Making concessions in academic writing: A corpus study of patterns and semantic sequences

Maggie Charles
Oxford University Language Centre

1. Introduction

The concessive relation is defined by Aarts (1988: 40) as follows:

> 'The concessive relation expresses a semantic discrepancy between the matrix and subordinate clauses: what is said in one clause is surprising and unexpected in the light of what is said in the other clause.'

Aarts' definition highlights the semantic elements that characterise this logical relation, focusing particularly on the mismatch between the information given in the two clauses. Thus, in example (1) taken from the corpus described in Section 2, the expectation implied in the *although* clause is that an estimate would lead to inaccuracy, but the information in the matrix clause shows that this expectation is not fulfilled:

(1) *Although* we had to use an estimate for the Timer 1 it was still accurate to 1 sec. (0258i: Computer Science)

Biber et al. (1999) refer to this logical relation as 'contrast/concession' and draw attention to its role in discourse structure, making the point that such clauses are important in the construction of arguments. In academic writing, contrast/concession statements allow the limitations on facts, events or claims to be presented. This restrictive effect is evident in the example above, where the claim of accuracy in the matrix clause is limited by the use of the estimate mentioned in the subordinate clause.

The construction of argumentation is a key function of academic discourse and can be a source of difficulty for international students. Studies of academic writing in this area have primarily focused on the use of linking adverbials (e.g. Charles 2011a, b; Gardezi & Nesi 2009) with a wealth of research comparing the written production of native (NS) with non-native speakers of English (NNS) (e.g. Altenberg & Tapper 1998; Field & Yip 1992; Granger & Tyson 1996; Tankó 2004). Little attention has been paid to the role of subordinators, although it is likely that they are of equal importance to the development of convincing arguments and their use may well also pose problems for the student writer.

Much of the work on linking adverb use focuses on the ways in which learner writing deviates from professional norms by over- under- or misusing certain linking adverbials (Bolton et al. 2002; Milton and Tsang 1993; Shaw 2009). Relatively less research has been devoted to the writing of successful students. However, as I have argued elsewhere (Charles 2011a, b), the characteristics of professional writing (e.g. the research article), are not necessarily the same as those we would expect to find in student writing since the genres and purposes of the texts are very different. Nor can we assume that all student writers aspire to write research articles

in the future. There is a pedagogical argument, then, for investigating the features of successful student texts, since such data could inform materials and courses on academic writing, particularly at undergraduate and Master's level.

The aim of this paper, then, is to use corpus tools to investigate the five subordinators of contrast/concession attested by Biber et al. (1999) ('although', 'though', 'while', 'whilst' and 'whereas'), and to shed light on the ways in which they contribute to successful student argumentation. The paper identifies the patterns and semantic sequences with which these subordinators are associated and suggests that they are important for understanding and teaching the construction of arguments in academic writing.

2. Semantic Sequences

In examining the role of contrast/concession subordinators in successful student writing, it is helpful to make use of the notion of semantic sequence developed by Groom (2007) and Hunston (2008, 2009). Hunston (2008: 271) defines semantic sequences as 'recurring sequences of words and phrases' which constitute 'sequences of meaning elements' and stresses that they cannot be considered as formal sequences because of their diversity of form. Using a corpus of articles from New Scientist and starting with the grammar pattern: Noun + *that*-clause, she shows how the noun *observation* occurs predominantly as part of two longer sequences of meaning elements:

> **THEORY/ARGUMENT + ARISES FROM + *the observation* + *that*-clause**[1]
> ***the observation* + *that*-clause + CONSISTENCY + THEORY/ARGUMENT.**

She makes the point that these semantic sequences reflect the way in which natural science disciplines make use of observation in constructing knowledge.

Groom (2007) examines two corpora of research articles in history and literary criticism. Drawing on methods developed by Gledhill (2000), he uses salient grammatical words (e.g. prepositions, pronouns) as probes to identify semantic sequences. He notes both similarities and differences in the typical semantic sequences that occur and links these to the epistemology of their respective disciplines. For example, investigating the preposition *among*, Groom finds that the following sequence occurs only in the history corpus:

> **STATISTICAL INDICATORS + *among* + SOCIAL GROUP OR INSTITUTION**

By contrast, only the literary corpus contains the following sequence:

> **RELATIONSHIP + *among* + CONCEPTUAL PHENOMENA**

Such data show evidence of the greater interest in social processes in historical discourse and the tendency in literary studies to focus on the abstract relations between concepts. The value of semantic sequences, then, particularly for pedagogical applications in academic writing, is that they enable regularities in phraseology to be linked to extended units of meaning which

[1] I follow the orthographic conventions established by Groom (2007): bold is used for the whole semantic sequence; small caps for semantic elements; italics for word forms; + is used to show the sequence of elements.

can be used to teach students how longer stretches of text are frequently structured and expressed.

3. Corpus Data and Method

The present study draws on data from the British Academic Written English Corpus (BAWE), which contains assignments written by NS and NNS students for assessment at three universities in the UK and awarded good grades. The whole corpus contains data from 35 disciplines, amounting to roughly 6 million words and assignments cover the work of undergraduates in years 1, 2 and 3 and graduate students at Master's level.

In this paper, four of the disciplinary subcorpora are selected for detailed examination. The data come from four contrasting disciplines, representing each of the knowledge groupings identified by Becher and Trowler (2001): Business Studies (soft-applied); Chemistry (hard-pure); Computer Science (hard-applied); and Politics (soft-pure). This corpus amounts to just under a million words and consists of over 400 assignments. Details are given in Table 1.

	Business Studies	Chemistry	Computer Science	Politics	Total
No words	321,116	128,298	202,890	320,020	972,324
No files	146	89	87	110	432

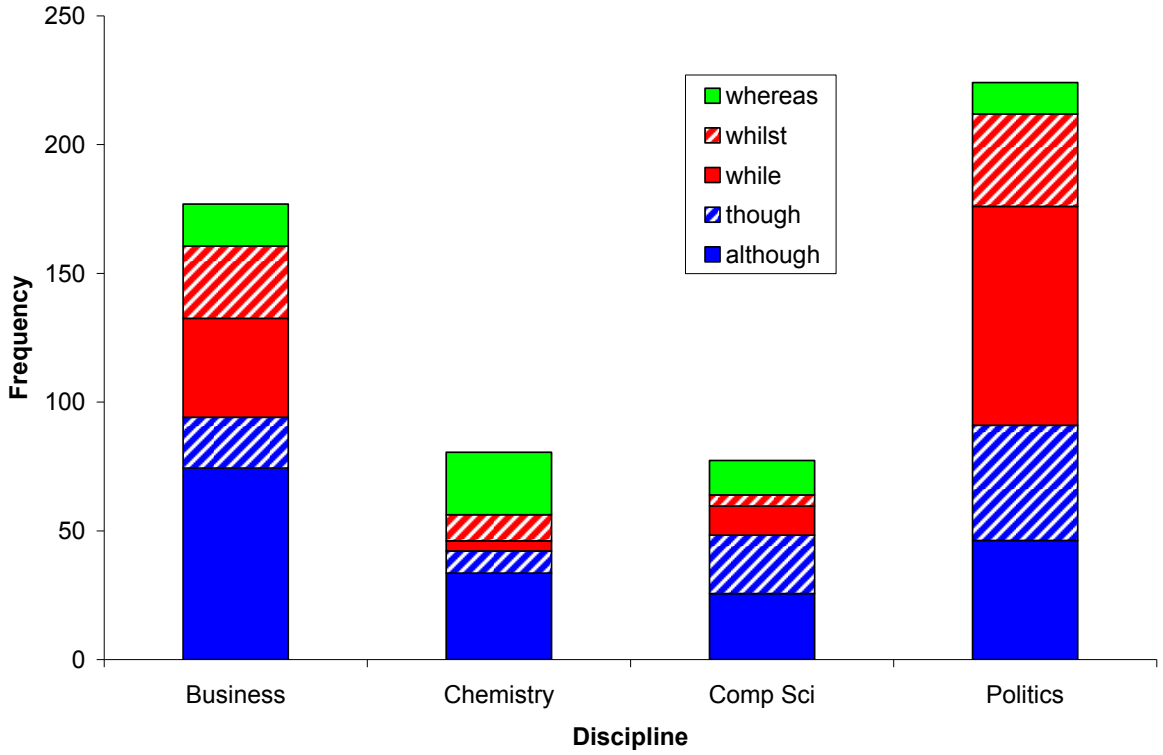
Table 1. Data on subcorpora

The two social sciences, Politics and Business Studies, are considerably larger in terms of the total number of words (around 300,000) than those in the natural sciences, with Chemistry and Computer Science containing only around 100-200,000 words. The social science subcorpora also contain more files, over 100 in both cases, while for the natural sciences, the totals are closer to 90. It is probable that the discrepancy in these figures generally reflect the greater amount of writing that is required in social sciences as opposed to natural sciences.

The four subcorpora were examined using WordSmith Tools software (Scott 2005) for all instances of the five subordinators of contrast/concession attested in Biber et al. (1999), *although, though, while, whilst* and *whereas*. Concordances were retrieved for each subordinator. Occurrences of *even though* were included in the analysis, but instances of *though* as a linking adverb were eliminated. Although subordinators of contrast/concession can introduce non-finite and verbless clauses, numbers are very low in these subcorpora and thus only instances which introduce finite clauses are discussed here.

4. Overall results and discussion

The results for each of the five subordinators are given in Figure 1. Figures are normalised per 100,000 words to allow comparison of different sized subcorpora.



The two social science fields show total figures that are at least twice as high per 100,000 words as those found for the hard disciplines (Business Studies: 176.9; Politics: 224.1; Chemistry: 80.5; Computer Science: 77.3). This finding is in the expected direction, since the higher frequencies in the soft disciplines reflect the prevalence of recursive knowledge construction in those fields (Becher & Trowler 2001). Thus disciplines such as Politics and Business Studies tend to develop ideas and knowledge by taking issue with the work of other scholars, which is liable to lead to more argumentation and hence the need for more markers of contrast/concession. Becher and Trowler (2001) note, however, that work in the hard natural sciences tends to proceed cumulatively, with one researcher building on the work of another. This type of knowledge construction is therefore likely to require lower use of contrast/concession subordinators.

4.1 Patterns

There are three patterns associated with the use of contrast/concession subordinators to introduce finite clauses: initial position of the subordinate clause (i.e. before the matrix clause); final position of the subordinate clause (i.e. after the matrix clause) and medial position, in which the subordinate clause interrupts the matrix clause. These are exemplified below:

(2) *While* there have been increasing demands and net contributions for both Sonites and Vodites products, the sales of SELF had been decreasing gradually. (0237b: Business Studies Initial)

(3) Alternative software does not receive enough advertising to make people switch - *even though* many are more secure. (3072a: Computer Science Final)

(4) The turnout of the 1997 general election, *although* it now seems like a relatively high percentage, 71%, is seen as marking the start of a downwards trend… (0275c: Politics Medial)

In these subcorpora, the subordinators are also sometimes used in a non-standard way in order to introduce a main clause:

(5) *Whereas* Polymer 5D involved addition of BzMA to an already forming MMA polymer after 4 hours. (0382a: Chemistry)

However medial instances and non-standard uses are few and it was therefore considered more worthwhile to focus on the two major patterns associated with the subordinators: initial position and final position.

Data on academic prose from the Longman Spoken and Written English Corpus (LSWE) given in Biber et al. (1999: 833) show a slight preference for final position. Totals for these BAWE subcorpora, however, give the opposite finding, with 289.5 occurrences in initial and 200.2 in final position (see Table 2). Thus writers in these subcorpora are roughly a third more likely to put the subordinate clause before the matrix clause than after it. The LSWE corpus contains research articles and extracts from books, mostly written for a technical audience and is thus a corpus of skilled professional writing (Biber et al. 1999: 32). This opens the question of the extent to which the difference between the two sets of findings is due to the status of the writer, an expert in the LSWE corpus, but an apprentice or student in the BAWE corpus.

When we look at the data by individual discipline, however, the picture becomes more complex. The normalised figures for each pattern in the four disciplines are given in Table 2.

Subcorpus	Initial Position	Final Position
Business	83.2	64.8
Chemistry	32	37.5
Comp Sci	39	33.5
Politics	135.3	64.4
Total	289.5	200.2

Table 2. Patterns of contrast/concession subordinators: Frequency per 100,000 words

There is a marked difference between social and natural sciences in terms of preference for initial or final position. Thus Chemistry shows results which are similar to those of the LSWE Corpus, with slightly more instances of final than initial position. In Computer Science the use of the two positions is also fairly equal, although in this case the preference is slightly for initial position. However when we examine the social sciences, the preference for initial position is clear. The tendency is most marked in Politics, which has over twice as many clauses in initial as in final position, but it is also noticeable in Business Studies, which has about a third more initial instances. In disciplinary terms, the Chemistry and Computer Science subcorpora may well be the closest to the LSWE Corpus, which contains a preponderance of natural science writing. In genre terms, too, the methodology recount which predominates in the Chemistry corpus most resembles the classic research article IMRD

structure (Introduction-Method-Results-Discussion). These two factors may thus account for the closer agreement in pattern use seen between Chemistry and Computer Science on the one hand and the LSWE Corpus on the other.

However, focusing on subordinator use as a whole conceals the considerable differences in the patterns associated with individual markers. In the next section, I will investigate individual subordinators along with some of their typical semantic sequences and will argue that the student assignments use contrast/concession markers to perform functions that are necessitated by the specific requirements of assessed student writing and that these requirements differ considerably from those of professional writing.

5. Results and Discussion of Individual Subordinators

In commenting on the individual subordinators, I group together *although* and *though*, since they can be considered as alternants (Huddleston & Pullum 2002). They vary only in level of formality, with the more formal *although* the preferred option in academic writing (Aarts 1988; Biber et al. 1999). Similarly I discuss *while* and *whilst* together, considering them as variants chosen according to writers' individual preferences.

5.1. *Although* and *Though*: Patterns and Semantic Sequences

Table 3 shows that the predominant pattern associated with *although/though* in all four subcorpora is the initial position of the clause. In Chemistry and Computer Science there are roughly twice as many instances of initial than final position, while in Business Studies and Politics, initial position is around three times as frequent.

Subcorpus	Initial Position	Final Position
Business	58.2	18.8
Chemistry	24.2	13.3
Comp Sci	28.1	14.8
Politics	57.5	19.4

Table 3. *Although/Though* patterns: Frequency per 100,000 words

The position of the subordinate clause has an effect on information structure (Biber et al. 1999: 835). Thus the use of initial position presents the subordinate clause as given information and puts emphasis on the more important new information in the subsequent matrix clause. The question then arises as to why this arrangement of clauses is so clearly preferred in these BAWE subcorpora. In seeking to offer a possible explanation, it is useful to examine the semantic sequences that are associated with *although/though*, since these could give an indication of the functions which the pattern is being used to perform.

It is noticeable that the both matrix and subordinate clauses often contain signals of evaluation, including, for example, lexical markers, modal verbs, negation or comparatives. Further, the subordinate clause tends to be marked by positive evaluation, while the matrix clause tends to evaluate negatively:

(6) *Although* the **figures** given in Table 4 are **useful for** obtaining a feel for the results, in the absence of any significance tests, the **statistical validity** of any comparisons made is **not assured**. (0232b: Business Studies)

(7) *Although* this **experiment** presents a **rigorous method** for assessing fluoride concentrations, there is **scope for improvement**. (0415c: Chemistry)

(8) **Though** the **conclusions** drawn when using the model are **reliable** to a certain degree, the few variables used when constructing the initial hypothesis **allow for criticisms to be made as to the precision of the findings**. (0269b: Politics)

This sequence is sometimes extended with the addition of a reason, as exemplified below:

(9) *Although* the **first three tests were successful**, the **fourth was not**. **The reason for this was** that when using the strictly greater than sign, neither of the threes is the maximum. (6101e: Computer Science)

The whole sequence can be expressed as follows:

although/though + **POSITIVE EVALUATION** + **NEGATIVE EVALUATION** + **REASON (optional)**

Such sequences are more frequent in the two natural science corpora, occurring with 58% of the instances of *although/though* in Chemistry and 43% in Computer Science, while the figures for Politics and Business Studies are 32% and 26% respectively. The examples cited so far all deal with the writer's work and it is the positive evaluation of their own work which is presented in the given information slot, as a conceded proposition, while the negative evaluation is highlighted by occurring as new information in the matrix clause. The higher frequencies seen in Chemistry and Computer Science may reflect the fact that these disciplines involve practical or experimental tasks, which can easily go wrong and where problems must then be acknowledged. Although this may help account for the greater extent of negative evaluation in these disciplines, it does not provide an explanation for the semantic sequence as a whole. Indeed the order of information may seem rather unexpected, emphasising, as it does, the negative rather than the positive aspects of the writer's work. However I would argue that it is explicable in terms of the requirements of assessed student writing.

Assigned tasks often require students to evaluate their own work and even if not explicitly asked to do so, such evaluation is often prized by assessors as a means of demonstrating the student's understanding of disciplinary standards and their ability to be self-critical. In particular, being upfront about problems and difficulties can create academic credibility, giving the impression of a careful, honest and reliable worker. However, students are generally keen to obtain the highest grade possible for a given piece of work. Thus if a student only points out the negative aspects of an assignment, this would tend to undermine its value and might lead to a lower mark. In order to redress the balance, then, students may find it necessary to note the positive features as well. One of the ways of doing this unobtrusively is to introduce these positive elements as given information in the subordinate clause, while reserving the new information position for the negative aspects. Thus the positive information serves to restrict the force of the negative assessment. At the same time, the sequence enables students to show appropriate academic modesty: by downgrading the importance of their

positive evaluation, the student encroaches less on the assessor's role and leaves them space to reach their own conclusions.

The semantic sequence described above is also used in these corpora to evaluate the work of other researchers:

(10) *Although* the **prior studies on business - IT alignment** has been **helpful** in general, **many previous researches** have **ignored the notion** of context dependency… (0291b: Business)

(11) *Although* **this paper** contains some **interesting chemistry** overall it is **fundamentally flawed.** (0388f: Chemistry)

(12) *Although* trial software / prototyping is a **very successful** methodology **in many studies**… it seems **not a good choice** of design methodology in safety-critical system development… (3039a: Computer Science)

Again, we can offer an explanation based on the requirements of assessed assignments. Student writers are frequently urged to be critical; at the same time making criticisms of well-established researchers or research approaches is potentially a face-threatening act and may be perceived by students as inappropriate or risky. The use of the above-described semantic sequence enables the student to add a positive evaluation to their critical comment, which limits and mitigates the effect of the threat and renders the criticism more acceptable. Putting the positive evaluation into the subordinate clause presents it as given information and allows the writer to emphasise their criticism in the new information of the matrix clause. In this way, the contributions of established experts are acknowledged, academic modesty is maintained and the two-part sequence of positive and negative comment constructs a balanced and reasonable argument. Thus in both uses of this semantic sequence, writers show awareness of their relatively lowly status in the field and respond to demands which are characteristic of student writing: the need for critical evaluation of self and others while maintaining academic modesty.

5.2. *Whereas*: Patterns and Semantic Sequences

The prevalent pattern found for *whereas* offers a striking contrast to that for *although/though*, as seen in Table 4. In all four corpora final position is preferred, with considerable disparity in the use of the two patterns: in Politics final position is almost twice as frequent; in Chemistry, almost 10 times as frequent, while in Business Studies and Computer Science, final position is over 20 times as frequent, with initial position scarcely used at all.

Subcorpus	Initial Position	Final Position
Business	0.6	13.4
Chemistry	2.3	20.3
Comp Sci	0.5	12.8
Politics	4.4	7.8

Table 4. *Whereas* patterns: Frequency per 100,000 words

Whereas often signals a two-part parallel structure in which one entity is contrasted with another. Here the element of unexpectedness or surprise noted by Aarts (1988) as typical of concessions seems to be missing and this subordinator can be considered only peripherally concessive (Aarts 1988; Huddleston & Pullum 2002). Examples are given below:

(13) …**finance** is essential to the small enterprise **during the initial stage of growth**, *whereas* **networks** are fundamental to the firm's **long-term sustainability**. (0271b: Business)

(14) **Encryption** is the process of **coding the plaintext into ciphertext**, *whereas* **decryption** is the **reverse** process. (6170e: Computer Science)

The semantic sequence can be written as follows:

ENTITY + STATEMENT + *whereas* + CONTRASTING ENTITY + CONTRASTING STATEMENT

A variation on the sequence is also found in which the prepositions *in* or *for* introduce a context, which is then contrasted with a second context.

i*n/for* + CONTEXT + STATEMENT + *whereas* + *in/for* + CONTRASTING CONTEXT + CONTRASTING STATEMENT

This sequence is exemplified below:

(15) **For a tetrahedron** the bond angles are **109º**, *whereas* **for the square planar** there is an angle of **90º**. (6123c: Chemistry)

(16) **In their theory of EU governance**, it is the **supranational level** that is emphasised, *whereas* **in the traditional literature** it is the **nation state**. (0255c: Politics)

These semantic sequences are very frequent in all four subcorpora, with over 70% of all clauses in final position showing these two-part parallel contrasts (77% in Business Studies; 85% in Chemistry; 76% in Computer Science; 72% in Politics).

A similar semantic sequence is used with initial pattern:

***whereas* + ENTITY + STATEMENT + CONTRASTING ENTITY + CONTRASTING STATEMENT**

This is illustrated here in Politics and Chemistry, the two disciplines that make reasonable use of initial subordinate clauses:

(17) *Whereas* the protected **arylimine precursors** reported here are **stable**, **alkyl ones** are sometimes **inaccessible**… (0415h: Chemistry)

(18) *Whereas* **Realism** therefore "paints a **rather grim picture** of world politics", **Liberal notions** convey a **significantly more optimistic view**. (0138c: Politics)

The semantic sequences in both patterns tend to construct straightforward contrasts relating to the subject matter rather than concessions to other potential or actual views. In this regard it is interesting to note that the subcorpus that uses *whereas* by far the most frequently is

Chemistry. This holds true both for the normalised figures (Chemistry: 24.2; Business Studies: 16.3; Computer Science: 13.3; Politics: 12.2) and as a percentage of the total use of contrast/concession subordinators in each subcorpus (Chemistry: 30%; Computer Science: 17%; Business Studies: 9%; Politics: 5%). This relatively high use of *whereas* may well reflect the fact that, as a hard-pure science, Chemistry bases its knowledge construction primarily on experimental data rather than argument and may therefore be more concerned with establishing contrastive distinctions between entities than with constructing arguments based on differing views.

Both semantic sequences are used to give detailed explanations and descriptions and their primary function is to display the student's knowledge. In the case of *whereas* too, then, the prevalence of these sequences can be seen as a response to the requirements of assessed student writing. In order to get good grades, writers have to set out what they have learnt, describing and differentiating between key entities and concepts in order to establish the distinctions necessary to show their acquisition of disciplinary knowledge.

5.3. *While/Whilst* Patterns and Semantic Sequences

In all the subcorpora except for Business Studies, the predominant pattern is the same as for *although/though*: initial position. Frequencies per 100,000 words are given in Table 5 and reveal large differences in use between the two patterns. Thus Computer Science and Politics show almost twice as many instances of initial as final position, while Chemistry has almost 50% more initial occurrences. By contrast, Business Studies uses almost 50% more instances of final than initial position.

Subcorpus	Initial Position	Final Position
Business	24.4	32.2
Chemistry	5.5	3.9
Comp Sci	10.4	5.4
Politics	73.4	37.2

Table 5. *While/Whilst* patterns: Frequency per 100,000 words

In initial position we find usage of a similar evaluative sequence to that prevalent for *although/though*:

while/whilst + **POSITIVE EVALUATION** + **NEGATIVE EVALUATION**

As with *although/though*, the sequence can be used to evaluate the student's own work in a way that makes clear its flaws, while also noting its good points:

(19) *Whilst* comparison to the data available in literature shows **good agreement** with **our rate law** (Okhanowic and Williams, 1974) **our value** for the rate constant however, is **somewhat deviant**. (6123f: Chemistry)

Similarly, the research of others can also be evaluated, balancing positive comments against negative:

(20) *While* **Cha's contention** is **admirable**, it still **fails to recognise** that the D.P.R.K will continue the same mode of coercive negotiation… (0255b: Politics)

Such uses of the sequence may be seen as operating in a similar way to those discussed for *although/though* in that they show how successful students negotiate and shape their responses to the demands of their assignments by being critical, but not too critical, of self and others.

The sequence is also applied to the subject matter of the field, where the evaluation may be of an individual action or entity, as seen below in Business Studies, or of the discipline as a whole, as in the following example from Computer Science:

(21) *Whilst* their chairman **backs change** he **fails** to do so visibly to the organisation… (0172a: Business Studies)

(22) *While* **developments are being made**, **there is still some way to go** before Weiser's vision of ubiquitous computing is fully realised. (6160a: Computer Science)

Again, the writer sets out both sides of the argument, weighing them up and coming to a balanced and reasonable conclusion. This is a quality that is likely to be highly valued by assessors and students could benefit from seeing how it is achieved through the use of this semantic sequence.

I now turn to the final position pattern, focusing particularly on Business Studies, in which its use is predominant. Two semantic sequences can be distinguished, both of which serve to develop the student's argument by expanding or commenting on generalisations. The first consists of an initial generalisation followed by the matrix clause which contains a statement of quantity and the final *while/whilst* clause, which includes a contrasting statement of quantity:

GENERALISATION + QUANTITATIVE STATEMENT + *while/whilst* + CONTRASTING QUANTITATIVE STATEMENT

As with the *whereas* sequences, a two-part parallel contrast is set up, but here it functions explicitly to give supporting detail to the earlier generalisation. This sequence is illustrated by the following examples:

(23) …it is clear that unionisation has fallen. **In 1980**, approximately **65%** were union members, *while* **in 1998** it was down to **36%** (Cully et al 1998:235). (0202l: Business Studies)

(24) The employer lobby opposes union involvement on the pretext of time constraints and a preference for informal employee relations. Between 1984 and 1998, the number of workplaces with only **representative voice mechanisms halved**, *while* those adopting **direct voice trebled**. (0178c: Business Studies)

(25) As the time horizon increases so do the future prospects for Santa Clara. In 10 years and using a 10% discount rate **Santa Clara** will be worth **$40.9** *whilst* **Waltham** will be worth **$40.1**. (0225b: Business Studies)

The second semantic sequence, which is characteristic of *while,* serves the same general function of adding supporting detail to a generalisation, but in this case, the detail occurs in the form of an example and is signalled by *for example* or *for instance.* The semantic sequence is thus as follows:

> **GENERALISATION + *for example/for instance* + EXAMPLE + *while* + CONTRASTING EXAMPLE**

This is exemplified below:

> (26) Firstly, different forms of direct EI techniques have different impacts on employee attitudes. **For example**, Cotton (1999) suggests that **self-directed work teams** have a **strong effect** on employee attitudes *while* **job enrichment** has an **intermediate effect** and **quality circles** have a **weak effect**. (0264c)

> (27) Unfortunately, the management action of fostering collaboration within informal groups and focusing on social relations does not seem to bring the predicted outcomes. **For instance**, **managers** assume that a two-way communication attachment could **promote co-operation with workers**, *while* **employees** treat it as information which **management can use to eliminate misunderstanding** (Bendix, 1963). (0124a: Business Studies)

These semantic sequences function in a similar way to those described above in that they reflect the specific needs and requirements of assessed assignments. They enable the student writer to display disciplinary knowledge in sufficient detail to gain a good grade and they show how successful arguments use quantitative data and specific examples to support generalisations. Many of the semantic sequences with *while/whilst* also provide information which is attributed to cited authorities, as seen in examples (23), (26) and (27) above. Such uses of the sequence also illustrate one of the ways in which high grade assignments incorporate appropriate source use into their argumentation.

6. Conclusions

This paper has examined four subcorpora of student writing from the BAWE corpus in relation to the patterns and semantic sequences that are associated with subordinators of contrast/concession. Pattern data show an overall preference for the initial position of the subordinate clause, which is not in agreement with results reported in Biber et al. (1999) for expert writing, where a slight predominance of final position is found. The BAWE data also show that pattern preferences vary according to individual subordinator and discipline. Thus *although/though* and *while/whilst* are generally more frequent in initial position and *whereas* in final position; comparison of disciplines shows that final position *whereas* has a relatively high frequency in Chemistry, in contrast to final position *while/whilst,* which occurs frequently in Business Studies.

The examination of semantic sequences has been used in order to shed light on such discrepancies and I have suggested that the sequences seen here reflect key characteristics of successful student writing. Thus the initial pattern of *although/though* provides students with a means of satisfying the requirement to be critical, both of their own work and that of others, but at the same time allows them to mitigate their criticisms and display academic modesty.

The two-part parallel contrasts constructed by *whereas* can be used to display students' in-depth understanding of distinctions between disciplinary entities or concepts and the use of *while/whilst* in sequences involving generalisation enables them to expand and support their statements with contrastive examples and quantitative data. In employing such semantic sequences, then, student writers are responding to the specific demands of assignments written for purpose of assessment.

However, this study has only focused on high quality student assignments; what it has not been able to show is whether there are systematic differences in the use of individual subordinators between successful and less successful student writing on the one hand or between successful student and expert professional writing, on the other. The further examination of student writing, both of high and low quality, could help establish the discourse characteristics that contribute to student success. Such work would greatly assist both materials writers and teachers of academic writing courses.

Similarly, establishing the extent and nature of the differences between expert and student writing would be of great value to course providers in deciding on the most appropriate models and examples to offer to students, particularly at undergraduate and Master's level. Currently, much of the analysis of academic writing has dealt with the research article. However, as noted in the introduction, it is by no means clear that the characteristics identified for this genre are relevant or useful for students at the early stages of their academic writing careers. It remains to be determined whether sequences such as those identified in the BAWE subcorpora can also be found in expert writing and if so, what their role might be.

In her discussion of the value of corpus-based descriptions for learners of English, Hunston (2009) casts a measure of doubt upon the usefulness of semantic sequences for pedagogical applications, since they do not distinguish between what is correct and incorrect. However, for university level learners, satisfying assessors in terms of well-structured arguments is likely to be at least as important as linguistic accuracy. By revealing the phraseology and discourse structure of longer stretches of text, the study of semantic sequences has the potential to raise student awareness both of what is often said and how it is usually expressed.

Acknowledgement

The data in this study come from the British Academic Written English (BAWE) corpus, which was developed at the Universities of Warwick, Reading and Oxford Brookes under the directorship of Hilary Nesi and Sheena Gardner (formerly of the Centre for Applied Linguistics [previously called CELTE], Warwick), Paul Thompson (formerly of Department of Applied Linguistics, Reading) and Paul Wickens (Westminster Institute of Education, Oxford Brookes), with funding from the ESRC (RES-000-23-0800). (http://www.coventry.ac.uk/bawe)

References

Aarts, B. (1988). Clauses of concession in written present-day British English. *Journal of English Linguistics, 21*, 39-58.

Altenberg B. & M. Tapper (1998) The use of adverbial connectors in advanced Swedish learners' written English. In Granger S. (ed) *Learner English on computer.* London: Addison Wesley Longman, 80-93.

BAWE Corpus http://wwwm.coventry.ac.uk/researchnet/BAWE/Pages/BAWE.aspx

Becher, T., & Trowler, P. (2001). *Academic Tribes and Territories*. (2nd ed.). Buckingham: SRHE and Open University Press.

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman Grammar of Spoken and Written English*. Harlow: Pearson Education.

Bolton K., G. Nelson & J. Hung (2002). "A corpus-based study of connectors in student writing. Research from the International Corpus of English in Hong Kong (ICE-HK)". *International Journal of Corpus Linguistics* 7: 165-182.

Charles M. (2011a). "Corpus evidence for teaching adverbial connectors of contrast: *however, yet, rather, instead* and *in contrast"*. In N. Kübler (ed), *Corpora, Language Teaching and Resources: From Theory to Practice*. Bern: Peter Lang.

Charles, M. (2011b). Text connection at tertiary level: Patterns and functions. In E. Panourgia (Ed.), *Proceedings of the 2nd Conference on ESP/EAP* (pp. 22-31). Kavala, Greece: Kavala Institute of Technology.

Field Y. & L. M. O. Yip (1992). "A comparison of internal conjunctive cohesion in the English essay writing of Cantonese speakers and native speakers of English". *RELC Journal* 23: 15-28.

Gardezi A. & H. Nesi (2009). "Variation in the writing of economics students in Britain and Pakistan: the case of conjunctive ties". In M. Charles, D. Pecorari & S. Hunston (eds), *Academic Writing: At the Interface of Corpus and Discourse*. London: Continuum, 236-250.

Gledhill, C. J. (2000). *Collocations in science writing*. Tübingen: Gunter Narr.

Granger S. & S. Tyson (1996). "Connector usage in the English essay writing of native and non-native EFL speakers of English". *World Englishes* 15: 17-27.

Groom, N. W. (2007). *Phraseology and Epistemology in Humanities Writing: A Corpus-driven Study.* Unpublished PhD thesis, University of Birmingham.

Huddleston, R., & Pullum, G. (2002). *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.

Hunston, S. (2008). Starting with the small words: Patterns, lexis and semantic sequences. *International Journal of Corpus Linguistics, 13*(3), 271-295.

Hunston, S. (2009). The usefulness of corpus-based descriptions of English for learners. In K. Aijmer (Ed.), *Corpora and Language Teaching* (pp. 141-154). Amsterdam: Benjamins.

Milton J. & E. S. Tsang (1993). "A corpus-based study of logical connectors in EFL students' writing: directions for future research". In R. Pemberton & E. S. Tsang (eds), *Studies in Lexis*. Hong Kong: HKUST Language Centre, 215-246.

Scott, M. (2005). WordSmith Tools (Version 4). Oxford: Oxford University Press.

Shaw P. (2009). "Linking adverbials in student and professional writing in literary studies: what makes writing mature". In M. Charles, D. Pecorari & S. Hunston (eds), *Academic Writing: At the Interface of Corpus and Discourse.* London: Continuum, 215-235.

Tankó G. (2004). "The use of adverbial connectors in Hungarian university students' argumentative essays". In J. M. Sinclair (ed), *How to Use Corpora in Language Teaching.* Amsterdam: Benjamins, 157-181.