Understanding Metadiscoursal Use: Lessons from a 'Local' Corpus of Learner Academic Writing

Wendy Noble, The Australian National University, Australia

Abstract
Classroom-based corpus research into learner writing using both computational and manual text analysis can provide a starting point for shaping students' understanding of academic argumentation including metadiscourse. Following Mauranen (1993), Ädel's (2003) model delineates metadiscourse from evaluation, refocusing attention on explicit reflexive language. This study examines the use of metadiscourse in high-scoring and low-scoring essays written by undergraduate L2 business students. The subcategories studied are connectives, frame markers, code glosses and self-mentions. The results show striking differences in the range and use of markers to structure text. The wider aim of this study is to identify accessible models and methods to enable teachers to analyse metadiscourse in academic literacy.

1. Introduction

This paper on learner use of metadiscourse came about from a larger study on the ways texts reveal writers' strategies to create argumentation in academic texts (Noble 2006). It is motivated by questions arising from this researcher's experience teaching university student writers in English as a second or other language and the experience of compiling a corpus of learner academic writing. These experiences have highlighted two areas of need in academic writing pedagogy. First, teachers need to be able to determine how argumentation in writing has been attempted by a particular cohort of students so that their writing can be further developed. Second, teachers need to be able to articulate and demonstrate the rhetorical strategies available to learner academic writers in a developmental approach appropriate to their students' individual levels of understanding.

In terms of analysis, in contrast to the current trend of studying larger and larger corpora, it will be argued that we must not lose sight of the value of small, classroom-based corpus research, here referred to as 'local' corpus research. Using a combination of simple computational and manual methods of text analysis enables teachers to identify competence levels as the starting point for building and shaping their

students' understanding of academic argumentation. One aspect of particular interest in the literature on metadiscourse in learner writing (Crismore et al. 1993; Cheng and Steffensen 1996; Ädel 2003) is the use of metadiscourse which is the focus of this paper.

The goal of this study is a practical, pedagogical outcome useful to teachers of academic English or subject lecturers. If they are to make use of the methods used here, the comprehensibility of each method is paramount. The data collection replicates university tutorial conditions where a teacher or lecturer has very limited information about the students' background, let alone writing experience. What the educator does possess, however, is essay texts themselves and all that text analysis can reveal about the writer and their notions of how to use metadiscourse in a particular academic context.

2. Metadiscourse

Variously referred to as metadiscourse, metatext, or text reflexivity, this concept is used to refer to a variety of non-propositional elements that help to organise text, in other words, "the text's commentary on itself" (Mauranen 1993:113). These elements may include, for example, linking words that express contrast, sequence or additional information (e.g. *however, secondly, also*) as well as asides made by the writer to the reader (e.g. *as we will see*) or guidance to different parts of the text (e.g. *in the next section*).

Metadiscourse is a highly relevant area to examine in learner academic writing as it helps to organise text, establish relations between writer and reader, and in its broadest definition, to convey a writer's attitudes towards text or the readers themselves (Hyland 2000). In other words, metadiscourse signals the presence of the writer in a text (Vande Kopple 1988). The term, though, is problematic due to two main strands that have arisen; one focusing on text organisation; the other on this and writer attitudes.

The development of a classification system for metadiscourse may be attributed to Vande Kopple (1985, 1988) and later adapted for academic discourse by Crismore and Farnsworth (1990). Their approach brought together Halliday's textual and interpersonal functions with Vande Kopple's seven categories of metadiscourse. Of these, Crismore and Farnsworth chose five to examine: those that are textual—code

glosses, modality markers—and those that are interpersonal—hedging/emphatics, attitude/evaluative markers, and commentaries. Other researchers base their analysis on this taxonomy to varying degrees (Crismore et al. 1993; Hyland 1998, 2000, 2004). Application of the whole model, both textual and interpersonal components, has come to be known as 'the integrative approach' (including, for example, studies by Vande Kopple 1985; Crismore & Farnsworth 1990), whereas implementation of textual or organisational functions is known as 'the non-integrative approach' (including, for example, the study by Mauranen 1993); alternative labels are the 'interactive approach' and the 'reflexive approach' (Ädel, this volume).

In order to establish the metadiscoursal framework best suited to the analytical and pedagogical requirements of this study, the issue to be addressed is whether or not the writer's attitude towards the content and readership of a text be included, or should it be excluded and classified elsewhere as stance, evaluation (Hunston & Thompson 2000) or appraisal (Martin & White 2005).

An alternative theoretical model of metadiscourse to the interpersonal metadiscourse model is proposed by Ädel (2003) based on three of Roman Jakobson's six functions of language (1980, 1998) namely: the metalinguistic, expressive, and directive. Thus, Ädel's model, as seen in Figure 1, focuses on these functions played out in discourse within a three-way reflexive relationship between the text, writer, and reader (2003:73).
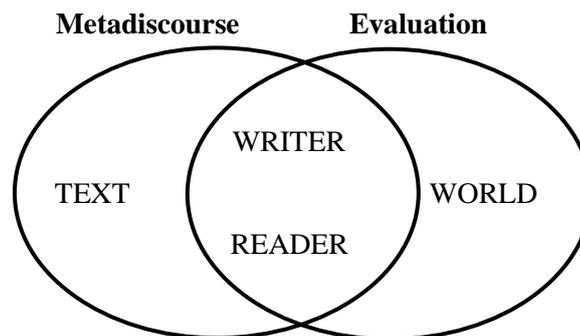

Figure 1. Overlap between metadiscourse and evaluation (Ädel, 2003:90)

While in agreement with Mauranen (1993), who argues that the most important feature of metadiscourse is that it refers to the current text or the writing process, Ädel adds that reference is needed not only to the reader of a text but the writer as well (2003:76). This clearly delineates metadiscourse from evaluation without which, Ädel asserts, too many features have been housed under this one term for historic reasons, making the definition of metadiscourse vague as it has moved away from a focus on explicit reflexive language. To be meaningful in a pedagogical context, it may be wise to adhere to the advice of Swales (2002:67) regarding structural models:

> One seemingly predisposing feature for the acceptance of structural models is certain simplicity […] It looks as though being simple engenders being memorable, and this in turn engenders being useful, quotable, and perhaps teachable.

In order to maintain the practical pedagogical purposes of this study, the definition of metadiscourse applied here consists of a simplified, restricted model, adapted from Ädel (2003) in which self-reflexive language is the defining feature of the type of metadiscourse examined. The term metadiscourse may be thought of as a writer's "commentary on the running text" (Ädel 2003:74) referring to references made by the writer about him- or herself, to the reader or about the text at hand, but not about the world 'outside' the text. As seen in Figure 1, the model delineates between metadiscourse and evaluation, as well as metadiscourse and attribution to outside sources.

Metadiscourse here will focus on what Mauranen (1993) calls 'text reflexivity' consisting of features that explicitly guide the reader through the text (e.g. *This essay will...; Firstly...*), and writer-reader interaction, meaning the writer discloses his/her intentions or includes the reader with reference to the text (e.g. *I will discuss...; as we have seen...*). Connectors are included insofar as their scope is textual (*however*, *first of all*), that is, intersentential, but sentence-internal connectors are not. The model will not include intertextual references, that is, comments made by the writer about other texts (e.g. *As McKenzie points out...; Some of their results...*), nor stance markers which express the writer's opinions or attitudes to topic other than the text itself (e.g. *This trend seems to indicate...; There is no evidence to suggest...*).

Questions used to guide this research project focus on observations to be made about learner use of metadiscourse markers. Possible

practical outcomes to improve the teaching of metadiscourse in academic institutions will be suggested in the discussion. The research questions are: Which metadiscourse markers are mainly employed by learner writers and which are not? What, if any, are the main differences in the use of metadiscourse in high and low scoring undergraduate essays?

3. Method

One purpose of this study is to find an approach not only to analyse texts for the illumination of the profession of linguistics, but also to provide possibilities for the development of an accessible text-based teaching model based on a local corpus of texts produced by a class of student writers. Just as a teacher examines his or her students' texts to assess their level of competence in a given skill, close textual analysis allows the teacher/researcher insight into not only the target linguistic items chosen by a student, but also the absence of what might have clarified the writer's meaning and thus potentially the next step the writer could take to improve his or her writing. In this study, the combined force of two types of textual analysis, computational and manual, provide different insights into the rhetorical choices made by student writers. Thus this study is textual, rhetorical and practically oriented for a pedagogical outcome.

3.1 Corpus-Based Research

A significant contribution to text analysis has been the development of computer-assisted corpus analysis. Due to the limits of human observation often influenced by preconceptions and intuition, "much deep patterning is beyond observation and memory" (Stubbs 1996:21). A concordancing program, however, identifies every example in the data whether or not they fit expectations. While the differences between learner language and native-speaker language have been much debated, now corpus-based research can substantiate claims with hard evidence (Bowker & Pearson 2002:211). The study of learner corpora can provide insights into how certain groups of students tend to express certain meanings, and to what extent certain items are overused or underused (Granger 1998). Corpus analysis tools allow the analyst to study broadly or narrowly focused interests. For example, these tools are able to find

the highest frequency terms in a corpus or find an individual token within the context of a sentence fragment or string.

The creation of a local corpus consisting of the output of students in the same class provides a rich source of learner text for a variety of purposes. Such a corpus may be used to teach, model, discuss, challenge and improve texts that are familiar and attainable by students. Moreover, the use of terms found in a learner corpus may be contrasted with the way they are employed in a corpus of expert writing on the same topic.[1] In addition, a corpus of successful learner texts may be used by students in the role of researcher in an experiential investigative approach to learning to write (Johns 1997:92). In my study a learner corpus is not only used to examine the use of metadiscourse devices in student essays, but also to create a database appropriate to the local context of the course from which to draw pedagogical materials.

When used for class work, all identifiers, such as the file specification used below (*Essay36A*), are removed from sentence fragments to retain the anonymity of students.

> effects of heavy television viewing. <p> In addition, the field experiment (Essay 36A)

The learner corpus used in this study is deliberately small, as it mimics the size of a typical first-year cohort. As Fillmore states, "every corpus I have had the chance to examine, however small, has taught me facts I couldn't imagine finding out any other way" (1992:35). Moreover, recent research recommends the development of small corpora for use in EAP programs (Hyland 2000; Tribble 2002). While large corpora are desirable to determine the frequency of particular words to generalise findings, analysing a small corpus using a concordancing program combined with manual analysis allows a deeper understanding of individual learner writer strategies.

---

[1] In the larger study this paper has come from, an additional corpus comprised of the scholarly texts from the class reading list was used for comparison with the learner corpus.

4. Materials

The essays analysed in this study were written by first-year university students who had English as a second or other language (L2). These students were majoring in business studies and computing, but were entering university through a bridging program, consisting of one or two years' study in a private college, Sydney Institute of Business and Technology (SIBT), before being admitted into a regular second year university degree program at Macquarie University. The level of English required for entry into SIBT was an overall IELTS score of 5, while direct entry Macquarie international students were required to have an overall IELTS[2] score of 6. SIBT students whose first language was not English, whether Australian-born, immigrants or international students and who had not attended an English-speaking high school were required to pass a credit-bearing course named "English for Academic Purposes 100" (EAP 100) before being admitted to Macquarie University. During the course, students received instruction in both global and discrete aspects of academic writing including: taking a critical approach to reading, summarizing, structuring an essay and using reporting verbs.[3] In this course, students wrote the essays that have been used to form the learner corpus in this study. Students came from a range of language backgrounds including Mandarin, Cantonese, and Korean.

The data for this study consist of a 120,000 word corpus comprising 80 essays approximately 1,500 words in length ranging in grading from high (28/30) to low (15/30) with a high score roughly equivalent to an IELTS score of 5.5. The scores where determined by course tutors based on these criteria: clear argument; logical structure; appropriate use of evidence; strong cohesion; and grammatical accuracy. Grading in most Australian universities uses the cline of High Distinction (HD) for excellent, Distinction (D) for very good, Credit (CR) for good and Pass (P). The corpus contains essays from two classes of approximately 40 students in each. These classes are represented by A or B in the file

---

[2] International English Language Testing System (IELTS) is an English proficiency test accepted by most Australian, British and Canadian universities. The four skills test a nine-point band with 5.5 overall needed for undergraduate university entry and 6.5 overall for postgraduate entry.
[3] Materials from EAP 100 have since been published (cf. Brick 2006).

specification seen at the end of each concordanced sentence fragment, as in *Essay 34A*.

The essays are argumentative (rather than expository) and were written using scholarly readings from psychological and sociological studies on the topic "Does television violence affects children's behaviour." The texts are authentic student texts so no corrections have been made to spelling, grammar or other errors. Therefore, examples from the learner corpus discussed in this study are intentionally left in their original state and may contain usage errors. Noting the types of errors made is pedagogically informative.

5. Analysis

The present study used three stages of analysis. First, the computational technique of concordancing was used to analyse the learner corpus to identify the range and number of metadiscoursal items. The concordancing software used was MicroConcord (Scott & Johns 1993). Second, two subcorpora, extracted from this learner corpus, consisting of high and low scoring texts, were analysed to compare difference in usage based on score. Third, manual text analysis of several essays added a deeper view of patterns of use than frequency lists by revealing how metadiscoursal choices affect the structure of a complete text.

The analysis compares and contrasts the use of metadiscourse markers in all texts and then high (HD) and low (P) scoring texts specifically. The meaning of "metadiscourse marker" as used here is that of an element that helps the writer to talk about or organise the text, to engage the attention of the reader, or to reflect on the text. The categories of metadiscourse markers investigated in this study are listed below.

    Connectives
      Logical connectors: e.g. *therefore, in addition, however*
    Frame Markers
      Sequencing: e.g. *first, second, then*
      Label stages: e.g. *finally, to conclude*
    Code Glosses: e.g. CALL, DEFINE, MEAN, *i.e.*
    Self-mention: *I, we, my, our*

Corpus analysis of the whole learner corpus[4] (80 essays) was conducted in order to ascertain which types occur and how frequently. Firstly the learner corpus was examined for patterns of interest to explore in depth. Next the two subcorpora were compared, one consisting of 10 high scoring essays and the other 10 low scoring essays (approximately 14,000 words each). These subcorpora allowed for observation and comparison of the kinds of choices learner writers make when attempting to clarify their textual intentions to the reader, and also the tracking of relationships between metadiscourse markers. Clearly the goal here was not to accumulate statistically accurate frequencies as in large corpus studies. As Tribble states, "the large corpus […] provides either too much data across too large a spectrum, or too little focused data, to be directly helpful to learners with specific learning purposes" (2002:132). It is important to note that, in general, learner writing does not present a wide range of types used with the most frequent type being connectors. Finally two essays were examined to observe how metadiscoursal connectors, as the most prevalent type of tokens, operate in context. Particular attention was given to the number of tokens employed, how they were distributed, and where they were placed in the sentence.

6. Results

Internal text structuring in the learner corpus using metadiscoursal markers is examined focusing on connectives, framing, code glosses and self-mentions. Overall the learner writers show heavy reliance on a narrow range of connectors, but differences are seen in the types employed. Also a comparison of high and low scoring essays highlights differences in the, albeit infrequent, use of framing, code gloss and self mentions.

6.1 Connectives: Whole Corpus

Connective use is particularly prevalent in learner writing (Field 1994) as these linking words are relatively easy to use, and require little or no

---

[4] Here "learner corpus" refers to the students being both learners of English and learners of academic writing in a university context.

grammatical adjustment when added to a sentence. The most frequent linking words found in the learner corpus can be seen in Table 1 below.

Table 1. Frequency of Connectives in the learner corpus

Rank	Connectives	Total
1.	*Also*	435
2.	*But*	316
3.	*However*	299
4.	*Therefore*	212
5.	*on the other hand*	69
6.	*Moreover*	62
7.	*in addition*	53
8.	*in conclusion*	44
9.	*In fact*	30
10.	*as a result*	20

The marker *also* was by far the most frequently used, perhaps because it is commonly used in spoken text as well as written text and thus would be very familiar to students. In fact, some essays used this connector almost exclusively.[5] A student's reliance on one connective is not *per se* an indicator of writing quality as indicated by the essay grades listed in Table 2. Yet, this result may indicate students' lack of knowledge or confidence to apply other types.

Table 2. Predominance of *also* in Four Essays

Essay	*also*	other connectives	grade
9A	21	0	CR
46A	14	0	D
41A	13	0	F
45B	12	1	CR

A more informative result is found when the total number of uses is compared to the range of terms or the token/type ratio employed. Figure 2 shows a comparison of High Distinction, Distinction, Credit and Pass

---

[5] From the sample set of connectives that were tested.

level essays. It is important to note that most HD essays had twelve or more occurrences of connectives, and utilized five or more different types of markers. In contrast, in ten Pass essays, only one used more than twelve markers. Thus higher scoring essays contained more connectives and a greater variety of different types.
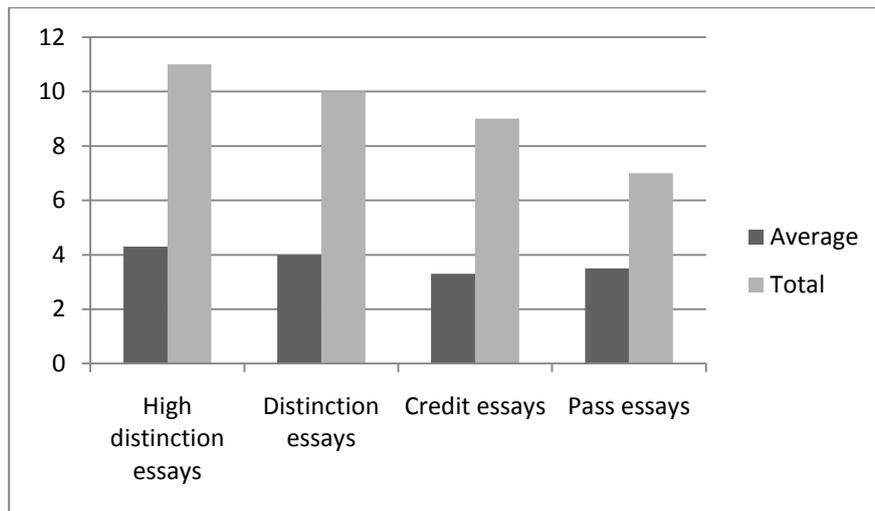

Figure 2. Average number of metadiscourse markers used and total frequency

In terms of placement in the sentence, some connectors appeared in the initial position very often (*in addition*), and others very rarely (*and*) (see Table 3). This result may have been influenced by exercises the students completed during EAP 100 on the placement of connectors (cf. Thurstun & Candlin 1997). Moreover, research has shown that Cantonese speaking students from Hong Kong tend to have a distinctive way of organising essays characterised by very frequently prefacing points with connectors possibly entrenched through Hong Kong school textbooks (Field 1994; Field & Yip 1992). A large cohort within the EAP 100 course consisted of students from Hong Kong, so a residual educational effect could have also contributed to this result.

Table 3. Markers in initial sentence position

Marker	Initial	%	Total
in addition	51	96.2	53
therefore	42	76.4	55
however	13	4.3	299
on the other hand	12	1.7	69
and	11	0.7	1,590

In examining the placement of connectives, a useful method for identifying the common problem of overusing connectives in initial position by learner writers was found. By searching the learner for a full stop followed the connective in question, for instance *'. In addition'*, overuse is immediately evident. For instance, in concordance Extract 1 the concordanced sentence strings readily show which essays contain *in addition* in initial position in the sentence. Even more problematic are those instances where connectives with an additive function (e.g. *in addition, further, also, moreover*) or a contrastive function (e.g. *however, yet, instead, in contrast*) are inappropriately used because they refer back across paragraph boundaries (<p>), as seen in lines 2, 4, 8 and 9 below.

Concordance Extract 1 from the learner corpus

```
1)  still observable 2 years later. In addition, researchers were led
2) f heavy television viewing. <p>  In addition, the field experiment
3) iment is criticised by Freedman. In addition, timing difference of
4) en have VCRs in their house. <p>  In addition, there are all uncut,
5) ate so to children's aggression. In addition, we should regard TV
6)  increased aggressive behavior.  In addition, the cognitive develo
7) ad to a decrease in aggression.  In addition, many of the published
8) n Singer & Singer, 1988: 5). <p> In addition, children's perception
9) ve sign pattern disappeared. <p> In addition, it should add one stu
10)tical significance is distorted. In addition, the reasons which cou
11)effects on his aggression level. In addition, Freedman's studies on
12)children's aggressive behaviour. In addition, Freedman (1988: 3) ha
13)y were angry in the first place. In addition, a recent study conduc
14) and violent television program. In addition, Milavsky also conduct
15)in a whole full of TV influence. In addition, parents and teachers
16) really exist or was very small. In addition, other similar studies
```

6.2 Connectives in Subcorpora

As is characteristic of student writing, a narrow range of types of markers is represented here. Table 4 shows a comparison between connectives found in the subcorpora of low and high scoring essays. Low

scoring texts contained a smaller range of connectives meaning that a few types were used repeatedly. Moreover connectives often associated with spoken English were prevalent in the low scoring essays, such as *but* and *so*. High scoring texts, on the other hand, exhibited more types of connectives often associated with the more formal written register, e.g. *in addition* and *thus*. Although an overlap of constituents is evident in the main, the presence of connector types from the academic register in higher scoring essays indicates a developing emergence of register awareness.

Table 4. Frequency of Connectives in the subcorpora[6]

Low Score	F	High Score	F
but	42	*also*	52
because	40	*however*	38
however	37	*but*	25
also	37	*because*	25
therefore	28	*therefore*	20
although	19	*still*	12
still	15	*although*	11
so	13	*in addition*	8
since	12	*thus*	8

6.3 Comparison of Two Essays

When few connectives are used in an essay, it tends to be either an indication that there is a problem, or that another method of linking ideas has been employed. Example 1 consists of text segments from two essays for comparison. Both essays 3A and 33B exhibit fewer occurrences of connector use overall compared with other essays, yet they are structured quite differently. In essay 3A, the writer's focus is on Freedman (the researcher) and his actions. This connector use is rather descriptive and narrative-like in the way that ideas are treated like events that unfold

---

[6] Table 4 compares two subcopora consisting of high and low scoring essays. Thus the frequencies differ from those found in Table 2 representing the entire learner corpus.

sequentially. In contrast, the writer of essay 33B focuses on the problematic results of particular experiments, thus demonstrating a more analytical understanding. In essay 33B, as seen in Example 2 below, links are made between sentences by topical or lexical connections (e.g. weak, contradictory results) rather than by connectives.

Example 1: Few Connectives in Essays 3A and 33B

Essay 3A	Essay 33B
<p>In ***Freedman*** article, ***he*** argues that there is a little or no solid evidence to support the assertion that TV violence lends to increased aggression, that is opposite to Singer & Singer's idea. ***Freedman*** look at the three experiment as well, ***he*** says that these studies vary in terms of role of anger in determine aggressive behavior, the argues that if anger is a precondition aggressive behavior, **Then** the effects of TV violence are in doubt **and** *he* disagree that what happen in the laboratory is what happen in the real world. **On the other hand** ***Freedman*** showed that the field experiment are too small in number **and** some must be disregarded, **because** they are case studies rather than real experiment or the results were weak **and** in consistent. ***Freedman*** **also** compared the rate of crime in some cities that had television to tee rates of crime in those that did not. It found that TV had no effect on serious crimes or other serious crimes, **but** only on petty theft. (J.L. Freedman 1988)	p>***The*** ***results*** of control experiments conducted in the field are too weak **and** inconsistent to show that there is a positive correlation between TV violence and aggression. This is not surprising **since** ***this kind of experiment*** is difficult to control **and** requires a lot of time, effort and money. ***An experiment*** by Fesbach and Singer (1971) shows that the children who watch non-violent programs tend to be more aggressive than those who watch violent program. ***This finding*** contradicts the hypothesis that there is a correlation between TV violence and aggression. ***Research*** conducted by the same method by Wells (1973) showed a weak opposite result to the one by Fesbach and Singer (1971). ***These contradictory results*** **also** raise questions about the consistency of the findings and the methodology of this research. ***An experiment*** by Friedrich and Stein (1973) convincingly shows that there is no correlation between TV violence and aggression (Freedman,1988;Singer&Singer,1988).

Example 2: Linking in Essay 33B

> S1 results . . . weak and inconsistent
> S3 findings contradict the hypothesis
> S4 Research showed . . . weak opposite results
> S5 These contradictory results . . . raise questions about consistency
> S6 convincingly show . . . no correlation

The use of lexical chains to establish cohesion is outside the scope of this paper. Suffice it to say here that the presence of more connectors in a text does not necessarily create a more cohesive text.

Another aspect of connectives to consider is the rhetorical weight attached to any particular marker. In Example 3 below, *therefore* seems to function at a significant juncture of closure to a point of argument. The first sentence (S4) is the writer's main point (that television does not affect children) and the last sentence (S27) confirms this idea with evidence that watching television may indeed result in certain reactions (imitation of aggression, arousal) that could be construed as an effect, in spite of counter evidence.

Example 3: Rhetorical use of *therefore* (extract from Essay 4A)

> S4 It may be said that television do have some effects on children but the evidence is not very strong.
> S16 Freedman further states that aggression may be due to initial anger.
> S17 However, the research has not done that part. The inconsistent results are concluded because such research has not concern the prior anger.
> S19 Secondly, some studies consider anger but some do not.
> S20 Thirdly, if anger is the cause of aggression then there is no direct relationship between television violence and aggressive behavior because it may due to arousal.
> S23 Freedman focus on a few studies with slight or no effect but he has not considered that there are many other studies produce positive effects.
> S25 In fact, social scientists have examined parental influence before making conclusion.
> S26 To answer to Freedman's arousal factor aggression is defined as stimulated behavior.
> S27 <u>Therefore</u>, imitation of aggression, arousal and reduced inhibition of aggression are all possible results of watching television violence.

In Example 4, however, *therefore* plays quite a different role. The marker is used twice in this paragraph, but in neither case is there a sense of summative argumentation. In the first instance *but* or *however* would

have served better to construct a connection between the contrasted ideas. In the second instance, no linking word is required, because *since* is already in operation to show how these ideas are related.

Example 4: Incorrect use of *therefore* (extract from Essay 42A)

> <p>It seems to be true that laboratory experiment is relevant just in a certain circumstances since everything is undercontrolled. To explain this statement, we need to know what is laboratory experiment first. It is to ask children to watch a violent program in a laboratory room and then to compare their behaviour before and after watching that program. "This study show that violence can produce a momentary increase in scores on measures of aggression in the laboratory" (Freedman, 1988). <u>Therefore</u>, some psychologist suggested that laboratory measures of aggression have "field validity." On the other hand, this measurement can't reflect the long-term effect. Since the researchers can't observe the children after they getting out from the laboratory, <u>therefore</u>, some psychologists critize that the laboratory experiments are less consistent than they are usually thought to be. In an attempt to be more accurate to measure the effect to children, another experiment has applied, which is generalising measure the effect from laboratory research.

This section has examined the frequency, placement, rhetorical weight and incorrect use of linking words in the learner corpus and subcorpora. As connectives are the largest group of metadiscoursal markers, extended attention has been given to this category. The next three sections will briefly discuss other markers that appear in the learner corpus. These are frame markers, code glosses and self-mentions.

6.4 Frame markers

Very few frame markers were found in the learner corpus, although a few were used to good effect: *finally, first*, second*, last*, then, conclud*, conclusion*. The low frequency of frame markers in the corpus seemed to indicate that students were either avoiding the typical framing sequencers (*First…; Second…; Third…*) that they had been warned against overusing, did not understand how to use frame markers effectively or were employing other strategies to organise their texts.

In terms of expert writing, fewer frame markers may indicate a level of sophistication in text organization. As Hyland (2000:190) notes, "where texts are for specialist audiences, we find fewer textual devices" because the writer will rely on the reader's understanding of lexical relations. To demonstrate text organisation that does not rely on metadiscourse markers, I have compared two texts on the same topic but

from different genres as seen in Teaching Example 1. The text on the left is an extract from a journal of social psychology, and on the right is an extract from an undergraduate level textbook on child psychology. While the second relies on sequencing labels (*in a number of ways; first; second; third*), as is typical of a pedagogical genre, the first depends on more subtle conceptual labels, that take the form of abstract nouns (e.g. *several kinds of influence; passive form; active form*). Here we see differences in organisation dependent on the target audience.

Teaching Example 1: Same topic but different genre

Social psychology journal	Early childhood textbook
Television might <u>have several kinds of influence</u> on the child. For example, a more <u>passive form</u> of learning may occur simply by exposure, involving the inherent human tendency to categorize even fleeting or incidental experience. A more <u>active form</u> of learning may also involve the child's attempting to form schemas or categories and to organize ultimate action scripts based on the kinds of experiences encountered with the television medium. (…) <u>With respect to content</u>, heavy exposure to the medium also means heavy exposure to an inordinate amount of aggressive behavior carried out by both "good guys" and "bad guys". As children actively seek to organize scripts about human actions, either through direct imitation or through the formation of ongoing cognitive structures, they must inevitably – if…heavy TV viewers – begin to see violence as the major form of problem solution (Singer & Singer 1988:169)	Television may influence children's aggressiveness <u>in a number of ways</u> (Liebert and Sprafkin, 1988). <u>First</u>, some children may directly imitate; they simply copy what they see on television. Obviously other factors are involved, as most children do not imitate such behavior. Aggressive children, though, may learn different ways to aggress by watching television. <u>Second</u>, televised violence disinhibits aggression. People have certain inhibitions against violence, and witnessing aggression may reduce these inhibitors. <u>Third</u>, television violence may lead to antisocial attitudes and encourage children to accept violence as a way of dealing with problems. Children become desensitized to violence on television and come to accept it as a normal part of life, not taking it seriously (Cole & Didge, 1998). (Kaplan 2000:455)

Lexical relations are highly valued in scholarly writing. The target audience for student essays is the academic who marks them, thus the

target register is more closely aligned to scholarly rather than pedagogical genres. To begin with, learner writers of undergraduate essays would be wise to create an explicit organisational framework, but as they become more advanced writers they should be encouraged to rely more and more on implicit lexical cohesion in order to more closely approximate scholarly texts.

6.5 Code glosses

Code glosses were rarely used by learners but appeared slightly more in High Distinction than in Pass essays as indicated in Table 6.

Table 6. Frequency of Code Glosses in the subcorpora

Marker	HD Essays	P Essays
such as	24	18
in fact	7	0
for example	5	7
DEFINE	1	3

The most frequently used code gloss in the subcorpora was *such as*, followed by *for example* and *in fact*. These markers are useful for defining new terms, giving examples, and reworking a complex idea into a simpler form. It is therefore interesting that students rarely used them, and further research may reveal a lack of topical development as the cause. Where ideas in sentences are touched on but not developed, there would be little need to give examples or elaborate on complexity.

6.6 Self-mentions

The last category of explicit metadiscourse markers to be discussed is self-mentions. Within this category, a plethora of discourse functions are possible. In this section the markers *I, my* and *we* will be discussed in terms of examples from concordances and individual high and low scoring essays. When the first person pronoun *I* is used it clearly refers to the writer of the text, as does the possessive pronoun *my*.[7] The other

---

[7] Unless these terms are used in quoted material in which case they are not metadiscoursal.

constituent in this category, *we*, is, however, more difficult to identify. Only those cases where *we* is used reflexively as well as including the reader, can be considered to be metadiscoursal. Most instances in the corpus refer to ideas and events external to the text, as seen in Example 5 below.

> Example 5: Also as adults, <u>we</u> have responsibility for our children (Essay 9A)

With each of these terms, care needs to be taken to ensure that the persona referred to by *I* and *we* are indeed the writer or the writer and reader. Thus an example can only qualify as metadiscourse if the actor (the writer for example) is carrying out a task "within the world of discourse" according to Ädel (2003:91), and as seen in Example 6.

> Example 6: Here <u>I</u> would like to point out some shortcomings about the study (Essay 43B)

Learner writers, particularly non-native English speakers, often find it difficult to know when it is appropriate to use *I* in academic writing. In his study of learner writing, Myers (2001) found that students were, for the most part, confused about how to present personal views for two reasons: (a) expressions of opinion can have different functions in text and (b) expressions of opinion have different constraints in different genres (Myers 2001:77). The students who wrote the essays collected in the learner corpus for the current study were warned by their teachers against using *I*-statements, which may explain why there are so few examples in the corpus.

The use of the self-mention *my* was also avoided by the student writers. Concordancing showed that several tokens belong to the same text. It seems that some students felt at ease using their personal voice (Leki 2001) even though most did not.

The self-mention marker *we* is widely represented in the learner corpus. In order to determine which occurrences are metadiscoursal, two principles were applied. Firstly, the metadiscoursal meaning of *we* should be inclusive, as in 'you and me' because it refers to the writer and reader. Employment of this term as a powerful argumentative device often used by expert writers may be lost on learner writers. Thus a test was applied to determine whether each use of *we* could be replaced by

the term 'people'. In other words, does *we* in this case mean 'it could be anyone', 'society in general' or does it mean 'you and me'?

Concordance Extract 2: *we* from the learner corpus

```
1)    Milavsky concludes that we should concentrate more on real world
2)   y. That's the reason why we should look and study closely about
3)    e connection. Therefore we should take seriously about this issu
4)  t is questionable whether we should believe his thesis or not sinc
5)   aggression. In addition, we should regard TV programs which have
6)  nts send the message that we should buy a product and documentaries
7)  e or not to children. <p> We should be concentrated on children's
8)   aggression. In addition, we should regard TV programs which have
9)  ods without it. Actually, we should objectively accept that a rise
10) r our children, therefore we should set rules about TV and  show
11) seems to be obvious that  we should believe what have been found
12) an is very logical, since we sometimes met someone expectation jus
13) ence to show the linking, we still need more studies in different
14) is inconclusive. Finally, we suggest that parents teach their chil
15) nderlined critically when we think children in the world who spend
16) ult to do so.  Therefore, we will discuss every experiments done
17) elevision. In this essay  we will discuss, what kind of effects
18) y used for research. Then we will compare the major laboratory exp
19) hey support the evidence, we will briefly compare what the authors
20) out the situation in USA, we would easily to realize the people in
```

For example, in Concordance Extract 2, sentence fragment 1 (Milarsky concludes that we should concentrate more on real word issues (30B)) could be reworked as: Milarsky concludes that people should concentrate more on real world issues (30B adapted). However, sentence fragment 18 (Then we will compare the major laboratory experiment (13A)) could not be changed to: Then people will compare the major laboratory experiments (13A adapted).

A related term to *we* is *people*; an overly general term for the register of academic writing. For example in Essay 34B the marker *people* is used 13 times where, in most cases, the writer could have identified the types of people mentioned by their specific roles. For example, in *Williams and her group were the <u>people</u> who did this experiment*, *people* could be replaced with *researchers*. This example is not metadiscoursal because it is a reference to people outside the world of the text, but perhaps the last line, *However, by the research results from Singer and Singer, <u>people</u> still can conclude that violence television program does affect most of the children behavior becomes aggressive*, would benefit by changing from *people* to *we* as by this point in the essay (the

conclusion) the writer would hope to have persuaded the reader to his or her stated point of view.

As seen in Ädel's (2003) study of personal metadiscourse, there is much potential scope for writer-reader interaction. Possible discourse functions she identifies are: anticipating the reader's reaction, clarifying, persuading the reader, imagining scenarios, hypothesising about the reader, and appealing to the reader (2003:147; see also Ädel, this volume, and Pérez-Llantada, this volume). Thus self-mentions are identified as important to writer-reader interaction, yet in the learner corpus were poorly applied or avoided possibly because of confusion about the constraints of using expressions of opinion in academic writing. These results highlight the need for more explicit attention to be paid to the teaching of self-mentions to learner writers.

7. Conclusion

This study of metadiscoursal markers has examined internal text structuring and reflexivity in a learner corpus focusing on connectives, frame markers, code glosses and self-mention use by learner student writers. As predicted, the corpus analysis recorded a high frequency of connectives used by the learners within the main learner corpus as compared to other markers.

The finer text analysis between high and low scoring essays was possible through the use of subcorpora extracted from the main corpus. The results showed clear differences between high and low scoring essays. High scoring essays used a higher frequency and range of metadiscoursal markers. Low scoring essays tended to rely more heavily on markers common in spoken English (rather than written English texts) making these essays more casual in register. In addition, close text analysis of two essays revealed that the way markers are used to structure texts can differ markedly resulting in narrative rather than analytical structuring.

Students' lack of knowledge, confidence or instruction in applying certain devices was indicated by the lack of attempts to use frame markers, code glosses and self-mentions in the learner corpus. It was noted that sequencing labels commonly used in student writing (*firstly; secondly*, etc.) do not necessarily mark good writing and in fact high scoring essays showed more use of lexical cohesion. Students also

seemed to avoid self-mentions perhaps heeding warnings about the use of first person, or lacking confidence in the use of their own voices in academic writing.

In terms of methods applied in this study, analysis of the whole corpus, subcorpora of high and low scoring texts and close manual text analysis of two essays proved to be complementary, with each method adding to the overall picture of student use of metadiscoursal use.

While the debate over how to define metadiscourse wages on, from an analyst's perspective, the sheer volume of data generated by this topic has warranted focusing on a narrower definition than Vande Kopple's original taxonomy. But more importantly, this distinction provides useful in teaching, as students need to learn to manage text organisation, reflexivity and the evaluation of ideas.

7.1 Implications for teaching

The intention of this study has been to demonstrate the value of localised classroom-based corpus research using computational and manual methods of text analysis. Computer-assisted concordancing provides more direct evidence of how language is used than grammar books or dictionaries (Murison-Bowie 1993). If teachers are able to identify competence levels of metadiscoursal use by students through the analysis of their texts, a local corpus is an excellent starting point from which to build up and apply targeted development their students' academic argumentation. A learner corpus is not only useful to examine the use of rhetorical devices in student essays, but also to create a database appropriate to the local context of the course from which to draw pedagogical materials for current and future class work. Interrogation of a local learner corpus is fertile ground for class discussion, given the anonymity of computer assisted concordanced sentence fragments when file specifications have been removed.

Other strategies complementary to the study of a local corpus are found in the literature. For some time, teachers have been encouraged to find relevant examples of markers used in real texts. Even more powerful, though, is the use of a concordancing program with a corpus of authentic texts (either a learner or expert corpus). Concordancing authentic texts provide teachers and students with the tools to analyse texts, manipulate texts, and understand different audiences and registers (Hyland 2005). Another strategy is cognitive mapping to help students

improve the quality of their ideas through adequate elaboration (Crismore 1989) which would help to address the paucity of code gloss markers used by student writers. Also, as argued in this paper, the delineation of functions into metadiscourse, evaluation and intertextuality through the use of clear modeling (Mauranen 1993; Ädel 2003) is likely to enhance pedagogical practices.

Finally, it is not this researcher's intention that the results found in this small, localised study should be generalised, as this would miss the point. Rather other educators are encouraged to do as this study has shown is possible; to examine learner writing for evidence of metadiscoursal use and target teaching from that point of departure. Also it is possible to engage students themselves in uncovering patterns and features in writing using a variety of analytical tools such as corpus analysis to enhance comparison, and to encourage discussion of authentic texts, thereby addressing authentic student needs and interests.

References

Ädel, Annelie. (2003). *The Use of Metadiscourse in Argumentative Writing by Advanced Learners and Native Speakers of English*. Ph.D. dissertation. Göteborg, Sweden: Göteborg University.

Bowker, Lynne & Jennifer Pearson. (2002). *Working with Specialized Language: A Practical Guide to Using Corpora*. London: Routledge.

Brick, Jean. (2006). Academic Culture: A Student's Guide to Study at University. Sydney: NCELTR.

Cheng, Xiaoguang & Margaret Steffensen. (1996). "Metadiscourse: A Technique for Improving Student Writing." *Research in the Teaching of English* 30(2): 149-181.

Crismore, Avon. (1989). *Talking with Readers: Metadiscourse as Rhetorical Act*. New York: Peter Lang.

Crismore, Avon & Rodney Farnsworth. (1990). "Metadiscourse in Popular and Professional Science Discourse." *The Writing Scholar: Studies in Academic Discourse* [Written Communication Annual Vol. 3]. Ed. W. Nash. Newbury Park: Sage. 118-136.

Crismore, Avon, Raija Markkanen & Margaret Steffensen. (1993). "Metadiscourse in Persuasive Writing: A Study of Texts Written by American and Finnish University Students." *Written Communication* 10: 39-71.

Field, Y. (1994). "Cohesive Conjunctions in the English Writing of Cantonese Speaking Students from Hong Kong." *Australian Review of Applied Linguistics* 17(1): 125-139.

Field, Y. & L.M.O. Yip (1992). "A Comparison of Internal Conjunctive Cohesion in English Essay Writing of Cantonese Speakers and Native Speakers of English." *RELC Journal* 23(1): 15-28.

Fillmore, Charles. (1992). "'Corpus Linguistics' or 'Computer-Aided Armchair Linguistics'." *Directions in Corpus Linguistics. Proceedings of Nobel Symposium 82, Stockholm, 4- 8 August 1991.* Ed. J. Svartvik. The Hague: Mouton de Gruyter. 35-66.

Granger, Sylviane (Ed.). (1998) *Learner English on Computer*. London/NewYork: Addison-Wesley-Longman.

Hunston, Susan & Geoff Thompson (Eds). (2000). *Evaluation in Text: Authorial Stance and the Construction of Discourse*. Oxford: Oxford University Press.

Hyland, Ken. (2005). *Metadiscourse*. London: Continuum.

Hyland, Ken. (2004). "Disciplinary Interactions: Metadiscourse in L2 Postgraduate Writing." *Journal of Second Language Writing* 13: 133–151.

Hyland, Ken. (2000) *Disciplinary Discourses: Social Interactions in Academic Writing*. Harlow, UK: Pearson Education.

Hyland, Ken. (1998). "Persuasion and Context: The Pragmatics of Academic Metadiscourse." *Journal of Pragmatics* 30: 437-455.

Jakobson, Roman. (1998). *On Language: Roman Jakobson*. Eds L. R. Waugh, & M. Monville-Burston. Cambridge, Massachusetts: Harvard University Press.

Jakobson, Roman. (1980). *The Framework of Language*. Michigan: University of Michigan.

Johns. Ann M. (1997). *Text, Role, and Context: Developing Academic Literacies*. Cambridge: Cambridge University Press.

Leki, Ilona. (2001). "Hearing Voices: L2 Students Experiences in L2 Courses." *On Second Language Writing*. Eds T. Silva & P. Matsuda. Mahwah, NJ: Laurence Erlbaum. 17-28.

Martin, J.R. & White, P.R.R. (2005). *The Language of Evaluation: Appraisal in English*. London: Palgrave Macmillan.

Mauranen, Anna. (1993). *Cultural Differences in Academic Rhetoric: A Text Linguistic Study*. Frankfurt am Main: Peter Lang.

Murison-Bowie, Simon. (1993). *MicroConcord Manual: An Introduction to the Practices and Principles of Concordancing in Language Teaching*. Oxford: Oxford University Press.

Myers, Greg. (2001). "'In My Opinion': The Place of Personal Views in Undergraduate Essays." *Academic writing in Context.* Ed. M. Hewings. Birmingham: Birmingham University Press. 63-78.

Noble, Wendy. (2006). *An Analysis of Argument in Learner Academic Writing in a Second Language*. Ph.D. dissertation. Sydney: Macquarie University.

Thurston, Jennifer & Candlin, Christopher N. (1997). *Exploring Academic English: A Workbook for Student Essay Writing*. Sydney: NCELTR Macquarie University.

Tribble, Chris. (2002). "Corpora and Corpus Analysis: New Windows on Academic Writing." *Academic Discourse.* Ed. J. Flowerdew. Harlow: Longman. 131-149.

Scott, Mike & Tim Johns. (1993). *MicroConcord*. Oxford: Oxford University Press.

Stubbs, Michael. (1996). *Text and Corpus Analysis*. Oxford: Blackwell.

Swales, John M. (2002). "Integrated and Fragmented Worlds: EAP Materials and Corpus Linguistics." *Academic Discourse.* Ed. J. Flowerdew. London: Longman. 153-167.

Vande Kopple, William J. (1988). "Metadiscourse and the Recall of Modality Markers." *Visible Language* 22: 233-272.

Vande Kopple, William J. (1985). "Some Exploratory Discourse on Metadiscourse." *College Composition and Communication* 36: 82-93.

Primary sources

Kaplan, Paul S. (2000). *The Child's Odyssey: Child and Adolescent Development*. 3rd edition. Belmont CA: Wadsworth-Thompson Learning.

Singer, Jerome L. & Dorothy G. Singer. (1988). "Some hazards of growing up in a television environment: children's aggression and restlessness". *Television as a Social Issue*. *Social Psychology Annual 8*. Ed. S. Oskamp. Newbury Park, CA, Sage Publications. 169-180.