes with different detection systems, and *found* that in-lens collection enhanced dopant contrast levels. (mat8)

Most FIND verbs appear as part of an extended description or discussion of the cited author's work. This is shown by their co-occurrence with cohesive devices such as *and*, *in addition*, *also*, personal pronouns and lexical repetition. These citations display the writer's detailed knowledge of the research:

- (19) *Boland (1991) studied* the role of hydrogen on the silicon surface in the adsorption and dissociation of disilane, and *showed* that growth at high hydrogen coverage is governed by the reactions of disilane with the hydrogenated surface. *He also observed* that the result of annealing a disilane-saturated surface was strips of epitaxial dimer rows. (mat1)

All the SHOW verbs are factives; for example, the non-factive *indicate* does not appear with a human subject. These verbs enable the thesis writer to construct a consensus with the cited author by signalling a clear acceptance of the findings:

- (20) *Cairns et al. (1975) have shown* that the ability of a material to undergo successful zone annealing is governed by a number of factors... (mat7)

The cited work can then be used as the basis for the writer's arguments or research procedures, thus contributing to the cumulative construction of knowledge in which one researcher's work builds on that of another (Becher & Trowler, 2001). This is the 'support' function identified in Dudley-Evans (1986) and Hopkins and Dudley-Evans (1988).

- (21) *Sun (1990) showed* that crystals of Ni₃Ga and Ni₃Al with the same orientation deformed by similar dislocation processes... It is therefore reasonable to assume that the mechanisms will be the same in specimens tested under creep conditions. (mat6)

4.4. The ARGUE verb group pattern

Turning now to ARGUE verbs, I have suggested above that human subjects with integral citation are used to present the relative positions of other researchers, which creates a context for the thesis writer to situate their own work within the field. There are several possible reasons why this function is often performed by ARGUE verbs. First, since the major source of information about others' research is the cited text, we would expect many citations in both corpora to refer to the written account:

- (22) *Eden's official biographer, R.R. James, writes* that "more had been achieved in twelve months to resolve critical issues than in the entire post-war period"⁴⁵. (pol3)
- (23) However, *Toy and Whittemore (1989) reported* that Gitzen et al. (1956) observed no such loss in strength... (mat3)

A further reason stems, I think, from the nature of the thesis itself. Since theses are based on research questions or problems which arise from differing views or lack of knowledge, they are located at sites of conflict or uncertainty in the field. Thus, thesis writers must deal with potentially controversial issues, but at the same time, as candidate members of the community, they need to exercise care in the way they handle them. Thus, it may be more appropriate for a thesis writer to use an ARGUE verb when dealing with aspects of the literature where there is conflict or uncertainty. This hypothesis would fit in with the general finding that the recursivity of knowledge construction in politics leads to greater discussion of alternative views and would also enable us to account for the higher than expected figure in materials.

Without detailed disciplinary knowledge it is not possible to be precise, but examination of the context suggests that in both corpora over half of the instances refer to a specific situation of conflict or uncertainty, signalled, for example, by verb choice (e.g., *speculate, propose*), conjunctions (e.g., *but, yet, though*) and modal forms. In the following examples, the thesis writer takes up a position which is in conflict with that of the cited author:

- (24) ...*Spijkerboer suggests* that women who flee persecution because they have transgressed social mores should base their claims on political and/or religious grounds. *He suggests* that... Also, *he argues* that...⁶² *However, there are several problems with this analysis*. (pol4)
- (25) *Susan Woodward mentions* that 'principles' were a major motivating factor in the international community's decision to intervene in Bosnia, but fails to elaborate either what is meant by principles or why states should feel compelled to defend them.²⁰ (pol7)

- (26) *Ling and Anderson (1992) comment* that non-random particle distributions are to be expected in view of the enhanced reduction in the interfacial energy of the system. However, since non-random particle distributions lead to smaller pinned grain sizes it can be argued that they have precisely the opposite effect. (mat7)
- (27) *Kapustin also asserts* that critical current changes within the region of 1–10 A cm⁻² have no effect, this is not consistent with the experimental results presented here. (mat4)

Conflict and uncertainty may also occur among other researchers, as shown in the following examples:

- (28) *Knight and Epstein argue* that Segal and Spaeth's test of the effects of precedent is too limited. . . (pol5)
- (29) But where some observers universalize non-hegemonic conditions and *claim* that these constitute the characteristics of the international system per se, *Cox asserts* that this involves a degree of historical distortion. (pol1)
- (30) Normally a value of t/h^3 10 is assumed to be required though Manika, Maniks and Teteris (1988) quote t/h^3 1.5 for amorphous As–S and As–Se films on quartz substrates. *Aisenberg and Kimock (1989) claim* that t/h^3 5 is sufficient for diamond-like carbon films. (mat3)
- (31) Vittadini, Selloni and Casarin used LDA, and a repeated slab for their calculations, and allowed the top four layers of the slab to relax. *They claim* that the major difference between their calculations and those of Wu and Carter (1992) is the lack of surface relaxation in Wu and Carter. (mat1)

In materials, about 25% of the instances use verbs which denote uncertainty, including *suggest*, *propose*, *hypothesise* and *speculate*. As *Thompson and Hunston point out* (2000, p. 25), such evaluations 'are not neutral with respect to cultural value'. In a discipline which studies the properties of physical matter, truth and certainty are assumed to be attainable: the 'truth will out device' (*Gilbert & Mulkay, 1984*). Thus, it is the task of materials science to reveal the truth about the substances under investigation. Uncertainty gives rise to conflict and demands resolution. Hence, an evaluation of uncertainty is related to a negative evaluation and can constitute a face-threatening act. One way to moderate this threat is to use an ARGUE verb:

- (32) This conflicts with Kapustin, who attempts to explain the signal maxima either side of peak field as the result of a "dynamic resistance" (Kapustin 1992)... *Kapustin conjectures* that at low temperatures (4.2 K) flux lines are strongly pinned... (mat4)
- (33) The value of the beam current in a SEM is not usually discussed in detail in SE-theory, but may affect the surface charge of a specimen which could promote or suppress SE-emission. For example, *Yablonovitch et al. (1986) suggest* that the injected carrier density will affect the extent of band bending at the specimen surface. (mat8)

We should note here, however, that the uncertainty expressed in the verb may be the ‘writer’s interpretation’ in [Thompson and Ye’s terms \(1991\)](#). It is attributed to the cited author by the writer and it is possible that the uncertainty is not present in the original text. It could be that writers attribute uncertainty to a cited author with whom they wish to disagree, since a lack of certainty on the part of the cited author would make the disagreement less face-threatening. Detailed disciplinary knowledge and an examination of the cited original would be required, however, to verify this suggestion.

Of course we would not expect all examples of ARGUE verbs to be used in a situation of uncertainty or conflict, since citation is also used for comparison with and in support of the writer’s own research ([Dudley-Evans, 1986](#); [Hopkins & Dudley-Evans, 1988](#)). In the following examples, the thesis writer makes use of the work of the cited author:

- (34) *Martin Wight has argued* that international political theory can be construed as a debate between three broad traditions. . . The three notions of state socialisation to be evaluated in this thesis represent each of these traditions. (pol1)
- (35) . . . *Okuda (Okuda et al. 1985) asserts* that below a thickness of 0.1 mm eddy current heating, even in metals, is negligible. Assuming this is true, the thickness of these samples was in the range 130–300 mm and thus large metallic phases within the film could suffer significant heating. (mat4)
- (36) All subsequent experimental papers rely on the analysis of Swartzentruber et al. (1990), whose formalism will be examined in detail, and a useful extension derived. To summarise their assertions, *they said* that. . . there are two distributions. . . (mat1)

Moreover, in presenting the research, particularly in a cumulative discipline such as materials, the thesis writer needs to show how one researcher provides the basis for another’s work:

- (37) *Anderson (1986) has noted* that. . . the interface mobility will [be] independent of driving force. Chen (1987) has taken this further. . . (mat7)
- (38) Also, Korner’s (1992) observations of incomplete KW locks indicate an agreement with Hirsch’s model. *Korner argues* that the screw dislocations may have cross-slipped onto (010). . . (mat6)
- (39) Pethica introduced the concept of an area function for the tip to correct for this. . . This concept has since been extended by Doerner and Nix (1986), *who argued* that the total displacement should be separated into elastic and plastic components. (mat3)

Thus, although a major function of ARGUE verbs in these corpora is to moderate potentially face-threatening comments, they are also used to create consensus between the thesis writer and others working in the field, as well as within the disciplinary community as a whole.

5. Conclusions

This paper has shown that for both corpora there is considerable use of reporting clauses with a *that*-clause complement in reporting others' research. Although there is some evidence of disciplinary differences in grammatical subject and verb choice, the most frequently occurring pattern is common to the two corpora: an integral citation, a human subject and a present tense ARGUE verb. I have suggested that one of the prime functions of this pattern is to moderate potentially face-threatening comments on other researchers, particularly in situations of conflict or uncertainty. This function stands in contrast to that of the other pattern identified: an integral citation, a human subject and a past tense FIND/SHOW verb in materials. This allows the writer to give credit to other researchers and to use their work in the cumulative construction of knowledge.

The key pattern identified here, integral citation + human subject + present tense ARGUE verb + *that*, is used by thesis writers to situate the research of others within the field so that they can position themselves in relation to it. There are certain features which make this pattern especially well-suited to the positioning of self and others. First, the use of the reporting clause opens up a space which allows comment on the cited proposition. Second, the use of integral citation with a human subject facilitates comment on individual researchers and studies. Third, the use of a present tense ARGUE verb conveys comment on the cited author's text. Finally, ARGUE verbs clearly show Thompson and Ye's (1991) three forms of evaluative potential: they can indicate the writer's stance towards the cited research; they allow the writer to attribute a stance to the cited author and they enable the writer to give an interpretation of the author's behaviour. As the biggest verb group, both in Francis et al. (1996) and in my data, ARGUE verbs provide a particularly rich evaluative resource. I would argue that it is all these characteristics which account for the predominance of integral citation with a human subject and a present tense ARGUE verb in both corpora. Clearly, however, further research on theses, using much larger corpora, is necessary in order to verify the importance of this pattern in a wider range of disciplines and to provide a more detailed account of the phraseology of citation.

6. Pedagogical implications

As noted earlier, the ability to cite appropriately is of key importance in academic writing, but presents considerable difficulties for the novice writer. I suggest that the investigation of citational patterns is of particular value, since it can reveal recurrent forms of expression, which can be linked to genre and/or disciplinary purposes. Attention to this phraseology in academic writing courses would encourage students to examine the wider context of citations, to become aware of the different functions of citation within the text and to relate this knowledge to sets of specific linguistic choices. It is helpful, for example, for students to realise that one way of incorporating citation into text is to use a reporting verb followed by a

that-clause and that in this pattern a human subject is likely to be used. Following on from this, discussion of contrastive examples, e.g. a comparison of the use of SHOW/FIND verbs with that of ARGUE verbs would enable students to recognise the importance of reporting verb choice and the association between verb group and tense choice. This, in turn, could lead to consideration of the wider context of the citation and the reasons for using a particular phraseology within that context. Working with patterns, then, can be beneficial in raising student awareness of contextual factors and in enhancing their understanding of what lies behind the language choices evident on the page. It is important for students to gain this insight into the extent and type of patterning found in written academic discourse, since this will help them to write in a way which is not just grammatically correct, but also appropriate both for their specific purposes and within their disciplinary community.

Acknowledgement

The permission of the thesis writers to use their texts for linguistic research is gratefully acknowledged.

Appendix 1. The Corpora

Politics/International Relations Corpus

Alderson, K. H. (1995). *Convergence under anarchy: Socialisation and the state system*. (pol1)

Bays, J. (1996). *Fireproof house to middle power: Metaphor, identity and Canadian foreign policy, 1939–1950*. (pol2)

Hobbs, P. W. (1996). *The US and the UK in the Middle East: ‘Alpha’ and the search for a Palestine settlement in 1955*. (pol3)

Kuttner, S. A. (1996). *Crossing new boundaries: Refugee regimes and gender-related persecution*. (pol4)

Martin, P. J. (1996). *Attitudes and coalition-building: The importance of significance in Burger court civil liberties cases*. (pol5)

McMenamin, K. I. (1996). *Giuseppe Mazzini’s nationalism in European context*. (pol6)

McQueen, C. V. (1996). *International relations and international law: Safe areas in Bosnia*. (pol7)

Webb, M. C. (1996). *Abandoning the world role: British foreign policy making, 1945–1967*. (pol8)

Materials Science Corpus

Bowler, D. R. (1997). *A theoretical study of gas source growth of the Si(001) surface*. (mat1)

Burden, A. P. (1996). *Electron microscopy techniques to further the understanding of conductive polymer composites.* (mat2)

Cock, A. (1998). *The high temperature erosion of coated thermal barrier tiles.* (mat3)

Hole, C. (1995). *Pulsed magnetic field characterisation of technological high temperature superconductors.* (mat4)

Hull, P. J. (1996). *Synthesis and characterisation of quantum dots.* (mat5)

Lunt, M. J. (1998). *The yield stress anomaly and inverse creep in Li_2 single crystals.* (mat6)

Miodownik, M. A. (1996). *Fundamentals of grain growth phenomena in ODS alloys.* (mat7)

Sealy, C. (1997). *SEM characterisation of semiconductor dopant distribution.* (mat8)

Appendix 2. Reporting Verbs in the Corpora: ARGUE, SHOW, FIND and THINK verb groups

Table 8

All ARGUE verbs in the corpora: frequencies per 100,000 words

Politics		Materials	
ARGUE Verb	Frequency	ARGUE Verb	Frequency
Argue	15.3	Suggest	7.0
Suggest	13.7	Argue	5.3
Assert	4.2	Conclude	4.0
Note	4.2	Report	3.7
Predict	4.2	Predict	2.7
Write	3.7	Propose	1.7
Explain	2.6	Assert	1.3
Conclude	2.1	Claim	1.3
Mention	2.1	State	1.3
Admit	1.6	Point out	1.0
Observe	1.6	Postulate	1.0
Accept	1.1	Accept	0.7
Claim	1.1	Estimate	0.7
Complain	1.1	Hypothesise	0.7
Imply	1.1	Acknowledge	0.3
Point out	1.1	Comment	0.3
Posit	1.1	Conjecture	0.3
Say	1.1	Contend	0.3
Add	0.5	Document	0.3
Hypothesise	0.5	Emphasise	0.3
Insist	0.5	Mention	0.3
Maintain	0.5	Say	0.3
Propose	0.5	Speculate	0.3
Remark	0.5	Stress	0.3
Reply	0.5		
Total	66.5	Total	35.4

Table 9

All SHOW verbs in the corpora: frequencies per 100,000 words

Politics		Materials	
SHOW Verb	Frequency	SHOW Verb	Frequency
Reveal	2.1	Show	14.0
Show	2.1	Indicate	3.7
Mean	0.5	Demonstrate	2.3
		Reveal	1.0
		Confirm	0.3
		Imply	0.3
Total	4.7	Total	21.6

Table 10

All FIND verbs in the corpora: frequencies per 100,000 words

Politics		Materials	
FIND Verb	Frequency	FIND Verb	Frequency
Find	1.1	Find	10.7
Realise	0.5	Note	2.3
		Observe	2.3
		Discover	1.3
		Calculate	0.7
		Establish	0.7
		Deduce	0.3
		Identify	0.3
		Infer	0.3
		Recognise	0.3
Total	1.6	Total	19.2

Table 11

All THINK verbs in the corpora: frequencies per 100,000 words

Politics		Materials	
THINK Verb	Frequency	THINK Verb	Frequency
Think	3.7	Assume	5.3
Hold	1.6	Know	2.0
Assume	1.1	Think	1.0
Feel	0.5		
Hope	0.5		
Total	7.4	Total	8.3

References

- Becher, T., & Trowler, P. (2001). *Academic tribes and territories* (2nd ed.). Buckingham: The Society for Research into Higher Education and Open University Press.
- Berkenkotter, C., & Huckin, T. (1995). *Genre knowledge in disciplinary communication: Cognition/culture/power*. Hillsdale, NJ: Erlbaum.

- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Harlow: Pearson Education.
- Bloch, J., & Chi, L. (1995). A comparison of the use of citations in Chinese and English academic discourse. In D. Belcher & G. Braine (Eds.), *Academic writing in a second language: Essays on research and pedagogy* (pp. 231–274). Norwood, NJ: Ablex.
- Dong, Y. R. (1996). Learning how to use citations for knowledge transformation: non-native doctoral students' dissertation writing in science. *Research in the Teaching of English*, 30(4), 428–457.
- Dudley-Evans, T. (1986). Genre analysis: an investigation of the introduction and discussion sections of MSc dissertations. In M. Coulthard (Ed.), *Talking about text* (pp. 128–145). University of Birmingham: English Language Research.
- Francis, G., Hunston, S., & Manning, E. (1996). *Collins COBUILD grammar patterns 1: Verbs*. London: HarperCollins.
- Gilbert, G. N., & Mulkay, M. (1984). *Opening Pandora's box*. Cambridge: Cambridge University Press.
- Groom, N. (2000). Attribution and averral revisited: three perspectives on manifest intertextuality in academic writing. In P. Thompson (Ed.), *Patterns and perspectives: Insights into EAP writing practice* (pp. 14–25). Reading: University of Reading.
- Harwood, N. (2004). Citation analysis: a multidisciplinary perspective on academic literacy. In M. Baynham, A. Deignan, & G. White (Eds.), *Applied linguistics at the interface* (pp. 79–89). London: BAAL and Equinox.
- Hoey, M. (2001). *Textual interaction*. London: Routledge.
- Hopkins, A., & Dudley-Evans, T. (1988). A genre-based investigation of the discussion sections in articles and dissertations. *English for Specific Purposes*, 7, 113–121.
- Hunston, S. (1993). Professional conflict: disagreement in academic discourse. In M. Baker, G. Francis, & E. Tognini-Bognelli (Eds.), *Text and technology: In honour of John Sinclair* (pp. 115–134). Amsterdam: Benjamins.
- Hunston, S., & Francis, G. (2000). *Pattern grammar*. Amsterdam: Benjamins.
- Hyland, K. (1999). Academic attribution: citation and the construction of disciplinary knowledge. *Applied Linguistics*, 20(3), 341–367.
- Hyland, K. (2002a). Activity and evaluation: reporting practices in academic writing. In J. Flowerdew (Ed.), *Academic discourse* (pp. 115–130). London: Longman.
- Hyland, K. (2002b). What do they mean? Questions in academic writing. *Text*, 22(4), 529–557.
- Jacoby, S. (1987). References to other researchers in literary research articles. *ELR Journal*, 1, 33–78.
- Myers, G. (1990). *Writing Biology: Texts in the social construction of scientific knowledge*. Madison, WI: University of Wisconsin Press.
- Pecorari, D. (2003). Good and original: plagiarism and patchwriting in academic second-language writing. *Journal of Second Language Writing*, 12, 317–345.
- Scott, M. (1999). *WordSmith tools (version 3.0)*. Oxford: Oxford University Press.
- Shaw, P. (1992). Reasons for the correlation of voice, tense, and sentence function in reporting verbs. *Applied Linguistics*, 13(3), 302–319.
- Swales, J. M. (1981). *Aspects of article introductions*. Aston ESP Research Reports. Birmingham: The University of Aston Language Studies Unit.
- Swales, J. M. (1986). Citation analysis and discourse analysis. *Applied Linguistics*, 7(1), 39–55.
- Swales, J. M. (1990). *Genre analysis*. Cambridge: Cambridge University Press.
- Thomas, S., & Hawes, T. P. (1994). Reporting verbs in medical journal articles. *English for Specific Purposes*, 13(2).
- Thompson, G., & Hunston, S. (2000). Evaluation: an introduction. In S. Hunston & G. Thompson (Eds.), *Evaluation in text: Authorial stance and the construction of discourse* (pp. 1–27). Oxford: Oxford University Press.
- Thompson, G., & Ye, Y. (1991). Evaluation in the reporting verbs used in academic papers. *Applied Linguistics*, 12(4), 365–382.
- Thompson, P. (2001). *A pedagogically-motivated corpus-based examination of PhD theses: macrostructure, citation practices and uses of modal verbs*. Unpublished PhD thesis, University of Reading.

- Thompson, P., & Tribble, C. (2001). Looking at citations: using corpora in English for academic purposes. *Language Learning and Technology*, 5(3), 91–105.
- Weissberg, R., & Buker, S. (1990). *Writing up research*. Englewood Cliffs, NJ: Prentice-Hall.

Maggie Charles is a Tutor in EAP at Oxford University Language Centre, where she teaches academic writing. She has recently completed a doctorate at Birmingham University. Her interests include the study of evaluation, discipline-specific EAP and the application of insights from corpus linguistics to the teaching of academic writing.