

Should learning developers provide instruction in the use of metadiscourse?

Abstract

Metadiscourse is the language writers use to guide their readers through their texts and organise their arguments. This can take the form of phrases, for example, *this essay will discuss*, or *in conclusion*; or individual words such as *firstly* or *therefore*. This study aims to determine how undergraduate students develop their use of metadiscourse over their first two years of study at a UK university and to investigate whether use of metadiscourse is related to the grade that a text receives from subject tutors.

To achieve this, a corpus of summative written assignments was collected from 67 undergraduates studying a health discipline. This is the writing that we as learning developers are most closely involved with: assignments written as part of a course of study. The assignments were analysed using software developed for the field of corpus linguistics to identify how students used metadiscourse. The results of this study suggest that including explicit instruction in Learning Development sessions in the use of some aspects of metadiscourse could be of value. This supports an 'academic literacies' (Lea and Street, 1998) approach in that it recognises the need to make clear the implied assumptions that surround academic writing and the inherent variation between disciplines.

Keywords:

metadiscourse; student writing; corpus analysis; academic discourse

Introduction

One of the roles of Learning Development is to aid students in acquiring the writing conventions of academia and one of the ways that writers realise the functions of academic writing is in the use of metadiscourse. Metadiscourse has been defined as “discourse about discourse or communication about communication” (Vande-Koppel, 1985, p83) and the reason for its importance is that it makes a text “reader friendly” (Cheng and Steffensen, 1996, p154), thus facilitating comprehension. Some key functions of metadiscourse are to guide readers through a text and to organise and develop the writer’s arguments (Hyland, 2005); these include the use of phrases such as *this essay will discuss*, or *in conclusion*; or individual words such as *firstly* or *therefore*.

There is evidence that undergraduate essays containing a higher frequency of such metadiscoursal features achieve higher marks due to the increased readability of the essays and greater consideration of the needs of the reader. This is largely based on two studies: the first carried out with non-native speaking students in English for Academic Purposes (EAP) classes (Intaraprawat and Steffensen, 1995) and the second involving native-speaking undergraduates attending a ‘composition class’, a class which teaches generic academic writing and largely features non-subject specific argumentative essays (Cheng and Steffensen, 1996). Neither of these situations are representative of the writing that we as learning developers are most closely involved with, i.e. formative and summative assignments.

In 2019, this study of metadiscourse use in student writing was undertaken, looking at how student use of metadiscourse varied between different genres of assignment, how students developed their use of metadiscourse between their first and second years of university education, and to determine whether there is a relationship between metadiscourse use and writing success as measured by the grade an assignment receives.

This study uses a corpus analysis of aspects of metadiscourse in the writing of undergraduates studying a health discipline at a British university. This discipline is considered a ‘hard’ as opposed to a ‘soft’ discipline (Becher and Trowler, 2001) due to its epistemology (Holmes and Nesi, 2009).

The motivation for the study was pedagogic; if an increased awareness of the needs of the reader, as evidenced by metadiscourse use, is shown to lead to greater student success, there would be clear value in learning developers providing support to students in this area. Here I report on those findings of the study which I believe will be of interest to learning developers.

Metadiscourse

Metadiscourse has been described as a “fuzzy” category of language (Ädel 2006, p.4; Hyland, 2017, p.17). There are two elements to the fuzziness. Firstly, and non-controversially, metadiscourse can be realised by a wide range of lexical features from single words to sentences (Hyland, 2017, p.18). The second element concerns which rhetorical functions can be considered metadiscoursal and is subject to much debate. Several researchers have proposed taxonomies (e.g. Vande Koppel, 1985; Crismore and Farnsworth, 1990; Mauranen, 1993; Hyland, 2005; Ädel, 2006) but one that has been widely adopted is that of Hyland (2005). As this enables a degree of comparison with previous studies, this was the taxonomy adopted for this study: specifically Hyland’s category of interactive metadiscourse as that includes the rhetorical functions concerned with how writers use metadiscourse to help the reader navigate the text, which is my primary research interest.

Table 1 - Hyland’s taxonomy of interactive metadiscourse (Hyland 2004:139)

Category	Function	Examples
<i>Interactive resources</i>	<i>Help to guide reader through the text</i>	
Transitions	Express semantic relation between main clauses	<i>In addition; but; thus; and</i>
Frame markers	Refer to discourse acts, sequences, or text stages	<i>Finally; to conclude; my purpose is to</i>
Endophoric markers	Refer to information in other parts of the text	<i>Noted above; see Fig.; in Section 2</i>
Evidentials	Refer to sources of information from other texts	<i>According to X; (Y, 1990); Z states</i>
Code glosses	Help readers grasp meanings of content	<i>Namely; e.g.; such as; in other words</i>

In this study, Hyland's (2004) category of code glosses has been sub-categorised into exemplifiers and reformulations (Hyland, 2007). This study introduces a third sub-category, abbreviations, as this type of reformulation was salient in the Health corpus but was not explicitly mentioned in Hyland's taxonomy. The category of endophoric markers has been sub-categorised into linear and non-linear, depending on whether they refer to parts of the text itself (linear) or to items outside the text such as figures or appendices (non-linear)(Cao and Hu, 2014).

One category was excluded from the study: evidentials. How to reference the work of others is routinely part of academic writing skills sessions, and as a result the use of evidentials is already well explored.

Methodology

Data Collection and Preparation

A corpus (collection of texts) of summative assignments was collected from 67 students: 31 students in their second year of study and 36 in their third year, which was 74 per cent of the total cohorts. The students were all female and were all native speakers of English. As they were all native speakers, there was no need to take into account any variation due to level of English proficiency. The students each contributed three assignments, which were completed between 2016 and 2019: a first year (level 4) essay, a second year (level 5) essay and a second year (level 5) case study. (N.B the results of the analysis comparing the essay genre with the case study genre are not reported here.) Word counts for the assignments varied between approximately 1500 and 2500 words and all of the assignments had received a pass, with grades awarded varying from A to D. It was decided to exclude assignments which did not reach the standard of a pass as any amount of metadiscourse will not compensate for a lack of subject-related content.

Ethics approval for the study was granted by the university and all participants gave informed consent. The assignments collected were anonymised and stored securely in accordance with data privacy regulations.

The use of a relatively large sample of student academic writing gives results that are more generalisable than those of a small-scale qualitative study. The Level 4 essay sub-corpus consisted of 173,543 words and the Level 5 essay sub-corpus was smaller at 140,527 words. To investigate whether there was a relationship between metadiscourse use and the grade an assignment received, two additional sub-

corpora were created, one containing Level 4 and 5 assignments which were graded above 65% (76,824 words) and a second containing those graded 51% and below (70,798 words). The proportion of Level 4 and 5 assignments were comparable in each. Using these criteria identified the highest graded sixth and the lowest graded sixth of the assignments and provided a suitable balance between obtaining sub-corpora of sufficient size to allow reliable analysis and maintaining a large enough differential between grades awarded to reveal any salient differences in metadiscourse use. The sizes of these sub-corpora compare favourably with published corpus-based metadiscourse studies, which generally range from 50,000 to 160,000 words (e.g. Gardezi and Nesi, 2009; Shaw, 2009; Noble, 2010; Zhang, 2016). All sub-corpora should therefore be of sufficient size to yield reliable results.

It should be noted that the larger size of the Higher-graded sub-corpus compared to the Lower-graded sub-corpus is due to the difference in required word count for the different assignments and is not related to the grade awarded. In fact, the average assignment word length exceeded the specified word count by seven per cent in the Higher-graded sub-corpus and six per cent in the Lower-graded sub-corpus, a difference unlikely to be statistically significant.

Data Analysis

Although metadiscourse is a functional category which can be realised structurally by both words and phrases, it is common in corpus linguistic research to use words or very brief phrases as search terms to identify instances of metadiscourse in a corpus (e.g. Hyland and Tse, 2004; Ädel, 2006). This is considered reasonable as a large proportion of the linguistic features used to realise metadiscourse are adverbials, of which approximately 70 per cent are single words (Biber *et al.*, 1999, p.769).

First, the target lexical items for the study needed to be identified. As is common practice (e.g. Hyland and Tse, 2004; Ädel, 2006; Aull and Lancaster, 2014), target lexical items within each category were initially identified by manual inspection of a small sample of assignments. This initial list was supplemented by several lexical items commonly used to perform the chosen metadiscoursal functions, such as *secondly*, to mitigate the risk that the sample was unrepresentative. The final list of target lexical items is shown in the appendix. The list of target lexical items is neither expected nor intended to retrieve every instance of metadiscourse present in the corpus. It does, however, contain a similar number of target lexical items to other

comparable studies (e.g. Ädel, 2006, p98; Aull and Lancaster, 2014, p176) and should therefore return informative results.

Once the list of target lexical items was finalised, each of the sub-corpora were searched for those items using the concordancer in WordSmith Tools 5 (Scott, 2007). Once identified, concordance lines were manually inspected to exclude those uses which were not metadiscoursal. This is illustrated below using *overall*. As an adverbial (example (1)), it functions as a frame marker, introducing a conclusion:

(1) *Overall*, it is clear that there are many risks to Joanna's pregnancy.

As an adjective (example (2)) it functions as a modifier, is not metadiscoursal, and is thus excluded from the study:

(2) ... must be given due consideration when assessing her *overall* wellbeing.

When lexical items occurred more than 50 times in a sub-corpus, an overall estimate of the proportion of metadiscourse use for that item was obtained by examining a random sample of 50 concordance lines: a common approach in corpus linguistics (e.g. Hyland, 2004).

Differentiating metadiscoursal use from non-metadiscoursal use for some transition markers was problematic: for example, in deciding whether *in addition* was adding experiential information (non-metadiscoursal) or adding to an argument (metadiscoursal). Rather than risk introducing errors into the data, it was decided to exclude those transition markers whose function was frequently ambiguous. This led to transition markers signalling addition being excluded from the study. Some frequently used items such as *but* and *so* were also excluded to allow the study to be completed in the time available. Therefore, in considering the findings of this study, it must be remembered that the list of transition markers included was not exhaustive (see appendix). A future analysis may investigate the frequency of transition markers without differentiating between metadiscoursal and non-metadiscoursal use.

Inspection of the concordance lines showed several instances where lexical items had been used incorrectly. These instances were included as they signalled an intention of the author to use metadiscourse and the error was usually restricted to using an incorrect word from the same metadiscourse category, for example, by using a connective (transition marker) inappropriately. The study takes no account of

lexical items which have been misspelled as these were not retrieved by the search. The effect of this cannot be quantified but is expected to be small due to the prevalence of spell checkers.

To facilitate comparison, frequencies of target lexical items were normalised to frequency per 10,000 words. Multi-word items such as 'as a result' were treated as a single unit. To test whether any differences obtained were statistically significant, the log likelihood test (Rayson, 2008, p.527) was applied to the actual frequencies. Differences were considered significant if the log likelihood was 3.84 or above ($p < 0.05$). For differences which were significant, the effect size was calculated (Rayson, n.d.) in the form of %DIFF (Gabrielatos and Marchi, 2012) to give a measure of the size of the difference.

Findings and Discussion

Metadiscourse use in student writing

Before discussing differences detected between sub-corpora of student writing, it is useful to consider the characteristics of the corpus as a whole. The findings are summarised in table 2 and are detailed in the appendix. The most frequently used metadiscoursal features were code glosses and transition markers, which is consistent with previous research on both student and expert academic writing (Hyland and Tse, 2004).

Considering the categories of code gloss, this study found more reformulation than exemplification. This is typical of a hard discipline (Hyland, 2007, p.273) and is contrary to the situation in soft disciplines, where exemplification is more common than reformulation (Hyland, 2007; Yüksel and Kavanoz, 2018). This is most likely due to the increased need in the hard disciplines to define and explain specific details clearly.

The most common form of reformulation in the current study was abbreviation, all in parentheses, and mostly taking the form of initialisms and acronyms referring to organisations and technical medical terms and procedures, for example, *National Health Service (NHS)*. The large number of organisations is indicative of the degree of regulation and external oversight to which the discipline is subject. The number of abbreviations related to technical terms and procedures would be expected to be replicated in other scientific disciplines.

Table 2 – Frequency of metadiscourse in the Health corpus by category

Category	Frequency per 10,000 words
Code Glosses	
Exemplifiers	22.56
Reformulators	6.46
Reformulators - abbreviations	25.58
Total	54.60
Transition Markers	
Similarity	0.74
Contrast/concession	15.78
Consequence	12.65
Total	29.17
Endophoric Markers	
Linear	8.00
Non-linear	3.78
Total	11.78
Frame Markers	
Sequence	1.59
Announce goals	4.98
Label Stages	2.39
Total	8.96

Examination of a sample of abbreviations used demonstrated that frequently students did not follow the accepted guideline for using abbreviations, namely that on its first occurrence a term should be written in full followed by the abbreviation in parentheses and that on subsequent occasions the abbreviation should be used alone (Bailey, 2018, p.187). In over 25 per cent of cases, students frequently reformulated the full term with an abbreviation in parentheses despite not subsequently reusing the term. This may be because, in practice, it is the abbreviation which is in common use, as in example (3);

(3) observations will be ... documented on the Newborn early warning trigger and track (NEWTT) chart.

By stating the abbreviation, although superfluous for writing style, students are demonstrating and claiming membership of their disciplinary community.

Of other reformulations, two-thirds were in parentheses. Often, these reformulations provided a definition, see example (4), a rewording of a technical term in less specialised language, see example (5) or a statistic, see example (6).

(4) ...if these are within normal range (6-8mmol/l)...

(5) If diabetic nephropathy (Chronic loss of kidney function) is present ...

(6) The majority of all users of the scheme (68%) were aged 15-17 ...

While reformulations such as examples (4) and (6) provide a solution to integrating necessary numerical data into a narrative sentence structure, reformulations such as example (5) would be unlikely to appear in a non-pedagogic genre. The student is concerned to use the terminology of the discipline, but is equally concerned with demonstrating to the reader, i.e. the tutor, that the terminology is understood. Placing reformulations in parentheses is a practice common in hard disciplines (Hyland, 2007). It can also be seen as a sensible strategy in a pedagogic genre where word counts are strictly limited, as providing code glosses in parentheses requires fewer words than incorporating the code gloss into a sentence.

The remaining reformulators each occurred with a frequency of less than 0.5 per 10,000 words. This indicates significant underuse when compared with the published articles investigated by Hyland (2007), particularly in the use of *i.e.*, which accounts for 25 per cent of the reformulators used in published articles (Hyland, 2007, p.273). In this study, *i.e.* was used in only five per cent of texts, and its unabbreviated counterpart, *that is*, was used only once. This suggests students would benefit from a wider repertoire of reformulators.

Over 80 per cent of the exemplifications were performed by one lexical item, *such as*, with the next common item, *for example*, being used on a further 14 per cent of occasions. This preference is consistent with both student writing and expert academic writing although the dominance of *such as* is less pronounced in other contexts (Yüksel and Kavanoz, 2018). In expert writing, the use of *e.g.* is almost as common as *for example* (Hyland, 2007, p.278), whereas in the Health corpus, the use of *e.g.* accounts for only just over one per cent of exemplifiers.

Transition markers of contrast and consequence were similarly prevalent with a frequency of nearly 16 and nearly 13 per 10,000 words respectively. Markers of

similarity were rarely used with a frequency of less than one word per 10,000, although this category contained a smaller number of lexical items.

The students showed a strong preference for *however* to mark contrast. While this preference for *however* has previously been found in both student writing (Gardner and Han, 2018, p870) and expert writing (Aull and Lancaster, 2014), the strength of the preference is much more marked in this study, with *however* being used to mark contrast in 70 per cent of occurrences. The situation is similar with markers of consequence in that the marker *therefore* was used on over 80 per cent of occasions. This suggests that this is another area where students could enlarge their lexical resources.

The frequency of endophoric markers and frame markers is low compared with the findings of Hyland and Tse (2004). This is not unexpected as several researchers (e.g. Bax *et al.*, 2019) have commented that the frequency of these features increases with text length and at 2,000 words, a typical text in the Health corpus is significantly shorter than the dissertations and theses investigated by Hyland and Tse (2004).

The non-linear endophoric markers all occur in the texts for one of the three assignments, referring the reader to an appendix which was a compulsory element. The linear endophoric markers, however, are in large part expressions where the text itself is referred to (see example (7)), which occurred at a frequency of 6 per 10,000 words, or just over one per text.

(7) The vulnerable group chosen for this essay are teenagers.

The large majority of these expressions occur in the early stage of the assignment as part of a sentence announcing goals, as in example (8).

(8) This essay will discuss local and national public health initiatives...

Williams (1990, p.128) cautions against these text-referential expressions, preferring more sophisticated means of communicating aims. The assignment of agency to the text is, however, a common strategy to avoid self-mention (Tang and John, 1999) and occurs in this study in over 80 per cent of assignments, using items such as *essay*, *assignment* and *report*. Whether self-mention should be discouraged when announcing goals is debateable.

Given the findings above, it is unsurprising that just over half of the frame markers used in the current study were used to announce goals. Frame markers signalling the labelling of stages and sequencing were much rarer, with a frequency of just under 4 per 10,000 words. Sequencing features occurred at a frequency of 0.36 per text and only 55 per cent of students chose to signal their conclusion with an explicit marker, such as *in conclusion*, *to conclude* or *overall*. These very low frequencies could suggest a lack of coherence and a lack of awareness of the needs of readers. Alternatively, it is possible that the highly prescriptive assignment briefs supplied to the students influenced the students to believe that explicit guidance to signpost readers through the assignment, in the form of metadiscourse, was superfluous as the intended reader, the tutor, would be familiar with the direction and stages of the assignment. It should be noted that the assignment briefs frequently direct students to provide an explicit statement of the goals of the assignment which has most probably influenced a high proportion of (interestingly, not all) students to include this.

When looking at the corpus as a whole, it appears that metadiscourse use is consistent with that in a hard discipline, with some typical features such as the heavy use of parentheses for reformulations. This suggests that the students have to a large extent been successful in adopting the practices of their academic discourse community. There is, however, a paucity of frame markers to signpost readers through the assignment which could impact coherence. In addition, there is scope for providing pedagogic intervention to increase the range of metadiscourse markers available to the students, particularly to mark code glosses and markers of contrast and consequence.

Metadiscourse use and level

Comparing metadiscourse use between level 4 and level 5 essays allows the development of academic writing during the course of a student's university career to be investigated. Overall, Level 5 texts contained significantly more metadiscourse markers than level 4 texts ($p < 0.01$), although the magnitude of the difference was modest (see table 3). This suggests students are successful in assimilating the academic writing conventions to which they are exposed as students and is consistent with other studies which found that first year undergraduates underused

both code glosses and some transition markers when compared with both more experienced student writers and expert writers (Aull and Lancaster, 2014).

Table 3 – Comparison of metadiscourse use between the Level 4 essays and Level 5 essays

	Level 4 Essay Freq. per 10,000 words	Level 5 Essay Freq. per 10,000 words	Effect size %DIFF
Code Glosses			
Exemplifiers	24.14	23.20	
Reformulators**	4.90	7.47	52.55
Reformulators - abbreviations**	12.85	24.41	89.95
Total	41.89	55.08	
Transition Markers			
Similarity	0.69	0.57	
Contrast/concession**	12.85	16.22	26.26
Consequence	9.80	11.95	
Total	23.34	28.75	
Endophoric Markers			
Linear	7.84	7.47	
Non-linear	0.00	0.00	
Total	7.84	7.47	
Frame Markers			
Sequence	1.67	1.07	
Announce goals**	3.63	6.33	74.46
Label Stages*	1.73	2.92	68.78
Total	7.03	10.32	
Overall Total**	80.10	101.62	32.10

* difference significant at $p < 0.05$

** difference significant at $p < 0.01$

The largest increase was seen with the use of abbreviations. This could reflect a higher informational load at level 5 or a greater awareness among students of the need to include information and references to outside agencies in their assignments. Considering other code glosses, the frequency of reformulations increased from level 4 to level 5 but at a much lower level. This increase was due to an increase in other markers of reformulation such as *i.e.*, *is where*, and the use of parentheses. This indicates a broadening of the students' lexical resources in this area. This

broadening was not apparent with exemplifications: both level 4 and level 5 essays remained heavily reliant on *such as* and *for example*.

The variation in the use of transition markers was most noticeable for markers of contrast. There was a moderate increase at level 5, driven by an increase in the use of *however*. While this does indicate that a broadening of the student's lexical repertoire from level 4 is not apparent, it demonstrates that students at level 5 are developing their ability to take into account a wider range of viewpoints by comparing and contrasting information and sources (Aull and Lancaster, 2014) as in example (9).

- (9) Fasting for long periods should be discouraged [...] *however* this could present a religious and ethical issue for women who observe these practices during Ramadan.

Consequence markers did not increase significantly in frequency from level 4 to level 5 and both sub-corpora showed a marked preference for *therefore*. There was a small increase in formality, however, with the second most popular consequence marker changing from *this means* at level 4 to *thus* at level 5. This is a small effect, however, as despite being the second most popular markers, these items represent barely 10 per cent of occurrences.

Frame markers which announce goals and label stages were significantly more frequent in the level 5 essays. This indicates more attention being paid to coherence and to guiding the reader. It is still the case, however, that only two-thirds of students at level 5 chose to mark their conclusions explicitly.

Although more students at level 5 announced their goals, this was increasingly formulaic. At level 4, seven assignments (10%) included a phrase which followed the structure:

This	essay/assignment	will	discuss/be discussing/look at/ explore/highlight/explain ...
------	------------------	------	---

At level 5, this had increased to 22 assignments (33%) with 16 preferring the verb *discuss*. It seems likely that this increase is in response to an intervention in some form from academic staff.

The general increase in use of code glosses, frame markers and transition markers with increasing level of study points to an increasing development in academic writing skills and a gradually widening repertoire of lexical items in most areas. These factors demonstrate the ability of the students to adopt the conventions and requirements of the academic discourse community in their writing as they spend time as part of that community, although in several areas the progress could be hastened by targeted pedagogic intervention.

Metadiscourse use and grade

Variation within the cohort was examined by comparing the Higher-graded and Lower-graded sub-corpora. There was no significant difference in the overall quantity of metadiscoursal features found in the assignments given a higher grade and those given a lower grade (see table 4). Within categories, the one significant difference was in the use of reformulators.

Table 4 – Comparison of metadiscourse use between the Lower-graded and Higher-graded sub-corpora

	Lower-graded sub-corpus Freq. per 10,000 words	Higher-graded sub-corpus Freq. per 10,000 words	Effect Size %DIFF
Code Glosses			
Exemplifiers	20.48	24.86	111.07
Reformulators**	4.38	9.24	
Reformulators - abbreviations	25.28	25.38	
Total	50.14	59.49	
Transition Markers			
Similarity	0.56	1.30	
Contrast/concession	11.72	14.58	
Consequence	16.24	15.36	
Total	28.53	31.24	
Endophoric Markers			
Linear	9.60	10.80	
Non-linear	2.82	3.77	
Total	12.43	14.58	
Frame Markers			
Sequence	4.24	3.25	
Announce goals	8.05	7.42	
Label Stages	2.54	2.99	
Total	14.83	13.67	
Overall Total	105.94	118.97	

* difference significant at $p < 0.05$; ** difference significant at $p < 0.01$

There was a more frequent use of parentheses in the Higher-graded sub-corpus. In the Lower-graded sub-corpus, parentheses were used 23 times for reformulation and this rose to 48 times in the Higher-graded sub-corpus. The Higher-graded sub-corpus contained many more examples of technical terms being explained (16 compared with 3) as in example (10), and more specifications (8 compared with 4) as in example (11). It is possible that this is due to an increased awareness of either the needs of the reader for elaboration, or the need to demonstrate knowledge, and may have contributed to the higher grades obtained.

(10) Neonatal hyperbilirubinaemia (*jaundice*) is a common condition ...

(11) This is important as early feeding (*within the first hour of birth*) is essential ...

Several researchers (e.g. Intaraprawat and Steffensen, 1995; Cheng and Steffensen, 1996) reported that writers receiving higher grades not only used metadiscourse more frequently, but also used a wider range of lexical items to do so. There is little evidence to support that in this study; the range of lexical items used was similarly narrow in both sub-corpora.

From this study, there is little evidence that an awareness of the reader and ability to create an argument, evidenced by metadiscourse, results in a student conveying their propositional content more successfully leading to the award of a higher grade. The reality seems to be that findings from studies in EAP or from writing centres should be applied with caution to the context of this study. It is possible that in hard disciplines such as the health discipline, which are heavily evidence based and require students to demonstrate knowledge and understanding, the rhetorical functions realised through metadiscourse are less important than they may be in a soft discipline. The more frequent use of reformulations in essays given a higher grade underlines this emphasis on demonstrating understanding, and should be a pedagogic focus.

Conclusions

The development of student writing between Level 4 and Level 5 was shown by the significant but modest increase in the metadiscoursal features found. This is indicative of a growing awareness by the writers of the need to engage with a range of views, and the need to guide the reader, both in understanding and navigating the text. There was, however, very little difference in the use of the metadiscoursal

features between assignments receiving the highest grades and those receiving the lowest, contrary to the findings of previous studies. Whether this holds true in a wider range of disciplines would be worthy of further investigation.

One area where the students differed from student academic writers in other contexts was in their overreliance on a small number of lexical items, particularly for code glosses and transition markers of consequence.

While it is fully appreciated that metadiscourse is only one aspect of successful writing, the findings of this study suggest that there would be value in providing students with targeted activities in two areas. The first is in using metadiscourse to help guide the reader through the text. Students should be encouraged to explore the relationship between the writer and the reader, ideally during students' early terms at university, identifying the needs of the reader when approaching an unfamiliar text. Students could explore successful texts, identifying those metadiscoursal aspects which contribute to their effectiveness. This could allow students to identify the importance of using, for example, frame markers to sequence their texts and label their stages. Secondly, students should be introduced to a broader range of lexical resources to perform metadiscoursal functions, particularly for code glosses and transition markers. This could be achieved by providing model sentences using a range of vocabulary that students could refer to while writing. This could be supplemented by interactive online activities giving students an opportunity to rewrite sentences using alternative lexical items.

Learning Developers are in an ideal position to provide such learning opportunities, ideally embedded within teaching programmes so that the features specific to each discipline can be explored.

Appendix – Actual number of metadiscourse occurrences in the Health corpus of each lexical item included in the study

Category	No. of occurrences	Category	No. of occurrences
CODE GLOSSES	2512	ENDOPHORIC MARKERS	542
Exemplifiers	1038	Linear	368
such as	844	earlier	7
example (excl, 'for')	8	above	25
examples	7	essay	221
for example	143	care study	19
e.g.	13	case study	8
(*	17	discussed	25
for instance	6	below	4
		as follows	2
Reformulators	297	aforementioned	4
labelled	1	assignment	25
specifically	16	previously	26
thought of	1	report	2
referred to as	18	introduction	0
precisely	1		
namely	1	Non-linear	174
which is	23	appendix	174
is where	15		
;	5	FRAME MARKERS	412
:	1	Sequence	73
(*	200	following	25
i.e.	10	initially	1
that is	1	then	7
this means	4	lastly	9
		firstly	12
Reformulators - abbreviations	1177	numbering (1, 2, 3...)	4
(*	1177	finally	13
		secondly	2
TRANSITION MARKERS	1342		
Similarity	34	Announce goals	229
likewise	4	discuss	105
again	16	discussed	44
similarly	9	explained	3
equally	5	outlined	2
		discussion	21
Contrast and concession	726	focus on	39
yet	9	focus(s)ed on	5
alternatively	5	focus(s)ing on	7
on the other hand	10	focus(s)es on	1
conversely	13	introduction	1
in contrast	4	conclusion	1
nevertheless	13		
however	514	Label Stages	110
although	108	to conclude	43
even though	8	in conclusion	50
though	7	overall	17
despite	34		
in spite of	1		
Consequence	582		
therefore	481		
as a result	13		
thus	34		
for this reason	12		
which means	2		
this means	25		
consequently	15		

References

- Ädel, A. (2006) *Metadiscourse in L1 and L2 English*. Amsterdam: John Benjamins.
- Aull, L. L. and Lancaster, Z. (2014) 'Linguistic Markers of Stance in Early and Advanced Academic Writing: A Corpus-based Comparison', *Written Communication* 31 (2), pp.151-183. Available at: <https://doi.org/10.1177/0741088314527055>.
- Bailey, S. (2018) *Academic Writing: A Handbook for International Students*. 5th edn. Abingdon: Routledge.
- Bax, S., Nakatsuhara, F. and Waller, D. (2019) 'Researching L2 writers' use of metadiscourse markers at intermediate and advanced levels', *System* 83, pp.79–95. Available at: <https://doi.org/10.1016/j.system.2019.02.010>.
- Becher, T. and Trowler, P. (2001) *Academic tribes and territories: intellectual enquiry and the cultures of disciplines*. 2nd edn. Buckingham: Society for Research into Higher Education and Open University Press.
- Biber, D., Johansson, S., Leech, G., Conrad, S. and Finegan, E. (1999) *Longman Grammar of Spoken and Written English*. London: Longman.
- Cao, F. and Hu, G. (2014) 'Interactive metadiscourse in research articles: A comparative study of paradigmatic and disciplinary influences', *Journal of Pragmatics* 66, pp.15-31. Available at: <https://doi.org/10.1016/j.pragma.2014.02.007>.
- Cheng, X. and Steffensen, M. S. (1996) 'Metadiscourse: A Technique for Improving Student Writing', *Research in the Teaching of English* 30(2), pp.149–181. Available at: <https://www.jstor.org/stable/40171358>.
- Crismore, A. and Farnsworth, R. (1990) 'Metadiscourse in popular and professional science discourse', pp. 118-136, in Nash, W. (ed.) *The writing scholar: studies in academic discourse*. London: Sage.
- Gabrielatos, C. and Marchi, A. (2012) 'Keyness: Appropriate metrics and practical issues', *Corpus-assisted Discourse Studies: More than the sum of Discourse Analysis and computing? CADS International Conference 2012*. University of Bologna, Italy 13-14 September.
- Gardezi, S. A. and Nesi, H. (2009) 'Variation in the Writing of Economics Students in Britain and Pakistan: The Case of Conjunctive Ties', pp. 236–250 in Charles, M., Pecorari, D., and Hunston, S. (eds) *Academic Writing: At the Interface of Corpus and Discourse*. London: Continuum.
- Gardner, S. and Han, C. (2018) 'Transitions of contrast in Chinese and English university student writing', *Educational Sciences: Theory and Practice* 18(4), pp.861–882. Available at: <https://doi.org/10.12738/estp.2018.4.0067>.
- Holmes, J. and Nesi, H. (2009) 'Verbal and Mental Processes in Academic Disciplines', pp. 58–72, in Charles, M., Pecorari, D., and Hunston, S. (eds) *Academic Writing: At the Interface of Corpus and Discourse*. London: Continuum.
- Hyland, K. (2004) 'Disciplinary interactions: Metadiscourse in L2 postgraduate writing', *Journal of Second Language Writing* 13(2), pp.133–151. Available at:

<https://doi.org/10.1016/j.jslw.2004.02.001> .

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

Hyland, K. (2007) 'Applying a Gloss: Exemplifying and Reformulating in Academic Discourse', *Applied Linguistics* 28(2), pp.266–285. Available at: <https://doi.org/10.1093/applin/amm011> .

Hyland, K. (2017) 'Metadiscourse: What is it and where is it going?', *Journal of Pragmatics* 113, pp.16-29. Available at: <https://doi.org/10.1016/j.pragma.2017.03.007> .

Hyland, K. and Tse, P. (2004) 'Metadiscourse in academic writing: A reappraisal', *Applied Linguistics* 25(2), pp.156-177. Available at: <https://doi.org/10.1093/applin/25.2.156> .

Intaraprawat, P. and Steffensen, M. S. (1995) 'The use of metadiscourse in good and poor ESL essays', *Journal of Second Language Writing* 4(3), pp.253–272. Available at: [https://doi.org/10.1016/1060-3743\(95\)90012-8](https://doi.org/10.1016/1060-3743(95)90012-8) .

Lea, M. R. and Street, B. V. (1998) 'Student Writing in Higher Education: An academic literacies approach', *Studies in Higher Education* 23(2), pp.157–172. Available at: <https://doi.org/10.1080/03075079812331380364> .

Mauranen, A. (1993) 'Contrastive ESP Rhetoric : Metatext in Finnish-English Economics Texts', *English for Specific Purposes* 12, pp.3–22.

Noble, W. (2010) 'Understanding Metadiscourse Use : Lessons from a “Local” Corpus of Learner Academic Writing', *Nordic Journal of Applied Linguistics* 9(2), pp.145–169. Available at: <https://doi.org/10.35360/njes.221> .

Rayson, P. (2008) 'From key words to key semantic domains', *International Journal of Corpus Linguistics* 13(4), pp.519–549. Available at: <https://doi.org/10.1075/ijcl.13.4.06ray>

Rayson, P. (no date) *Log Likelihood and Effect Size Calculator*. Available at: <http://ucrel.lancs.ac.uk/llwizard.html> (Accessed: 17 August 2019).

Scott, M. (2007) 'WordSmith Tools version 5'. Liverpool: Lexical Analysis Software.

Shaw, P. (2009) 'Linking Adverbials in Student and Professional Writing in Literary Studies: What Makes Writing Mature', pp. 215–235 in Charles, M., Pecorari, D., and Hunston, S. (eds) *Academic Writing: At the Interface of Corpus and Discourse*. London: Continuum,.

Tang, R. and John, S. (1999) 'The “I” in identity: Exploring writer identity in student academic writing through the first person pronoun', *English for Specific Purposes*, 18, pp.S23-S39. Available at: [https://doi.org/10.1016/s0889-4906\(99\)00009-5](https://doi.org/10.1016/s0889-4906(99)00009-5) .

Vande-Koppel, W. (1985) 'Some Exploratory Discourse on Metadiscourse', *College Composition and Communication* 36, pp.82–93. Available at: <https://doi.org/10.2307/357609> .

Williams, J. M. (1990) *Style: Toward Clarity and Grace*. Chicago: University of Chicago Press.

Yüksel, H. G. and Kavanoz, S. (2018) 'Dimension of Experience: Metadiscourse in

the Texts of Novice Non-Native, Novice Native and Expert Native Speaker', *Advances in Language and Literary Studies* [online], 9(3), pp.104-112. Available at: <https://doi.org/10.7575/aiall.v.9n.3p.104>.

Zhang, M. (2016) 'A multidimensional analysis of metadiscourse markers across written registers', *Discourse Studies*, 18(2), pp.204–222. Available at: <https://doi.org/10.1177/1461445615623907>.

