



University of HUDDERSFIELD

University of Huddersfield Repository

Walker, Brian and McIntyre, Dan

Thinking about the news : thought presentation in early modern English news writing

Original Citation

Walker, Brian and McIntyre, Dan (2015) Thinking about the news : thought presentation in early modern English news writing. In: Corpora and Discourse Studies. Palgrave Advances in Language and Linguistics . Palgrave Macmillan, London, UK, pp. 175-191. ISBN 9781137431721

This version is available at <http://eprints.hud.ac.uk/id/eprint/24158/>

The University Repository is a digital collection of the research output of the University, available on Open Access. Copyright and Moral Rights for the items on this site are retained by the individual author and/or other copyright owners. Users may access full items free of charge; copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational or not-for-profit purposes without prior permission or charge, provided:

- The authors, title and full bibliographic details is credited in any copy;
- A hyperlink and/or URL is included for the original metadata page; and
- The content is not changed in any way.

For more information, including our policy and submission procedure, please contact the Repository Team at: E.mailbox@hud.ac.uk.

<http://eprints.hud.ac.uk/>

Title:

Thinking about the news: thought presentation in Early Modern English news writing.

Authors:

Brian Walker and Dan McIntyre (University of Huddersfield, UK)

Contact details:

Linguistics and Modern Languages, University of Huddersfield, Queensgate,
Huddersfield, HD1 3DH, UK

Tel. + 44 (0) 1484 478444

Email: b.d.walker@hud.ac.uk; d.mcintyre@hud.ac.uk

Abstract

In this chapter we describe an investigation into the forms and functions of discourse presentation in Early Modern English (EModE) news reporting, concentrating particularly on the presentation of thought. In so doing, we consider the methodological issues behind the manual annotation of a small corpus for categories that are function rather than form-based. Our corpus is composed of 40,000 words of EModE prose fiction and journalistic writing, with the latter subsection consisting of approximately 20,000 words of EModE news; it is results from this section of the corpus that we concentrate on in this chapter. Our corpus was manually annotated using categories of discourse presentation developed from a model originally presented in Leech and Short (1981 [2007]). We compared quantitative findings from our corpus against a similarly annotated corpus of present day English (PDE) news journalism (details of which can be found in Semino and Short 2004) in order to determine diachronic changes in the relative frequencies of discourse presentation categories. Alongside the presentation of our quantitative findings we offer a qualitative analysis of the function of thought presentation in EModE news and an explanation of how we have ensured the replicability of our annotation. The latter is especially important considering the criticisms that have been levelled at the suitability of the Leech and Short/Semino and Short model for corpus analysis. In addressing this issue, we suggest that the principles we have followed are valuable for anyone manually annotating a corpus for function-based categories.

1. Introduction

This chapter reports on our continuing research into discourse presentation (DP) in a corpus of Early Modern English (EModE) writing, manually annotated for categories of DP originally proposed in Leech and Short (1981 [2007]) and later developed by Semino and Short (2004). Our focus in this chapter is on EModE news texts and follows on from McIntyre and Walker (2011, 2012), where we found

particular DP categories to be over-represented in our EModE data when compared against Present Day English (PDE) news journalism. For example, we found that two categories of thought presentation (indirect thought and the narrator's presentation of a thought act) were used significantly more in the EModE data than in PDE. Much of the indirect thought in the EModE data was concerned with constructing the hypothetical thoughts of others. Consequently, we hypothesised that EModE news writers were particularly concerned with speculating about reactions to events rather than giving their own opinions. In this chapter we examine in more detail the forms and functions of thought presentation in EModE news writing, since the presentation of the thoughts of others by journalists in news texts seems rather odd because it is impossible to access other people's thoughts. Indeed, presenting one's own thoughts from a past situation is highly problematic. In this chapter, we will look at the extent to which this phenomenon occurs, and suggest some possible explanations for why.

2. Discourse presentation

Prototypically, discourse presentation refers to the presentation of speech, writing or thought from an anterior discourse in a posterior discourse. A person can report the speech and/or writing and/or thoughts of a third party or themselves using a variety of different forms. Hence, the original utterance (the anterior discourse) *Away, away, every man shift for his life!* can be reported by a third party (in the posterior discourse) using any of the following structures.

- (i) 'Away, away, every man shift for his life!'
- (ii) 'Away, away, every man shift for his life!' he cried.
- (iii) He said that every man should away and shift for his life.
- (iv) He wanted them away, away; every man should shift for his life!
- (v) He expressed his concern.
- (vi) He shouted frantically.

Example (i) expresses the exact words of the original utterance; (ii) includes the exact words plus a reporting clause indicating the presence of a narrator; (iii) presents the original utterance in an indirect form, with the original speaker's words subjected to a backshift in tense, and contained within a subordinate clause; (iv) is a free indirect rendering that blends aspects of a narratorial report with a flavour of the original speaker's utterance; (v) reports only the speech act of the original speaker (none of the propositional content of the original utterance can be reconstructed); and (vi) reports only the fact that speech occurred. Each of the different forms used in examples (i) to (vi) expresses varying degrees of narrator "interference" (Leech and Short 2007: 260, 276), as well as claims to the faithfulness of the reporting of the supposed original utterance.

Perhaps the most comprehensive model of discourse presentation is that introduced by Leech and Short (1981 [2007]). This has been developed and extensively researched over a number of years through corpus-based research projects at Lancaster University (see Semino and Short 2004). The model provides a framework for the analysis of discourse presentation that has been successfully applied to both written data (see, for example, Semino and Short 2004) and spoken data (see McIntyre *et al.* 2004). Table 1 outlines the model of discourse presentation that we used in our project, which is based on the one described in Short (2007). The categories shown in Table 1 form a continuum ordered according to "[...] the amount of 'involvement' of (i) the original speaker in the anterior discourse and (ii) the person in the posterior discourse presenting what was said in the anterior discourse [...]" (Semino and Short 2004: 10). The continuum ranges from categories in the grey-shaded cells at the bottom of the table (N, NPS, NPW, NPT), where no discourse presentation is involved, to the direct discourse presentation categories at the top of the table, which apparently involve only the original speaker. Anything lying in between these two extremes is a combination, in varying proportions, of the original speaker and the person presenting the discourse. Moving up the table through the categories coincides with a gradual transition in viewpoint, shifting

more and more from the point of view of the person presenting the discourse, to the viewpoint of the original speaker/writer/thinker.

In later versions of the model, the term discourse *presentation* is used in preference to discourse *report* or *representation* (see Short *et al.* 2002: 336; Short 2012). This is because hypothetical and forward-facing discourse presentation does not involve the report or representation of something already said, written or thought. Therefore, we use “P” (for “presentation”) in some of the acronyms in Table 1 rather than “R” (for “representation” or “report”), which is used with earlier publications on the subject. We note this in order to avoid any confusion that might arise from changing the acronyms. A consequence of applying this change consistently across the model is that the category descriptions for NV and NPS, NW and NPW and NT and NPT become the same. So, for example, what used to be NRT–Narrator’s Report of Thought (i.e. parts of the narration that signal DP, such as reporting clauses) has become NPT–Narrator’s Presentation of Thought, which has the same description as NT, but clearly describes a functionally different phenomenon. Although this definition ideally needs revising, we do not deal with this issue in this chapter for two reasons: (i) careful examination of more data is needed in order to determine a suitable definition for the presentation of thought in what are prototypically called reporting clauses; and (ii) for now, it is possible to proceed with our use of the model on the grounds that the acronyms in the table below make clear the differences between thought and writing as a ‘reporting’ clause (NPT, NPW) and thought and writing as a ‘reported’ clause (NT, NW).

Speech presentation		Writing presentation		Thought presentation	
(F)DS	(Free) Direct Speech	(F)DW	Free Direct Writing	(F)DT	Free Direct Thought
FIS	Free Indirect Speech	FIW	Free Indirect Writing	FIT	Free Indirect Thought
IS	Indirect Speech	IW	Indirect Writing	IT	Indirect Thought

NPSA	Narrator's Presentation of a Speech Act	NPWA	Narrator's Presentation of a Writing Act	NPTA	Narrator's Presentation of a Thought Act
NV	Narrator's Presentation of Voice	NW	Narrator's Presentation of Writing	NT	Narrator's Presentation of Thought
				NI	Internal Narration
NPS	Narrator's Presentation of Speech	NPW	Narrator's Presentation of Writing	NPT	Narrator's Presentation of Thought
N	Narration	N	Narration	N	Narration

Table 1 Speech, writing and thought presentation model based on the description in Short 2007

Note that compared with the speech and writing presentation clines, thought presentation has an extra category, NI, or Internal Narration. This category relates to the internal (mental) state of the characters/people in a text, and has no clear equivalent category on the speech and writing scales. It has been suggested by Semino and Short (2004), Toolan (2001: 141-2) and Short (2007) that NI could be considered a variant of narration rather than discourse presentation. Hence, this category has been placed in a light-shaded cell to denote that it lies somewhere between discourse presentation and narration. Apart from that one exception, Table 1 shows three parallel discourse presentation scales. This means that while examples (i) to (vi), shown above, constitute speech presentation, the same principle can be applied to the presentation of a third party's writing or thoughts. So, if *Away, away, every man shift for his life!* was originally thought and not uttered, it could be reported using the following structures:

- (I) 'Away, away, every man shift for his life!' (FDT)
- (II) 'Away, away, every man shift for his life!' he thought. (DT)
- (III) He thought that every man should away and shift for his life. (IT)
- (IV) He wanted them away, away; every man should shift for his life! (FIT)
- (V) He feared for the lives of his companions. (NPTA)
- (VI) He thought for a long time. (NT)
- (VII) He was frantic. (NI)

Of course, presenting the thoughts of another in a posterior discourse requires access to the original thoughts and this is altogether more difficult than having access to original speech or writing. We take this point up later in the chapter. Note that the underlined words in examples II and III are instances of NPT–Narrators Presentation of Thought, meaning clauses that introduce or signal DP, but are not part of the DP.

3. Corpus Construction

Our aim was to build a corpus of texts that represented the Early Modern period, and then manually tag the corpus for discourse presentation categories so that we could quantify the usage of these categories, and then compare our findings of those from a corpus of PDE (see Semino and Short 2004) . There were a number of issues to deal with, both methodologically and analytically, in order to achieve this. First, there is no common consensus among historical linguists about when the Early Modern period starts and finishes (Crystal 2005). We opted to define the period fairly broadly, running from approximately 1500 (following, for example, Burnley 1992) to approximately 1750 (following, for example Crystal 2005). Secondly, the labour intensive nature of discourse presentation annotation meant that we were unable to construct a corpus of equivalent size to the news section of the Lancaster corpus, our comparator.

We collected our data from a variety of sources using, wherever possible, texts that were already in electronic, machine readable format. However, many of

the texts in the news section of our corpus were manually transcribed from electronic facsimiles obtained from The Burney Collection, which is available on-line via the British Library. We found that while we were able to collect sufficient texts for the middle and latter part of the EModE period represented by our corpus, the choice of news texts for the sixteenth century was limited. It is worth noting that our earliest examples of news journalism are somewhat different from PDE newspapers, since the newspaper as a text-type did not emerge until mid-way through the Early Modern period. The first print news periodical did not appear until 1620 (Brownlees 2014: 1), with the first newspaper as we might recognise it (*The London Gazette*) not appearing until 1666 (Raymond 2003: 107, 157). Before then, news was distributed in the form of occasional pamphlets or *relations*, which first started to appear towards the end of the 16th Century and became more frequent in the early 17th Century. Often, pamphlets contained what Brownlees (2014: 13) terms “epistolary news”, or simply copies of letters containing personal accounts of newsworthy events. Therefore, our data is unavoidably not absolutely equivalent to the news data in the Lancaster corpus of discourse presentation in contemporary writing. However, it does afford an opportunity to develop hypotheses as to how the genre develops across the period.

4. Approaches to Early Modern news discourse

Early Modern news reporting has been the focus of studies from a range of disciplines, spanning history, literary studies and sociology. The suggestion in the work of scholars such as Clarke (2004) and Sutherland (1986) is that news reporting practices in the Early Modern period were different from present day practices. Conboy (2007: 6) elaborates on this, claiming for instance that the language of news reporting “reflected the letter-writing style of the time”, a claim that mirrors Brownlees’s (2014: 13) comments about “epistolary news”. However, Conboy neither explains what “the letter-writing style of the time” was or describes how this was reflected in news reports. This lack of specific detail with regard to the

linguistic structure of early news reports is a common issue in work on the development of the newspaper. Consequently, we carried out a pilot project (reported in McIntyre and Walker 2011) which investigated reporting strategies in Early Modern English writing generally (including both news and prose fiction) with the aim of (i) investigating some of the qualitative claims of historians and literary critics and (ii) providing a greater degree of linguistic detail concerning the strategies employed by EModE journalists. We focused on how speech, writing and thought was presented in EModE writing by annotating our 40,000 word corpus of EModE writing using the model of discourse presentation (DP) outlined in Semino and Short (2004). By comparing our findings against Semino and Short's similarly annotated corpus of present day English writing, we were able to determine which categories of DP were significantly over- and under-represented in EModE in statistical terms. Our findings took both fiction and news reports into account, and with regard to the latter we were able to generate a number of hypotheses concerning reporting strategies in the period. Chief among these was that EModE writers tended to favour minimal forms of speech and writing presentation, compared against the present day proclivity for reporting using direct forms of speech and writing. For example, EModE favoured indirect summaries of news reports such as 'Middleton also writes to them out of Holland, that Colonel Dezmond was shipped away ...' as opposed to reporting the exact content of the original discourse. We suggested that this was partly to do with the fact that news was often delivered in letters which needed to be summarised fairly quickly. Using indirect forms allowed (i) the source (and thus authenticity) of the news to be established, and (ii) the news editor's voice, as well as the original letter writer's, to be visible in the report.

Our interest in the development of DP in EModE news journalism is shared by a number of scholars, though none have used the exact methodological framework we employ. Moore (2002), for instance, explores the phenomenon from a non-corpus-based angle. She focuses on Early Modern slander depositions and

suggests that in this text-type reported speech was often presented in English, despite the primary language of the depositions being Latin. Jucker (2006), on the other hand, takes a corpus approach to the issue that is similar to ours. Nonetheless, despite its value as an exploratory study, we would argue that there are a number of issues with Jucker's work that weaken the force of the claims he is able to make. For example, Jucker does not discuss the criteria by which he identified particular DP categories, nor does he provide frequency information concerning the distribution of the different categories in his data. This means that his study is not replicable from the information contained in his article; and this is problematic since it means his findings are not open to falsification. More reliable in methodological terms is Włodarczyk's (2007) investigation of DP in Early Modern English courtroom discourse. However, Włodarczyk's calculation of category frequencies includes both reporting clauses and reported clauses, meaning that it is difficult to separate out the elements of a sentence that are discourse presentation and those that are simply narratorial clauses that introduce the phenomenon. We aimed to address the problems associated with these other studies of discourse presentation by ensuring that the decision-making process for determining categories is as clear and precise as we can make it. We describe the principles that underpin this process in the next section.

5. Annotation

The annotation scheme we used was a development of that described in McIntyre *et al.* (2004) and McIntyre and Walker (2012). The annotation was carried out by hand since no computer software exists that can automatically distinguish the different categories of DP in a text (particularly an EModE text), since, for the most part, DP is not form-based. We used xml-style tags that comprise an element *dptag* and an attribute *cat*. The *cat* attribute consists of up to ten fields (see Table 3) into which pre-designated alphanumeric codes are entered detailing the discourse presentation categories outlined in Table 1.

Field	Possible constituents	Definition of constituent
1	x N F	Narrator's; Narration; Free
2	x P I D	Presentation; Indirect; Direct
3	x S T W V I	Speech; Thought; Writing; Voice; Internal state
4	x A	Act
5	x p	Topic
6	x h	hypothetical
7	x i	inferred
8	x y	discourse summary
9	x 1 2 3 4	numbers = DP split into sections
10	x #	# = odd/interesting cases

Table 2 Constituents of the fields of the *cat* attribute

The possible constituents designated to the first four fields relate to the major DP categories (outlined in Table 1) and are always capital letters. The constituents designated to the remaining fields relate to DP sub-categories and provide further details about the DP. These are generally lower-case letters, but the hash symbol (#) and numbers are also possible in certain fields. Example [1] shows a sample of annotated text.

[1]

```
<dptag cat="N"> But though that general has taken the crown from a head
that was not worthy to wear it, and placed it on his own for good of the
publick, and the honour of that mighty empire, </dptag>
<dptag cat="xIWxxxxx1"> he has, </dptag>
<dptag cat="NPW"> as they write from Vienna, </dptag>
<dptag cat="xIWxxxxx2"> settled the succession </dptag>
```

<dptag cat="N"> so, that after his death the crown shall return to the family again in the person of the young prince and deposed Sophi's son, of whose education he takes particular care. </dptag>

Notice in [1] that an instance of DP can be split into more than one part by, for example, being interspersed with a reporting clause. This situation is indicated in attribute 9 of a tag. We use *x* as a placeholder for empty field positions in the *cat* attribute, but do not mark empty positions following the final attribute value. This means that NT is marked <dptag cat="NxT"> and not <dptag cat="NxTxxxxxxx">. By doing this, *cat* attribute constituents always occur in the same field position, making searches of the annotated corpus for particular DP categories using computer tools more straightforward. Once the corpus was annotated we used Multi-Lingual Corpus Toolkit (MLCT) (Piao 2014) to count the numbers of instances of each DP category. We did that using regular expressions to search for individual DP categories using our tagging format. For example, we searched for Indirect Thought (IT) using the following regular expression:

```
<dptag cat="xIT[a-z1-9]*?">(.*)</dptag>
```

With MLCT, it is also possible to extract the words in between particular tags into separate files. In this way we were able to extract all instances of, say, IT into a file for further analysis using, for example, AntConc (Anthony 2014). Furthermore, by creating files for every category used in our analysis and then checking that the sum of the words in the individual files matched the sum of the word in the original corpus, we were assured that all our data had been accounted for.

Systematically annotating a corpus forces the analyst to be clear about the decisions they make and to be consistent in their application. In the case of annotating for function-based categories this issue is even more important since form cannot necessarily be used as an indicator of what category an item should belong to. This is especially true when annotating for categories of discourse presentation because some categories are not form-based (e.g. free indirect

discourse presentation), while others are realised by a number of different forms, none of which are unique to DP. For example, indirect speech/writing/thought is prototypically signalled by a reporting clause, that introduces the discourse, and a subordinated reported clause that contains the discourse. However, subordination can take numerous different forms and indirect presentation can take numerous non-prototypical forms (see Thompson 1996). We follow Semino and Short (2004) who, based on their findings from the Lancaster SW&TP projects, conclude that indirect speech, thought and writing should include both finite and non-finite reported clauses (Semino and Short 2004: 86). For example in [2], the reporting clause is finite ('he commaunded'), while the reported clause is non-finite ('to speke and shewe').

[2]

he commaunded the sayde Ruge Cros to speke and shewe the seyde werdes
of his message

Also, we included instances where both the reporting and reported clause were non-finite. For example:

[3]

We hear, some mandates are arrived from the imperial court, ordering the landtgrave of Hess d'Armstadt to forbear his hostilities against the Landtgrave of Hesse Homburg.

Again, following Semino and Short (2004), we also counted nominalised forms that introduce reported discourse, for example:

[4]

But whilst these Harbours are so closely watched by the united Squadron of her Majesty and the states, commanded by the Lord Henry Seymour, it is the general Opinion, that his Highness will finde it impossible to put to sea,

In [4], the reported discourse is a relative clause that modifies the NP 'the general Opinon'.

The examples above help to show why no automatic annotation software is yet possible for DP. They also demonstrate that it is crucial to have a principled method for deciding on what category a particular stretch of text belongs to. This is especially true when multiple people are involved in the annotation process because everyone involved must be making reliable and consistent annotation decisions. For this project, two people were involved in the annotation, whereby each text in the corpus was annotated by one person and then checked by the other. Any differences were highlighted and resolved through (sometimes lengthy) discussion. Where decisions could not be reached due to, for example, ambiguities in the text, multiple tags were allocated and a note tag added explaining why. The resolutions to differences were used to improve the annotation criteria for subsequent annotation and in this way, our annotation guide became gradually more sophisticated.

Notice that being clear about how categories are applied is not to say that annotation resulting from such a method is impervious to criticism; but it is retrievable and open to falsification, and that is what makes it valuable. Also, the process of systematically working through data can reveal interesting cases that are not accounted for by the framework or model being applied. For instance, we noticed a number of instances where *hear* implied discourse presentation, as in the following example:

[5]

We hear from Dijon, that they have lately seised above 30 high-way men in those parts: Abundance of murders and robberies are also committed in the

neighbourhood of this city, especially in the forests of Villers, Gotrets and senlis, where some days ago five persons were found murdered. We have great numbers of bankrupts here of late, one within this 5 or 6 days is broke for 500000 livres.

Our view for this case was not to tag *hear* + [CONTENT] as DP. Instead, we marked it using the narration (N) constituent in field 2 of the tag attribute and added the 'summary' constituent in field 8, so that all instances can be retrieved if we need to return to this issue at a later date. The decision taken here is a pragmatic one: our project is interested in major DP categories, and while there are still interesting and unresolved issues concerning DP, these were beyond the scope of the project. In addition to this, it would make sense to investigate such issues using a corpus of contemporary written English first, before investigating them in EModE.

6. Results and discussion of thought presentation.

Chart 1 shows the percentages of discourse presentation categories in the PDE and EModE news corpora. In PDE news, nearly 90% of the discourse presentation is speech, whereas in EModE, the amount of speech presentation is around 55% of the total DP. However, in our sample of EModE news, around 15% of the DP is writing presentation, and almost 30% is thought presentation. This is almost three times more than the amount in the PDE news corpus. The difference in percentage totals for writing presentation is, to a large extent, due to news writers relying heavily on letters and publications from abroad, which they summarised or re-reported, often using reporting clauses (e.g. "the port letters say") that clearly identify the source of the news (see McIntyre and Walker 2011).

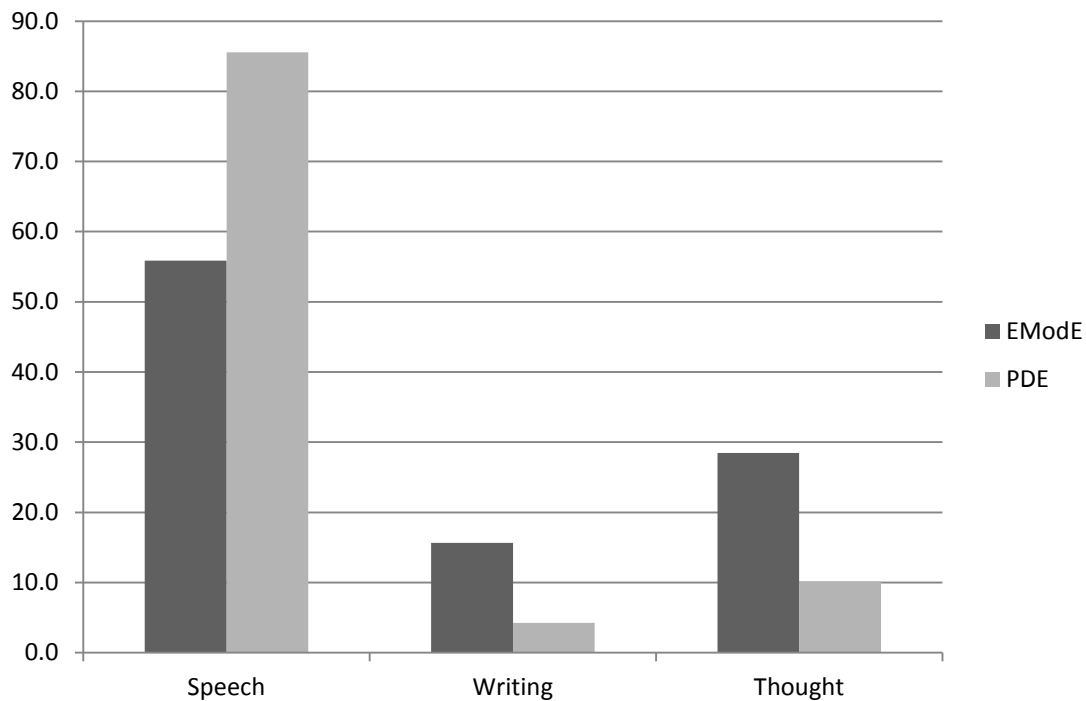


Chart 1. A comparison of the percentage composition of DP in PDE and EModE news.

All the differences shown in Chart 1 are statistically significant ($p < 0.001$, or LL critical value of 10.83). Our focus in this chapter, though, is thought presentation, and Chart 2 shows the categories in the thought cline in more detail in comparison to the PDE data. It shows that there is around twice as much NI, just over seven times more NRTA and over eight times more IT in EModE news than there is in PDE news. These differences in frequencies are statistically significant ($p < 0.001$, or LL critical value of 10.83). For the remainder of this section, we discuss in more detail examples from the two most significant categories, indirect thought and internal narration.

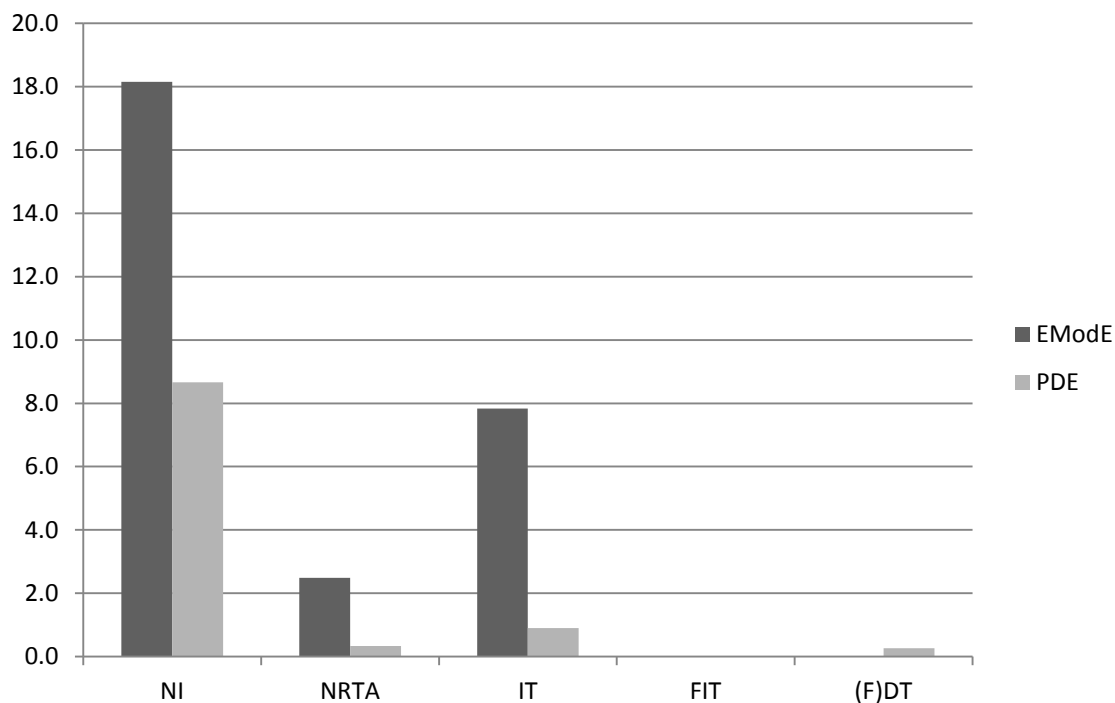


Chart 2. Comparison between the percentages of individual thought presentation categories in PDE and EModE news.

Given that news reports, unlike works of fiction, report real events and the activities and interactions of real people, it is surprising that there is any report of thoughts at all, because, unlike fictional narrators, news reporters and editors do not have access to the minds of the people they are writing about. This does not go unnoticed by Semino and Short (2004: 138) who also find thought presentation in news report, and comment that any such thought presentation “[...] must logically be based on inferences, rather than direct access”. Consequently, as with Semino and Short’s (2004) work, much of the thought presentation tagged in our corpus includes the constituent ‘i’ in field 7 of the tag to indicate that the thought must be inferred (see Table 2). Key to DP being tagged as thought are the reporting verbs which include *trusting*, *thought*, *seeming*, *believed*, *feeling*, *doubting*, *expected*, *supposed*, and *reckon’d*. Since most presented thought is necessarily inferred, we

turn now to the functional effects of the stylistic choices involved in thought presentation in our news data, drawing together examples to make some generalisations which will serve as hypotheses for further investigation.

6.1 Thoughts as news summary or general opinion

Producers of print news in England during the Early Modern period often heavily relied on material from abroad, such as letters and publications from Holland and Germany known as *corantos*. News also came via the embassy in London, where according to Atherton (quoted in Raymond 1999: 44), the diplomatic staff of the time were responsible for ‘... collecting, sifting and recirculating news’ The writing of newsletters often involved summarizing such sources. A number of instances of reported thought found in our corpus reflect that summary process. For example:

[6]

'Tis reckon'd that the King's journey to Holstein will not go forward till the new Year, and that the Queen Dowager will then remove into her palace which she has bought.

London post with intelligence foreign and domestic 1700

The indirect thought used in [6] represents a form of discourse summary (Short 2012). We noted in McIntyre and Walker (2012) that the use of indirect writing allows (potentially) for the reconstruction of the original written words, which could be connected to news producers showing faithfulness to the original sources for reason of demonstrating a truthful rendering of the news. However, this cannot be said of indirect thought because, unlike writing, it is not possible to access the original anterior discourse. Therefore, there is an issue with regard to faithfulness claims when presenting thought. In addition, we can note that the use of an agentless passive means that the original source cannot be identified anyway.

Having identified a number of instances of IT being presented in passive constructions, we were interested to know how widespread this practice was across the EModE period generally. Here, however, is where the disadvantages of a small corpus are keenly felt. Manual annotation is labour intensive and costly, meaning that it is difficult to tag large samples of text. However, it is possible to use the interpretative insights generated through the analysis of a small tagged corpus, such as our EModE sample, to target a search of a larger, unannotated corpus. To this end, we used a corpus of just over 500,000 words of EModE news report covering the years 1620-1720 in order to investigate this phenomenon further. This larger corpus was originally constructed for another project that investigated modality in EModE news, and while it was annotated for modal function, it was not tagged for DP. Based on the findings from our smaller, intensively tagged corpus, we used AntConc (Anthony 2014) to search for instances of *it is*, *it's* and *tis*, the results of which we filtered to include only those involved in passive constructions. We found that there were numerous instances of passive constructions consisting of [it is / it's / tis] + a past tense verb of cognition. Examples include:

[7]

It is beleevd now, that his Majesty will continue longer there than was supposed

It is coniectured, and is also very probable that all Sauoy will fall vnder the obedience of the French King

It is feared that irruption may cause some obstruction

it is hoped that they lye now about Norway

It is supposed that Tilly hath his ayme at Saxony

It is thought that they would set vpon Hulst,

it's believed the rest will do the like.

'tis hoped his Brother will suddenly do the like

This grammatical pattern extends to other discourse presentation categories (for example, “it is rumoured ...” and “It is said ...”) and persists across the full breadth of the corpus. The use of *'tis* in these constructions increases in the latter part of our corpus, with over half the instances occurring in the last two decades. Thus, we found further evidence to suggest that the passive construction is a stylistic feature of EModE news report that omits the source of the original depiction of events.

The use of cognition verbs, though, could be related to the original content of the letters (i.e. writing) being summarised as a general thought. It is also possible that they reflect or repeat cognition verbs used by the original writers, particularly if they were personal letters containing the writer’s own opinions. However, the agentless passive construction means that it is not possible to recover who is thinking these general opinions, or where these opinions come from. This, again, can be related to the summary nature of these reports, but could also be a way of reporting opinion anonymously, or in a way that shifts responsibility for any such opinions away from the news editor. This latter possibility might have been desirable for news producers, given the harsh penalties imposed during some parts of the Early Modern period for dissenting voices in news reporting. This was, after all, a period in which press freedom was not yet established.

6.2 Thoughts as propaganda

Other instances of indirect thought reflect that some news consisted of letters containing eye-witness accounts of events, and it is the eye-witness that is making generalisations. These might be for reasons of propaganda, particularly, during the English Civil War, as in [8]. This is from a report about the relief of Gloucester, which

had been placed under siege by Roylist forces (the enemy in this account), by a large Parliamentary force that marched from London:

[8]

These two poore regiments were the very objects of the enemies battery that day, and they have since made their boast of it. It is conjectured by most, that the enemy lost four for one

A true and exact RELATION of the Marchings of the Two Regiments 1643

The thoughts being reported here are a convenient way to add credibility to the account, even though the ratio presented is merely conjecture. Notice again that the reporting clause that introduces the thoughts of 'most' is a passive construction containing the present tense singular form of BE, which suggests ongoing and continuous action (or conjecturing). It does not necessarily describe a situation where a posterior discourse is reported in an anterior discourse, because there does not seem to be a posterior discourse. It nevertheless purports to present the thoughts of others, and thus counts as discourse presentation.

Other instances of thought presentation for purposes of propaganda relate to the papist plot. In particular, one report of the execution of eight 'papists' involved in the Gunpowder Plot has a clear protestant ideological stance, and suggests, through indirect thought presentation (underlined), the manner in which Catholics should be viewed by people.

[9]

the following account is written of the carrage of the eight papists herein named, of their little show of sorrow, their usage in prison, and their obstinacy to their end. First for their offence--it is so odious in the ears of all human creatures that it could hardly be believed that so many monsters in nature should carry the shapes of men

The weekly newes numb. 19 London 1606

The ideological stance in this writing is unashamedly pro-protestant, and intended “to dissuade the idolatrously blind from seeking their own destruction.” Casting Catholics as ‘monsters’ was part of that strategy and would have been viewed favourably by those with authority over the press.

6.3 Thoughts as a fictionalising of news stories

Looking now at instances of NI, or Internal Narration, this category occurs in all texts in our corpus, and captures points in the news report where the reader is given an insight into the mind or the emotional state of a person or persons that is/are being talked about. As we mentioned in section 2, this category is contentious because it can either be seen as the most minimal form of thought presentation, or as narration (therefore not DP). However, that contention aside, it seems rather odd that the news writer should report what was happening in the mind of another person, because there is no way that this could be known.

The news-sheet titled *A True Designe Of The Late Eruption Of Movnt Ætna In Sicily*, contains a letter written by the “Right Honourable the Earl of Winchilsea”, which contains a relation of the events given to the Earl by the Bishop of Catalania. In this relation the author, whoever that is, seems to be at pains to tell the reader about the fear of the people caught up in the eruption and earthquake.

[10]

But on Friday the 18/8th of March 1669 the Sun was observed before its setting to appear of a pale and dead colour, which (being contrary to what it ever before appeared to us) struck no small terror into the inhabitants, [...]. The same night happened in this City as well as the whole Countrey hereabouts, a terrible and unusual Earthquake, whose strong and unequal

motions, joynd with horrible Roarings from Monte Gibello exceedingly
frighted the Inhabitants,

He goes on to describe “distressed People” who look on “with grief and astonishment” and flee the town “with much trouble and amazement”. While such emotional reactions to a life threatening natural disaster are what one would expect, the writer is, nevertheless, reporting what he cannot possibly know for sure, only what he can infer from, say, people’s facial expressions or vocalizations (assuming, of course, that the writer was even there). This “true relation” therefore contains elements that are subjective, but add dramatic effect. In effect, this is a fictionalisation of real events.

In [11] (*A Courante of news from the East India, 1622*), the news is presented as an eye-witness account.

[11]

the Dutch Generall threatned to doe the like unto Polaroone wherefore our principall Factor of Polaroone being there but newly arrived, went unto the Dutch Generall unto the castle of Nera, and told him that he heard that he purposed to take Polaroone by force, which hee could not beleeve, although his owne messenger sent to Polaroone to speake with our Factors, had given it out also, that the Generall himselfe should say, that if the English did not come presently unto him & yeeld the sayd fort of Polaroone unto him, he would send his forces, and over-run all the countrey,

This example begins to demonstrate some of the complexities of the discourse presentation in the news data, where discourse presentation can be embedded in other DP. This is partly due to the nature of news gathering at that time. This news is advertised as “A true relation of the taking of the Ilands of Lantore and Polaroone in the parts of Banda in the East Indies by the Hollanders, which Ilands had yeilded

themselves subject unto the King of England. Written to the East India company in England from their factors there.” This is, therefore, a reprint of a report in a letter written by employees of the East India Company. While there is little doubt that these events in the East Indies (*Indonesia*) occurred, and the news sheet promises a true relation of them, it nevertheless reports the intentions of a Dutch General. These thoughts are embedded in the reported speech of an official of Polaroone (*Pulau Run*). In both instances, the faithfulness to the original seems likely to be highly dubious. The writer of the report would have needed to be present when the Factor of Polaroone spoke to the General. Even if he was there, he would have needed to take very careful notes very speedily in order to record what was said, in order to report it faithfully. Therefore, there appears to be some flexibility about what counts as true, but the eye-witness nature of the account suggests faithfulness to the depiction of the events, while including DP that can never be verified and might never have happened that help with the telling of the story.

7. Conclusion

Our aim in this chapter has been to shed light on the nature of Early Modern news reporting by using a corpus of news data annotated for discourse presentation to explore the stylistic tendencies of writers of the time. In particular, we have suggested some of the functions of the most significant thought presentation categories deployed by news writers. By carrying out more detailed qualitative analyses of the quantitative patterns on both our tagged corpus and another larger, untagged one we have shown how thought presentation in our data is used for a number of purposes, including summary, propaganda and fictionalisation. The manual annotation of our corpus for categories of discourse presentation allowed us to compare EModE journalistic practices with those of the present day, and also to identify the relationship between particular stylistic forms and functions. We also hope to have shown that, in methodological terms, a small, intensively tagged corpus can be used to discover patterns and generate hypotheses that can then be

explored in a larger untagged dataset. This is a potentially useful method of procedure for anyone working with categories that are not form-based, since it is often not practically possible to annotate large corpora for function-based categories. While small corpora may not be able to produce generalisable findings, their methodological value lies in their capacity to generate robust hypotheses for testing against larger datasets.

References

- Anthony, L. (2014) *AntConc* 3.2.4w. Available at: <http://www.laurenceanthony.net/software.html>.
- Brownlees, N. (2014) *The Language of Periodical News in Seventeenth-Century England*. Newcastle upon Tyne: Cambridge Scholars Publishing.
- Burnley, D. (1992) *The History of the English Language: A Source Book*. London: Longman.
- Clarke, B. (2004) *From Grub Street to Fleet Street*. London: Ashgate.
- Conboy, M. (2007) *The Language of the News*. London: Routledge.
- Crystal, D. (2005). *The Stories of English*. London: Penguin.
- Jucker, A. H. (2006) "'But 'tis believed that...': speech and thought presentation in Early English newspapers", in Brownlees, N. (ed) *News Discourse in Early Modern Britain: Selected Papers of CHINED 2004*, pp. 105-25 Bern: Peter Lang.
- Frank, J. (1961) *The Beginnings of the English Newspaper 1620-1660*. Harvard University Press: Cambridge, Massachusetts.
- Leech, G. and Short, M. (1981). *Style in Fiction*. London: Longman.
- McIntyre, D. and Walker, B. (2011) 'A corpus based approach to discourse presentation in Early Modern English writing: A preliminary corpus-based investigation', *International Journal of Corpus Linguistics* 16(1): 101-30.
- McIntyre, D. and Walker, B. (2012) 'Annotating a corpus of Early Modern English writing for categories of discourse presentation', in Manzano, F. (ed.) *Unité et Diversité de la linguistique*, pp. 87-107. Les Cahiers du Centre d'Etudes

- Linguistiques. Lyon: Atelier intégré de publication de l'Université Jean Moulin – Lyon 3.
- Moore, C. (2002) 'Reporting direct speech in Early Modern slander depositions', in Minkova, D. and Stockwell, R. (eds) *Studies in the History of the English Language: A Millennial Perspective*, pp. 399-416. Berlin: Mouton de Gruyter.
- Piao, S. (2014) *Multilingual Corpus Toolkit*.
<https://sites.google.com/site/scottpiaosite/software/mlct> Accessed 20/11/2014.
- Raymond, J. (2003) *Pamphlets and Pamphleteering in Early Modern Britain*. Cambridge: Cambridge University Press.
- Semino, E. and Short, M. (2004) *Corpus Stylistics: Speech, Writing and Thought Presentation in a Corpus of English Writing*. London: Routledge.
- Short, M., Semino, E. and Wynne, M. (2002) 'Revisiting the notion of faithfulness in discourse presentation using a corpus approach', *Language and Literature* 11(4): 325-55.
- Short, M. (2007) 'Thought presentation 25 years on', *Style*, 41 (2), 227-257.
- Short, M. (2012) 'Discourse Presentation of Speech (and writing but not Thought) Summary.' *Language and Literature*, 21(1): 18-32
- Sutherland, J. (1986) *The Restoration Newspaper and its Development*. Cambridge: Cambridge University Press.
- Thompson, P. (1996) 'Voices in the Text: Discourse Perspectives on Language Reports', *Applied Linguistics*, 17 (4): 501-530
- Toolan, M. (2001) *Narrative: A Critical Linguistic Introduction (2nd Ed.)*. New York: Routledge
- Włodarczyk, M. (2007) *Pragmatic Aspects of Reported Speech: The Case of Early Modern English Courtroom Discourse*. Frankfurt: Peter Lang.