

Revisiting the Notion of Faithfulness in Discourse Report/(Re)presentation Using a Corpus Approach¹

*Mick Short and Elena Semino, Lancaster University and Martin Wynne,
Institut für Deutsche Sprache, Mannheim*

Abstract

Of late a number of theorists have argued that the notion of faithfulness to an original should be abandoned when trying to account for direct speech presentation in particular and hence, by extrapolation, the rest of the scale(s) of discourse (re)presentation. This paper discusses critically this increasingly fashionable view, and incorporates the observations which have prompted it in an argument for a more sophisticated, pragmatically based account of the notion of faithfulness, and its relation to the various clines of discourse report/(re)presentation and their categories. Our account relies on the results of a corpus-based approach to the study of discourse (re)presentation and a qualitative analysis of a set of newspaper articles on a particular news story from outside this corpus. Also important is the development of an understanding of discourse (re)presentation which is based on a clear and consistent application of the various descriptive terms which have been applied to it during the 20th century, in particular (a) 'discourse', (b) 'speech', 'thought' and 'writing'; and (c) 'report', 'presentation' and 'representation'.

Keywords: direct discourse, discourse presentation, discourse report, discourse representation, faithfulness, indirect discourse, speech, thought, writing, verbatim quotation

1 Introduction

This paper is concerned with the notion of faithfulness to an original in the representation of other people's words and thoughts. The issues we are going to discuss are best introduced by means of an example. The extract below is taken from a newspaper article discussing the closure in 1994 of a newly-opened children's theme park in Bare, a suburb of a seaside resort in Northern England called Morecambe:

A woman walking her Yorkshire terrier along the empty paths said: 'I'm very sad. This place would have brought a lot of people to Morecambe.' The town's newspaper said: 'Imagine how we look now. Morecambe should hang its head in shame' (*Independent on Sunday*, 'Bare ladies' protest puts end to Crinkley Bottom', 4/12/1994).

The extract contains two instances of what is traditionally known as 'direct speech'. The first ('I'm very sad. This place would have brought a lot of people to Morecambe.') relates to the words *uttered* by an unnamed local resident, whereas the second one ('Imagine how we look now. Morecambe should hang its head in shame.') reports the *written* words of what is described as 'the town's newspaper'. Following the most traditional view of this kind of reporting, one would have to conclude that, because quotation marks are used in both cases, both stretches of text are quotations: that is, faithful verbatim reproductions of their respective originals. On the other hand,

some recent radical revisions of the traditional view (e.g. Fludernik 1993; Sternberg 1982a, 1982b; Tannen 1989) would object that, because direct speech can be used when faithful reproduction is not possible (e.g. in casual conversation), its presence can never be associated with faithfulness to an original. Within this view, neither of the stretches within quotation marks in the above extract can be seen as verbatim reproductions of the original texts, but just as attempts on the part of the reporter to dramatise the voices of different participants in the debate over the closure of the theme park.

In this paper we argue that there is a middle way between these two positions, where the focus is on the many different factors that affect the possibility or necessity to be faithful to an original stretch of discourse when reporting it. Within our approach, the two instances of 'direct speech' in the example above are actually quite different from each other. The first one is indeed a report of speech in direct speech (DS) form, although the source is an unnamed member of the public, whose *exact* original words are unrecoverable and are less important for the article than the general attitude that is attributed to her. On the other hand, the second instance is best described as 'direct writing' (DW), since the source is in the written medium, a newspaper which, although unnamed, is identifiable to at least part of the readers, who could in principle check the accuracy of the quotation. This checkability will lead to stronger expectations on the part of readers that the quotation will be verbatim. Whereas the first reported extract could never cause the reporter legal problems, it is possible to imagine a situation where an inaccuracy in the second quotation may result in a court case.

In general terms, then, we wish to take account of, and try systematically to describe the variety and complexity of the phenomenon which is made to seem deceptively simple when it is referred to as 'discourse report'. We agree with Clark and Gerrig (1990) that the verbatim assumption for direct discourse arises from what Linnell (1982) has called the 'written language bias' of linguistics and philosophy, and we are supportive of the present attempts to show in detail how direct presentations of discourse in spoken language are signalled and function (e.g. Tannen 1989, Clark and Gerrig 1990, McCarthy 1998, Myers 1999). But to notice that the verbatim assumption does not apply in some forms and contexts of direct discourse report does not imply that it should be abandoned in *all* forms and contexts, as Sternberg (1982a, 1982b), Tannen (1989) and Fludernik (1993) suggest. Our position is essentially similar to that of Clark and Gerrig in the quotation below (though we are unsure as to their exact final position and its consequences (see our discussion below) and so do not want, for present purposes, to assume that their position is the same as the more straightforward views of Fludernik, Sternberg and Tannen). Clark and Gerrig point out that whereas indirect report describes what was said, direct report *demonstrates* it:

Suppose you want to represent a person's locutionary acts verbatim. If you use indirect quotation, your addressee cannot be sure what sentence really was uttered. The mapping from your words to your source sentence is undecidable. What you need is direct quotation. Even that isn't enough, since the locutionary act need not be a depictive aspect of the quotation. You must also say or imply that your wording is verbatim. It is CONVENTIONALLY implied that the wording is verbatim in newspapers, law courts, and literary essays, but, as we later argue, not elsewhere (Clark & Gerrig 1990:792).

One of the reasons that Clark and Gerrig use the word ‘conventionally’ above is because they acknowledge that verbatim quotations do not have to be an *exact* mimicry of an original locution *in all respects*. They point out, for example, an American newspaper report, in which direct speech is used for an English quotation from the Pope which, from the context, must have originally been in Italian. Clearly this is not complete mimicry of the original, otherwise it would have had to be reproduced in a language which most of the newspaper’s readers would not have been able to understand. But it is also clear that the writer was claiming a greater faithfulness to the original than an indirect speech version would have suggested, and that the extent of the claim is made clear contextually. It can be seen from this example, as well as from that in relation to the Bare ladies above, that the notion of verbatim quotation is complex and, in varying degrees, context-dependent. And, as Clark and Gerrig say at the end of the above quotation, the assumption of a verbatim quality in a report may well vary from one situation to another. We try to spell out some aspects of this complexity towards the end of our paper.

Clark and Gerrig (1990) point out the kinds of issues which Fludernik, Sternberg and Tannen have noted in promoting their anti-faithfulness position, and, indeed, in general rhetorical terms their article does appear to push in the same direction as the work of Fludernik, Sternberg and Tannen. But they are rather more careful to be less simplistic and direct in their discussion. For example, they point out that there are unresolved issues over what exactly counts as verbatim quotation, and in their conclusion they point out that ‘speakers *aren’t necessarily committed* to trying to reproduce a source utterance verbatim’ (our italics), which seems to admit the possibility that on some occasions they might be so committed. In some ways, then, Clark and Gerrig appear to leave the door open to the more complex view of faithfulness that we want to argue in favour of. Moreover, they propose that we should see quotation as a kind of demonstration (or language-mimicry). This proposal needs separate careful consideration, and it is arguable that demonstration itself presupposes faithfulness to an original in some form. Because it is difficult to be sure exactly where Clark and Gerrig (1990) stand on the proposal to do away with faithfulness as a criterion in theories of discourse (re)presentation we will not discuss their paper further in this article.

Below we begin by reviewing the debate over faithfulness in reporting, indicating the problems we have in general terms with the various anti-faithfulness arguments which have so far been put forward. We go on to explain why we think that the notion of faithfulness is important for discourse (re)presentation theory, and make what we believe are some crucial terminological and conceptual distinctions which are necessary to make sense of this very complex area. Next we introduce a corpus-based study of discourse representation in narrative texts in order to show how its results contribute (in both quantitative and qualitative terms) to the faithfulness debate, and then examine faithfulness in a series of DW representations of a particular newspaper article in a series of news reports which followed it. We finish by proposing the outlines of a context-sensitive approach to faithfulness in reporting.

2 The ‘anti-faithfulness’ approach to discourse representation: the baby and the bath-water

In recent years, a number of scholars have questioned the notion that different modes of discourse presentation can be differentiated on the basis of their faithfulness claims. In particular, there have been several attacks on the idea that the direct forms, and direct speech in particular, can be defined in terms of the expectation that they faithfully reproduce an original. The following two quotations give a representative flavour of the anti-faithfulness view:

. . . we have here an instance of the 'direct discourse fallacy' that Meir Sternberg (1982a, 1982b) has so admirably exposed, i.e. the mistaken (ingrained) belief that direct discourse is in every sense of the word *primary* and *originary* to the other types of quotation. As Sternberg has shown, this is at best a convention (in the real world as well as in fiction), and it is a convention which fiction in fact frequently subverts, either by inventing words which were not actually spoken or by distorting what was spoken to serve the ulterior ends of the narrative . . . (Fludernik 1993: 281)

. . . the term 'reported speech' is a misnomer, an abstraction with no basis in the reality of interaction. When speakers cast the words of others in dialogue, they are not reporting so much as constructing dialogue (Tannen 1989: 133).

At least four different kinds of argument have been put forward in support of this 'anti-faithfulness' approach to discourse representation (Fludernik 1993: 22). Below we sketch out each of these arguments in turn and indicate why, in general terms, we do not feel that the arguments are decisive.

Firstly, it has been pointed out that the direct forms can be used even in cases where there is no such thing as a straightforward original which can be faithfully reproduced. This is the case, for example, when one refers to a hypothetical utterance (e.g. If she was here, she would say 'Over my dead body!'), when a person's non-verbal behaviour is 'translated' into verbal form (e.g. The expression on his face said 'What the hell is this?'), when the reporter spells out what is implied by someone else's speech (e.g. What she really meant was 'Get lost!'), when what is reported relates to multiple originals (e.g. My students are always saying 'The books are not in the library. '), when the original utterance is in a different language from that used by the reporter, and so on (see Fludernik 1993: 409-414; Sternberg 1982a, 1982b; Tannen 1989: 110-119). We accept that this kind of thing occurs (for a detailed discussion of our approach to hypothetical discourse see Semino, Short and Wynne forthcoming). But we do not see that it necessitates the dismissal of the criterion of faithfulness in direct discourse report (a) because such cases only represent a small minority of total cases (see the quantitative evidence from our corpus presented in section 6 below) and (b) because this minority of cases can be identified as a separate and distinct category as a result of contextual cues which induce the reader to cancel the conventional assumption that the DS form reports verbatim a specific stretch of anterior discourse. In other words, it is clear to both addresser and addressee that *presentation* (what Tannen above calls 'constructing dialogue'), but not *report* is involved.

Secondly, it has been noted that the direct forms may contain expressions which explicitly suggest that the reporter is missing out some elements of the original (e.g. 'so-and-so', 'what's-his-name', 'something or other'), or may be accompanied by

expressions which undermine the reliability of the quotation (e.g. She shouted something like 'I am not prepared to put up with this'). Our objections to this kind of 'counter-example' are in line with what we said about hypothetical utterances, namely (i) that the proportion of such cases is very small and (ii) it is almost always made clear contextually that the claim of making a verbatim report is not being made, but merely, in this case, the claim of *presenting* a rough approximation. Hence this kind of 'counter-example' is not really relevant to the faithfulness issue.

Thirdly, much has been made of the many features of original utterances which are lost in any report, including direct report, especially when the anterior discourse is in one medium and the posterior discourse in another. So, for example, the use of DS in written texts inevitably fails to represent a range of aspects of spoken utterances, such as voice quality, accent, intonation, paralinguistic features, and so on. Similarly, it may not be possible, particularly in spoken contexts, to reproduce exactly all the characteristics of a written original, including layout, typeface and orthographic features (Sternberg 1982a: 134, 1982b: 77-83). But in our view what this kind of example shows is that not all aspects of an original utterance are equally important in assessing faithfulness to an original, and different aspects of the original will not be equally important in different reporting contexts.

It is only with the modern technologies of the computer, the tape recorder and the photocopying machine that it has become possible to reproduce exact versions of an original when reproducing it (and indeed it is arguable that even here variations occur in the copying process). Even if the discourse medium is held constant in the anterior and posterior contexts, variations are allowable. Hence, if someone re-says exactly what someone else said we would not expect the voice quality to be the same, simply because we know that the original speaker and the reporter are not the same person. Similarly, no-one would quibble with a direct writing representation of something in the Times font being reproduced in the New York or Chicago font, except in very particular circumstances. This is because we all automatically make adjustments in accordance with the type/token distinction when we compare one thing with another. Similarly, for discourse reports which are in a different medium from the original, we would not expect an accurate representation of features which cannot survive the medium change.² What can be common to speech and writing are the words and grammatical structures used to produce some propositional structure, and so it is not surprising that more carefully-worded distinctions between the direct and indirect forms refer specifically to these two linguistic levels.

Last but not least, it has been pointed out that in conversational contexts, where speech is the reporting medium, it is simply impossible, due to the limitations of human memory, that speakers reproduce previous utterances word-for-word, even when they adopt a form of reporting which is the oral equivalent of direct speech presentation (Fludernik 1993: 414ff.). Our reaction here, as before, is to accept the evidence but to give it a different weighting in the theoretical discussion. Speech reports of previous speech in very informal settings, where the need to be accurate is usually low, are the most likely direct discourse representations not to be accurate verbatim reports, especially where the speech reported is also very distant in time, so that memory limitations are at their most severe. This clearly argues for the faithfulness criterion being relaxed in such cases in favour of presentational factors like the need for

vividness in presentation. But to see whether the faithfulness criterion should be abandoned across all the discourse presentation (speech, thought, writing) scales we would first need to see whether similar non-faithful behaviour was general in *all* cases of direct discourse. Sternberg, Tannen and Fludernik have, not unreasonably, chosen the most convenient kind of examples for their argument. However, the situation in which one would expect the faithfulness criterion to be at its strongest, and so constitute the best test for their view, would be in *formal written reports of writing* using the direct presentational form. If faithfulness is problematic to any significant degree in this sort of context then the ‘out with faithfulness’ position would have considerably more force. We examine an example of this kind of context in section 7 below, and do not find much in the way of evidence to support the anti-faithfulness view.

In summary, then, an alternative to the anti-faithfulness view is that *sometimes* people *report* and *sometimes* they merely *present*, and that it is usually clear contextually which activity they are involved in. Under this view, when they are reporting (which, by definition, also involves presenting) the criterion of faithfulness would still apply, but when they are merely presenting, it would not. However, the *effects* associated with the varying faithfulness claims and expectations for different categories of report along the (re)presentational scales (e.g. vividness for DS vs. measuredness for IS) will still be perceivable by receivers and so manipulable by language producers though analogy with the effects of particular presentational categories in situations of report. Thus, while we do not aim to propose a ‘naive’ notion of faithfulness, we do feel that, by eliminating completely any idea of reproduction of an original, some recent studies have, as it were, thrown the baby out with the bath-water. In the rest of this paper we argue for a more balanced, context-sensitive view of faithfulness (see sections 6-8 below and Short *et al.* 1999). Our argument is partly based on the findings of a corpus-based study of discourse presentation in narrative which we have been involved in since 1995 (see sections 5 and 6 below and, for example, Short *et al.* 1996, Semino *et al.* 1997, Wynne *et al.* 1998). First, however, it will be helpful to make it clear why the criterion of faithfulness is important (section 3) and to draw attention to some terminological and conceptual confusions, some of which have been alluded to above, which have arisen in the discussion of discourse presentation theory (section 4).

3 Why is Faithfulness Such a Big Deal?

It may first be helpful to spell out clearly the traditional account of discourse report. This assumes that the different reporting categories have identifiable linguistic forms which are associated with different sets of faithfulness claims, which increase as one moves along the report scale from the most indirect towards the direct end (see, for example Leech and Short 1981 and Short 1996). It is easiest to see these varying faithfulness claims in written reports of writing (see also section 5 below):

Table 1 - Writing (Re)presentation Categories and Faithfulness Claims

FORM	FAITHFULNESS CLAIM	EXAMPLE
Narrator's report of writing (NW)	Writing occurred.	She wrote furiously.
Narrator's representation of a writing act (NwRA(T))	Writing occurred + indication of the writing act involved (+ optional indication of the topic).	She wrote her letter of resignation (from her position as chair of the committee).
Indirect Writing (IW)	Writing occurred + indication of the writing act involved (+ optional indication of the topic) + propositional content of the original in the words of the reporter.	She wrote that because the committee had passed a motion concerning deer hunting which contradicted her own beliefs she was resigning.
Direct Writing (DW)	Writing occurred + indication of the writing act involved (+ optional indication of the topic) + propositional content of the original in the words of the reporter + the words and structures used to produce that propositional content.	She wrote, 'Dear committee members, the committee has today passed a motion concerning deer hunting which is contrary to my own deeply-held beliefs, and so I hereby resign as chair.'

The free indirect writing (FIW) form, which is ambiguous between the direct and indirect faithfulness claims is then situated between IW and DW on the above scale. The free direct form does not appear to have an additional faithfulness claim associated with it, and so it is not included above. For this reason we have argued elsewhere (Short 1988 and Semino, Short and Culpeper 1997) that free direct speech (FDS) and free direct writing (FDW) would perhaps best be seen as sub-variants of DS or DW rather than being categories in their own right. We have not yet tested out this suggestion with respect to thought presentation.

Given that the 'anti-faithfulness' school of thought in the study of discourse report/(re)presentation has demonstrated that faithfulness in direct discourse report does not always apply, particularly in 'hypothetical' and speech-reporting-speech

situations, the reader may legitimately ask why we are so keen on hanging on to the notion of faithfulness.

- Firstly, outside linguistics, stylistics and critical discourse analysis considerable store is set by the differing faithfulness claims of direct and indirect speech and writing. For example, (i) the laws of libel and slander are used precisely in cases where the exact words which someone uses, or is reported to have used are crucial, (ii) in educational and academic contexts unacknowledged use of what others have written or said constitutes the offence of plagiarism (and unacknowledged direct quotation is more heinous than unacknowledged indirect forms), (iii) in academic contexts, to misquote someone in order better to attack their position would be seen as a grave offence and (iv) newspaper style books, manuals of style and grammars for lay people often make much of the faithfulness distinction between direct and indirect speech/writing report. In American newspapers, unlike those in the UK, free indirect speech (FIS) is not allowable as a form to be used by reporters, precisely because there is ambiguity over which words belong to the person quoted and which to the reporter. We have also reported in Short *et al.* (1999) evidence of the fact that writers of newspaper articles, although not entirely consistent in their quoting behaviour do, by and large, try to quote accurately and to acknowledge direct quotation even when their source is an unattributed press release. All these kinds of behaviour would be difficult, and perhaps impossible, to describe and explain without recourse to the notion of faithfulness.
- Secondly, although it is technically possible to make the various distinctions among the speech, thought and writing (re)presentational categories on formal linguistic grounds alone (along with some additional contextual information in some cases), if we only take linguistic factors into account we have no *reason* for making the linguistic distinctions that we do. This is parallel to other uses of the form/function distinction in linguistic description. For example, the grammatical distinction between nouns and verbs can be made on purely formal grounds, but the distinction is of interest to most linguists because of the different semantic roles and functions which the different grammatical forms have when they are used (this point would not be disputed, we believe, by the ‘anti-faithfulness’ scholars). The reason for the various discourse (re)presentation category distinctions, in our view, relates systematically to the varying prototypical faithfulness claims of the different (re)presentational forms (see Semino, Short and Culpeper 1997 for a discussion of this matter).
- Thirdly, we need an explanation for how the presentational *effects* associated with the various categories come about. For example, the presentational qualities of directness, vividness etc. of direct speech and the more reserved and measured qualities of indirect speech don’t come from nowhere. They themselves depend on decisions made by reporters in prototypical cases to be more, or less, faithful to the original, and consequently to be less or more integrative when incorporating a stretch of anterior discourse into their posterior discourse. In other words, while the ‘anti-faithfulness’ school wishes, reasonably, to promote the presentational aspects of the discourse presentational categories, these presentational aspects are themselves determined to a large degree by the very faithfulness assumptions which they want us to abandon.

In general terms, in assessing arguments for and against faithfulness in discourse (re)presentation the status of the counter-examples we take notice of is important. In qualitative terms, counter-examples need to be relevant. In our view, many of the hypothetical 'counter-examples' discussed in 2 above fail on these qualitative grounds. The same is true of the observation often made that IS reports may contain words produced in the anterior discourse situation. Although this is true, the observation is strictly irrelevant to the issue, as IS does not make a faithfulness claim with respect to the words and structures used. Finally, counter-examples also need to be weighed against statistical trends in order for their representativeness to be assessed. We will report relevant statistical evidence from our corpus in section 6.

4 Some useful terminological and conceptual distinctions

It will be clear from the discussion in sections 2 and 3 that in order to arrive at well-balanced conclusions, it is necessary to be very clear concerning exactly what we are talking about. Unfortunately, the discussion of discourse (re)presentation has suffered from a number of important terminological and conceptual confusions. In this section we try to make clear some of these issues in order to create a firm conceptual base for what we want to say later.

1 Discourse

The first thing which we need to beware of is the increasing fashionability of the term 'discourse' across a range of areas of study (e.g. stylistics, critical discourse analysis, literary theory). This has resulted in a range of possible meanings for the term which are not always carefully distinguished. For some writers 'discourse' means exclusively *spoken* language (cf. the well-known text vs. discourse distinction) but for others it means the whole of language:

Discourse - language beyond the sentence - is simply *language* as it occurs, in any context, (including the context of linguistic analysis), in any form (including two made-up sentences in sequence, a tape recorded conversation, meeting or interview; a novel or a play) (Tannen 1989:6).

It is presumably this latter view which gives rise to expressions like 'literary discourse', 'written discourse' 'political discourse' 'medical discourse', 'the discourse of AIDS', 'the discourse of ante-natal care', 'media discourse' 'advertising discourse' 'institutional discourse', 'colonial discourse' and so on. The problem, of course, is that the term 'discourse' is being used of such a wide range of different kinds of phenomena, in different areas of study, at different levels of generalisation and in more or less well-defined ways, that it is rapidly becoming so vague as to be almost meaningless. We do not want to address this vexed issue in any detail here, as we wish to concentrate on matters related to speech, thought and writing presentation, but merely wish to alert our readers to the difficulties concerning the term 'discourse'.

Discourse (re)presentation theory, unlike other uses of the term 'discourse' include thought presentation as part of its remit. This is because much work in this area has concentrated on the analysis of fictional prose. The most straightforward use of 'discourse' in the area with which we are concerned, then, is as a generic term which

has (re)presented speech, thought (and writing) as its hyponyms.³ However, it is easy for writers in this area to move from using it in this inclusive sense to using it as a synonym for one of its hyponyms:

Qua representational discourse, therefore, each act of quotation serves two masters. One is the original speech or thought that it represents, pulling in the direction of maximal accuracy. The other is the frame that encloses and regulates it, pulling in the direction of maximal efficacy. Reported discourse thus presents a classic case of divided allegiance, between original-oriented representation (with its face to the world) and frame-oriented communication (with its face to the reader) (Sternberg 1982a: 152).

There is much that we would be happy to agree with in Sternberg's characterisation of the janus-faced character of the report of anterior speech, but note how the use of 'representational discourse' at the beginning of this quotation leads Sternberg to equate 'original speech' and 'original thought' in his 'the original speech or thought that it represents'. However, the thoughts of other people are never directly available to us, and even the 'quoting' of one's own anterior thoughts is problematic in that it is difficult to be sure what words, if any, were used to think those thoughts. The notion of faithfulness is therefore fundamentally problematic in relation to thought presentation, in a way that does not apply with speech or writing presentation. Speech and thought (re)presentation are thus very different in important ways, which the use of the term 'discourse' here tends to disguise.⁴

For clarity's sake we believe it is important within discourse (re)presentation theory that we distinguish systematically among speech, thought and writing, and only use 'discourse' when it is genuinely inclusive of its hyponyms. This latter practice is the one we adopt in this paper, and as a consequence, following on from the systematic distinction between speech and thought presentation in Leech and Short (1981: Ch. 10) we distinguish, for example, direct speech (DS) from direct thought (DT) and direct writing (DW), and so on along the three separate, but related, clines of discourse presentation possibilities (see 2 below and section 5).

2 *Speech*

The term 'speech' is often used as a general term in discourse (re)presentation theory, sometimes including thought presentation and sometimes writing report/presentation. A clear example would be Jones (1968), who uses 'speech presentation' as an inclusive term referring to the presentation of both speech and thought in Conrad's *The Secret Agent*. This presumably results from the fact that in ordinary language people often use locutions like 'I said' when they are referring to what they have written. But if we are to be clearer about differing functions, meanings and effects in discourse presentation we need to distinguish systematically between speech, thought and writing (a) with respect to the anterior context and (b) with respect to the posterior context. For example we noticed in section 1 that the written report from *The Independent* represented both spoken and written anterior sources. We cannot assume, other things being equal, that a written report of speech as direct speech (DS) will automatically have the same functions, meanings and effects as (a) a spoken report of the same DS or (b) a written report of the same content in the form of direct writing (DW). It may, but

it may not. These remarks apply, *pari passu*, for the various combinations of speech, thought and writing in anterior and posterior discourse contexts (where they both apply - see 3 below) and across all the categories of the speech, thought and writing presentation scales.

It is only when these various combinations have been adequately charted that it makes sense to examine 'mixed' cases, which might report discourse using a mix of written and spoken means.⁵ Examples of such 'mixed' cases include multimedia web sites which use text, sound, video and pictures of documents, and TV news reporting, in which a reporter may summarise or quote the original utterance, as well as showing a sound or video recording or transcript of the original.

3 *Fiction*

We need to distinguish systematically between fictional and non-fictional discourse (re)presentation. Until recently, most of the work in discourse (re)presentation theory was carried out on invented examples (in grammatical study) or fictional texts (in stylistics and literary theory). But invented and fictional examples are fundamentally different from discourse report in non-fictional language use, given that they do not, by definition, refer back to an antecedent occasion when the language being re-presented was originally produced. When we read novels we pretend that the reports are real, and so use the canonical assumptions associated with the various categories on the discourse presentation scales analogically, and with effectively guaranteed faithfulness results. The over-use of invented and fictional examples has, in our view, led to an unclear account of how discourse (re)presentation works and some confusion over the meaning of the terms we discuss in 4 below. A fully explanatory account of discourse presentation in fictional texts needs to be undertaken within a more general theory of discourse presentation.⁶

4 *Report, presentation and representation*

We need to be aware of the distinctions and overlaps among the terms 'presentation', 'representation' and 'report'. Because they have overlapping meanings they are sometimes used interchangeably (indeed, to avoid over-repetition of the same word we have sometimes used them in this way in the discussion so far, although only when we thought they were properly interchangeable). But the three terms emphasise different aspects of what is involved, and should only be used interchangeably with considerable care. Which of the terms is used as 'basic' also varies from one branch of linguistic study to another. The term 'presentation' is mainly used in stylistic analysis because stylisticians concentrate on fictional texts which, as we pointed out in 3 above, do not have an anterior discourse situation independent of the 'reporting' discourse. The novelist makes it all up. The term 'representation' on the other hand, is most often used in critical discourse analysis because such studies are often at pains to point out mismatches between the anterior discourse and its posterior representation (e.g. Slembrouck 1992). The term 'report' appears to be the basic term in linguistic traditions where the analyst invents representative examples to argue positions (e.g. generative linguistics, logic, philosophy of language), and where it is therefore assumed that the relation between, for example DS and the speech in the anterior situation which it reports is unproblematic. Interestingly, Tannen (1989), who comes originally

from such a tradition, uses the term ‘report’ even when she is attacking the assumption of identity between DS and the anterior speech which it supposedly reports.

5 Our discourse presentation corpus and associated annotation system

In order to study systematically the forms and functions of speech, thought and writing presentation in written English, we have constructed an electronic corpus which contains approximately 260,000 words of contemporary written British English, divided almost equally between three different narrative genres: prose fiction, newspaper news stories, and biographies and autobiographies. Each genre is roughly equally divided into a ‘serious’ and a ‘popular’ section, in order to allow a comparison between texts produced for different audiences and/or purposes. We made a point of using what we took to be central examples to keep the serious/popular divide as clear as we could make it. As far as newspapers were concerned, broadsheets (e.g. *The Independent*) were classified as serious and tabloids (e.g. *The Sun*) were classified as popular. When it came to fiction, we selected authors and texts which we thought could be uncontroversially included in one category or the other (e.g. Virginia Woolf’s *Night and Day* as opposed to Elizabeth Adler’s *Peach*), and also consulted a group of stylistics researchers at Lancaster University. For biographies and autobiographies, we took into account the profession of the protagonist (e.g. politician vs. sports personality), the status of the author (e.g. historian vs. tabloid journalist), and the style of writing (high- vs. lowbrow). So, for example, the serious section contains Margaret Thatcher’s autobiography *The Downing Street Years* and Peter Ackroyd’s biography of T. S. Eliot, while the popular section contains the autobiography of Cilla Black, an English ex-pop singer and light entertainment TV show host and Alan Henry’s biography of Damon Hill, the Formula 1 racing car driver.

The whole of our corpus was annotated (or ‘tagged’) manually for speech, thought and writing presentation. All texts were checked and discussed by all three authors of this paper. We initially adopted the categories of speech and thought presentation proposed by Leech and Short (1981). This model was developed and extended during the pilot phase of the project (see Short *et al.* 1996, Semino *et al.* 1997 and Leech *et al.* 1997), and further revised in the process of tagging our current corpus, during which we made our coding system conformant to the Standardised General Mark-up Language, or SGML (Wynne *et al.* 1998). The following is a list of our basic categories for speech and thought presentation:⁷

NV = Narrator’s Report of Voice
NRSA = Narrator’s Representation of Speech Act
IS = Indirect Speech
FIS = Free Indirect Speech
DS = Direct Speech
FDS = Free Direct Speech

NI = Narration of Internal States
NRTA = Narrator’s Representation of Thought Act
IT = Indirect Thought
FIT = Free Indirect Thought
DT = Direct Thought

FDT = Free Direct Thought

The above lists include the categories proposed by Leech and Short (1981) plus two new categories, NV and NI, which lie at the boundary between narration and, respectively, speech and thought presentation. NV (Narrator's report of Voice) captures minimal references to talk or speech events, where no indication is provided as to the words, content or even speech act value of the utterance (e.g. 'Voices fretted along the murmuring wire' in J. G. Ballard's *The Empire of the Sun*, or 'The winner [. . .] was yesterday being interviewed by officials' in *The Independent*, 12/12/1994). NI (Narration of Internal States) is similar to what Cohn (1978) has called 'psychonarration', in that it captures references to cognitive or emotional experiences that do not amount to the report of specific thoughts (e.g. 'My mind was racing.' in Michael Caine's autobiography). For a detailed discussion of NI, see Short *et al.* 1996 and Semino *et al.* 1997. In the process of annotating our data, we soon realised that some of the stretches of text which we had begun to code as speech presentation were in fact explicitly presented as relating to *written* texts rather than talk. We therefore decided to adopt a set of tags for writing presentation which were parallel to those for speech and thought presentation (see also table 1 above):

NW = Narrator's Report of Writing

NRWA = Narrator's Representation of Writing Act

IW = Indirect Writing

FIW = Free Indirect Writing

DW = Direct Writing

FDW = Free Direct Writing

All stretches of text which we felt did not represent speech, thought or writing were tagged as narration (N). When we thought that a particular stretch of text was ambiguous between two (or more) categories, we used 'portmanteau' tags to include each of the categories involved (e.g. the tag N-FIT indicates an ambiguity between narration and free indirect thought). In addition, we added a number of further letters to our annotations to highlight specific variants of our main categories. For example, an 'e' was placed at the beginning of a tag if the relevant category was embedded within another category (e.g. example IS embedded within DS was coded as 'eIS'). We appended a 'q' to the end of a tag if a non-direct category contained a stretch of quotation embedded within it ('e.g. a stretch of IS containing some words within quotation marks was coded as ISq'). Finally, we highlighted, by adding 'h', those instances of ST&WP which were *not* presented as having occurred in the world of the text prior to the act of narration, but rather as parts of hypotheses, wishes, fantasies or predictions about the future. We have discussed this phenomenon of hypothetical discourse in detail in Semino *et al.* (forthcoming), where we have related it to possible-world theory. For more detail on our entire annotation system, see Wynne, Short and Semino (1998).

6 An analysis of the direct forms in our corpus

We will now show how an analysis of our corpus can help to arrive at a more context-sensitive approach to the notion of faithfulness in reporting. A preliminary example of this is the way in which, by adding the 'h' tag to hypothetical instances of ST&WP, we

have been able to gain some idea of the frequency of a phenomenon which, as we mentioned earlier, is often quoted as evidence against the applicability of the notion of faithfulness to direct reports (Fludernik 1993, Sternberg 1982a and 1982b, Tannen 1989). We have found that hypotheticals only account for 4% of all tags in our corpus, most of which (62%) are embedded inside other non-hypothetical categories of ST&WP. As for direct speech, only 1.5% of all instances turned out to be hypothetical. Hence, this phenomenon can be seen as an interesting exception, rather than as a massive trend capable of dealing a fatal blow to the notion of faithfulness in reporting.

One of the advantages of the availability of a tagged electronic corpus is precisely the possibility of easy access to statistical information about the different modes of discourse presentation. Given that the issue of faithfulness of reproduction relates particularly to the direct modes, below we only provide figures for direct speech (table 2), direct thought (table 3) and direct writing (table 4). The columns in each table relate to different sub-sections of our corpus. Moving from left to right, we first provide figures for the corpus as a whole and then figures for each of the three genres within the corpus, starting in each case with the serious section (in **bold**), followed by the popular section (in *italics*). For the sake of completeness, we have separated the direct and free direct forms, even though, as we mentioned earlier, we regard the free direct variants as sub-categories of the direct forms, rather than as categories in their own right. In each case we have provided raw figures for the number of occurrences of each category and, in brackets, percentages relating to the total number of tags in the relevant section of the corpus. For example, the cell immediately below **Fiction Serious** in table 2 indicates there are 347 instances of DS in the serious fiction section of our corpus, which represent 12.9 % of the tags in that section of the corpus (percentages have been rounded up to the nearest decimal point).

Table 2 - Direct Speech and Free Direct Speech

	Whole corpus	Fiction Serious	<i>Fiction Popular</i>	Press Serious	<i>Press Popular</i>	Biog Serious	<i>Biog Popular</i>
DS	2047 (12.4%)	347 (12.9%)	485 (16.2%)	286 (10.5%)	398 (14.8%)	121 (4.4 %)	410 (15.3%)
FDS	927 (5.6 %)	282 (10.5%)	455 (15.2%)	28 (1 %)	58 (2 %)	24 (0.9 %)	80 (2.9 %)
Total	2974 (18%)	629 (23.4%)	940 (31.4%)	314 (11.5%)	456 (16.8%)	145 (5.3 %)	490 (18.2%)

Table 3 - Direct Writing and Free Direct Writing

	Whole corpus	Fiction Serious	Fiction Popular	Press Serious	Press Popular	Biog Serious	Biog Popular
DW	109 (0.6 %)	6 (0.2 %)	5 (0.2 %)	11 (0.4 %)	11 (0.4 %)	60 (2.2 %)	16 (0.6 %)
FDW	32 (0.2 %)	4 (0.1 %)	4 (0.1 %)	0 (0.0 %)	2 (0.1 %)	19 (0.7 %)	3 (0.1 %)
Total	141 (0.8 %)	10 (0.3 %)	9 (0.3 %)	11 (0.4 %)	13 (0.5 %)	79 (2.9 %)	19 (0.7 %)

Table 4 - Direct Thought and Free Direct Thought

	Whole corpus	Fiction Serious	Fiction Popular	Press Serious	Press Popular	Biog Serious	Biog Popular
DT	38 (0.2 %)	12 (0.4 %)	7 (0.2 %)	2 (0.1 %)	5 (0.2 %)	0 (0.0 %)	12 (0.4 %)
FDT	69 (0.4 %)	28 (1 %)	30 (1 %)	0 (0.0 %)	0 (0.0 %)	5 (0.2 %)	6 (0.2 %)
Total	107 (0.6 %)	40 (1.4%)	37 (1.2 %)	2 (0.1 %)	5 (0.2 %)	5 (0.2 %)	18 (0.6 %)

The overall totals in each table show that direct speech is a significant category in our corpus, representing 18% of all tags (note that the total number of tags includes the N tags for stretches of narration). On the other hand, direct writing and direct thought each represent less than 1% of all tags. This helps to explain why the word ‘speech’ is often used to refer to all the (re)presentational media when discourse (re)presentation is discussed. It is also easy to notice that the direct presentation of the three different media is not equally divided across the three genres in our corpus, nor across the serious and popular section in each genre (more on this below). A closer analysis of our data has also highlighted an important difference in the form of the free-direct sub-categories between speech, thought and writing. As far as speech and writing are

concerned, the free-direct forms are mostly characterised by the absence of a reporting clause and the presence of quotation marks. On the other hand, free direct thought may or may not have a reporting clause, but usually has no quotation marks (which, incidentally, suggests that quotation marks are associated with the reporting of phenomena where the original can, in principle at least, be reproduced verbatim). All this begins to highlight the importance of distinguishing not just speech presentation from thought presentation, but also writing presentation from the other two. The trend in the literature to subsume all three modes under the label 'discourse presentation', sometimes results in confusion, especially when claims are made about the notion of faithfulness. We shall now focus on each mode of presentation in turn.

Direct speech

If we consider DS and FDS together, we can see that they are considerably more frequent in fiction than in the other two text-types. (Free) direct speech represents 27.4% of all tags in fiction, as opposed to 14.2% in the press and 11.8% in (auto)biography. This may be due to a number of factors. Direct speech contributes to produce effects of immediacy, drama and involvement that are particularly important in the telling of fictional stories. But this is not necessarily the only explanation. In spite of the fact that the anti-faithfulness argument has been advocated with most vigour by literary scholars, fiction is, in our view, one of the genres where the reader needs to assume that, other things being equal, direct speech *does* indeed provide a faithful verbatim reproduction of a character's words (and our corpus suggests that things *are* equal in the vast majority of cases). This is because, as we pointed out above, in fiction the 'original' speech has no independent existence whatsoever, and is only accessible *via* the report itself. In other words, the pretence of faithful reproduction is one of the consequences of the suspension of disbelief that applies to the reading of fiction. Indeed, it would be difficult to talk about some important aspects of the process of characterisation if it was never possible for readers to have the illusion of direct access to the words of fictional characters. In our view, therefore, the high frequency of direct speech in fiction can in fact be seen as supporting rather than undermining the notion that the direct forms carry with them conventional expectations to do with faithful reproduction.

As for the other two genres, overall the newspaper data has slightly more (free) direct speech than (auto)biography, but the difference is not great (14.2% as opposed to 11.8%). What is interesting is that in all three genres, but particularly in (auto)biography and the press, the popular section has considerably more direct speech than the serious section. Again, this will, in part, be due to the greater wish for effects like dramatisation and immediacy in the popular texts. We would argue, however, that in serious non-fictional writing there may be a greater reluctance to use the direct forms where the original cannot be reproduced. Indeed, it is likely that readers will have different expectations of the status of material within quotation marks depending on their perception of the text-type they are reading. This is an issue that could be investigated by means of informant testing.

While a quantitative analysis has enabled us to discuss general trends in our data, only a more qualitative focus on specific examples will allow us to highlight the contextual factors which affect the issue of faithfulness in direct reporting. Our first example is

taken from the autobiography of the runner Linford Christie (co-written with Tony Ward), and specifically from a chapter entitled 'Harassment', where Christie narrates a series of racially motivated attacks on his family.

Dad was awoken by a telephone call about an hour later, around five-thirty. 'Mr Christie,' said a voice, 'you don't know me but I know you. We have met and I know that you are a good man and a religious man. I am therefore sorry to tell you that I am going to kill your son Russell.' Dad said nothing. He was too stunned (p. 43).

Here direct speech is used in writing to represent an utterance which was not experienced at first hand by the narrator, but which he must have heard second-hand from his father. It is of course unreasonable to assume that Christie's father would have been able to remember word-by-word the threats he received in the middle of the night, and equally impossible that Linford Christie would remember exactly what his father said to him (and the father may not even have told the story of this accident using direct speech in his narration of it). In any case, what matters here is not the individual words which were uttered, but the general tone and content of the threats directed towards Christie's family. As a consequence, the choice of DS here cannot be related to a claim of faithful verbatim reproduction, but is presumably aimed at involving the reader and creating drama and suspense.

Quite different considerations apply to the following extract from the autobiography of British politician Kenneth Baker, who is at this point talking about the period that led to Margaret Thatcher's resignation as Prime Minister of the UK government:

The BBC's intrepid political reporter, John Sargeant, who had been addressing the camera and had his back to Margaret, was alerted to her arrival and achieved his place in history by spinning round, defying Bernard Ingham and thrusting the BBC's microphone into her face, crying out, 'Here's the microphone Prime Minister.' Margaret then made her prepared statement, saying, 'I am naturally very pleased that I got more than half the Parliamentary Party, and disappointed that's not quite enough to win on the first ballot. So I confirm it is my intention to let my name go forward for the second ballot.'

Here there are two instances of DS. The first one relates to John Sargeant's invitation to Thatcher to speak into his microphone, 'Here's the microphone Prime Minister.' This utterance is clearly rather trivial, so that its faithful verbatim reproduction is not at all crucial. On the other hand, Sargeant's words were probably captured on camera, so that it is possible that they will indeed have been reported verbatim. The second instance of direct speech relates to what Thatcher herself said. This is a crucial statement, both because it was produced by the Prime Minister, and because every word was important in conveying her attitude to the results of the first ballot of her Parliamentary Party voting on whether she should continue as leader and Prime Minister. In addition, her statement would have been recorded, and subsequently broadcast on numerous occasions on both radio and television. As a consequence, it is both possible and important that her utterance was faithfully reproduced word by word. The fact that the written text cannot represent Thatcher's tone of voice, paralinguistic features and so on, does not detract from the high faithfulness stakes in this kind of

context. Indeed, if it turned out that the quote was not verbatim in relation to the words and structures used, we would probably have to assume that the writer made a mistake or, if the discrepancy was significant, that he had the intention to deceive his audience.

Direct writing

Table 3 shows that, overall, DW and FDW are about four times more frequent in (auto)biography than anywhere else. This is actually due to the high frequency of direct writing in serious (auto)biography, which alone contains more than half of all instances of direct writing in the corpus (79 instances out of a total of 141). The explanation for this is related to a series of connected factors. Serious (auto)biographies are less concerned with the dramatisation of the lives of the protagonists than popular ones, and tend to focus much more on spelling out the documentary evidence on which claims are based (e.g. diaries, journals, letters). In addition, those whose (auto)biographies we classified as serious tend to be the kind of people who produce considerable amounts of written texts (writers, politicians, artists), whereas those whose (auto)biographies we classified as popular are less likely to have produced a great deal of written material (e.g. sports people, singers, TV presenters).

Consider the following example from Humphrey Carpenter's biography of W. H. Auden:

Auden had the experience, some time in the early 1930s, of being invited to dinner by Eliot and his first wife Vivien, of whose strangeness he was made aware on arrival: 'I told Mrs T. S. E. that I was glad to be there and she said: "Well, Tom's not glad"' (footnote: *Tribute*, p. 155).

Here a footnote provides a reference to the source of the written quotation, which amounts to a claim of verbatim reproduction. It is also worth noticing that the DS contained within the quotation marks ('Well, Tom's not glad') also has fairly high faithfulness claims, since the whole point of the anecdote revolves around the unmitigated rudeness of the words used by Eliot's wife, which were thus likely to be highly memorable. Unlike the case of the written quotation, however, there is no way of checking the accuracy of the DS report against the original.

The use of DW in (auto)biographies and the press is almost invariably accompanied by similar references to the source, so that the reader is entitled to assume complete faithfulness to the words of the original (even though typographical characteristics are unlikely to be replicated). In fiction, on the other hand, there are cases where the accuracy of DW may be less likely and less relevant. The following extract is from the first-person narration novel *Behind the scenes at the museum* by Kate Atkinson:

Bunty holds up the dress against herself, in a sitting-position, like an invalid. She turns to me, 'What do you think?'
I sigh and shake my head in envy and longing, 'It's lovely.' (Extracts from Ruby Lennox's school report, summer term, 1966 - *Ruby has a real talent for acting . . . Ruby was the star of the school play.*)

Here we have to conclude that the stretch of text in italics is the narrator's DW representation of what she remembers a school report said about her performance in the dramatic activities of the school. More precisely, the DW is embedded inside the homodiegetic narrator's free direct thought (FDT). As such the chances that it is to be interpreted as a faithful reproduction of the report within the fictional world are fairly low. In this case, however, faithful reproduction is irrelevant even within the fictional world: the point of the quotation is to suggest that the narrator's previous utterance ('It's lovely.') was a lie, and that this lie was easily disguised as a sincere response thanks to her acting talents.

Direct thought

With thought reporting there is no such thing as an accessible original in verbal form, so that, as we suggested earlier, the notion of faithfulness is strictly irrelevant. Table 4 shows that, not surprisingly, most of the cases of direct thought presentation occur in fiction, where the reader's suspension of disbelief allows for the presence of omniscient narrators who have direct access to characters' minds. Provided that the narrator is not marked in some way as unreliable, thought reports in fiction have a high degree of reliability, even if the notion of *verbatim* reproduction may not apply in the same way as it does in the report of speech or writing. Consider the following example:

She looked round the room; [. . .] 'It's homely,' she thought with satisfaction, 'It's what I like' (Graham Greene, *Brighton Rock*).

It has often been pointed out that the use of the direct form for thought presentation creates the impression of highly conscious and articulate thought, so that it is quite possible that readers will imagine that the character in the extract above 'says' the quoted words to herself in her own head (e.g. Leech and Short 1981: 342-3). Even if this were not so, the reading convention of the omniscience and reliability of the third person narrator would lead to the conclusion that the extract within quotation marks is a faithful verbalisation of the character's thoughts, within the context of the suspension of disbelief that the reading of a novel involves. In fact there is a double suspension of disbelief here – firstly, the reader is, in general, prepared to believe that the narrator is telling the truth in writing that the character thought something (when in fact we know it is all made up) and, secondly, the reader is prepared to believe that the character is thinking in a verbal (and often very articulate and well-organised) way.

DT and FDT are much less frequent in (auto)biography and rare in the press (where only DT occurs). Indeed, all instances of DT in the press are embedded within another category, usually to do with speech presentation. In the following example from the *News of the World* a direct thought report occurs within free direct speech. The article tells the story of a British aid worker who was lost in the jungle in Africa for several days:

Day five brought the hunger pains.
'I was very weak,' Don says. 'I found some berries which looked like small grapes, and there were ape droppings with bits of these berries in them.'

‘So I thought, “If they’re good enough for them, they’re good enough for me”’
(*News of the World*, ‘I was knee deep in mud, bullets and hippos ... and all I
could think was “The missus isn’t going to like this”’, 4/12/1994).

All of the instances of DT in the press are similar to the above example in that they involve someone reporting what *they themselves* thought. Exactly the same applies to our (auto)biography data, since all cases of direct thought representation occur within autobiographies and relate to the thoughts of the protagonists themselves. So, even if we cannot assume that they originally produced their thoughts using exactly those words, nor that they thought in words at all, there seems to be a strong association between the use of the direct form and the most reliable possible source, namely the ‘thinker’ him/herself. This, again, suggests a connection between the direct forms and faithful reporting.

7 A faithfulness analysis of reports of the article ‘Don’t scare tycoons away from politics’ by Lord Simon, *The Times*, 1 August 1997

The most testing case for the anti-faithfulness position (see section 2) will be to examine cases of direct writing (DW) in written reports in clearly non-casual contexts. We examine such a case below. We acknowledge that an examination of the treatment of just one story does not constitute a complete refutation of the anti-faithfulness stance. That would require comparative analysis of a representative range of relevant texts. But in qualitative terms it does constitute a significant counter-example to the anti-faithfulness position.

The newspaper articles analysed below are not part of our corpus, but were collected later, specifically to test the anti-faithfulness view. The context is that of a political row in the UK in the summer of 1997. Lord Simon of Highbury, a prominent businessman, had joined Tony Blair’s new Labour government as Minister for Trade and Competitiveness. His appointment was clearly an attempt by the new government to show that it supported business, and the Conservative opposition mounted a smear campaign against him. This campaign eventually resulted in him resigning from the government, as a consequence of the fact that he had not divested himself of his £2 million share holding in BP, a company which he had chaired before joining the government. In the argumentative to and fro, Lord Simon wrote an article in *The Times* on 1 August 1997, defending his position. A number of newspapers picked up the story the same day, when the first edition of *The Times* came out, and quoted from it in their articles. *The Times* itself also carried a story about the article in the issues of the paper in which the article appeared.

Below is a tabular account of the degree of faithfulness in Direct Writing quotation in the newspapers which carried the story and used direct quotation. All of the newspapers except one, *The Express*, are broadsheet newspapers, and so represent the serious end of the poplar/serious divide in the British newspaper industry. *The Express* is a tabloid which was a broadsheet earlier in its history, and so is to the serious end of the tabloid spectrum.

Table 5 - Faithfulness in the Lord Simon Report

Newspaper	Total words in DW quoting from original article	Comments on the character of any non-verbatim aspects of the DW reports
<i>The Times</i>	133	'It was decided' becomes 'It was therefore decided', 'next January' becomes 'in January next year' and 'Until then' becomes 'During this period'.
<i>The Daily Telegraph</i>	73	All accurate, but 'In the meantime' is missing from the beginning of a lead sentence in one quotation, with no indication that it is missing. This is also the case with an intervening sentence in the middle of one quotation
<i>The Guardian</i>	29	All accurate
<i>The Independent</i>	76	All accurate
<i>The Express</i>	122	All accurate, but one intervening sentence in the middle of a quotation is not indicated as missing. There is also one added paragraph boundary in the middle of another quotation.
TOTAL	433	

It will be clear from the above table that faithfulness in DW in this context is being adhered to with some considerable consistency. If we first of all consider all of the newspapers except *The Times* we can see that all 300 words involved are quoted completely accurately. The only deviations from exact verbatim quotation relate to the fact that some material which is omitted from the original is not indicated as omitted by the use of continuation marks. Oddly enough, the only newspaper which alters the wording is *The Times*, and then in what are clearly three very minor cases involving a total of 7 words in the original and 11 in the report. It is, of course, very strange that only *The Times* has these inaccuracies when it was the very newspaper which carried the original article. We are not privy to exactly what happened on the day that the article and accompanying story appeared, and have no easy way of checking, but we suspect that, given that the news story appeared in the first edition of *The Times*, the *Times* journalist probably based his first-edition article, which was repeated in later editions, on a pre-final draft of Lord Simon's article, to which only the *Times* journalists would have had access. The discrepancies would then, of course, be accidental, and connected to the rush of newspaper production. There would be no break in faithfulness on the part of the journalist as he would have been quoting from a slightly different version of the article.

This evidence from written sources containing direct writing representation in a serious political context is thus contrary to the evidence produced by the anti-faithfulness school, and suggests that the move to remove the faithfulness criterion from discourse presentation theory is, at the very least, premature. Our view is that the variations can be explained on pragmatic grounds depending on (a) the medium of the anterior source, (b) the medium of the posterior report and (c) contextual factors like the

casualness or seriousness of the discourse being reported. If this is the case, we effectively need a more complex notion of faithfulness than has been hitherto assumed, one where its assumed level of importance in discourse report is related to pragmatic knowledge and assumptions. In the final section of this paper we sketch out how such a pragmatically-based model of the salience of faithfulness in discourse (re)presentation might work.

8 Factors determining the salience of the faithfulness criterion in discourse report

Below we list a number of factors which have to be taken into account in determining the salience of the faithfulness criterion in particular discourse reporting contexts. There may well be other factors which we have not thought of, and we do not necessarily assume that each factor is equally weighted. Further research would need to be carried out (a) to investigate our faithfulness salience proposals in general terms and, if they find favour, (b) to arrive at a more complete list of such factors and (c) to develop a mechanism for weighting the factors appropriately.

1 *Anterior discourse accessibility*

In order for someone to report an anterior discourse accurately the reporter must have had direct access to the anterior discourse. Writing and speech are both in principle accessible to the reporter, but speech is only accessible at the moment of utterance unless it was recorded. Writing, on the other hand, because it does not degrade rapidly over time, is usually available for a reporter to check whether his or her report is accurate. However the thoughts of other people are never accessible directly, and even when people report their own previous thoughts there must be considerable doubt as to whether an ‘anterior thought discourse’ actually existed in linguistic form or even if it can be sensibly described as a discourse.

In order for someone other than the reporter to check whether a direct discourse report is accurate, access to the anterior discourse is needed after the posterior discourse has been produced, and the factors described in the previous paragraph apply in a parallel way.

Reporters are thus normally able to achieve accurate writing report more easily than accurate speech report, while it is not clear whether the faithfulness factor can be properly applied to thought report at all. Similar remarks relate to any pressure reporters may feel as a result of assessing whether others can check up on the accuracy of their reports. The net result of the anterior discourse accessibility factor can be graphically represented on a scale as follows:

Anterior discourse accessible *Anterior discourse inaccessible*

Writing<----->Speech<----->
>Thought

2. *Posterior discourse accessibility*

For someone, either the reporter or someone else, to check the accuracy of a discourse report the posterior discourse also has to be available. Again, this is easier when the

reporting medium is writing rather than speech, and the notion of a posterior discourse in thought mode does not properly apply - unless you can manage extra-sensory perception. For the posterior discourse accessibility factor we thus have a scale parallel to that for anterior discourse accessibility:

Posterior discourse accessible *Posterior discourse inaccessible*
 Writing<----->Speech<----->
 >Thought

3. *The importance of what is being reported*

It seems reasonable to assume that, other things being equal, the importance of the information being imparted (and/or the style in which it was produced) will affect a reporter's felt need to be accurate in report. Thus the views someone has declared about the weather as a phatic conversational move will usually be less important than views in relation to a political debate or an academic argument, and so we would expect the former to be reported less punctiliously than the latter.

4. *The status, social role and personality of the producer of the original discourse*

We are likely to be more careful when reporting what someone we regard as important has said than when we report the words of someone we do not think of as important. Similarly, we are likely to be more punctilious about reporting carefully the words of someone who has power over us, has a nit-picking personality, or is easily hurt.

5. *Attitude, social role and personality of the reporter*

Other things being equal, we might expect a reporter who is sympathetic to what is being reported, or to who said or wrote it, to be more careful in relation to faithfulness than someone who is antipathetic to what was 'discoursed' or who produced it. There may, however, also be circumstances where a sympathetic reporter may try to improve on the original, and thus be unfaithful to the original. The 'punctiliousness rating' of the character of the reporter could also be a factor.

6. *Text-type or speech context*

We suspect that, unless legal proceedings are thought possible, popular magazines and newspapers will probably be less careful about faithfulness than more serious journalistic organs. Similar remarks apply to what is said in a court of law vs. what is said in a conversation in a bar. Fictional texts are a special case with respect to this factor because in fictions no anterior discourse actually occurred. The writer makes it all up, but the reader pretends to him-or herself that the writer was a witness to the 'anterior discourse'.

7. *The part of text in which reporting occurs*

It may well be that position in a text might affect the faithfulness quotient. For example, there is some evidence that headlines in newspaper news reports are less restricted in terms of faithfulness for text in quotation marks than the main body of reports (cf. Short 1988).

8. *The memorability of the original*

Particular instances of speech or writing may be particularly memorable, for example because of the unusual context in which they occurred or the particular form of the language used.

There could well be other factors, and we would welcome the suggestions of other scholars.

9 Conclusion

In summary, we think that although the ‘anti-faithfulness’ school have provided some important data which needs to be taken into account, the move to remove faithfulness to the anterior discourse as a criterion in discourse report is premature, and has some unfortunate consequences in the attempt to arrive at a comprehensive theory of discourse (re)presentation. We also believe that to come to final conclusions about faithfulness in particular, and discourse (re)presentation in general, scholars will need to systematically distinguish among speech, thought and writing (re)presentation in their studies, and think of them as three separate but related scales of discourse presentation. It is clear that in terms of medium the notion of faithfulness applies most easily to writing presentation, and that readers’ interpretations of the meanings and effects of presentation categories in other media are related to canonical assumptions related to writing presentation. Indeed, the thought presentation scale is very different from the speech and writing scales, precisely because faithfulness does not sensibly apply to this scale. Nonetheless, the meanings and effects of the various thought presentation categories can, in our view, only be properly accounted for by relating it analogously to the other scales.

The distinctions we have made here, and our corpus-based research more generally, have led us to propose a set of factors which need to be taken into account when the salience of the criterion of faithfulness is assessed in accounts of discourse report. We do not expect our list of ‘faithfulness factors’ to be exhaustive. However, we would want to suggest that such factors will be important in developing a context-sensitive account of the role of faithfulness in discourse (re)presentation. We hope that other scholars will take up the challenge, along with us, of developing a fuller set of factors, and investigating (a) how these factors interrelate with one another in different reporting/presentation circumstances and (b) what counts as ‘verbatim’ in different reporting contexts. It is only when this has been achieved that we can hope to arrive at a truly general, yet discriminating, account of discourse (re)presentation and how it operates in different circumstances. The ‘tagged corpus’ approach would appear to be an important tool to use in arriving at such an account.

Notes

- 1 The research project referred to in this paper was supported by the British Academy’s Humanities Research Board (grant BA LRG M-AN2314/AON3489) and by pump-priming and project-completion grants from the Social Sciences Faculty and the Committee for Research of Lancaster University. We would like to thank John Heywood for his perceptive comments on an earlier draft of this paper.
- 2 It is interesting to note that Tannen (1989) uses inverted commas to mark off direct discourse reports in transcripts of speech, even though such punctuation devices, being writing devices, could not possibly have been present on tape. This

- indicates a tacit assumption that direct speech in spoken and written language are automatically and directly equivalent, an assumption that is by no means obvious. Others, for example Myers (1999) are more sparing in their use of such orthographical markings, indicating a more careful attitude to the issue of cross-medium 'equivalence'.
- 3 However, writing is usually disregarded, or treated in the same way as speech. In our Speech, Thought and Writing Presentation corpus-based project, writing presentation has been treated throughout as a separate discourse presentation cline, enabling the different forms and functions associated with writing (re)presentation to be separately identified and clarified.
 - 4 For an account of a quotation in Fludernik (1993) which has similar meaning shifts in relation to 'discourse' see Short, Semino and Wynne (1997).
 - 5 The same is true of the categories on the discourse presentation scales. It has been suggested to us verbally that, given the prevalence of 'blended categories' like FIS, FIT, FIW and the presentation of quotation ('q') phenomena in the more indirect discourse presentation categories which have been revealed by our corpus work, it would be better to abandon the clear canonical categories like DS and IS. But this seems not sensible to us, as, in logical terms, the blended categories can only be adequately accounted for through the establishment of theoretically prior non-blended categories.
 - 6 It is worth noting that the terms 'direct speech' and 'indirect speech' were in use well before the advent of the invention of equipment to record speech, when the chances of checking up on the faithfulness of speech report were very small indeed. The earliest instances quoted in the *OED* come from the mid-19th century, but are, of course part of a tradition which Fludernik (1993: 26ff) has traced back to Roman and Greek times. Indeed, the terms *oratio recta* and *oratio obliqua* are, as Fludernik notes, obvious antecedents of direct and indirect speech. From the earliest times examples of writing and speech, and history and fiction, were used together when exemplifying or discussing discourse categories.
 - 7 The direct/free-direct distinction was applied throughout, even though we already thought when we began the process of annotation that the latter was probably a minor variant of the former in order to allow us to gather data to test our hypothesis. This issue is discussed in detail in Semino, Short and Culpeper (1997).

References

- Clark, Herbert H. and Richard J. Gerrig (1990) 'Quotation as demonstration,' *Language* 66, 784-805.
- Cohn, Dorrit (1978) *Transparent Minds: Narrative Modes for Presenting Consciousness in Fiction*, Princeton NJ: Princeton University Press.
- Fludernik, Monika (1993) *The Fictions of Language and the Languages of Fiction*. London: Routledge.
- Jones, Charles (1968) 'Varieties of speech presentation in Conrad's *The Secret Agent*', *Lingua* 20, 2, 162-76.
- Leech, Geoffrey and Mick Short (1981) *Style in Fiction*. London: Longman.
- Leech, Geoffrey, Tony McEnery and Martin Wynne (1997), 'Further levels of annotation'. In Roger Garside, Geoffrey Leech and Anthony McEnery (eds) (1997) *Corpus Annotation*, London: Longman, pp. 85-101.

- Linnell, Per (1982) 'The written language bias in linguistics', *Studies in Communication* 2. Linköping, Sweden: University of Linköping.
- McCarthy, Mike (1998) *Spoken Language and Applied Linguistics*, Cambridge: Cambridge University Press.
- Myers, Greg (1999) 'Functions of reported speech in group discussions', *Applied Linguistics* 20, 3, 376-401.
- Semino, Elena, Mick Short and Jonathan Culpeper (1997) 'Using a computer corpus to test a model of speech and thought presentation', *Poetics* 25, 17-43.
- Semino, Elena, Mick Short and Martin Wynne (forthcoming) 'Hypothetical words and thoughts in contemporary British narratives', *Narrative*, Fall 1999.
- Short, Mick (1988) 'Speech presentation, the novel and the press'. In: Willie van Peer (ed.) (1988) *The Taming of The Text*, London: Routledge, pp. 61-81.
- Short, Mick (1996) *Exploring the Language of Poems, Plays and Prose*, London: Longman.
- Short, Mick, Elena Semino and Jonathan Culpeper (1996) 'Using a corpus for stylistics research: Speech and thought presentation'. In: Jenny Thomas and Mick Short (eds.) (1996) *Using Corpora in Language Research*. London: Longman, pp. 110-31.
- Short, Mick, Elena Semino and Martin Wynne (1997) 'A (free direct) reply to Paul Simpson's discourse', *Journal of Literary Semantics*, XXVI, 3, 1997, 219-28.
- Short, Mick, Elena Semino and Martin Wynne (1999) 'Reading reports: Discourse presentation in a corpus of narratives, with special reference to news reports'. In Hans-Jürgen Diller and Erwin Otto-Gert Stratmann (eds) (1999) *English Via Various Media*, Heidelberg: Winter, pp. 39-65.
- Slembrouck, Stefaan (1992) 'The parliamentary Hansard "verbatim" report: The written construction of spoken discourse', *Language and Literature* 1, 2, 101-119.
- Sternberg, Meir (1982a) 'Proteus in quotation-land: Mimesis and the forms of reported speech', *Poetics Today*, 3, 2, 107-56.
- Sternberg, Meir (1982b) 'Point of view and the indirections of direct speech', *Language and Style*, 15, 2, 67-117.
- Tannen, Deborah (1989) *Talking Voices: Repetition, Dialogue, and Imagery in Conversational Discourse*. Cambridge University Press: Cambridge.
- Waugh, Lynda (1995) 'Reported speech in journalistic discourse: The relation of function and text', *Text* 15, 1, 129-73.
- Wynne, Martin, Mick Short and Elena Semino (1998) 'A corpus-based investigation of speech, thought and writing presentation in English narrative texts'. In Antoinette Renouf (ed.) (1999) *Explorations in Corpus Linguistics*, Amsterdam: Rodopi pp. 231-45.