Jun 3rd, 9:00 AM - Jun 6th, 5:00 PM

Reading Group Discourse: A corpus-based analysis of argumentation and collaboration

Kieran O'Halloran
Open University

Reading Group Discourse: A corpus-based analysis of argumentation and collaboration

KIERAN O'HALLORAN

Centre for Language and Communication
Faculty of Education and Language Studies
Open University
Walton Hall
Milton Keynes
MK7 6AA
UK
k.a.ohalloran@open.ac.uk

ABSTRACT: This presentation will make a contribution to understanding the nature of debate in reading groups (i.e., people who meet to discuss books, usually novels). I will report on the findings of The Discourse of Reading Groups, a UK Arts and Humanities Research Council (AHRC) one-year, funded project (2008) that aims both to gather evidence about contemporary reading practices and, in part, to contribute to a sensitive understanding of social literary argumentation as a contemporary micro-culture.

KEYWORDS: argumentation, collaboration, corpus linguistics, keyness, linguistic form, linguistic function, literary evaluation, literary interpretation.

1. INTRODUCTION

1.1 Orientation

This paper will make a contribution to understanding the discourse of reading groups, shining light on the kind of argumentation used in evaluation and interpretation of books read in a variety of reading groups. Reading groups, book groups, book clubs - the terms are used interchangeably—regularly meet in members' houses, in pubs or restaurants, in book shops, workplaces, schools, libraries or prisons[1] to share their experiences of reading contemporary fiction and more established, canonical literature. I will report on one part of the findings of a one-year project—'The Discourse of Reading Groups'—funded by the UK Arts and Humanities Research Council and conducted in 2008. The project gathered evidence about contemporary reading practices in book groups, and contributes to an understanding of social literary discourse as a contemporary micro-

[1] There are also on-line reading groups (e.g. using salon.com). This being a one-year project, it was felt better to retain the focus on face-to-face groups.

O'Halloran, K. (2009). Reading Group Discourse: A corpus-based analysis of argumentation and collaboration. In: J. Ritola (Ed.), *Argument Cultures: Proceedings of OSSA 09*, CD-ROM (pp. 1-12), Windsor, ON: OSSA.

culture.[2] The research questions for this article are as follows: how are evaluations and interpretations constructed, justified, contested, developed between participants?

The article will respond to these questions through novel synergy between a *quantitative* tool of corpus linguistics (WMatrix) and *qualitative* coding software (Atlas-ti) which is used to map different discourse functions. Through active combination of these two tools, involving movement back and forth between them, insight is gained into patterns of co-occurrence between collocation[3], i.e., linguistic *form*, and discoursal *functions* in argumentation. In turn, such form and function regularities will illuminate relationships between time, space and reasoning in evaluative and interpretative discourse in reading groups.

When I refer to 'argumentation' in this article, I am signalling the process of arguing.

1.2 Why reading groups?

Reading groups have become an important cultural phenomenon in Britain and other countries. Some estimates put the number of such groups at around 50,000 in Britain, and up to 500,000 in the USA (Hartley, 2002). Reading groups are sites of critical literary debate: they provide evidence of how 'ordinary people'—as opposed to critics and academics—interpret and evaluate literature. Despite the cultural salience of reading group activity, reading groups are under-represented in empirical research.

1.3 The reading groups used for this research

This article reports on the investigation of 10 different groups in the UK. These were selected on a variety of factors: age, gender composition, geographical setting, sexuality, social setting (e.g. home, workplace, prison, school). A key reason why diversity was sought was because if dominant patterns of linguistic form and argumentational function could be ascertained, then they could be taken, with greater credibility, to be reasonably representative of reading groups writ large. A breakdown of the different groups can be seen in Table 1. Group abbreviations will be used in other tables.

1.4 Codes for discourse functions

In each reading group investigated, participants discussed different novels they had all read. Discussions were audio-recorded and then transcribed.[4] All the transcriptions were qualitatively coded for different discourse functions. Two types of discourse code were decided upon: 'Discussion' and 'Thematic.' 'Discussion' codes relate to argumentation,

[2] The project team consisted of 3 researchers, Joan Swann, Kieran O'Halloran and Daniel Allington, all working in the Centre for Language and Communication, Open University, UK.

[3] I use the term 'collocation' generally in this article to include both 'colligation' (the tendency for two or more words from different grammatical categories to co-occur in the same structural relationship, e.g., the verb 'want' colligates with the structure 'to+infinitive') and 'collocation' (lexical words which frequently occur together in texts e.g., 'small + amount').

[4] In the entire project, each group was recorded twice, discussing different novels on each occasion. However, because this was a 1 year project, and there were other research questions investigated (see footnote 5), only 10 recordings were coded for argumentation.

to the research questions indicated in Section 1.1, and will be introduced in Section 4. What we called 'Thematic' codes are those which signal other, less related, discourse functions in the transcriptions, which could be in one speaker turn or over several speaker turns. For instance, 2 Thematic codes 'Evaluation' and 'Interpretation' were attached to text where evaluation (e.g. 'I thought the book was great') and interpretation (e.g. 'I thought the main character represented Satan') took place respectively. Only these 2 thematic codes, though, will feature in this article given its focus towards argumentation. In total, there are 19 thematic codes[5] and these were largely determined by the project's 4 other research questions[6], observation of reading groups and immersion in the transcripts. (See Appendix 1 for Table 1 on the Reading groups.)

In the next section, I set out the research methods used for answering the questions indicated in Section 1.1.

2. RESEARCH METHODS

One research method for this study is corpus linguistics: the quantitative study of lexical and grammatical patterns on the basis of large databases known as corpora. Corpora have been used for a variety of purposes but their use for investigating argumentation has only been recent (e.g. Coffin and O'Halloran, 2008, forthcoming; Degano, 2007; O'Halloran, 2009; Reed, 2006; Zagar, 2007). The method employed in this article for the investigation of argumentation is novel in its active combination of a *quantitative* corpus-based tool—in this instance, WMatrix (Rayson, 2008), and *qualitative* semantic coding software—in this instance, Atlas-ti.[7]

Transcriptions of audio recordings of reading groups were qualitatively coded using Atlas-ti. This allows areas of text, or what Atlas-ti refers to as 'quotations,' to be highlighted according to the extent of a particular function or phase of discussion (see Figure 1 in Appendix 2) or, in fact, any coding deemed useful by a researcher. Thematic and Discussion code names reflecting different discourse functions (e.g. the Thematic codes 'Evaluation' or 'Interpretation') were attached to quotations. Quotations can be extracted on individual codes and corpora built on this basis. In this paper, I extract

[5] Other examples of Thematic codes include: 'Material Text,' where somebody either comments on the physical book (its cover, etc) or uses it in some way (e.g. by reading aloud); 'Paratext,' where a text in the book other than the main text is discussed, e.g. the blurb.

[6] The research questions relevant to this article (see section 1.1) form one set. The others for the project were as follows:

> a) What kinds of interpretations and evaluations do reading group members provide of contemporary fiction and other literary texts?
> b) How is such critical work interwoven with the social, interpersonal and affective demands of group interaction?
> c) What kinds of reader identities are negotiated by participants?
> d) How do interpretations/evaluations relate to the professional judgements made by literary critics and academics?
>
> On responses to these other research questions from the project, see Allington and Swann (forthcoming).

[7] Atlas-ti version 5.5 was used: http://www.atlasti.com. WMatrix is an on-line software tool, while Atlas-ti is not.

quotations on particular discourse functions relating to argumentation and then build individual function-based corpora from these quotations (see Sections 5-6). Via WMatrix, contrastive quantitative linguistic analysis of these corpora (e.g. examining collocations) will illuminate differences in linguistic behaviour between different argumentational functions. In turn, this will provide insight into relationships between time, space and reasoning in evaluative and interpretative discourse in reading groups.

WMatrix allows comparison of constructed corpora with different, large reference corpora (e.g. some are around 1 million words). Such comparison allows statistically salient words—*keywords*—to be revealed. These are defined as being statistically more salient in a text or set of texts as compared with a large reference corpus. 'Keyness' is established through the statistical measure, log likelihood (see Dunning, 1993). In WMatrix, a log likelihood value of ≥ 6.63 confers keyness ($p < 0.01$); keyness of a word is proportional to the size of the log likelihood value greater than 6.63. Calculation of log likelihood takes into account the different sizes of the reference corpus and the corpus being compared. This means that if the same reference corpus is used, corpora of different sizes can be meaningfully compared with one another for word frequencies.

Another useful function of WMatrix is that it can run text data through two different automated taggers: a part-of-speech tagger and a semantic tagger.[8] These taggers are software programs which respectively: label the linguistic content of texts in accordance with grammatical information; group words into semantic categories on the basis of a preconfigured lexicon. So, for example, the part of speech tagger would label 'book,' 'novel,' 'reader' as nouns and the semantic tagger would label these words with the superordinate semantic category, 'The Media Books.' With text data tagged in these ways, one is in a position to see to what extent particular grammatical phenomena, as well as particular semantic fields, feature in a corpus. Since the reference corpora in WMatrix have already been tagged for grammatical and semantic information, statistical comparisons are afforded. In other words, it is possible to see in the corpus or corpora being investigated whether particular parts of speech or particular semantic fields have keyness. For the function-based corpora investigated in Sections 5-6, parts of speech and semantic fields which exhibit keyness can be taken as characterising different types of argumentational discourse in the ten reading groups. (See Appendix 2 for Figure 1 for a screen shot of discourse function coding with Atlas-ti.)

3. 'ON-BOOK' AND 'OFF-BOOK' CODES

One thing that was noticeable in the majority of reading groups was a rhythm to the discussion. Engagement with the novel rocks between 'On-book' discussion - talk about the novel, its characters, its plot, its setting - and 'Off-book' discussion where the participants talk generally about a topic which relates to their lives—cultural, economic, philosophical, political—one that emerged from the specifics of their 'On-book'

[8] The taggers were developed at Lancaster University. The part of speech tagger is known as CLAWS (Constituent Likelihood Automatic Word-tagging System) and the semantic tagger is known as USAS (UCREL Semantic Analysis System). For more information on the taggers, the part-of-speech tagset (i.e., the labelling used by these taggers), and the full list of semantic categories, see the following (Lancaster) University Centre for Computer Corpus Research on Language (UCREL) websites: http://ucrel.lancs.ac.uk/claws/ and http://ucrel.lancs.ac.uk/usas/

discussions. So, for example, the North American setting of the novel *Lucky Star* by Joe Keenan led to the Gay group comparing United States with British / European culture. In order to really understand On-book argumentational discourse - the main focus of this article—separation of On-book from Off-book discourse was needed. While listening to the audio-recordings, the 10 transcripts were coded in Atlas-ti for On-book and Off-book.[9] Table 2 shows word amounts for each of the 10 groups in Off- and On-book discourse.

	Book	*Gay*	*Home*	*Knit*	*Lib*	*Pri*	*Pub*	*Sch*	*U3A*	*Work*	*Total*
OFF-BOOK	3449	5171	2445	6343	0	723	899	2183	1802	1864	24879
ON-BOOK	6221	6871	5698	4044	4595	2430	4688	3922	11374	3468	53311

Table 2 Word amounts—Off-book and On-book—for the 10 groups

The rest of this article relates only to the 53, 311 words of On-book discourse.

4. DISCUSSION CODES

In actual discourse, argumentational claims etc are not always easily identified, a point highlighted in Coffin (2007) and North et al. (2008). So, rather than trying to maintain a clear distinction between 'argumentation' and other discourse functions, Coffin (2007) created a looser category of 'Discussion' that incorporates potential argumentational moves together with those that are unambiguously argumentational. Taking a cue from Coffin (2007), and given that initial coding of the reading group data found areas of ambiguity, the looser category of 'Discussion' was regarded as essential. There are 12 Discussion codes in total.[10] Because of space constraints, only a handful of Discussion codes are referred to in this article; Table 3 contains these.

DISCUSSION CODE	*DESCRIPTION*
CHALLENGE	challenges a previous claim
CLAIM	a contestable proposition relating to how things are
CO-CONSTRUCTION	where claims from different speakers build on each other
EXTRINSIC REFERENCE	a claim / challenge which includes reference to outside the novel
INTRINSIC REFERENCE	a claim / challenge which includes reference to the novel

[9] Judgements as to where Off-book and On-book start and finish were sometimes tricky. At times, there were clear-cut distinctions; at others, the dividing line was blurred. This is because there is sometimes a gradual transition from On-book to Off-book; in contrast, there is a more marked shift from Off-book to On-book. However, since I am looking for dominant patterns across On-book discourse, this offsets occasions where it was difficult to place an exact dividing line.

[10] Qualitative coding of argumentation in On-book discourse included other codes such as Agreement, Concession, Discussion prompt which are not referred to in this article.

Table 3 Discussion Codes

The unit of analysis for 'Claim' was what was interpreted as evaluation / interpretation of the novel which contained distinctive propositional content; a speaker turn could consist of more than one claim. If there was any support for the claim (e.g. 'One reason I liked the novel was because the central character was amusing') then this was included under Atlas-ti quotations for 'Claim.' The code, 'Challenge,' indicates that a speaker confronts a previous claim. If there was supportive reasoning for the challenge then this was included under the 'Challenge' quotation. Sometimes Atlas-ti quotations for 'Challenge' did not include propositional content, e.g., 'I disagree completely.'

It was felt useful to track specifically where support reasoning occurred for claims and challenges. However, while many claims and challenges made reference to the novel or to outside the novel, it wasn't always clear if claim/challenge + support reasoning was taking place via such references. Ambiguity or implicitness in the dataset made it sometimes difficult to assign a code such as 'Support Reasoning.' As a result, the wider codes of 'Intrinsic Reference' and 'Extrinsic Reference' were adopted. These were attached to claims and challenges where there was: i) explicit support reasoning; ii) implicit support reasoning; iii) a reference to the novel or a reference to outside the novel where it was ambiguous whether claim / challenge + support reasoning occurred.

In a single example of a Co-construction, participants make claims which others build upon. Nevertheless, these claims were not coded as different instances of 'Claim.' This is because the code, 'Claim,' was attached to singular, non-collaborative claims only.

Table 4 shows code distributions across the 10 reading groups for On-book discourse.

	Book	Gay	Home	Knit	Lib	Pri	Pub	Sch	U3A	Work	Total
THEMATIC											
EVALUATION	18	24	10	11	11	7	7	7	17	9	121
INTERPRETATION	20	2	18	6	9	5	15	2	19	15	111
DISCUSSION											
CHALLENGE	10	45	9	11	10	5	11	16	26	1	144
CLAIM	26	49	49	23	48	13	25	15	31	20	299
CO-CONSTRUCTION	15	10	8	12	0	2	19	6	24	16	112
EXTRINSIC REFERENCE	4	14	0	8	7	3	1	3	8	4	52
INTRINSIC REFERENCE	15	44	5	10	18	2	5	14	23	5	141

Table 4 Code distributions for On-book Discourse

5. PATTERNS OF LINGUISTIC FORM AND ARGUMENTATIONAL FUNCTION

5.1 Orientation

Three function-based corpora were compiled: Claim (10, 003 words), Challenge (2,948 words), Co-construction (14, 640 words).[11] The corpora were examined for grammatical and semantic information using WMatrix. The reference corpus chosen was the British National Corpus (BNC) spoken sample of 1 million words. Given the focus on uncovering form and function patterns across the 10 reading groups, one has to take care with interpreting log likelihood values. These can be large even if the linguistic item under investigation features infrequently. This happens when the reference corpus does not contain, or barely contains, this linguistic item. So, I needed to pay heed not only to statistical salience but to numerical salience also. When I indicate quantitative results below, much of this is done via brackets containing 2 figures: the first refers to frequency and the second to log likelihood (LL) value at $\geq$ 6.63 (i.e., measure of keyness). One other part of the data exploration below involves examining co-occurrences between the above Discussion codes and the 2 Thematic codes used in this article: Evaluation and Interpretation.

5.2 Claim corpus

Out of 299 claims, 143 co-occur with Evaluation and 76 with Interpretation, showing a strong tendency towards evaluation in claims. From WMatrix analysis of the corpus, it was found that the first person subject pronoun, 'I,' has keyness (521 times; LL = 27.3). Furthermore, the *past tense* of lexical verbs in the Claim corpus has statistical salience (318; LL =88.7). In contrast, present tense has no statistical salience. Thus, in the main, the claim function is made from a first-person perspective in the past tense. Commonly, this is done with the collocation 'I thought' (91 times) such as in the following evaluative claim:

> *I thought* that would transfer very easily into a film as well
>
> (Workplace group)

The WMatrix semantic field category, 'Interested/excited/energetic,' has statistical salience (47; LL = 89.6). In other words, when people make claims, they often do so in a reasonably effusive way via adjectives such as 'interesting,' 'gripping,' 'exciting,' 'passionate.' I went back into Atlas-ti and coded for the 47 adjectives subsumed by the semantic field of 'Interested/excited/energetic' and then merged the codes into a supercode, 'interesting.' For this merged supercode, there was a good spread across the groups (such distribution can't be shown in WMatrix).

While people are often effusive in constructing evaluative claims, there would seem to be a fair amount of mitigation too: the semantic category of

[11] Header information on the type(s) of code accompanies each quotation. This was stripped out in all corpora compilations.

'Exclusivisers/Particularisers,' realised overwhelmingly in the data by 'just,' has statistical prominence (97; LL = 10.38), e.g.:

> there *was just*, there *was just* like no future whereas Mice and Men at least there *was* a dream, I don't think there *was* even a dream in this really I *just thought* that they *were* victimised for no reason whatsoever
>
> (Home group)

as does use of 'sort of' (58; LL = 123.0), e.g.:

> I I must admit I *sort of thought* it was like y'know eating candy floss
>
> (Gay group)

From coding in Atlas-ti, both 'sort of' and 'just' have good distribution across the 10 groups. Notice the use of past tense in these 2 extracts (in italics). Use of past tense does not necessarily signal past time; the past tense can be used to signal indirectness or tentativeness in expressing a proposition. The salience of past tense in the Claim corpus can thus be seen to support the analysis that this corpus is characterised by mitigation.

5.3 Challenge corpus

Similar to the picture for claims in Section 5.2, out of 144 challenges, more relate to Evaluation (61) than Interpretation (43), though this is less skewed towards evaluation than in the Claim corpus. Again, like the Claim corpus, the first person subject pronoun has keyness (132; LL=12.97) and so too does the past tense of lexical verbs (82; LL=31.6); present tense has no statistical salience. The challenge function is made, in the main, from a first-person and past tense perspective, and commonly with the collocation 'I thought' (67 times), such as in this example (past tense italicised). (Different numbers in this and other data extracts refer to different participants):

> 1: yeah um I actually *thought* she *killed* the baby but perhaps I've got that wrong (CLAIM)
>
> 2: oh you see that's what I *couldn't* I I actually (CHALLENGE)
>
> 1: I *thought* she *was trying* to stop the baby making a noise (CLAIM)
>
> 2: well you see *I thought* they *said* there *was* a gun shot (CHALLENGE continued)
>
> (Library)

Interestingly, there is no pronounced pattern of long chains of Claim / Challenge / Challenge etc over several speaker turns. The topic of most Claim / Challenge patterns fizzles out after 2-3 turns. In other words, there is little evidence of challenging leading to lengthy defences of positions through alternate exchange between a proposer and an opposer. In fact, there is a tendency for challenging to redirect discussion towards newer topics (77 instances) which in turn maintains the vitality of interpretative co-construction (see Section 5.4). Finally, none of the categories which were salient in the Claim corpus—'Exclusivisers / Particularisers' and 'sort of'—have log likelihoods $\geq$ 6.63. Challenges, it would seem, are expressed without this kind of mitigation, on the whole, though this is not to say that past tense is not realising mitigation.

5.4 Co-construction corpus

Out of 112 instances of co-construction, 80 co-occur with interpretation, but 32 co-occur with evaluation. There is, then, a strong tendency towards interpretation in co-construction. Let's have a look at one Atlas-ti 'quotation' which was coded for 'Co-construction' and 'Interpretation'; a character from the novel, *Snow*, by Orhan Pamuk, is being discussed (i.e., there is plenty of Intrinsic Reference in this extract):

> 1: *he is* alienated from his country
>
> 2: Yes *I think* there *is* the alienation *is* important yes
>
> 3: *He has been* away from his country *because* of some coup in the 1980s and *he* was required to flee
>
> 4: *He has been* a political exile *I* suppose
>
> 2: And *come back* invited by a newspaper in Istanbul to go to Kars to research suicides on the young girl side
>
> 5: Investigative
>
> 2: Yes, an Investigative Journalist, so you know *he is*, not only *is he* a writer but *he is* somebody who *investigates* like a detective
>
> (Bookshop group)

Notice only 2 instances of 'I,' the first person subject pronoun, but 7 instances of the third person subject pronoun, 'he' (italicised above). This is rather typical in Co-construction reflected in there being no keyness value for first person subject pronoun, but high keyness for 3rd person subject pronouns, 'he' or 'she' (451; LL = 271.7). These pronouns usually refer to a character (as in the above extract) or the author. In contrast, for the Claim and Challenge corpora, log likelihoods for 'he'/'she' are significantly lower: 14.0 and 26.4 respectively.

Another feature that characterises the Co-construction corpus, and in contrast with the Claim and Challenge corpora, is statistical salience for the *present tense* (e.g. 3rd person present tense for lexical verbs: 143; LL = 22.2). Indeed, a very common collocation is 'he/she + present tense' or 'he/she + present perfect.' As echoed in the above example (present tense and present perfect italicised), it is common in co-construction to find an assembly of statements about a character being made predominantly in the present tense or the present perfect.

Unlike in the Claim corpus, the semantic field category, 'Interested/excited/energetic,' does not feature at a log likelihood value of ≥ 6.63. In other words, when people are co-constructing, they are not doing so in as effusive way as in the Claim corpus. Nor are they expressing interpretations, and to a lesser extent evaluations, in as mitigated a way as in the Claim corpus. The semantic category, 'Exclusivers/Particularisers,' does not have statistical salience; secondly, 'sort of' (31; LL = 30.3) has a much lower log likelihood than in the Claim corpus (58; LL = 123.0).

5.5 Reflection

When participants make claims, these are predominantly evaluative and in the *past tense*. We can say that claims mostly represent experience of reading the novel in the past and / or this experience is mitigated through use of the past tense. In contrast, when participants are co-constructing, they are more likely to be interpreting the novel in the *present tense*; they are less likely to be reporting pre-formulated, holistic, individual interpretations which were generated during the original reading. This is because there is a strong tendency for joint interpretation to be *emergent* in group discussion, emergent from assemblies of individuals' 'interpretative bits.' By 'interpretative bits' I mean statements about a character(s) which have the potential to feed into a globally emergent interpretation. In the main, no one speaker 'owns' the interpretation since it is frequently more than the sum of its parts. Another interesting finding is that co-constructive discourse is less diffidently expressed than claim discourse. This might be because claim-based evaluation is more holistic, more complete, than interpretative co-construction; it thus has a greater risk to face since such discourse is easier to contest. In contrast, a participant who offers only an 'interpretative bit' during discussion is less vulnerable to dispute.

6. REASONING

6.1 Analysis

Subordinator conjunctions have statistical salience (LL = 26.2); by far the largest of these is 'because' at 303 instances.[12] I indicated, in Section 4, that the reasonably elastic categories of 'Intrinsic Reference' and 'Extrinsic Reference' were chosen so as to capture Claim/Challenge + Support Reasoning as well as data where this was more implicit or ambiguous. Given the strong tendency in co-construction towards interpretation, it is illuminating to contrast co-occurrences with 'because' of interpretative co-construction, interpretative claims and interpretative challenges. I went back to Atlas-ti and coded for all instances of 'because' (i.e., including 'cos'), and then merged into one super-code, 'because.' 61 interpretative co-constructions (out of 80) contain instances of 'because' (76.2%). In contrast, only 5 interpretative challenges (out of 43) contain 'because' (11.6%), and only 14 interpretative claims (out of 76) contain 'because' (18.4%).

To put some qualitative flesh on these quantitative bones, here are some illustrative examples exhibiting Intrinsic Reference. The first is an *interpretative* claim in relation to a narrative device in the novel:[13, 14]

> he disguised things too much I swear the whole first chapter was describing the morning (Secondary School group)

and here is an *interpretative* challenge:

[12] In contrast, 'since,' another subordinator conjunction which can have a similar function to 'because,' only occurs 3 times.

[13] Out of the 141 instances of Intrinsic Reference, 66 co-occur with Evaluation and 52 with Interpretation.

[14] Intrinsic Reference, rather than Extrinsic Reference, is explored in this section given the greater quantitative salience of the former (see Table 4).

> But they were on they were on the Aborigines land. Ahm and the patch having been dug over first which he which William tried to dismiss ahm but they knew didn't they, they knew
> (U3A group)

Notice that reasoning is not so explicit in the above examples. There is no explicit Claim/Challenge+Support structure where, say, support in the form of Intrinsic Reference is packaged in a 'because' subordinate clause. The same picture is found in *evaluative* claims:

> one of the things that I really really liked, I just found so actually just refreshing, is up until the point she decides to go on the diet was how she was just so you know, my fatness and I was fat and - and I loved it (Knitting Group)

and in *evaluative* challenges:

> And also I disagree with the translation being sort of clunky and um I felt it was entirely modern and entirely global if you like, the language.
> (Bookshop Group)

Let me contrast with interpretative co-construction exhibiting Intrinsic Reference. The context of the following representative example (from *The Secret River* by Kate Grenville) is a white settler child, 'Dick,' making contact with Australian aborigines:

> 1 He makes some contact
>
> 2 Yes and seems like he does speak some of the language
>
> 1 yeah
>
> 2 *cos* they suddenly become aware Dick says something that's not English or that's how it seems.
> (U3A group)

Participant 2's use of 'cos' in the above provides explicit support for their first interpretation ('seems like he does speak some of the language') which, in turn, has built on the first observation made by participant 1 ('He makes some contact'). There is evidence for this pattern in the Section 5.4 extract also.

6.2 Reflection

It would appear that once participants are interpretatively co-constructing in the present tense, reasoning is more explicit than in claims or challenges which are in the past tense. Claim and challenge reasoning - which would seem to be based on preformulation - is sometimes ambiguous. This would suggest that while evaluation has involved prior thinking in the past, nevertheless this has involved a less reasoned response than takes place in interpretative co-construction in the present tense. This could well be because claims and challenges, particularly evaluative ones, are often affectively and thus less logically realised as the examples above suggest. The explicitness of reasoning in interpretative co-construction in the present tense is especially interesting given that: i)

such reasoning is on-the-fly; ii) the interpretative bits upon which reasoning is based are embryonically formulated. It would appear then that the multiple perspectives in the reading groups, and their collaborative ethos, energise explicitly *reasoned* co-interpretation.

In sum, from the quantitative-based WMatrix corpora analyses of qualitative-based Atlas-ti function-coded data, this article has highlighted some salient form / function patterns in literary argumentation across 10 distinct reading groups in the UK. In turn, this has provided insight into relationships between time, space and reasoning in evaluative and interpretative reading group discourse.

Link to appendix 1 Link to commentary
Link to appendix 2

REFERENCES

Allington, D. and Swann, J. (eds) (forthcoming) Special Issue of Language and Literature, 18 (3).

Coffin, C. (2007). The Language and discourse of argumentation in computer conferencing and essays, Research Report, *Economic and Social Research Council*, UK.

Coffin, C. and K. A. O'Halloran (2008). New directions, new methods. In: C. Coffin & K.A. O'Halloran (Eds.), Researching argumentation in Educational Contexts: New Directions, New Methods, *International Journal of Research and Method in Education*, 31(3), 219-227.

Coffin, C. and K. A. O'Halloran (forthcoming). Argument reconceived? *Educational Review*, 61(3).

Degano, C. (2007). Presupposition and dissociation in discourse: a corpus study. *Argumentation*, 21, 361-378.

Dunning, T. (1993). Accurate methods for the statistics of surprise and coincidence. *Computational Linguistics*, 19(1), 61-74.

Hartley, J. (2002). *The Reading Groups Book: 2002-2003 edition*. Oxford: Oxford University Press.

North, S, C. Coffin and A. Hewings. (2008). Using exchange structure analysis to explore argument in text-based computer conferences. In: C. Coffin, K.A. O'Halloran (Eds.), Researching argumentation in Educational Contexts: New Directions, New Methods, *International Journal of Research and Method in Education*, 31(3), 257-276.

O'Halloran, K.A (2009) Implicit dialogical premises, explanation as argument: A corpus-based reconstruction, *Informal Logic* 29(1), 15-53.

Rayson, P. (2008). *Wmatrix: a web-based corpus processing environment*, Computing Department, Lancaster University. http://ucrel.lancs.ac.uk/wmatrix.

Reed, C. (2006). Preliminary Results from an Argument Corpus. In: E.M. Bermúdez and L.R. Miyares (Eds.), *Linguistics in the Twenty First Century* (pp.185-196, Ch. 17), 459 pages: Cambridge Scholars Press.

Zagar, I.Z. (2007). Arguing from large corpora: some epistemological and methodological dilemmas; experimental study aimed at quantifying political argumentation. In: F.H. van Eemeren, J.A. Blair, C.A. Willard, B. Garssen (Eds.), *Proceedings of the Sixth Conference of the International Society for the Study of Argumentation* (pp.1553-1558, Ch. 231), Amsterdam: SicSat, International Center for the Study of Argumentation.