If on a Winter's Night Two Researchers... A Challenge to Assumptions of Soundness of Interpretation

ANNA MARCHI

Lancaster University

marchi@unix.lancs.ac.uk

CHARLOTTE TAYLOR

University of Portsmouth

charlotte.taylor@port.ac.uk

Abstract

This paper reports on a quasi-experiment into triangulation, which is increasingly frequently cited as a guarantor of validity and reliability of findings. The methodology that we are exploring is the increasingly widely employed combination of corpus linguistics and (critical) discourse analysis. It has been argued that corpus approaches can offer greater objectivity because they are data-driven (or at least data-supported), more generalisable as they are based on larger samples, and more transparent given the research may be replicated on the same data. In order to explore the extent to which integrating corpus approaches may contribute to the stability of interpretations the authors set up an exploratory experiment. We attempt to answer the question: would two researchers starting with the same corpus and research question and (broadly) theoretical /methodological framework come to the same/similar conclusions?

Keywords: corpus linguistics, critical discourse analysis, methodology, triangulation, news discourse.

1. Introduction

This paper reports on a quasi-experiment into methodological reliability. We will explore and problematise the increasingly frequently employed combination of Corpus Linguistics and (Critical) Discourse Analysis (Baker 2006, Hardt-Mautner 1995, Partington 2004, 2009, Stubbs 1995, 1997, 2006). It has been argued that corpus approaches can offer greater objectivity because they are data-driven (or at least data-supported), more generalizable as they are based on larger samples, and more transparent given the research may be replicated on the same data. In order to explore the extent to which integrating qualitative and quantitative approaches may contribute to the stability of interpretations the authors attempted to answer the following question: would two researchers starting with the same corpus and research

question and (broadly) theoretical/methodological framework come to the same/similar conclusions? The specific topic chosen for investigation was: how do journalists talk about themselves/each other and their profession in a corpus of British media texts?

In the first part of this article we will list the characteristics and limitations commonly attributed to the two methodologies and discuss the implications of triangulation, we will then present our case study and finally we will draw some conclusions both regarding our specific experiment and more broadly concerning multi/mixed-method approaches.

2. (C)DA and Corpus Linguistics

The combination of quantitative and qualitative analysis of discourse has been praised as offering a solution to the limitations of the two approaches. On one side we have the 'greater precision and richness' (McEnery and Wilson 1996:77) of qualitative analysis, on the other we are also granted the 'statistically reliable and generalizable results' (ibid.) of quantitative studies. In particular Corpus Linguistics and CDA have been recently described as a 'natural match' (Mautner forthcoming) and a 'useful methodological synergy' (Baker et al. 2008). Orpin (2005) identifies the main contribution that CL tools can give to CDA in the ability to handle more texts, thus enhancing reliability, Van De Mieroop (2005) similarly associates the idea of using a large corpus with getting the big picture, but he also sees CL analysis (e.g. cluster analysis) as offering good starting points for qualitative analysis. O'Halloran and Coffin (2004) list the characteristics of a corpus approach that can prove useful to compensate methodological weaknesses of CDA and improve the research. Hardt-Mautner (1995) thoroughly discusses the pragmatic advantages of integrating the best tools and techniques available from the two approaches and describes a mixed method that could be seen as a *third approach* (Bryman 2006), where the very nature of interpretation is redefined.

We will now illustrate some popular claims made about CL and CDA, reviewing strengths and weaknesses (reciprocally) attributed to the two approaches (summarised in Table 1) and criticising some aspects of the definitions.

Table 1. Some popular claims about CL and CDA

CL	CDA
✓ Quantitative ✓ Data driven ✓ Representative samples ✓ Statistical relevance = representative ✓ Breadth ✓ Generalizability ✓ Replicability = greater objectivity ✓ Descriptive power	✓ Qualitative ✓ Theory driven ✓ Individual contextualised examples ✓ Social relevance = meaningful ✓ Depth ✓ Precision and richness ✓ Political intent = subjective interpretation ✓ Explanatory power

CL is a quantitative approach. A corpus is often defined in terms of quantity as a 'large collection of authentic text' (Bower and Pearson 2002: 9) and is characterised for being '*representative* of some language or text type' (Leech 1992: 116). The research aims at 'exemplify[ing] the dominant structural patterns of the language without recourse to abstraction' (Sinclair 1991: 103). CL puts the emphasis on frequency and adopts statistical tools in order to access stable, reliable and generalizable descriptions based on samples of millions of words. On the other hand frequency is not always a good or sufficient parameter to determine the relevance of a phenomenon and reveal mainstream discourse, and a common criticism posed to CL is that the data are semiotically impoverished and decontextualised and that findings risk being limited to 'counting only what is easy to count' (Stubbs and Gebirg 1993: 78). Corpus linguists themselves warn not to 'get caught in using corpora just to tell you more about what you already know' (Sinclair 2004: 185).

CDA is a qualitative approach. It relies on close reading of limited amounts of texts, selected for their relevance to the matter under investigation. Like CL, CDA 'is situated firmly in the field of Applied Linguistics' (Orpin, 2005: 38), but the data are considered in the wider social, cultural and historical context (Titsher et al. 2000) and are approached from a rich theoretical framework. Dealing with *authentic* data is integral to CDA as is 'a constant movement back and forth between theory and empirical data' (Wodak 2006: 6), but the data selection and treatment tend to be exemplificatory, rather than comprehensive. This allows for in depth finegrained analyses and thick explanations, but it also brings upon CDA analysts the accusation of *cherry picking* 'small and unrepresentative data samples in order to suit researchers' pre-conceived notions about hidden ideological meanings' (Mautner forthcoming), i.e. in order to suit the theory. CDA tackles important social issues from a declaredly critical position, its interpretations can be equally important and meaningful in terms of social relevance, but scarcity of empirical evidence and theoretical overload can result in finding just what we already expect to find (Stubbs 1997).

In CL results emerge bottom-up from the *observation* of the data, more precisely from the observation and interpretation of patterns in the data that are made visible and significant by their recurrence. Once we have reduced the complexity, 'the rich chaos of language' to its 'boiled down extract' of wordlists, keywords, collocations, concordance lines – explain Scott and Tribble (2006: 5) – imaginative power is released.

> Human beings are unable to see shapes, lists, displays or sets without insight, without seeing in them "patterns". It seems to be a characteristic of the *homo sapiens* mind that is often unable to see things "as they are" but imposes on them a tendency, a trend, a pattern. (Scott and Tribble 2006: 6)

Having access to large amounts of data grants greater representativeness (statistical relevance) and 'allows a greater distance to be preserved between observer and the data' (Partington 2006: 268), but a corpus approach is neither solely descriptive nor is it immune to subjectivity.

CL allows us to start 'from a position whereby the data itself has not been selected in order to confirm our conscious (or subconscious) biases' (Baker 2006: 12), but bias is not a prerogative of qualitative research. 'A

'disinterested' social science has never existed and, for logical reasons, can never exist' (Myrdal 1970: 55), the very object of research emerges from an interest, i.e. a sense of significance, i.e. a subjective judgement.

A corpus approach is not a neutral one, selection comes into play throughout the process and early stages of analysis (e.g. the choice of the lexical items to investigate) can heavily determine the progress of the research. The patterns we identify and the findings we generate implement and shape subsequent questions; each previous step informs the next, possibly excluding other threads. *Homo sapiens* (Scott and Tribble, ibid.) works for pattern recognition, but also tends to look for confirmatory evidence, the combination of these two tendencies can lead to what we call the 'corroboration drive', i.e. a systematic search for elements that validate previous findings.

The integration of CL and (C)DA can certainly improve our analysis, in a variety of ways. CL approach can implement 'non-theory-specific categories emerging from large scale data' (Baker et al. 2008: 7), thus expanding the ways we can interrogate the texts and the starting points we can take. This bottom-up nature, and the attention to statistical relevance, make it easier for the researcher to guard against overinterpretation and/or underinterpretation (O'Halloran and Coffin 2004). The quantitative techniques of CL give greater generalizability, allow for replicability and strengthen the overall reliability and validity of the research.

(C)DA's qualitative approach can offer an entry into the data and provide powerful explanations, grounded in its rich theoretical framework and in its interdisciplinary nature. Furthermore it can complementarily improve validity by offering a situated point of view, which means that thanks to contextual knowledge we can check the progress of the analysis and question / control potential 'corroboration drive'.

We are not questioning the fact that triangulation improves the soundness and the reach of analysis, on the contrary we support the idea that research greatly benefits from being 'shamelessly eclectic' (Lortie 1982, quoted in Rossman and Wilson 1994), but we want to emphasise that integrating and complementing methodologies is not a solution to the vexed problem of objectivity and that there is more to *triangulation* than combined methods.

3. Triangulation

Indeed, although triangulation is frequently used synonymously with multi-method approach, Denzin (1970, discussed in Bryman 2003) identifies at least four possible variables, and Thurmond (2001) in her literature review lists five types. In the following section, where possible these different types will be discussed with specific reference to discourse analysis.

3.1 Types of triangulation

3.1.1 Methodological triangulation

As noted above, using more than one methodological approach to gathering and interpreting the data is the most common variable intended when

research in our field reports the use of triangulation. However, equating triangulation with multi-method approaches not only distracts from the plurality of triangulation types available, but also risks creating confusion regarding the methodology adopted. For example, in a mixed methodology like Corpus-Assisted Discourse Studies (CADS, see Partington 2009) the two methodologies of corpus linguistics and discourse studies are deliberately intertwined, and the interaction is carefully exploited. This is not therefore triangulation in the sense of applying two methodologies separately to the same question in order to compare findings.

One recent example of methodological triangulation, within the field of discourse analysis, could be the Lancaster project led by Paul Baker and Ruth Wodak which investigated discourses of refugees and asylum seekers in the UK press. Baker specifies in the final report that '[t]wo researchers (one with a background in CL another with a background in CDA) were employed on the project. Both researchers mainly worked independently, although met at regular project meetings where they reported on their findings to date' (Baker 2007 :23). This would, therefore, be an example of what Denzin (1970) termed *between-method* triangulation, which is contrasted with *within-method* triangulation, the test of internal reliability of a given method.

3.1.2 Data triangulation

In Thurmond (2001) this type of triangulation is further divided into triangulation of data sources and of data analysis, but given that in the context of discourse analysis the data analysis is contained within the methodology, in this paper it is just used to refer to data sources. According to Bryman (2003), this form of triangulation entails collecting data through several sampling strategies, so that slices of data at different times and from different social situations, as well as on a variety of people, are gathered. For instance, in corpus linguistic analysis the repeated retrieval of random concordance lines as suggested by Sinclair (1999) would allow for some micro-triangulation in so far as hypotheses formed from the first 30 could be compared to those arising from subsequent sets.

With more specific reference to the kind of questions addressed in (C)DA, we might take Koller's recent volume *Lesbian Discourses: Images of a Community* as an example of research employing data triangulation, as she utilises data from the related but different discourse types of magazines, pamphlets and blogs.

Alternatively, the data sources involved could vary more widely, for example in our research to investigate how journalists talk about journalism and journalists, in addition to looking at newspaper discourse and taking into account published memoirs or diaries by journalists, research funding permitting, it would have be enlightening to carry out a series of interviews with newsmakers.

3.1.3 Investigator triangulation

This third form refers to the use of more than one researcher to collect and/or interpret data. Although it is frequently the case that several researchers will work on the same project, it is exceptionally rare for the variable to

operationalised as true triangulation, with the exception perhaps of research projects in which, for example, inter-rater reliability is important. An example of where investigator triangulation could be particularly relevant in Critical Discourse Analysis would be in the assignment of analytical categories, for example in the kind of framework being developed by KhosraviNik (2008).

3.1.4 Theoretical triangulation

Theoretical triangulation refers to the use of more than one theoretical position in interpreting data, and the explicit use of this type of triangulation is much less common than methodological or data variation. A notable exception being the study carried out by Stubbs et al (2003) in which the same stretch of spoken discourse was analysed from the perspectives of; conversation analysis, interactional sociolinguistics, politeness theory, critical discourse analysis, and discursive psychology.

Although theoretical triangulation is infrequently explicitly referred to, it could of course be argued that any methodological choice carries theoretical implications and therefore theoretical triangulation is a necessary consequence of methodological triangulation. Indeed, if we consider the criticisms made regarding the two approaches of corpus linguistics and critical discourse analysis, as outlined in section 2, we may see they are frequently based on different theoretical assumptions about the 'best' way of analysing language in use.

3.2 Results of triangulation

In addition to the range of triangulation types, we would like to draw attention to the presence of at least three separate outcomes (or even aims) of triangulation despite the common tendency to focus exclusively on confirmation, possibly as a result of the *corroboration drive* discussed in section 2. The three possible outcomes or objectives are:

3.2.1 Convergence

In this case, the results of the triangulation confirm one another. This is naturally what most researchers are looking for when implementing triangulation, but to focus on it to the exclusion of the others is somewhat limiting. With reference to methodological triangulation, Erzberger and Prein note that 'the underlying assumption is that the validity of research results is enhanced if different methodological approaches produce convergent findings about the same empirical domain' (1997:144). However, we would argue that we should be wary of assuming that converging results have a greater *validity*; the replication and confirmation essentially illustrate the *reliability* of the findings, but, for example, in the instance of investigator triangulation both researchers may be equally wrong! Alternatively, in methodological triangulation, it could be that neither instrument was measuring the intended target.

3.2.2 Dissonance

In this outcome, the findings of one methodology, researcher, data-set or

theoretical position are incompatible with the findings of another and therefore dissonance is essentially Popperian as an objective. Such falsificationist findings are extremely useful but rarely reported, though we may hypothesise that in reality falsification is common and frequently leads to the researcher/research team modifying the research process prior to public dissemination of the findings. Moving out of the field of discourse analysis, interesting work in this area has been carried out within other social sciences, such as Perlesz and Lindsay's (2003) paper on 'Methodological triangulation in researching families: making sense of dissonant data'.

3.2.3 Complementarity

In this third possibility, the triangulated findings are viewed as parts of a jigsaw puzzle (Erzberger & Prein 1997), which when put together may offer a more complete view of the construct which is being investigated, and as such is a highly productive aim.

4. The Design of the Experiment

In order to investigate the effects of researcher triangulation, while attempting to keep other variables constant, we set up a simple quasi-experiment involving two researchers, the co-authors of this paper. Both researchers had the same corpus, a sub-section of the IntUne[1] corpus, containing the complete output of four British newspapers (*Telegraph*, *Guardian*, *Western Mail* and *Scotsman*) over a three month period plus two months of TV news programmes (*BBC News* and *ITN*) from early 2007, amounting to nearly 23 million words. The corpus was XML[2]-valid and TEI[3]-conformant and was marked-up for a variety of parameters (e.g. type of news, page number, headlines, paragraphs) and searchable using *XAIRA*[4] software. The researchers adopted the same tools for corpus interrogation, i.e. *WordSmith Tools* and *XAIRA*.

Furthermore, the two researchers (henceforth Researcher A and Researcher B) had worked together previously and we were therefore relatively confident that our methodological and theoretical approaches were sufficiently similar.

Finally, of course, we each had the same research question: how do journalists talk about themselves/each other and their profession in a corpus of British media texts? During the research phase we shared reading material but did not compare or discuss our results or findings.

What was interesting to note was the difference even in the starting points chosen for the research process. While Researcher A started by exploring lexical items derived from wordlists, keywords and key-clusters as well as intuition, Researcher B initially selected lexical items via introspection and from additional background reading, for example from books in which journalists talk about journalism. Indeed, even the choice of *intuition* versus *introspection* to describe the research process illustrates differences between the researchers right from the beginning. With reference to the lexical items selected, Researcher A grouped them into: journalists explicitly talking about themselves, each other and their trade (e.g. *journalis*, reporter, news*...),* metadiscourse and journalistic jargon (e.g. *headline, front page, hack...),* and

journalistic stance (e.g. *TELL + us / our /* newspaper's name), while researcher B categorised the node terms as: general terms referring to the trade e.g. *journalis*, news*, the press, Fleet Street, British press,* references to other media outlets e.g. *Guardian, Grauniad,* and references to the people/roles involved e.g. *columnist*, political editor* etc.

However, both researchers noted a common problem from the early stages of the research process, which was the dominance of the *Guardian* newspaper in virtually every analysis, as illustrated in Figures 1 and 2, which were produced in the initial phase of the analysis by Researchers A and B respectively.

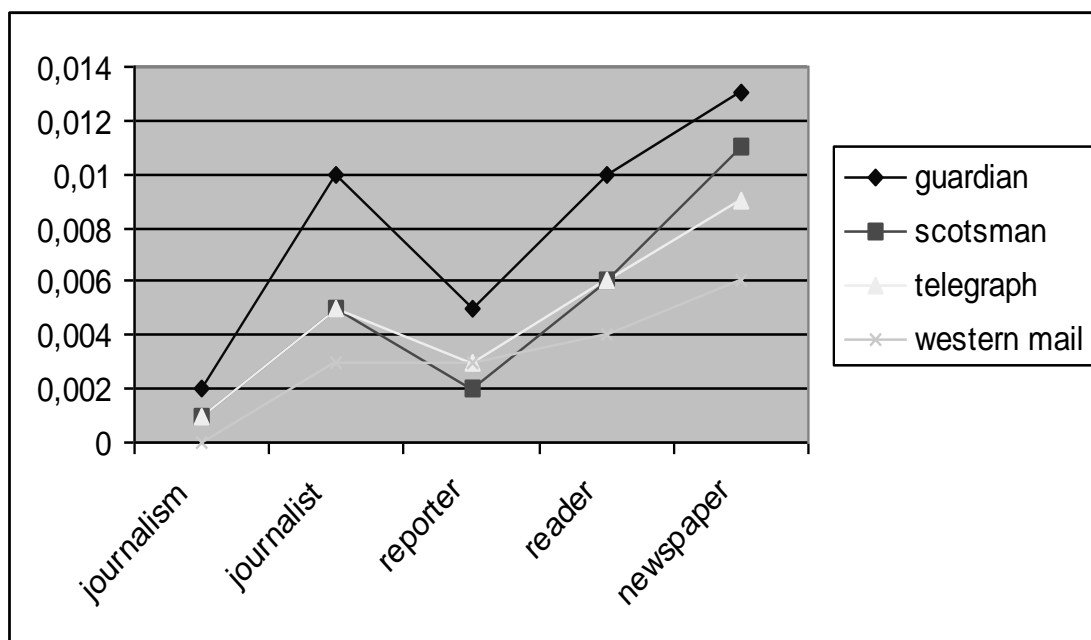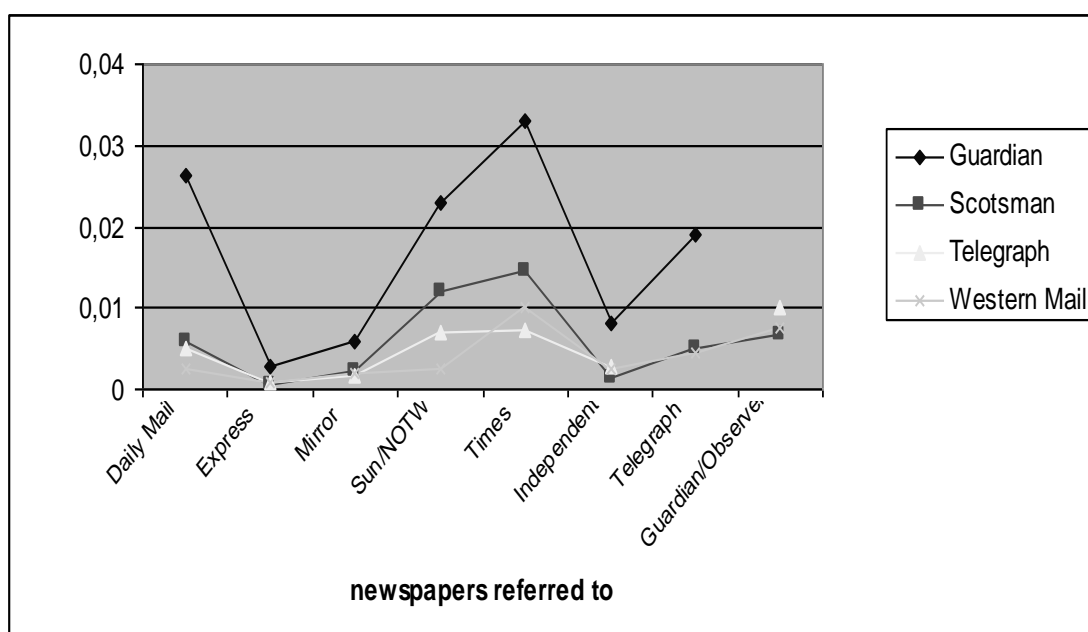**Figure 1.** Relative frequencies of news-related terms in the sub-corpora

Figure 2. Relative frequency of mentions of other newspapers (per thousand words)

As we had anticipated, the *Guardian* seemed to be by far the most interested in 'the trade'. This was largely due to the fact that the *Guardian* has a specific media section, but even the fact that such a section exists in the *Guardian* and not in the other papers is in itself significant. Comparing the individual newspaper corpora against each other by doing Keywords analysis, a wide range of Keywords and key-clusters related to media and newsmaking came up for *Guardian* sub-corpus e.g. *newsbites, newsroom, newspaper, reporter, journalist(s), reader(s), journalism, press* and *New York Times, Daily Mail and, in the Mirror, in an article, been written about*.

5. Three Orders of Findings

5.1 Dissonant findings

In a preliminary phase of the research, while going through the literature and browsing journalistic material, such as memoirs, we came across two interestingly diverging quotes by two well known journalists:

> Dog doesn't eat dog. That's always been the rule on Fleet Street. We dig into the world of politics and finance and sport and policing and entertainment. We dig wherever we like – but not in our own garden. (Davies - special correspondent for the Guardian, 2008: I)

> The ancient tradition of dog don't eat dog which once meant only Private Eye contained defamatory material about Fleet Street has long since been abandoned. (Ingrams - former editor of Private Eye, 1999: 139)

These quotes do not reflect different pre-assumptions of the two researchers, but they do give an accurate account of the difficulty in journalism as much as in social sciences of stating an objective fact. By looking at slightly different things we started with different initial findings which partially implemented different paths and different interpretations. For instance, Researcher A expected journalists to be prone to talking about themselves and each other, because of reflexivity would serve what 'the sociology of everyday life describes, a self-monitoring and reconfirmation of journalistic discourse' (Matherson 2003: 169).

The most dissonant finding from the two researchers was the conclusion by Researcher A that journalists are prone to talking about themselves, while Researcher B concluded that journalists do not talk about themselves but always about some other newsmakers. As an illustration of the latter, the analysis of collective nouns referring to members of the newsmaking trade shows a predominately negative evaluation.

Starting with the phrase *the media*, this was predominately used in quotes (and Letters), suggesting that it is more an *outsider* term than *insider*, and a more detailed reading of what *the media* actually does, as illustrated in Figure 3 with the phraseology *by the media*, shows that it was frequently the actor in negatively evaluated processes.

Figure 3. Concordance lines of *the media*

Teld	eving relatives are more often <u>appropriated</u>	*by the media* to advance an
sco	The couple were again <u>besieged</u>	*by the media* when their split was
teld	dd up to an accusation. Gina is <u>condemned</u>	*by the media* and soon afterwards
teld	How is your sport covered	*by the media*? I think they are
gua	st recruit was no more likely to be <u>derailed</u>	*by the media* than by an opponent
gua	shocking an image would be <u>played down</u>	*by the media*. Iconic pictures of
gua	Opik feels <u>drained</u>	*by the media*, however. "I
itv	rts, by British investigation, and endlessly	*by the media*.
teld	to knock them down which is so favoured	*by the media* when it comes to
gua	erated by me or the party. It was generated	*by the media*."
gua	tions which would be <u>transmitted gleefully</u>	*by the media* and stored up by the
Gua	ff, the "unfair treatment" meted out to him	*by the media* and the "blanket
Gua	omes public enemy number one - <u>hounded</u>	*by the media*, betrayed by friends,
gua	blem now is that there will be a <u>witch-hunt</u>	*by the media*. It's really too

Although, in the interest of balance, it should also be noticed that two other salient examples found in the corpus referred to the media as the victim of unpleasant processes such as *attack, blame*, and *manage, control*.

Similarly, *the papers* often referred to the negative evaluations of some Other, especially in the *Guardian.*

> (1)
> I rent a flat in London's Kentish Town on my modest junior doctor's salary (don't believe what you read in **the papers** about doctors' wages, either). (Guardian 12.02.07)

The irony of (1) of course being that this was published in a paper. Alternatively, *the papers* was often used in contexts which referred to the treatment of individuals by the press as seen in (2), and this was particularly common within reported speech.

> (2)
> Who the hell is this woman, <u>popping up</u> on daytime television and in **the papers** all the time, telling the world how sorry she feels for 'Big Brother's weeping Jo', whoever the hell that might be? (Telegraph 10.02.07)

Also noticeable in (2), in the phrase 'popping up', is the tendency to present *the papers* as mere carriers rather makers of news. *Fleet Street* too was generally critical, referring to some Other set of journalists as seen in (3) and (4).

> (3)
> This suggested a cover-up, and if there is one thing **Fleet Street** likes even more than a risk, it is a cover-up. (Guardian 14.02.07)

> (4)
> Yet apparently keen to stave off what is looking alarmingly like the inevitable, Steve is to hold what is traditionally referred to in <u>the breathless parlance of</u> **Fleet Street** as 'a series of crisis talks' with senior England players. (Guardian 15.02.07)

The *British Press* was also frequently presented as a force for ill, especially if preceded by a pre-modifier such as *elements of, sections of* or *other* which clearly demarcates the journalism being discussed from the (superior, favourably evaluated) text producer.

(5)
No lie is too fantastic, no invention too absurdly contrary to what is now common knowledge for employment by the more reckless <u>elements of</u> **the British press**. (Guardian 02.03.07)

(6)
But <u>sections of</u> **the British press** have been suckered into portraying the Iranian regime as bent on making nuclear weapons and wiping Israel off the map, while arming and largely controlling militias in Iraq.   (Guardian 09.04.07)

(7)
Fact Number Four: contrary to <u>some other</u> **British press** reports, America is 'talking' to Iran, or at any rate using diplomacy to deal with what is a nasty regime. (Telegraph 14.02.07)

The *press pack* also appeared to present a critical, outsider point of view of newsmaking, and, interestingly, was more frequent in the regional press than the national papers, suggesting that *press pack* offers a local view of the national newspapers, and, as illustrated in (9), signals the hostilities that emerge when the national press journalists are present on the regional press's territory.

(8)
As a result, **the press pack descended**, with about a dozen journalists and photographers all booking rooms in the Harris Hotel in the hope of interviewing the current 'cast-off' castaways. (Scotsman 08.03.07)

(9)
Worse, the question didn't come from **the press pack** that **trampled** all over Llanishen Fach primary school during his visit, but from a member of the school council. (Western Mail 21.04.07)

In the light of the findings briefly illustrated here, Research B drew the early conclusion that newspapers rarely discussed themselves, but predominately some (unfavourably evaluated) others. However, although this initially appeared to be a clear example of dissonance, it could also be interpreted as an instance of convergence as although Researcher A concluded that the press were frequently auto-referential, she also focussed on the way in which the journalists frequently absolved themselves of responsibility for the newsmaking process, and as briefly highlighted above, this discussion of the media from an external, outsider's perspective, with no recognition of the author's membership of the trade, would certainly seem to fit in with that view, which is, therefore, further discussed under the heading *convergent findings* below.

5.2 Convergent findings

5.2.1 'Power without responsibility'

One way adopted in order to delegate responsibility is portraying news as something that just comes into being by itself, rather than acknowledging its being crafted. News is made (see Fowler 1991 or Hartley 1982) and journalists are the material makers '[n]ot government officials, not cultural forces, not 'reality' magically transforming itself into alphabetic signs, but flesh-and-blood journalists literally compose the stories we call news' (Shudson, 1997: 8).

In the corpus we find a variety of expressions related to newsmaking where news is presented as an unavoidable self-crafting entity, independent from the reporters' intervention. Items such as: *headline(-)grabbing* (42 occurrences) / *grab the headline* (29 occurrences) / *headline(-)grabber* (7 occurrences), or *make / hit /dominate … the headlines / the news /front page* (180 occurrences).

In the *Telegraph* we find a series of examples of this pattern related to a specific piece of news, i.e. a tax-cut (see examples (10-13) below; the news is so headline-grabbing that the very 'grabbing' grabs the headline (11)).

> (10)
> This is a retrograde step and will do nothing to enhance confidence in pensions. Mr Brown has once again succeeded in burying the bad news behind **headline-grabbing** tax reductions. (Telegraph 24.03.07)
>
> (11)
> **Headline-grabbing** Budget that leaves the Tories floundering (Telegraph 22.03.07)
>
> (12)
> Brown is a Chancellor who can announce a **headline-grabbing** cut in income tax, from 22 pence in the pound to 20 pence, but which is really part of a cunning personal tax package that results in households paying about pounds 340m more tax in 2008-2009. (Telegraph 22.03.07)
>
> (13)
> Mr Brown will claw back nearly all the cost of his **headline-grabbing** move by scrapping the 10pc lowest rate of income tax for earned income and pensions. (Telegraph 24.03.07)

A close analysis of concordance lines shows how items such as the ones listed above achieve two types of delegation of responsibility, which we could label the 'spin-factor' and the 'mechanical newsworthiness factor'. Some events become news either as a result of the clever planning of politicians (as in the tax-cut examples) or by just automatically blooming out in print (see examples (14-16)).

> (14)
> Murders and transport stories **scream from the front page**: 'What the fog is the matter?' is one headline. (Guardian 05.02.07)

(15)
The prison system is in crisis. Our jails are bursting with convicts and crumbling with age. At least, **that's what the headlines tell us.** (Guardian 12.03.07)

(16)
The average British soldier, if killed in Iraq, is lucky to **make the news-in-brief columns** these days... (Guardian 27.04.07)

In our corpus, loud and misleading self-crafting news is something real-journalists despise. Grabbing the headlines or hitting the news is not something that depends on 'us' journalists and it is not a good thing. Negative evaluation is often made explicit in discourse and is, for instance, signaled by the frequent collocation with the adversative 'but' (examples (17-19)), followed by statements about the rightness or legitimacy of the sort of news that makes itself.

(17)
The cuts in income and corporation tax rates may **grab the headlines but** arguably the main macro-economic interest in the Budget lies in the public spending totals for the next three years. (Guardian 22.03.07)

(18)
The more garishly liveried sporting model will doubtless **grab the headlines, but** the two new Megane GTs - one petrol, one diesel - will be the volume sellers. (Guardian 10.02.07)

(19)
David Cameron drops his Thatcherite baggage just as Tony dropped Old Labour. Cameron can **grab the headlines but** he will not save the next generation (Guardian 02.02.07)

Ultimately, news does not seem to be the product of a process of professional selection and composition, but rather the effect of collision. It happens that events wandering through the air periodically *hit* the papers (see concordance lines in Table 2).

Table 2. Concordancing *hit the headlines*

The village of Trecynon	**hit the headlines**	last month when local vicar Father Paul Bennett was murdered on the doorstep of his vicarage.
The Mosquito alarm first	**hit the headlines**	in 2005 for using an annoying ultra-high tone, audible only to youngsters, to stop youths congregating
The Grosvenor Group	**hit the headlines**	earlier this year when its exclusive 77 Grosvenor Street property set a record for the most expensive
And the issue	**hit the headlines**	recently in the bird flu incident at the Bernard Matthews factory in Suffolk,
FRANCE Jacques Brunel	**Hit the headlines**	last November when he criticised the way New Zealand were treated by referees
The former stripper	**hit the headlines**	ever since tying the knot with an oil billionaire sixty three years her senior.
The Royal Free	**hit the headlines**	in 1998 when it provided a platform for Dr Andrew Wakefield to launch his campaign against the

		MMR
The original deal	**hit the headlines**	<u>after</u> Bank of Ireland pulled pounds 7m of financing for the deal following protest from its customers.
Mr Martin	**hit the headlines**	<u>earlier this week when</u> he single-handedly fought off a gang of muggers at a south London bus stop.
His Delta Three fund	**hit the headlines**	<u>last year when</u> it bought UK nursing-home operator Four Seasons Healthcare for pounds 1.4bn.
THE PUNK MALCOLM McLaren first	**hit the headlines**	<u>when</u> he and Vivienne Westwood, the designer, opened a London clothing shop called Let It Rock.
The show <u>recently</u>	**hit the headlines**	as one of a raft of programmes which have suffered apparent irregularities over premium rate phone
<u>Later the same year,</u> he	**hit the headlines**	again <u>when</u> it was disclosed that Nicola Stewart, a former prostitute and drug addict, had moved into his
One solution that has <u>recently</u>	**hit the headlines**	is the use of an ultrasonic deterrent aimed at young people.
The goats	**hit the headlines**	<u>last month when</u> someone apparently tried to poison them.
Then, <u>last summer,</u> the debate	**hit the headlines**	again <u>when</u> Dr Ian Walker, a traffic psychologist from Bath University, completed a study that
Mr Bonderman	**hit the headlines**	<u>over the weekend</u> following the agreed take-over of the US energy firm TXU Corp;

5.2.2 Good vs. Bad journalism

Analysis of collocates and concordance lines showed quite clear patterns of evaluation, defining two contrasting kinds of journalism: good journalism and bad journalism. Good journalism emerges as being *critical, committed, investigative, serious, independent, effective, objective, professional,* etc. and it *can achieve, can make a difference, has a battle to fight.* Conversely bad journalism is described as *inconsequential, lazy, dodgy* and it is identified as *tabloid / popular* journalism. The papers talk for example of *the murky practices of tabloid journalism* and the readers comment that *[t]his sort of journalism is surely more suited to a tabloid paper than a serious paper like the Scotsman.*

All the newspapers included in our corpus are broadsheet papers and they tend to strongly signal their difference from tabloids. When writing about their profession journalists do not avoid criticism, but they redefine the context, distinguishing between 'us' and 'them' and criticism is directed at a pseudo-journalism that is not 'us'. This, according to media scholars, serves the purpose of maintaining the 'definitional control of the field' (Dahlgren and Sparks 1992: 2) and of reproducing 'consensus about who [is] a 'real' journalist, and what (parts of) news media at any time would be considered examples of 'real' journalism' (Deuze 2005: 444).

(20)
As an honorary journalist, I feel a duty to turn the spotlight on my colleagues and shake the industry up from within. Like many teenagers, I watched All the President's Men and vowed to commit my life to uncovering political corruption. [...] Sadly, over the years, the word 'journalist' has become debased,

synonymous in my mind with fictions and half-truths [...] But as an admirer of what great **journalism** can achieve, I do wonder how we've reached a point where the Daily Mail... (Guardian 23.03.07)

By distancing from bad journalism (see also Frank 2003 on 'reflexive media criticism'), journalists reaffirm the values of real-journalism, they retain the virtues of the profession and divert from themselves responsibility for its vices.

5.3 Complementary findings

The research question: 'how do journalists talk about themselves/each other and their profession in a corpus of British media texts?', can be approached in many different ways, which probably is one of the main problems with this experiment. Yet, a narrower question may also restrict the range of angles we can adopt to interrogate the corpus. It is plausible that by restricting the breadth and reach of the topic and, consequently, the freedom of action of the researchers, we would also have limited the extent of 'complementary findings', which are discussed in this section.

In order to see how journalism was explicitly talked about, Researcher A retrieved collocates (R5 L5) of *journalism* and identified a list of topics related to the profession by the professionals. The list does not aim to be exhaustive, very often concepts were blurred and topics overlapping, therefore the chosen categories are subjective and reductive. Nevertheless they could help in getting an idea of what the main concerns of journalists are with respect to what they do.

The majority of collocates are found within the semantic field of education (e.g. University, schools, professor, training, etc.). A closer look at the concordance lines shows that the discussion is rarely about training itself, while academic titles and references to the academia are used in order to qualify people, i.e. they tend to serve the purpose of legitimation.

Journalism is also talked about in terms of its 'future' (as in (21)) and of its mission and ethics (as in (22)).

> (21)
> Newspapers are worried stiff that their future may be extinguished by the migration of news and advertising to the web and the rise of user-generated content usurping the hallowed tradition of in-house journalism. (Guardian 22.02.07)

> (22)
> I once spoke to a journalist who had covered the war in Bosnia in the early 1990s. He said that he and his colleagues kept heading into harm's way, because they believed that once the world knew of the horrors they had witnessed, the world would be stirred to act. They filed their reports and waited. Soon enough, they understood. The world knew what was going on - and yet it did nothing. For some of those reporters, this experience broke their faith in the power of journalism. (Guardian 14.03.07)

We get both celebration of the profession (23) and criticism for what we earlier called 'bad journalism' (24).

(23)
As even the most casual browse of photo bylines will reveal, most journalists have faces made for. . . journalism. Watergate reporter Woodward jowly, unshaven, maudlin-looking was no exception. (Guardian 17.03.07)

(24)
Everyone knows journalism can be a heartless trade. Last year it was the conductor Sir Simon Rattle's turn to be the victim of its relentless 'build 'em up, knock 'em down' syndrome. (Guardian 10.03.07)

And Researcher A had the impression that journalists are highly self-referential and that journalism appears to be 'an intensely reflexive occupation, which constantly talks to and about itself' (Aldridge and Evetts 2003: 560). Journalism is for example invoked as a comparative term for other realities (25-28).

(25)
Work has become more unpredictable. The kind of sudden deadlines that are typical of journalism are now symptomatic of a wide range of professions. (Guardian 01.03.07)

(26)
Since the golden rule of politics, like journalism, is repetition, let me begin by reminding everyone how far the Cameroons have moved. (Guardian 06.03.07)

(27)
For, dare I say it, writing - or editing, which is what much of journalism is about - for the theatre can be much more rewarding than contributing to the pages of newspapers, or to current affairs programmes on television or radio. So much gets lost and forgotten through the small screens and in the airwaves. (Guardian 16.04.07)

(28)
Surely fiction has as much right to explore such terrible contemporary phenomena as journalism has to report them. (Guardian 20.04.07)

As a brief example of complementary findings from the research carried out by Researcher B, we might look very briefly at the use of metaphor in the representation of journalists, which was an area of particular interest to the second researcher involved in the project. Here we will just focus on one group of metaphors. As noted above, the *press pack* was concordanced to see how collective nouns referring to journalists were used evaluatively, but the concordances also illustrated the way in which the press pack frequently becomes a metaphorical pack of (news) hounds as illustrated in the concordance lines in Figure 4:

Figure 4. Concordance lines illustrating the media as hounds metaphor

teld	A woman who can turn the **hounds** of Fleet Street into **tail-wagging poodles**
gua	public enemy number one - **hounded by the media**, betrayed by friends, the
teld	Woods **dogged** by **barking media**
gua	But the media have been **barking up the wrong tree** when blaming the ICC
gua	Don't be fooled by the louche image of the **tabloid hound** and models' squire

The news hound metaphor becomes extended and more strongly negatively evaluated in the JOURNALISTS AS BEASTS metaphor:

Figure 5. Concordance lines showing co-occurrences of media and beast*

gua   lised in populating the hinterland of the **media** with fairly terrifying **beasts**.
gua   ear from some of the big **beasts** in the **media** jungle about how they are respo
gua   'You can't control the **media beast**. What happens out there,
gua   free pass be given to the **beasts** of the **media**, so often happy to wave the flag

There were also six references to the media's *feeding frenzy* as illustrated in (29) in which, once again, some other set of journalists are evaluated unfavourably.

(29)
It's a fortunate British government that reaches the end of a third term without provoking the usual **media** and popular **feeding-frenzy,** in which ministers are jeered everywhere as obvious villains and fools, and the nation begins to enter into the collective delusion that the other lot are bound to be better. (Scotsman 28.04.07)

Further analysis showed that the motif of the media as feeding and eating was quite common and consistently used to negatively evaluate as in the examples below show:

(30)
When **the media** eventually gets around to **eating itself**, it will sound like this. (Scotsman 10.02.07)

(31)
It is with mounting irritation that commentators observe the rise of the Tory party in the opinion polls, and the continuing failure of David Cameron to make the kind of rash policy announcements that **the media** can **devour**, distort, denounce and destroy. (Guardian 06.03.07)

(32)
The majority of **the media** is corrupt, and corrupting. It **feeds off** conflict, crime and scandal. (Guardian 15.02.07)

However, although the ferocity of the media is consistently negatively evaluated, the metaphors also illustrate the dangers of domesticity as shown in (33).

(33)
The story was never printed. It showed, she said, that the Cameron Conservatives were **feeding** concocted anti-Bush stories **to the press**. (Telegraph 15.03.07).

The MEDIA AS HOUNDS metaphor is certainly not new, nor unexpected, though it is interesting to see the extensions and range of realizations, and of course, to see that the MEDIA AS BEASTS metaphor clearly pre-dated Tony Blair's June 2007 comments about 'feral beasts', and, most revealing, the metaphor was

frequently used by the media itself, although not, it should be noted, in any kind of self-reflective way.[5]

5. Conclusions

In conclusion we may say that, as researchers, we need to be aware that triangulation is richer than just the use of a multi-method approach, and offers more than the single objective of convergence. Given this wider range we would say that triangulation offers both analytical depth and creative potential in all its variations. As Jick (1979) noted early in the career of triangulation, it seems to demand imaginativeness from researchers, and should not only be about fine-tuning instruments but also stimulating creative research.

However, we would also caution that the implementation of triangulation within a research study in no way guarantees greater validity, nor can it be used to make claims for 'scientific' neutrality (and perhaps, in the social sciences, we would add, not *should* it). Finally, as Farmer at al (2006: 392) warned, we need to be aware that 'triangulation is only as strong as the study's underlying theoretical, methodological, and analytical paradigms and the researchers' skills and abilities. It is only one way of ensuring rigor'.

With reference to our particular project, we found it a fruitful experiment which may be (informally) repeated in future work. However, in terms of experimental design we would have needed a much more tightly focussed secondary research question in order to investigate with greater rigour the extent of the influence of the researcher, and it would have been more productive to have had a greater number of researchers involved in the project.

---

[1]   The *IntUne* corpus is a multilingual corpus of media texts compiled for the European project *Integrated and United? A Quest for Citizenship in an 'Ever Closer Europe'* (*IntUne*) www.intune.it.

[2]   eXtensible Mark-up Language (http://www.w3.org/XML/)

[3]   Text Encoding Initiative (http://www.tei-c.org/index.xml)

[4]   *Xaira* (XML Aware Indexing and Retrieval Architecture) developed at Oxford University is the XML version of *Sara*, the software originally developed for interrogating the British National Corpus. For further information: http://www.oucs.ox.ac.uk/rts/xaira/

[5]   Blair's comment that 'the fear of missing out means today's media, more than ever before, hunts in a pack. In these modes it is like a feral beast, just tearing people and reputations to bits' was made in a speech on 12 June 2007.

References

Aldridge, M. and Evetts, J. (2003) Rethinking the concept of professionalism: the case of journalism. *British Journal of Sociology* 54 (4): 547-64.

Baker, P. (2006) *Using Corpora in Discourse Analysis*. London: Continuum.

Baker, J.P. (2007) Discourses of refugees and asylum seekers in the UK press, 1996-2006: Full Research Report. ESRC End of Award Report, RES-000-22-1381. Swindon: ESRC.

Baker, P., Gabrielatos, C., Khosravinik, M., Krzyzanowski, M., McEnery, T. and Wodak, R. (2008) A useful methodological synergy? Combining critical discourse analysis and corpus linguistics to examine discourses of refugees and asylum seekers in the UK press. *Discourse & Society* 19 (3): 273-305.

Bowker, L. and Pearson, J. (2002). *Working with Specialized Language: A Practical Guide to Using Corpora*. London: Routledge.

Bryman, A. (2006) Editor's introduction: Mixed methods research. In A. Bryman (ed.), *Mixed Methods*. London: Sage, pp. XXV-LII

Bryman, A. (2003) Triangulation. In *Encyclopedia of Social Science Research Methods*. London: Sage.

Dahlgren, P. and Sparks, C. (1992) *Journalism and Popular Culture*. London: Sage.

Davies, N. (2008) *Flat Earth News*. London: Random House.

Denzin, N. K. (1970) *The Research Act in Sociology*. Chicago: Aldine.

Deuze, M. (2005) What is journalism? Professional identity and ideology of journalists reconsidered. *Journalism* 6 (4): 442-64.

Erzberger, C. and Prein, G. (1997) Triangulation: Validity and empirically-based hypothesis construction. *Quality & Quantity* 31: 141-154.

Farmer, T., Robinson, K., Elliott, S.J. and Eyles, J. (2006) Developing and implementing a triangulation protocol for qualitative health research. *Qualitative Health Research* 16: 377-394.

Frank, R. (2003) 'These crowded circumstances'. When pack journalists bash pack journalism. *Journalism* 4 (4): 441-58.

Fowler, R. (1991) *Language in the News: Discourse and Ideology in the Press*. London: Routledge.

Hartley, J. (1982) *Understanding News*. London: Methuen.

Hardt-Mautner, G. (1995) 'Only Connect': Critical Discourse Analysis and Corpus Linguistics. UCREL Technical Paper 6. Lancaster: University of Lancaster.

Jick, T. D. (1979). Mixing Qualitative and Quantitative Methods: Triangulation in Action. *Administrative Science Quarterly*. 24 (4): 602-611.

KhosraviNik, M. (2008). Actors, actions, argumentations: Towards an amalgamation of CDA methodological categories in representations of social actors. Paper presented at CADAAD 08. Hatfield: University of Hertfordshire.

Koller, V. (2008) *Lesbian Discourses: Images of a Community*. New York and London: Routledge.

Leech, J. (2001) The state of the art in corpus linguistics. In K. Ajmer and B. Altenbeg (eds.), *English Corpus Linguistics: Studies in Honour of Jan Svartvik*. London: Longman. pp. 8-29.

Matheson, D. (2003) Scowling at their notebooks. How British journalists understand their writing. *Journalism* 4(2): 165-83.

Mautner, G. (forthcoming) Corpora and critical discourse analysis. In P. Baker (ed.), *Contemporary Approaches to Corpus Linguistics*. London: Continuum.

McEnery, T. and Wilson, A. (1996) *Corpus Linguistics*. Edinburgh: Edinburgh University Press.

Myrdal, G. (1970) *Objectivity in Social Research*. London: Gerald Duckworth & Company Ltd.

O'Halloran, K. and Coffin, C. (2004) Checking overinterpretation and underinterpretation: Help from corpora in critical linguistics. In C. Coffin, A. Hewings, and K.A. O'Halloran, (eds.), *Applying English Grammar: Functional and Corpus Approaches*. London: Hodder Arnold. pp. 275-297.

Orpin, D. (2005) Corpus Linguistics and Critical Discourse Analysis. Examining the ideology of sleaze. *International Journal of Corpus Linguistics* 10 (1): 37-6.

Partington, A. (2004) Corpora and discourse: A most congruous beast. In A. Partington, J. Morley and L. Haarman (eds.), *Corpora and Discourse*. Bern: Peter Lang. pp. 11-20.

Partington, A. (2006) Metaphors, motifs and similes across discourse types: Corpus-assisted discourse studies (CADS) at work. In A. Stefanowitsch and S. Th. Gries (eds.), *Corpus-Based Approaches to Metaphor and Metonymy*. Berlin: Mouton de Gruyter. pp. 267-304.

Partington, A. (2009) Evaluating evaluation and some concluding thoughts on CADS. In J. Morley and P. Bayley (eds.), *Wordings of War. Corpus Assisted Discourse Studies on the War in Iraq*. London: Routledge.

Perlesz, A. and Lindsay J. (2003) Methodological triangulation in researching families: Making sense of dissonant data. *International Journal of Social Research Methodology* 6: 25-40.

Rossman, G.B. And Wilson, B.L. (1994) Numbers and words revisited: Being 'shamelessly eclectic'. *Quality and Quantity* 28(3): 315-27.

Scott, M. (2008) WordSmith Tools version 5. Liverpool: Lexical Analysis Software.

Scott, M. and Tribble, C. (2006) *Textual Patterns: Key Words and Corpus Analysis in Language Education*. Amsterdam: John Benjamins.

Shudson, M. (1997) The sociology of news production. In D.A. Berkowitz (ed.), *Social Meanings of News: A Text-reader*. London: Sage.

Sinclair, J.M. (1991) *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.

Sinclair, J.M. (1999) A way with common words. In H. Hasselgård and S. Oksefjell (eds.), *Out of Corpora: Studies in Honour of Stig Johansson*. Amsterdam: Rodopi. pp. 157–179.

Sinclair, J.M. (2004) *Trust the Text: Language, Corpus and Discourse*. London: Routledge.

Stubbe, M., Lane, C., Hilder, J., Vine, E., Vine, B., Marra, M., Holmes, J. and Weatherall, A. (2003) Multiple discourse analyses of a workplace interaction. *Discourse Studies* 5: 351-388.

Stubbs, M. and Gerbig, A. (1993) Human and inhuman geography: On the computer-assisted analysis of long texts. In M. Hoey (ed), *Data, Description, Discourse. Papers on the English Language in honour of John McH Sinclair on his Sixtieth Birthday*. London: Harper Collins. pp. 64-85.

Stubbs, M. (1995) Collocations and semantic profiles: On the cause of the trouble with quantitative studies. *Functions of Language* 2 (1): 23-55.

Stubbs, M. (1997) Whorf's children: critical comments on critical discourse analysis. In A. Wray and A. Ryan (eds.), *Evolving Models of Language*. Clevedon: Multilingual Matters. pp. 100-16 [Reprinted in M. Toolan (ed.) (2002). *Critical Discourse Analysis*. London: Routledge, pp. 202-18].

Stubbs, M. (2006) Corpus analysis: The state of the art and three types of unanswered questions. In G. Thompson and S. Hunston (eds.), *Style and Corpus: Exploring Connections*. London: Equinox, pp. 15-36.

Thurmond, V.A. (2001) The point of triangulation. *Journal of Nursing Scholarship* 33 (3): 253-258.

Titsher, S., Meyer, M., Wodak, R. and Vetter, E. (2000) *Methods of Text and Discourse Analysis*. Thousand Oaks, CA: Sage.

Van de Mieroop, D. (2005) An integrated approach of quantitative and qualitative analysis in the study of identity in speeches. *Discourse & Society* 16: 107-30.

Wodak, R. (2006) Critical Linguistics and Critical Discourse Analysis. In J. Verschueren and J. Östman (eds.) *Handbook of Pragmatics*. Amsterdam/Philadelphia: John Benjamins. pp. 1-24.