Personalisation in Mass Media Communication

British online news between public and private

Daniela Landert

University of Zurich

Author submitted manuscript

This is the author submitted manuscript version of the book that was published as:

Landert, Daniela. 2014.*Personalisation in Mass Media Communication: British online news between public and private* (Pragmatics & Beyond New Series 240). Amsterdam/Philadelphia: John Benjamins.

When citing this text, use the published version for the correct page numbers.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Contents

Author submitted manuscript

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Author submitted manuscript

List of Tables

Author submitted manuscript

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

List of Figures

Author submitted manuscript

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Author submitted manuscript

Author submitted manuscript

Acknowledgements

There are many people to whom I am indebted for their contributions to this project. First of all, I would like to thank Andreas H. Jucker. I could not have had a more supportive and inspiring supervisor. His approach to linguistics and to academic life more generally has profoundly influenced my own. I would also like to thank Marianne Hundt, who agreed to act as second reviewer and who provided many detailed and helpful comments on this study, as well as Christa Dürscheid, who shared her thoughts on computer-mediated communication in various meetings of the Pro*Doc programme "language as a social and cultural practice".

Earlier drafts of sections of this book were read by Nicole Studer-Joho, Julia Weber, and Gerold Schneider, who offered insightful comments and criticism. In addition, two anonymous reviewers gave me detailed and constructive feedback on the entire manuscript, and Shane Walshe's help was invaluable in polishing the language before publication. I would also like to acknowledge the help of Anna Weber, Ursula Weber and Gabriel Weber in typing the data from the *Times* from 1985. Additionally, Anna Weber and Julia Weber assisted me with the reliability testing of several categorisations in chapters 6 and 7. My thanks also go to Anita Fetzer, the editor of the Pragmatics & Beyond New Series, and to the team at John Benjamins, especially Isja Conen and Esther Roth for their very helpful assistance in the final steps of publication.

Working on a research project over several years is not possible without the support of friends and colleagues (a distinction, I might add, that has become increasingly blurred during this time). Manuela Neurauter-Kessels shared an office and research interests with me, was always happy to exchange ideas about online news sites, and taught me how to order beer in Dutch. Dank je wel! My colleagues at the English Department – especially those on the second floor of the PLH building – were patient listeners whenever I needed to talk to someone to clarify my thoughts. They were also indispensable in that they helped me keep my caffeine levels sufficiently high for productive writing. Marlise Landert will be almost as happy as I am (if not more) to hold this book in her hands, and she deserves special thanks for never getting tired of explaining the strange things I do in my professional life to everyone who asks her. Julia Weber has been there for me through all the ups and downs of this project and far beyond. She also played a crucial role, many years ago, when she suggested that I might find studying English more exciting than computer science. I guess she was right. Gianluca Miscione, finally, supported me in more ways than I can

Author submitted manuscript

name, such as by creating a most inspiring writing retreat for me whenever I needed one, by making me laugh even in the most unlikely moments, and – not least – by constantly reminding me of one of the most important steps in any research project: finishing it.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Author submitted manuscript

Author submitted manuscript

1. Introduction

From 17 December 2009 to 15 January 2010, the UK experienced a period of very cold temperatures and heavy snowfalls, which, according to the Met Office, was the "most widespread and prolonged spell of this type across the UK since December 1981/January 1982".[1] Unsurprisingly, this weather event attracted a great deal of media attention and it appeared prominently on the front pages of British newspapers and their online sites. News articles covered weather forecasts by the Met Office, reports about the traffic situation, warnings by the authorities and statistical information about the weather event and its consequences – in other words, they presented a great deal of factual information, abstracted from concrete events and issued through official sources. In addition to such impersonal accounts, however, news articles frequently contained stories, statements and surrounding material that provided more personal perspectives on the event. The following examples illustrate some of the ways in which the reporting of the heavy snow fall was personalised on British online news sites.

(1.1) WOMAN IN LABOUR WITH TWINS WALKS TO HOSPITAL IN BLIZZARD (mail-uk-100108)

(1.2) "I would have 13 exams in June instead of seven, this will obviously affect my results through NO FAULT of my own," wrote one on the Student Room website. Another said: "I started revising for my Jan exams since Sept ... and this is my last chance ... I do not wanna be doing my AS resits in May along side me A2s ... that's just too much pressure!!" (guardian-uk-100108)

(1.3) Are you in the affected areas? Send us your comments and experiences using the form below. Send your pictures and videos, telling us your stories, to yourpics@bbc.co.uk or text them to 61124. If you have a large file you can upload here. (bbc-uk-100106)

(1.4) ... a strange species, found only in the uplands of southern Shropshire and tamed after much struggle, as our picture shows, by our local snaterpillar whisperer... ninjawarrior (image caption appearing in a slide show with reader pictures on times-uk-100110)

(1.5) Villagers in remote Cow Ark, Lancs – cut off for three weeks since heavy snow on December 17 – were overjoyed yesterday when The Sun came to their rescue. Our mercy crew delivered bread, milk, cheese, ham and other vital provisions – including Britain's favourite newspaper. (sun-100108b)

As can be seen from the examples above, the personalisation of mass media communication took many forms. It was, for instance, realised through stories about individuals who were

[1] http://www.metoffice.gov.uk/climate/uk/interesting/jan2010, accessed on 19/3/2012.

Author submitted manuscript

directly affected by the severe weather, such as the pregnant woman mentioned in the subheadline of example 1.1. The news article in question contains a detailed account of her ordeal, complete with a picture of her and her partner together with one of the newborns. Example 1.2 comes from an article containing the experiences of a different group of directly affected individuals, namely students facing exam delays as a consequence of the weather. They are not presented in as much detail as the woman in the previous article, but their opinions are represented verbatim as direct speech quotes, including first person pronouns, colloquial forms (*wanna*) and emphasis (capitalisation of *NO FAULT*, double exclamation marks). Personalisation in examples 1.1 and 1.2 is, thus, directed at the news event, presenting it from the perspective of directly affected individuals, their experiences and emotions, which are sometimes represented in their own words. In contrast, personalisation in example 1.3 is directed at the audience. The online news site pretends to directly address individual readers, rather than a large, anonymous audience. Moreover, by asking readers to submit their comments, experiences and pictures, the text producers of *BBC News* invite their audience to engage in an interactive exchange that transgresses the unidirectionality of prototypical mass communication. Readers are not just asked to consume the news, but to contribute to it. This invitation is sometimes accepted. Example 1.4 presents the caption for an image showing four individuals posing on a large caterpillar made of snow. The image was included in an image gallery of reader pictures, accompanying a news article on the *Times Online* and shows how readers can literally become visible in personalised news reporting. Personalisation can not only affect news events and the audience, but also text producers and their media product. In example 1.5, text producers (*The Sun, Our mercy crew*) are turned into news actors who actively and visibly participate in events, rather than being an impersonal and objective reporting voice.

In this study, I propose a model that integrates these different aspects of personalisation. I argue that personalisation manifests itself in various forms and on various levels of mass media communication. Features of the communicative setting can create personalisation by giving readers more visibility and by increasing audience involvement, for instance by letting readers interact with text producers and with each other. On the level of content, the presence of private actors and private topics leads to more personal perspectives on news events. Last but not least, the linguistic realisation contributes to personalisation through features such as direct speech and first and second person pronouns, all of which create linguistic immediacy. As different as these features are, the common denominator of all of them is that one or more of the entities in the mass communication process – the news event as the message, the text producer as the sender, and the audience as the recipient – are presented in a personal rather than in an abstract way. I, therefore, suggest that all these features are related and that they can all be understood as manifestations of personalisation.

I apply this model to data from five British online news sites, the *Times Online*, the *Guardian*, *BBC News*, the *Mail Online* and the *Sun*. However, personalisation of mass media is not restricted to online news. In fact, many of the findings of this study could

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

probably be replicated by investigating the printed counterparts of the analysed online news sites. One of the reasons for studying personalisation in online rather than print news lies in its different communicative setting. While interaction between text producers and their audience is possible in older forms of news media (e.g. through letters to the editor), the Internet opens up entirely new dimensions for audience feedback and user-generated content. In addition, the online setting also provides more options for combining texts with images and videos, which can increase audience involvement in a different way.

A second reason for studying online news sites rather than print newspapers relates to their increasingly prominent role in accessing news. More and more portable electronic devices, such as smart phones and tablet computers, allow the consumption of online news from everywhere and around the clock. In the case of breaking news, continuously updated online sites have a particularly decisive advantage over other forms of mass communication. The importance of online news is also reflected in terms of production, where redesigns and re-launches of online news sites speak of considerable investments by media organisations. Three of the five investigated sites launched quite substantial redesigns within less than two years after I had collected my data. The most comprehensive redesign occurred in the case of the *Times Online*, which was replaced by the new platform *The Times*, a site restricting most of its content to paying customers.

However, despite the importance of online news sites in everyday life, they have not yet been studied in as much detail as their printed counterparts. Given the unpredictable persistence of online data, this is very regrettable. In contrast to printed newspapers, which can be made accessible for research long after their publication (albeit often only through the investment of a great deal of resources), online news sites do not have any foreseeable durability. While the texts of articles may in some cases be archived by the news site or Internet archiving services, the layout and many interactional elements are irretrievably lost as soon as a news site changes its design. We are currently witnessing an important media change in which the online news site is establishing itself as a new form of mass communication with its distinct characteristics and conventions. Unfortunately, this change is going largely undocumented. By studying personalisation with data from online news, I would, therefore, also like to throw light on one particular point in this process, documenting the state of five British online news sites in January 2010.

The aims of this study can thus be grouped into three different categories: theoretical, methodological, and descriptive. The main theoretical aim of this study is to develop a model that can account for the various realisations of personalisation in a systematic way. Personalisation cannot be reduced to a single feature of mass media texts, but it involves a broad range of characteristics which closely interact with each other. While each of them can and needs to be studied in their own right, the full extent of how news is presented in a personalised way only becomes visible when their combination and the intricate interplay between them is taken into account. I therefore propose a framework that integrates perspectives on the communicative setting, the content and the linguistic realisation of mass media texts. Personalisation is thereby related to shifts between the private and the

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

public sphere, notions of impersonality and objectivity in news reporting, and the role of linguistic features typical of conceptually oral language. At the same time, the model also accounts for the personalisation of all entities involved in the mass communication process: news events and news actors, the audience, and text producers and their media institutions.

From a methodological perspective, the study of online data is still relatively new and linguistic tools and methods to work with these data are still being developed. Online data is characterised by fluidity, non-linearity and multimodality (Jucker 2003), which means that the established practices of working with print data cannot in all respects be applied in a straightforward way. There have been suggestions for practices advancing the compilation of multimodal corpora (e.g. Baldry and Thibault 2006; Baldry 2007; Bateman 2008), but so far there has been little experience with putting these into practice for investigating genuine research questions. Furthermore, these approaches tend to aim for models that allow the representation of multimodal documents with a great deal of detail and in all their complexity, which means that the models are based on several layers of extensive descriptions of the data. As a result, they are difficult to apply to large amounts of data, which is also reflected in the fact that the corpora on which the models are based contain only a small number of documents (see also Santini 2010). It is certainly beyond the scope of this study to provide a comprehensive and universally applicable solution for compiling corpora of multimodal (online) data. Rather than providing a theoretical model for multimodal documents, I want to find practical solutions for storing and working with online news data, tailored to the specific research questions of this study. Nevertheless, by making my reflections transparent regarding data selection, collection and analysis, I aim to contribute to the ongoing search for methodologies and best practices in this area.

From a descriptive point of view, I am interested in finding out how personalisation is realised on online news sites. A first aim is to establish an inventory of personalisation features, document their effects and investigate how they are used. Of course, this inventory will necessarily be incomplete. While I discuss some of the most salient features on each level (communicative setting, content, and linguistic realisation), there are many more that could be taken into account. To name just two examples, the integration of news sites with social media platforms like Twitter is one of the aspects of the communicative setting which would certainly deserve more attention. Likewise, stance expressions are not covered in the current study, despite the fact that they would pose an interesting topic for studying linguistic personalisation features.

As a second descriptive aim, I want to characterise the personalisation strategies of each of the investigated sites in a contrastive way. Differences between up-market news sites, on the one hand, and mid- and down-market sites, on the other, will in particular be discussed. This continues a strong research tradition dealing with market differences of (mostly print) newspapers, variously also referred to as differences between broadsheet and tabloid newspapers, or between quality papers and the popular press.

Author submitted manuscript.

Author submitted manuscript. Consult published version for correct page numbers!

Discussing personalisation in online news is hardly possible without at least the occasional side glance at non-news articles, for news sites – like print newspapers – are far from restricted to only (hard) news articles. Depending on the site and its market orientation, they contain smaller or larger amounts of "soft news" such as feature articles with background stories on individuals and events, fashion reports and health advice. All news sites also contain opinionated texts like editorials, columns and letters to the editor. Both soft news and opinionated texts are usually much more personalised than news articles and one way to understand the personalisation of news is to look at it as an increase of soft news and opinion. It is for these reasons that a limited amount of data from soft news articles and opinion columns is included in this study. These data were only collected for two of the sites, the *Times Online* and the *Mail Online*, and they will be consulted only at selected points of the analysis.

Likewise, it is difficult to completely neglect the diachronic dimension of personalisation, given that personalisation is often understood as an increasing presence of personal aspects over time. While it is far beyond the scope of this study to provide a comprehensive analysis of the diachronic development of personalisation over any particular time frame, I still felt the need for a comparative data set that provides at least a very rough and highly selective indication of the main trends of personalisation over time. This data set consists of news articles from the printed *Times* from 1985. It was, thus, published a quarter of a century before I collected the online data, at a time before the World Wide Web came into existence and when access to the Internet was not yet widespread. The diachronic dimension would certainly deserve further attention, in particular by including earlier print editions of mid- and down-market newspapers. However, the comparison between the data from the *Times Online* and the printed edition of the *Times* from 1985 will provide interesting first insights into the recent development of personalisation.

I believe that in order to approach the many facets of personalisation, more than one perspective is needed. Instead of focussing on one aspect of personalisation and exploring it in a comprehensive and detailed manner, I will therefore pursue a number of different avenues into the object of my study. I will investigate a range of different features of personalisation and I will compare them across various dimensions of my data. I will also repeatedly switch between a birds-eye view of my data and detailed analyses of individual occurrences. The persuasive power of my study does not lie in one particular result, but instead in the cumulative evidence from a large range of different observations.

Chapter 2 develops a model of personalisation that integrates features of the communicative setting, the content and the linguistic realisation of mass media texts, and which can account for personalisation affecting news events and actors as well as text producers and readers. This model is developed against the background of various strands of research. In the context of research on mass media communication and journalism more generally, the relation of personalisation to norms of objectivity, on the one hand, and news values, on the other, is of particular interest. Research on linguistic features of conceptually oral language realised in the written mode is also consulted to relate personalisation to linguis-

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

tic immediacy. Furthermore, the connection between personalisation and the blurring of boundaries between the private and the public sphere is investigated, a topic in particular discussed in the context of communication studies and political science. After having thus introduced the model and its theoretical background, chapter 3 provides an overview of the data. It also discusses problems of working with online data and presents the principles and technologies that were applied in this study.

Chapters 4 to 8 are each devoted to one feature (or a small set of related features) of personalisation. Each chapter contains a theoretical section introducing the feature and creating a link between previous research and the model of personalisation presented in chapter 2. This is followed by the results and analysis of the empirical study based on my own data. Due to the different nature of the features, the empirical studies take different forms in the various chapters. In some cases, quantitative evaluations provide quite fruitful results, for instance for the analysis of direct speech in chapter 7. In other instances, case studies are needed to gain a deeper insight, for instance in chapter 5 on visual elements. The chapters can further be classified into three groups. Chapter 4 on feedback and interaction and chapter 5 on visual elements deal with features that closely depend on the communicative setting of online news sites. These – in particular chapter 4 – are therefore also the chapters most concerned with the impact of the online environment. Chapter 6 on news actors and private topics deals with the content dimension of news articles. By focussing on different types of news actors (i.e. individuals appearing in the news) and their presentation in news articles personalisation is shown to appear in the form of private content, i.e. information that does not directly affect the interests of the public at large. Chapters 7 and 8, finally, are devoted to linguistic features of personalisation, namely direct speech and first and second person pronouns. Given that many instances of first and second person pronouns appear within direct speech and given that their reference depends on whether or not they occur in direct speech, chapter 8 closely builds upon the results of chapter 7.

If I say that each of the chapters can be associated with one of the three dimensions of communicative setting, content, and linguistic realisation, this does not mean that this relation is exclusive. In fact, it is one of the aims of this study to show that these three dimensions are closely interrelated. Throughout the five chapters, I will discuss cases in which personalising features affect more than one level. For visual elements, not only the communicative setting is of interest, but also their content. Similarly, the sources of direct quotes provide a linking element between the content of news articles and their linguistic realisation. First and second person pronouns with reference to text producers and readers can also simulate a direct interaction and are thus related to interactive features of the communicative setting. Finally, user-generated content depends on an interactive environment while often leading to private content and linguistic immediacy, thus linking all three dimensions. The relations between these dimensions are, therefore, at the core of this study and also of its conclusion in chapter 9.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

2. Personalisation in mass media

Like any form of communication, mass media communication involves three entities, the sender, the message, and the recipient. The sender of mass media communication typically consists of a group of text producers including journalists, editors, type setters, correctors, etc. The message contains some form of information, such as reports about news events in the case of news media. The recipient is the audience, i.e. the readers of the mass media text. The three entities are present in mass media texts to various degrees and in various forms. For news media, at least, the entity which clearly has the strongest presence in the text is the message, i.e. the news event. However, text producers and the audience often are present as well. In the case of text producers, a news report may contain references to the news organisation (e.g. *the Guardian has uncovered evidence that ...*), to correspondents (e.g. *our correspondent said ...*), and to the journalist (e.g. *The camp's senior doctor [...] told* ***me*** *...*).[1] If the surrounding context of the news report is taken into account, references to text producers are quite common, for instance in the form of article bylines and contact information. A similar situation is given for the audience. Sometimes the readers are directly addressed in news reports and they are present in surrounding material such as letters to the editors, online comments and other audience contributions. In other words, mass media texts do not only contain the message, but they also include information about the text producers and the audience.

Personalisation can be defined as the foregrounding of persons who are part of any of the three entities of mass media communication. If news actors – i.e. individuals appearing in the news, such as decision makers, witnesses and individuals who are directly affected by a news event – are given a high presence, this personalises the news event. If readers are given a high presence, this means that the audience is personalised. Finally, if journalists, correspondents and news organisations are given a high presence in mass media publications, this amounts to the personalisation of text producers. These three forms of personalisation are not mutually exclusive; indeed, strong personalisation of several entities can often be observed in the same text.

However, the question is not only how much presence is given to each entity, but also in which way the entities are represented. The representation of each of the entities can be placed anywhere on a scale from highly abstract to highly personal. At the most abstract end of the scale, news events are represented as abstract processes (e.g. developments

[1] See chapter 8.5 for examples in context and full references.

Author submitted manuscript

at the stock markets), readers as an anonymous mass audience, and text producers as abstract media organisations (e.g. *The Times*). In a slightly more personalised way, news actors, readers, and text producers can be referred to as groups of individuals, defined through some common attributes. An example of such a group of news actors would be the members of a jury; similarly, readers living in a specific part of the country build a possible group of audience; and, finally, the editors of a news publication can form a group of text producers. Even more personalised are representations of individual named actors in professional roles, such as the judge of a news-worthy court case, a medical professional who writes a letter to the editor, in which she presents her professional opinion on a health topic, or one specific editor of a newspaper. At the most personalised end of the scale are representations of individuals in their private roles. Highly personalised news actors are often found in the role of victims who report how they personally experienced news events, focussing on their emotions and the effects on their private lives. Similarly, reader responses such as letters to the editors, online comments, and audience material often contain very private information, which results in a personalisation of the reader and, thus, the audience. For text producers, this can in particular be found in opinion columns, where columnists sometimes write about experiences they have had in their private lives, for instance as fathers interacting with their children.

These examples show that personalisation can be located on various levels of the communication process: the communicative setting; the content; and the linguistic realisation. On the level of the communicative setting, elements that allow interaction with and among readers provide opportunities for personalising the audience. On the level of content, a story can be told from the point of view of directly affected individuals, emphasising their personal experiences and emotions. On the level of linguistic realisation, finally, the first person singular pronoun can be used to give presence to the voice of the journalist in a news text. However, these are just some examples, and, as I will argue below, each of the three entities can be personalised by features on any of these levels.

In the following, I will explore the relation of personalisation to conceptualisations of mass media communication (section 2.1), its relation to linguistic immediacy (section 2.2), and the role it plays in the ongoing shifts between the spheres of private and public (section 2.3). Against the background of this discussion, I will then develop a model for categorising different forms of personalisation (section 2.4).

2.1 Mass media communication between impersonality and personalisation

One of the defining characteristics of mass media is that they have a large (i.e. mass) audience which consumes content that has been made publicly available. The transmission of such content can take place in various ways: via print, radio, television, or the Internet. News media are perhaps the most prototypical form of mass media, but the term equally includes books, magazines, talk shows, and advertisements and can in a broader understanding even cover paintings, architecture or street signs (Jucker 2011, 248). Moreover,

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

prototypical mass media communication is unidirectional from text producers to the audience. Even if the audience is able to respond to or even to interact with text producers (e.g. through letters to the editor), the speaking rights are still not symmetric (Jucker 2011, 255). The Internet has opened new possibilities, enabling the audience to take a more active role compared to older forms of mass media, but even in this setting text producers have invariably more control over the communication than the audience. Another characteristic feature of mass media texts is that they are usually produced by a large number of actors, only few of whom are visible to the audience. While an individual might be indicated as the producer of a message (e.g. a journalist as the author of a news article), this person is usually only one member of a large team which was involved in the production process and which might include editors, proofreaders, managers, sound technicians, layouters and printers (see also Bell 1991, 33-50). Although the Internet has recently also enabled individuals to address a potential mass audience by publishing blogs or websites, the term "mass media" is still usually reserved for content produced by formal institutions.

There is a certain tension between the notions of personalisation and mass media. In several respects, anonymity and impersonality seem to be exactly the properties that distinguish mass media communication from other forms of communication and interaction. This becomes quite clear in McQuail's description:

> The source is not a single person but a formal organization, and the 'sender' is often a professional communicator. [...] The relationship between sender and receiver is one-directional and rarely interactional, it is *necessarily impersonal* [...] The impersonality derives partly from the physical and social distance between sender and receiver, but also from the impersonality of the role of being a public communicator, often governed by norms of neutrality and detachment. (McQuail 1987, 31–32, my emphasis)

For McQuail impersonality is thus a defining criterion of mass media communication. He mentions two different aspects of impersonality. On the one hand, text producers and the audience typically do not know each other in mass media communication, and there is typically no direct interaction between the two. On the other hand, the communication is impersonal in the sense that the message should be formulated in a neutral and detached way.

McQuail also points out that these two aspects are related. According to him, the lack of interaction between text producers and audience is an inherent part of their roles in this communicative situation. In order to fulfil the "norms of neutrality and detachment", text producers do not appear as persons in the communicative process. Instead, they take on the depersonalised role of a "public communicator". Even though McQuail here discusses impersonality only with respect to the role of the text producers, the same could be said about the audience (see van Dijk 1988b, 74). The receiver of mass media communication is always addressed qua his or her status as a member of a large, anonymous and impersonal audience. Recipients of mass media communication may be able to picture themselves

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

as belonging to one of various segments of the entire audience: to the addressees if they expect to belong to the target audience of the media; to the eavesdroppers if the media content is clearly not designed for them (e.g. in the case of children consuming x-rated media content); or to any of the less clear-cut segments in between (see Bell 1991, 92-95). However, even recipients of the target audience will not perceive themselves as being addressed as interactants – at least not in the prototypical conception of mass media communication.

Such a depersonalised conception of mass media communication, however, results in a partial and somewhat idealised picture. It is a direct consequence of the assumed norms that mass media – in particular news media – have to fulfil: to provide neutral and detached information. Waugh (1995) states these norms of news media very clearly.

> The news text is an informative text. The ostensible purpose of a news article is to convey information from journalists who have the information to readers who do not, but who would like to have it. The information is assumed to consist of facts/truths about the objective/outside/referential world. Thus, the major function of newspapers is referential and epistemic. (Waugh 1995, 131-132)

According to Waugh, the purpose of newspapers (and other news media in extension) is to inform by reporting objective facts. Objectivity, neutrality and detachment are thus presented as properties that are intrinsically tied to the nature of news texts. This view of news reporting is, however, not uncontested. For one thing, the norm of objectivity is not universal. Hampton (2008), for instance, argues that British print journalism does not have an equally strong ideal of objectivity as American journalism. According to him, British newspaper journalists tend to embrace notions of independence, truthfulness and fair play, rather than objectivity (Hampton 2008, 478). The norm of objectivity has also received criticism from researchers in media sociology and critical discourse analysis, who have repeatedly demonstrated that news media certainly do not just report objective facts (e.g. Cohen and Young 1973; Fairclough 1995; Fowler 1991; Schudson 2003; Tuchman 1978; van Dijk 1988a,b). This line of research has argued that both the selection and the representation of news is always shaped by ideology, while at the same time supporting the reproduction of this ideology. Cotter (2010, 21) therefore calls the norm of objectivity in journalism a "myth", and Locher and Wortham (1994, 519) go as far as declaring it an "unattainable ideal". They argue that real objectivity would require that the text producer is in no way part of the world about which he or she reports in order to achieve complete "perspectivelessness".

In addition to the fact that the ideal of neutrality of objectivity of reporting has been challenged, it is also important to note that it is not the only norm that governs mass media production. Other norms and requirements, not least the demand for an economically successful product, tend to go against the ideal of impersonal and detached reporting. Text producers of news media try to maximise the appeal of their stories by taking into account a number of factors, called "news values", which determine how newsworthy an item is. Among these news values, Bell (1991, 156-160) lists at least two which stand in

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

conflict with McQuail's ideal of impersonal mass media communication. The first of these news values, "relevance", focuses on the effects of a news event "on the audience's own lives or closeness to their experience" (Bell 1991, 157). In order to maximise the relevance of their stories, text producers need to have a clear idea who their (intended) readers are and what lives and experiences they have. At the same time, news articles that are successful in describing the relevance of a news event for their audience let readers perceive themselves as being recognised and addressed. This may not go as far as inviting them to actually "talk back" to the newspaper, but it still serves to characterise the target audience and, thus, it lowers the degree of impersonality. Increasing the relevance of a news story, therefore, reduces the distance between text producer and audience, if only rhetorically.

The second news value of interest here, "personalization", states that "[s]omething which can be pictured in personal terms is more newsworthy than a concept, a process, the generalized or the mass" (Bell 1991, 158). In order to maximise this news value, a news article needs to focus on individuals who are responsible for or directly affected by an event and who report on their personal experiences. Especially when the person whose experiences are reported is an ordinary person (in contrast to a celebrity figure or an official actor), the audience is invited to identify with him or her, which increases the perceived immediacy of the news event and the involvement of the audience. According to Galtung and Ruge (1973, 67), the fact that the audience can more easily identify with a person than with an abstract structure may be one of the main reasons why news events are so frequently depicted in personal terms. However, by taking over the perspective of an individual, the norms of neutrality and detachment are necessarily violated. The appeal to readers stands once again in opposition to impersonal reporting.

Thus, mass media are not in all respects as impersonal as McQuail's quote above suggests. In contrast, Fairclough (2001) argues that mass media very often try to bridge the gap to the distant and anonymous audience by linguistic strategies that simulate a personal relation. Such strategies include, for instance, the use of imperatives and direct address of the audience with *you*, which individualise the addressee, but also features that individualise the text producer (e.g. self-reference, "snappy" syntax) as well as the personalisation of the object of communication (see Fairclough 2001, 168-170). Fairclough calls this process "synthetic personalisation" and states that: "[. . .] it seems that the more 'mass' the media become, and therefore the less in touch with individuals or particular groupings in their audiences, the more media workers and 'personalities' (including politicians) purport to relate to members of their audience as individuals who share large areas of common ground" (2001, 160).

According to Fairclough, synthetic personalisation is a common strategy in advertisement and political discourse (2001, 160, 168), but it can also be found in news media. Jucker (2000, 638-639), for example, names a range of linguistic strategies that are regularly used in print newspapers in order to address the audience. These include, for instance, first and second person pronouns, imperatives and rhetorical questions, which all simulate a certain degree of interaction between text producers and the audience (Jucker 2000, 638).

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

In a different way, emotional language and the use of direct quotes can create immediacy and lead to a higher involvement of the audience (Jucker 2000, 638-639).

News media are not homogenous, of course, and personalisation in its various forms is not evenly distributed across all publications. One factor which differentiates news media and which has generally received quite a great deal of attention from linguists studying print newspapers is their market orientation. The criterion of market orientation refers to the social distribution of the targeted audience, with up-market newspapers having a larger proportion of readers from the social classes A, B and C1 than mid- and down-market newspapers (Jucker 1992, 51; see also chapter 3.2 below). Various studies have pointed out systematic differences between up-market newspapers (also called qualities or broadsheets), on the one hand, and mid- and down-market newspapers (also called populars or tabloids), on the other, some of which relate to personalisation. For instance, the aforementioned linguistic strategies by which print newspapers address and involve their audience appear to be more frequently used by mid- and downmarket papers than by up-market papers (Jucker 2000, 638-640).

Several of the characteristics that Conboy (2003, 2006) identifies as typical features of tabloid journalism are closely connected to the personalisation of news. Tabloid newspapers are said to have a very explicit audience orientation, which is, for instance, expressed through the frequent use of direct address of the audience, simulating interaction (2006, 21-22). Furthermore, the "engaged and often enraged" presentation of news events and the "rhetoric of the vernacular" are described as ways in which tabloids align themselves with their target audience (2003, 47, 54). Both these strategies are not compatible with the norm of neutral and detached reporting. Similarly, self-referentialty, i.e. presenting the newspaper as an actor in the news is named as another typical feature of tabloid journalism (2006, 37), and has the effect of personalising the media institution. Finally, tabloids are said to frequently focus their reporting on individuals; on the one hand, on the lives of celebrities (2006, 185-206) and, on the other hand, on ordinary people who become involved in extraordinary events (2006, 31-34). All these observations indicate that the frequency and quality of personalisation strategies are dependent on the newspaper's market orientation.

However, personalisation features do not only differ synchronically across market-orientations. Their use also changes over time. Schneider (2002), for example, studies personalisation as one criterion to trace the emergence and development of popular journalism between 1700 and 2000. Looking at the ways in which text producers, readers and news actors are referred to and presented, she finds that popular newspapers have always used more personalisation strategies than quality papers. However, her analysis also shows that these strategies underwent profound changes during the 300-year period of her investigation, with a general tendency towards more personalised reporting. For example, references to members of the royal family changed from referring to their title only (e.g. *The King*), over combinations of titles and names (e.g. *Prince ERNEST*), to the use of first names and short forms at the end of the 20th century (e.g. *Charles*, *Di*) (Schneider 2002,

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

164-165). A similar development towards more informal and personal referring expressions can also be observed for ordinary people in news reports. While the earliest papers in Schneider's sample tend to refer to such news actors by last name, profession and town in which they live, later papers also add information about their age and appearance and include first person quotations. By the end of the 20th century, photographs are included and first names, shortened forms and colloquial kinship terms appear (Schneider 2002, 158-159). Furthermore, the reference to text producers was subject to several changes. According to Schneider, the most widespread form of reference to text producers around 1700 took the form of self-references to the newspaper (e.g. *We hear*), with correspondents starting to be referred to around 1800, albeit without indicating any names. Journalists have been identified by name since only about 1900, while towards the end of the 20th century the first photographs of journalists started to appear (2002, 143-144).

Thus on the basis of previous research results, two main observations can be made with respect to the distribution of personalisation strategies across news media. On the one hand, personalisation strategies are more apparent in down-market papers than in up-market papers; on the other hand, personalisation strategies are increasing overall, with popular papers leading the trend.

In combination, these two observations mean that personalisation also plays an important role in discussions of the so-called "tabloidisation" of news and of public discourse in general. According to Bird (2009, 41), "an increasing emphasis on the personal" is one of the main characteristics of tabloidisation, together with an increase of visual material, stylistic features such as decreasing sentence complexity, and the trivialisation of content (see also Turner 2010, 26-32). Such tendencies are often taken to signify a decrease of journalistic standards and even a potential danger to democracy (see, for instance, Esser 1999; Holly 2008; Ross 1998). However, it has also been argued that personalisation does not necessarily need to have negative effects and that the illustration of current affairs with personal case studies can facilitate understanding, political insight and participation in political processes (e.g. Greger 1998; Gripsrud 2000; Lauerbach 2010; Macdonald 2000; Wehner 1998). Moreover, personalisation, understood as the foregrounding of the person behind any one of the entities in mass media communication, has many facets. Adopting a one-dimensional view of identifying "tabloidisation" tendencies bears the danger of missing more subtle differences in how personalisation is realised across various publications. Rather than assessing the relative benefit or detriment of personalisation strategies for society, I therefore want to contribute to a more differentiated view of personalisation by developing a model for the systematic analysis of its various dimensions.

A very influential factor for the distribution of personalisation features across news media is the role of the technical medium through which news is transmitted to the audience. For instance, the way in which news actors are presented differs quite clearly between radio, television, and print news. Print newspapers can include quotes by news actors in their text, but unlike radio and television news they cannot reproduce their voice. The integration of (static) pictures of news actors is very common in present-day newspapers,

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

whereas this is not possible on the radio. Television, finally, can even provide dynamic pictures (videos), in addition to sound. Similarly, text producers are to different degrees visible in print newspapers, radio and television news. Print newspapers can indicate the name of their journalists and they can even publish a photograph of him or her. The voice of the presenter can, however, only be heard on radio and television, where it is almost necessarily present. Last but not least, the three communicative forms also vary with respect to their options for interacting with their audience. Print newspapers allow their readers to send in letters to the editor, some of which are regularly published. However, due to the production process of newspapers, such feedback can only appear with a time lag of at least one day. On the radio, and some TV programmes, phone-ins allow listeners (or viewers) to call the station and engage in a direct, real-time exchange with text producers or other members of the audience (Armstrong and Rubin 1989; Bierig and Dimmick 1979; Leitner 1983; Simonelli and Taggi 1985). While for newspapers space limitations mean that usually only a small quantity of all letters to the editor are published, time poses similar restrictions on audience interaction on radio and television. Time on air is expensive, especially for television, and therefore audience contributions are usually short and carefully selected.

The Internet has made it possible to combine and enhance the range of personalisation techniques of print newspapers, radio and television. On online news sites, static and dynamic pictures, text, and audio material can all be integrated into large article clusters. Moreover, the online setting provides a large range of options for direct interaction between text producers and their audience, which go far beyond the possibilities that are available in the earlier communicative forms. Since the time lag between the production of an audience contribution and its publication is minimal, the exchange with text producers and among members of the audience can (at least in principle) become much more interactive than in print newspapers. In addition, the space and time restrictions that limit the amount of published audience contributions in older forms play a very marginal role online. Consequently, there is usually neither a limit to the number of user comments that can be submitted, nor is there a selection process, apart from the deletion of contributions that violate the guidelines of the site. This potential of the online setting has led to the complementation of many old mass media forms with online resources (Burger 2005, 11), ranging from the simple guestbook that accompanies a specific television format (Burger 2005, 15), to sites that allow one to listen to radio stations online and which provide some additional content, and even complex websites containing all the content of a print newspaper alongside additional material. In chapters 4 and 5, I will investigate to what extent and in what ways such online news sites make use of the technical possibilities of the online setting for the personalisation of news.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

2.2 Personalisation and linguistic immediacy

Language plays an important role in the personalisation of mass media communication. Synthetic personalisation, the simulation of interaction with individual readers, is largely achieved through stylistic means, such as direct address of the audience, rhetorical questions and imperatives. If personalisation is seen as an attempt to overcome the unidirectionality and distance of the communicative situation, it is perhaps not surprising that linguistic personalisation strategies make use of language that shares characteristics of oral face-to-face conversations. Consequently, many of the features of personalising language have been discussed in linguistics literature as characteristic features of conceptionally oral language, referred to variously as features of "immediacy" (Wiener and Mehrabian 1968; Koch and Oesterreicher 1985; Koch 1999), "involvement" (Chafe 1982; Tannen 1982, 1986), "parlando" (Sieber 1998), and "linguistic expressions of affect" (Ochs 1986). Features with a strong personalising effect also score highly on Biber's dimension of involved versus informational production (1988, 102). According to his study, this dimension stands for one of the most fundamental differentiations across spoken and written texts (1988, 104). The most indicative features for involved production include private verbs such as *think* and *feel*, which are used to express personal opinions and emotions, first and second person pronouns, which tend to refer to speaker and addressee, as well as emphatics and amplifiers (1988, 102-108). All these features have an interactional or affective function, and of the genres included in Biber's study they can most frequently be found in telephone conversations and face-to-face interactions (1988, 105-106, 128).

Classic studies on linguistic immediacy and linguistic features of involvement often take their starting point from comparing prototypical written language (e.g. scientific writing) to prototypical spoken language (e.g. informal conversations), trying to identify features that distinguish the two. The studies differ with respect to the extent to which the differences are seen as being conditioned by the mediality of language (oral or written realisation). Chafe (1982), for instance, bases the interpretation of his analysis on observations of the speed of production and reception of spoken and written language. He argues that the slower speed of production (and the higher speed of reception) of written language is responsible for the higher degree of integration that he finds in his written data from academic papers. These data are characterised by a condensation of information through the use of nominalisations, attributive adjectives, complement clauses, etc. In the spoken data from informal dinner conversation, in contrast, he identifies a higher degree of involvement, created through features like first person references, expression of the speaker's mental processes, emphatic particles and the use of direct quotes. In the academic papers, such involvement features were rarely found. Moreover, Chafe identifies the frequent use of passive voice and nominalisations in the academic papers as features that create detachment. He explains the tendency for involvement in his spoken data and the tendency for detachment in his written data with the respective speech situations, especially the relation between speaker and audience:

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

> It is typically the case that a speaker has face to face contact with the person to whom he or she is speaking. [...] It means furthermore that the speaker is aware of an obligation to communicate what he or she has in mind in a way that reflects the richness of his or her thoughts — not to present a logically coherent but experientially stark skeleton, but to enrich it with the complex details of real experiences — to have less concern for consistency than for experiential involvement. The situation of the writer is fundamentally different. His or her readers are displaced in time and space, and he or she may not even know in any specific terms who the audience will be. The result is that the writer is less concerned with experiential richness, and more concerned with producing something that will be consistent and defensible when read by different people at different times in different places, something that will stand the test of time. (Chafe 1982, 45)

While this characterisation of different speech situations applies well to informal dinner conversations and academic papers, there is obviously a great deal of language produced in settings that differ from the above descriptions. Not all spoken language is produced in a face-to-face situation (e.g. telephone conversations, language on television), and there are face-to-face situations in which producing a logically coherent argument is much more valued than recounting real experiences (e.g. during an oral exam). Similarly, personal letters or, more recently, private e-mails and text messages, usually address a well-known reader and will in most cases not need to "stand the test of time".

The restriction of linguistic analyses to prototypical forms and, especially, the tendency to generalise these findings to all forms of spoken and written language have already been criticised by Lakoff (1982, 241) in an article published in the same volume as the above quoted article by Chafe.

> Actually, work on this topic [differences in written and oral discourse] tends to create a peculiar dichotomy: planned, nonspontaneous written discourse on the one hand, and spontaneous, direct oral communication on the other. It is unarguable that these represent the clearest cases, and the sharpest distinctions, and are worth studying therefore as ends of a continuum. But in order to understand how we utilize the various modes of communication to their fullest advantage, we must understand that there are other possibilities, and that some of [the] characteristics we have asscribed to 'oral' discourse, for example, are not necessarily characteristic of the oral medium per se, but rather their choice has more to do with immediate personal contact — eye contact, for instance — or the usefulness of an appearance of spontaneity, rather than to the use of the vocal channel itself. (Lakoff 1982, 241)

Lakoff here suggests a continuum of discourse styles, with prototypical spoken language, on the one hand, and prototypical written language, on the other. Written texts or instances of spoken discourse could then be placed on this scale, depending on the specific linguistic features they show.

Tannen, in the same volume, also proposes a continuum, which she calls the "oral / literate continuum" (1982, 4). She argues that the oral end of the continuum is characterised by a "relative focus on involvement" and interpersonal interaction with the audience, whereas a focus on content would be typical of the literate end of the continuum (1982, 3-4). She

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

also adds the "tendency to talk in terms of personal experience and to instantiate rather than talk in abstract or general terms" as one of the characteristics of involved language use. This is particularly relevant for this study, since such a focus on personal experience and specific examples are elements that create personalisation.

While the continua suggested by Lakoff (1982) and Tannen (1982) allow the placement of non-prototypical realisations of language, there is still a strong association of the two poles with spoken and written language through the terminology that is used (e.g. "oral / literate continuum"). On the one hand, this bears a certain danger to regard non-prototypical forms as (negligible) deviations from normal language use; on the other hand, it creates terminological difficulties, if not confusion, when informal written language is characterised as "oral" and formal spoken language as "literate". If it is true that the "characteristics we have ascribed to 'oral' discourse [...] are not necessarily characteristic of the oral medium per se" (Lakoff 1982, 241), then it is sensible to systematically distinguish, also terminologically, the level of linguistic characteristics from the medial realisation of language.

Koch and Oesterreicher propose a model of communicative immediacy and distance that fulfils this aim (Koch and Oesterreicher 1985, 2011; Koch 1999). Their model, reproduced in Figure 2.1, distinguishes clearly between two dimensions, medium and conception, a distinction which was previously proposed by Söll (1974). The medium dimension differentiates between graphically and phonically realised language. This is an absolute dichotomy; language is always produced either in spoken or in written form and there is no in between. In contrast, the dimension of conception is a continuum. Individual texts and utterances – or communicative acts, to use the cover term proposed by Jucker and Dürscheid (2012) – are close to the pole of communicative distance if they share many of the characteristics of prototypical written language, and they are close to the pole of communicative immediacy if they are very similar to prototypical spoken language. High communicative immediacy is thus, among other features, characterised by a non-public setting in which the communicative partners are close and the communication is spontaneous and interactive. Koch and Oesterreicher (1985, 21-23) argue that such a setting typically leads to communicative acts which are less planned, which have a low information density and which are linguistically less complex and elaborated. High communicative distance refers to the opposite end of the scale.

Communicative acts with a high degree of linguistic immediacy are more typically realised as spoken language, while communicative acts with a high degree of linguistic distance can more typically be found in written language. This is reflected by the triangles in Figure 2.1, which leave more room for prototypical constellations than for the less typical constellations. A coffee chat between close friends would be an example of spoken interaction with high communicative immediacy, which can be placed at the position marked as "A" in Figure 2.1. In contrast, a written contract between an employer and an employee would be close to the pole of communicative distance and could be placed at position "D". Less typical constellations are written language with high communicative

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

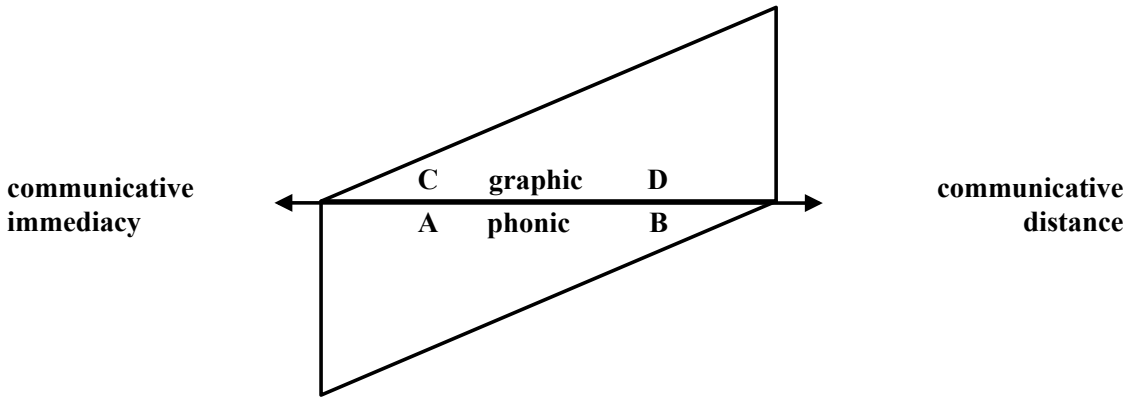


Figure 2.1: Koch and Oesterreicher's model of communicative immediacy and distance (Koch 1999, 400)

Author submitted manuscript

immediacy (position C) and spoken language with high communicative distance (position B). An example of spoken language with high communicative distance would be a lecture or a sermon, while informal private letters are mentioned as examples of written language with high communicative immediacy (Koch and Oesterreicher 1985, 18).

The possibility to also categorise less typical constellations is one of the main advantages of Koch and Oesterreicher's model. For this reason, their model has been applied to new communicative forms that have recently developed: web logs, online chats, web advertising, instant messages, live text commentaries, short text messages, and new forms of communication on social networking sites (e.g. Beißwenger 2000, 2010; Dürscheid 2007a; Frehner 2008; Gruber 2008; Janoschka 2004; Jucker 2010; Schlobinski 2006; Storrer 2009; Tuor 2009). These new forms of communication have in common that they are often used to produce communicative acts with a high degree of immediacy, while still being realised as written language. Applying Koch and Oesterreicher's distinction between the medial realisation and the communicative conception has helped avoid characterisations of such data as representing a "third medium" of language, which is "identical to neither speech nor writing" (Crystal 2006, 51-52). Such characterisations of new communicative forms are misleading and they have been criticised for their lack of distinction between the medial and the conceptual dimension (Dürscheid 2004; Jucker and Dürscheid 2012).

However, Koch and Oesterreicher's model has also been subject to modifications and criticism. Dürscheid (2003, 38-39), for instance, criticises Koch and Oesterreicher's choice of the term medium to refer to the distinction between speech and writing. The term medium in their model does not refer to the technology that is used to produce a message (print, handwriting, radio, etc.), but to the code in which it is realised and which is either graphic or phonic. The term medium is generally used in a variety of ways, especially across different disciplines. For example, from a semiotic perspective, Posner (2003, 43-46)

Author submitted manuscript. Consult published version for correct page numbers!

distinguishes six different conceptions of medium, including distinctions based on physics (e.g. acoustic vs. visual mediums), technology (e.g. pencil, typewriter, CD), functions (e.g. report, advertisement), and social factors (referring to different types of media institutions). For linguistic studies of mediated communication it has, therefore, been proposed reserving the term medium for the technical means by which texts are produced, transmitted and received (e.g. Dürscheid 2003, 2005; Herring 2003; Holly 1997). It is this sense of the term that I use in this study when referring to the distinction between news transmitted in print, via radio, television or the Internet.[2] These different forms of communication are realised with different technical mediums of production (e.g. the printing press), reception (e.g. the radio) and transmission (e.g. the Internet). Referring to speech and writing as different mediums of language – as Koch and Oesterreicher do – can thus easily lead to terminological problems. Jucker and Dürscheid (2012) therefore adopt the terminology used by Stöckl (2004b, 14) and refer to speech, writing and signing (a third option not discussed by Koch and Oesterreicher) as different medial variants of language, where medial is meant to be the adjective of mediality, rather than of medium (see Dürscheid 2003, 39).

Ágel and Hennig (2006a,b) propose quite far-reaching changes to Koch and Oesterreicher's model. On the theoretical level, they criticise the fact that the dimension of communicative immediacy and distance does not clearly distinguish between conditioning factors of the communicative setting, on the one hand, and resulting strategies of linguistic realisation, on the other (Ágel and Hennig 2006b, 13-14). They argue, for instance, that several of the strategies for linguistic realisation mentioned by Koch and Oesterreicher (1985, 23) could at least as plausibly be regarded as factors of the communicative setting, e.g. "tentativeness" and "processual character". With respect to its practical application, they call for a model that does not only allow for the relative positioning of communicative acts as more or less immediate than other communicative acts. Instead, they propose quantifying the axis of communicative immediacy and distance. For this purpose, they develop a large inventory of features that they see as indicative of communicative immediacy. These are organised into several hierarchical levels and various parameters and include features such as deixis, hesitation markers, repetitions, free choice of tense, etc. With this inventory, they claim to be able to determine the communicative immediacy of texts as a numerical factor, constituted by the frequency of features of immediacy divided by the number of words in a text. Several of the articles in the same volume illustrate the applicability of their model to a range of written texts with high communicative immediacy.

However, their approach raises a number of questions with respect to the quantification of individual features. For instance, it does not seem obvious why all features should be counted as having the same weight, given that some features certainly contribute more to the overall impression of immediacy than others. In a later publication, Hennig (2010, 320-321) herself observes that not all features have the same effect for immediacy and

[2] In addition, I use the plural form media when talking about mass media (communication) and when referring to media institutions. This use is distinct from the technical conception of medium (with the plural form mediums) which is discussed in this paragraph.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

that, therefore, more attention should be paid to the role individual features play, instead of focussing only on the overall frequency of features of immediacy. Furthermore, it is not clear whether the list of features provided by Ágel and Hennig is indeed complete, or whether additional features might also play a role. In sum, while their framework results in seemingly precise values of immediacy for each text, it is questionable whether such a degree of precision is adequate for representing a phenomenon like linguistic immediacy. Apart from these theoretical reservations, their model is also not suitable for the application to large amounts of texts, since it relies on a very detailed manual analysis. It is for these reasons that their model will not be used in this study. Instead, Koch and Oesterreicher's dimension of communicative immediacy and distance is one central element of the model of personalisation that I will further develop in the following.

2.3 Public, private, involving, and immediate

One of the most frequent contexts in which personalisation is discussed concerns the blurring of boundaries between the domains of the private and the public. The collection of articles edited by Imhof and Schulz (1998a) deals with such processes. The starting point for the volume is the observation of a twofold shift between the two domains. The first development concerns topics that were historically dealt with only in the private context of the family, but which have increasingly been gaining public relevance through the introduction of the social question into politics (Imhof and Schulz 1998b, 9-10). As an example of this process, the contribution by Keppler (1998) in the same volume mentions the attempt of the women's movement to politicise gender relations. While topics such as domestic violence and gender-based division of labour had long been treated as private issues, they started to gain visibility in the public sphere after the late 1960s (Keppler 1998, 157-158). The second development concerns the fact that the public sphere is increasingly dominated by private topics, which can in particular be observed in the tendency of mass media to give more room to human interest stories and other non-political content (Imhof and Schulz 1998b, 9-10). It is this second development that is of interest for personalisation.

The personalisation of politics, understood in particular as the focus on political actors as (private) persons, is one aspect of this development that has received considerable attention. Often, this has been interpreted negatively as a tendency to replace (important) political issues with (less important) private angles. Ross (1998), for instance, sees the personalisation of the public sphere and the depoliticisation of politics as two related tendencies. He argues that the increasing dominance of the personal lives of politicians in media supplant relevant factual information. According to Ross (1998), political decisions become less based on rational arguments and instead they are increasingly the result of emotions and personal sympathies for individual political actors. Similarly, Weiss (2002, 86-87) argues that the focus on individuals and private topics in mass media leads to a decay of the political public sphere. He sees the personalisation of politics as a process that

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

presents political content in a form that invokes subjective speculation and judgements of character based on lay psychology.

This view is not uncontested and it has been suggested that personalisation, if employed in a sensible and strategic way, can even facilitate the understanding of political issues and provide new perspectives on ongoing debates (Macdonald 2000, 264). Wehner (1998, 321) refers to personalised actors as navigational aids in the flood of information in mass media. By providing entry points into complex political processes, personalisation can, thus, also have positive effects. Likewise, Greger (1998) argues for a more differentiated view on personalisation, avoiding a simplified equation of personalisation with depoliticisation. According to him, a strong presence of non-political news actors in political reporting can be one way for mass media to respond to the diversity of social interests and needs (Greger 1998, 251, 280).

When discussing the shifting boundaries between the domains of the private and the public, one needs to address the fact that the terms "private" and "public" are used with various meanings and that these meanings are not always clearly distinguished (see also Benn and Gaus 1983; Heller 2006; Landert and Jucker 2011; Weintraub 1997). Various suggestions for classifications of private and public aspects and corresponding terminology have been made. Weintraub (1997, 4-5) distinguishes between two main dimensions of the public/private distinction, which he calls visibility and collectivity. Visibility indicates how easily accessible something is. The public end of the scale stands for "open, revealed, or accessible", whereas the opposite end of the scale stands for "hidden or withdrawn" (1997, 5). Collectivity gives an indication of who is affected by an event or an item. The private end of the scale stands for things that affect the interests of an individual, whereas the opposite end of the scale stands for what affects the interests of a collective (1997, 5). Visibility and collectivity are both continua and, while they are related to each other in several ways, they can still be clearly distinguished analytically.

Similar distinctions are also proposed by Dürscheid (2007a) and Heller (2006), who are both concerned with private and public aspects of communicative forms on the Internet. Dürscheid (2007a) adopts a two-dimensional distinction between accessibility and content of communication. The degree of accessibility is defined through the ease of access to the communication. This is independent of the content, which can be more or less private. In order to minimise terminological problems, Dürscheid proposes using the terms public and non-public to refer to the two ends of the accessibility dimension, and the terms private and non-private to refer to the content dimension.

Heller (2006) uses a similar distinction between the accessiblity and the content of communication, but proposes adding a third dimension that distinguishes between direct and mediated communication. This suggestion has been criticised on the basis that the distinction between direct and mediated communication does not constitute a continuum in the same way as this is the case for the dimensions of accessibility and content (Landert and Jucker 2011, 1425). While accessibility and content are matters of degree, the distinction between mediated and non-mediated communication is a dichotomy. Heller

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

justifies her view of a continuum by claiming that different mediums have different degrees of complexity, depending on the number of channels (audio, visual) they can use (Heller 2006, 324). Even though both television and radio produce mediated communication, the former is, in her view, closer to the pole of mediated communication than the latter, due to its use of the visual channel.

Collapsing the criterion of mediated versus non-mediated communication with the criterion of the number of channels is problematic, however. It raises the question as to why other characteristics of the medium should not be included as well, such as synchronicity, 1-way versus 2-way transmission, and persistence (see Herring 2007). Moreover, the integration of a third dimension for mediated versus non-mediated communication poses theoretical problems, since it is a distinction which applies on a different level than accessibility and content. Like all the other medium characteristics, mediated versus non-mediated communication refers to communicative forms as a whole. We can, for instance, say that face-to-face conversations are non-mediated while messages posted in an online forum are mediated. In contrast, accessibility and content characterise individual communicative acts. A conversation in a crowded bus is accessible to a larger audience than a conversation in a confessional, and the same difference can be observed for forum messages posted in a public forum and forum messages in a forum restricted to a small number of registered participants. The argument for the content dimension is even more obvious – not all conversations deal with content of the same degree of privacy, nor do all messages posted in an online forum. Therefore, it seems reasonable not to integrate characteristics of the medium into the same model that is used to characterise the privacy of content and the degree of public accessibility of communicative acts.

Landert and Jucker (2011) propose a three-dimensional model to characterise communicative acts with respect to public and private characteristics. The first two dimensions adopt Dürscheid's distinction between public and non-public accessibility, on the one hand, and private and non-private content, on the other. They stress the importance of defining the two aspects independently from each other and, consequently, they reject a definition of private topics as topics that are usually discussed in a non-public context. Instead, they adopt Weintraub's distinction between visibility and collectivity (1997), which was introduced above. Visibility refers to the ease of accessibility, and, therefore, communicative acts are public if they can be easily accessed by a large audience. In contrast, collectivity indicates how large the group of people is who is affected by an event or item. Consequently, a private topic is a topic that only affects a single individual or a small group, whereas a non-private topic is of interest to a very large group or to people in general. Examples of private topics that can be found in news are, for instance, personal emotions which are only directly relevant to the person who has them, experiences of the everyday life of individuals such as walking one's dogs, and roles connected to one's immediate family life, like being a mother (see also chapter 6). In all these cases, only a very small group of individuals is directly affected. Nevertheless, it is still possible that these topics are discussed in a very public context, often when they are embedded into topics that are of

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

relevance to a larger audience (and thus less private), such as extreme weather situations or health policies. Keeping the definition of private and public independent from each other, thus, allows one to classify such instances as private *and* public communicative acts, without losing the distinction between the two categories.

As a third dimension, Landert and Jucker added the degree of linguistic immediacy to their model, based on Koch and Oesterreicher (1985; see also section 2.2 above). Linguistic immediacy has repeatedly been associated with private and public aspects. Fairclough (1995), for example, mentions the use of features of linguistic immediacy as a way in which mass media deal with the fact that their products are mostly consumed in the private sphere of the homes of individuals, even though they are produced in a non-private context. Similarly, Imhof and Schulz (1998b, 11) name the increase of informal and emotional language on local radio broadcasting as an important element of how private aspects increasingly dominate the public domain. Kress (1986) demonstrates how linguistic features associated with written and spoken language are used in mass media to construct the domains of the public and the private. Linke (2000) reads changes towards more informal greeting conventions in German as an indication of more general socio-cultural shifts between the domains of private and public. Last but not least, Koch and Oesterreicher use public and private as characteristics of language of distance and immediacy, respectively (although without the terminological distinction between private/non-private and public/non-public introduced above).[3] Landert and Jucker (2011), therefore, integrate linguistic aspects into their model for characterising private and public aspects of texts.

Landert and Jucker see their own model as an extension of Koch and Oesterreicher's model of communicative immediacy and distance. Koch and Oesterreicher's model contains a single dimension that includes linguistic features as well as features of the communicative setting and of the content of communicative acts. This poses problems for the characterisation of communicative acts for which the three levels do not correspond in the prototypical manner, in particular texts that mix public and private features. This is why Landert and Jucker propose clearly differentiating between these three levels by using a separate axis for each. They demonstrate the applicability of their model with a case study comparing letters to the editor in the printed *Times* from 1985 with user comments on the *Times Online* in 2008 (see Figure 2.2). The letters to the editor in the *Times* from 1985 are publicly accessible, tend to deal with non-private topics and contain very formal and distant language. This is a prototypical situation and these texts could be classified as distant/public/non-private without differentiating between the various dimensions. The user comments on the *Times Online* however show a different pattern. While they are

[3] In their 1985 article, only *Öffentlichkeit* 'publicness' and *keine Öffentlichkeit* 'non-publicness' are explicitly mentioned as characteristics of language of distance and immediacy. Unfortunately, they do not specify clearly in which sense they understand this feature. They also mention private letters as examples of a communicative form that can be classified with their model. The fact that they specify the letters as being private suggests that this feature is implicitly treated as a criterion for language of immediacy. In their later publication of 2011, they explicitly refer to privacy as a feature of linguistic immediacy (e.g. Koch and Oesterreicher 2011, 7-9, 63, 73). See also Koch (1999, 400-401), who lists privacy as a feature of immediacy and publicness as a feature of distance.

Author submitted manuscript.

Author submitted manuscript. Consult published version for correct page numbers!

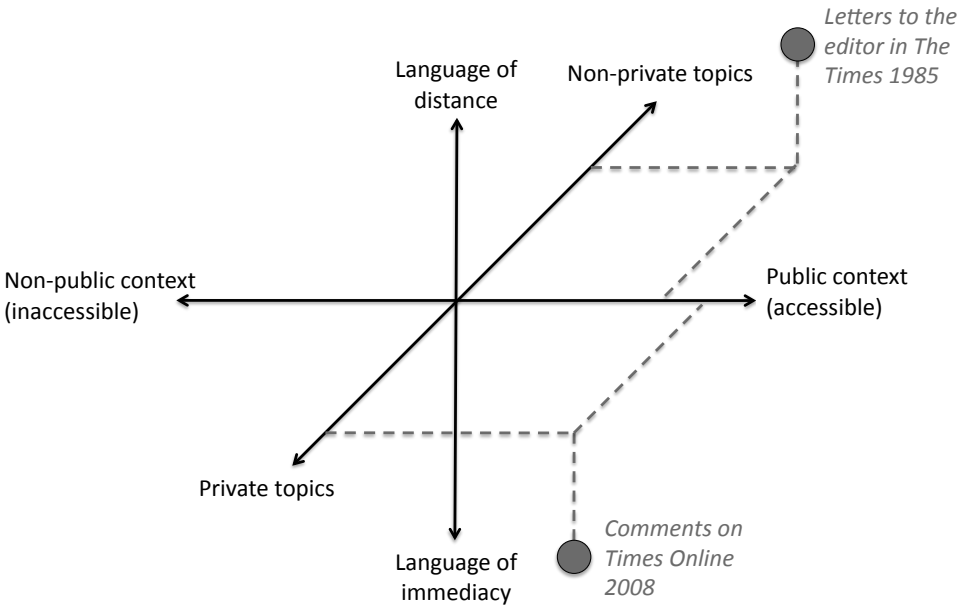


Figure 2.2: Letters to the editor in the *Times* from 1985 and user comments on the *Times Online* in 2008 according to Landert and Jucker (2011, 1432, fig. 4)

Author-submitted manu

also publicly accessible (possibly even more so than the printed letters to the editor), their content frequently contains private topics in the form of personal experiences and emotions (Landert and Jucker 2011, 1429-1430). The language of the user comments also shows more features of linguistic immediacy than the printed letters, e.g. colloquialisms and informal terms of address (Landert and Jucker 2011, 1431-1432). Thus, differentiating between the three levels of communicative setting, content and linguistic realisation allows one to describe such differences in a more precise way than simply saying that the boundaries between the private and the public are blurred.

The model by Landert and Jucker (2011) is useful for describing forms of personalisation which are created through the use of personal content and personalising language in public texts. If texts contain private content such as personal experiences and emotions, they are close to the private pole on the content scale. If texts contain linguistic features that simulate closeness or interaction (e.g. first and second person pronouns, directives, emphatics), they are close to the pole of immediacy on the scale of linguistic realisation. Personalising effects can be caused by either private content or language of immediacy, but the two can also co-occur. All these constellations can be categorised with the model in Figure 2.2.

Forms of personalisation that cannot be captured with this model are effects that are created through features of the communicative setting. These include, in particular, the interaction between text producers and the audience, but also the use of additional

Author submitted manuscript. Consult published version for correct page numbers!

modalities of communication (images, sound). The scale of the communicative setting in Landert and Jucker's model is defined as the ease of access to a communicative act. The larger the audience that has (potential) access to a communicative act, the more public is the communicative act. Although this definition is suitable for characterising public and private features of communicative acts, for investigating personalisation, ease of access is not the most important feature of the communicative setting. After all, personalisation to some extent presupposes a public context. If communicative acts with private topics and language of immediacy take place in a non-public context, this is considered personal communication, but not necessarily personalisation. In fact, it is a characteristic feature of personalisation that personal elements are introduced or emphasised in a setting that is to some extent impersonal, i.e. public. Moreover, the data of this study all have a very high public accessibility in common and variations in the degree of accessibility will therefore not play an important role for my analysis. Consequently, a special dimension for capturing public accessibility will not be needed.

For this reason, I will replace the dimension of public accessibility with a dimension that is better suited for representing personalising features of the communicative setting. I call this dimension involving versus non-involving communicative setting. At the involving end, there are settings which invite their audience to interact with text producers and to contribute their own content. There are also more subtle ways in which the communicative setting can be employed to create involvement, for instance by providing direct access to supporting evidence of what is communicated. In the context of news reporting, such evidence is often presented in the form of images, videos or sound files, which allow the audience to "see for themselves". At the non-involving end, there are settings which do not allow for interaction and which also do not create involvement through other ways of exploiting the communicative setting. The fact that this dimension concerns the communicative setting means that it is closely related to characteristics of the medium. Printed newspapers, for instance, have much more limited possibilities of involving the audience than online news sites, given that features like real-time feedback and the presentation of videos are not possible. However, it is important to note that the dimension of involving versus non-involving setting is still understood as a continuum that can be applied to individual communicative acts (e.g. individual news articles), in the same way as the other two dimensions. Even within one particular communicative form, communicative acts differ with respect to how they employ the options for involving the audience. One online news article may include video reports from eye witnesses and forms for submitting user comments and audience material, while another article on the same news site does not. The first article would then be closer to the pole of involving setting than the second.

Combining the dimension of involving versus non-involving setting with the dimensions of the privacy of content and the immediacy of the linguistic realisation results in the model depicted in Figure 2.3. The letters A, B, C, D, and E stand for some of the possible configurations of personalisation features in communicative acts. A is the position of a prototypical non-personalised communicative act, such as the mass media texts envisioned

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

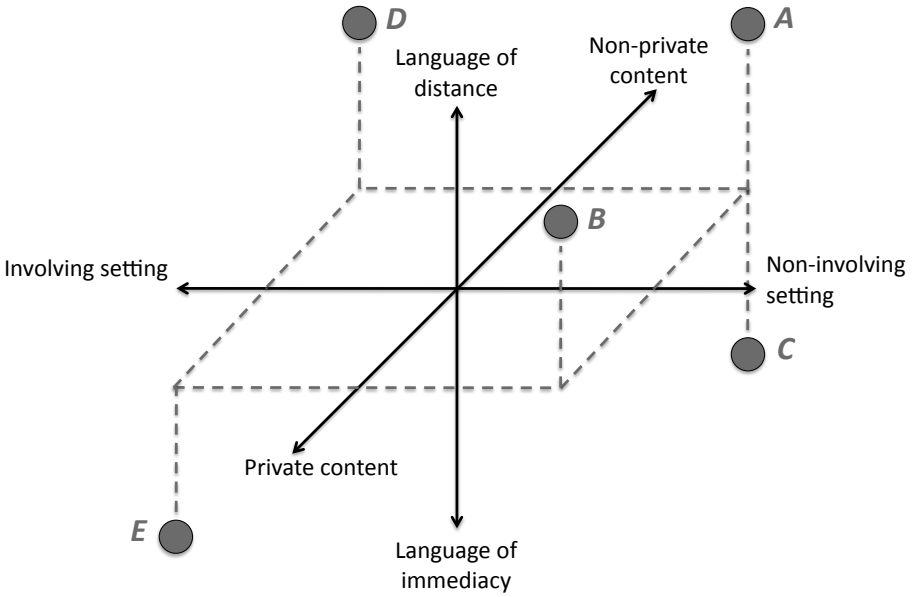


Figure 2.3: Locating texts with different personalisation strategies in a three-dimensional model

by McQuail and Waugh in section 2.1 above. Laws would constitute an even more extreme example of position A. Communicative acts at this position deal with non-private topics, they are formulated in very formal and distant language, and their setting does not contain any characteristics that lead to involvement. Position E is located at the other extreme of all scales. This stands for communicative acts which are highly personalised on all levels, thus dealing with private topics, using language of immediacy and being presented in a highly involving setting. Many communicative acts on social networking platforms like Facebook come close to this position, even though generalisations to all communicative acts on such platforms are, of course, not valid (see also Jucker and Dürscheid 2012). B, C, and D illustrate positions in which personalisation features are only present on one of the three dimensions. Communicative acts at position B contain private topics, but they occur in a non-involving setting and are formulated in language of distance. For position C, the linguistic realisation shows a high degree of immediacy, while the topics are non-private and the setting is non-involving. Communicative acts located at position D occur in an involving setting, but deal with non-private content and are formulated in language of distance.

More positions could, of course, be added to Figure 2.3, for instance positions with strong personalising features for two of the three dimensions. In addition, all dimensions are continua, so that positions anywhere on the scales are possible. There are no absolute reference points for the three scales and, thus, the model in particular allows one to visualise the relative difference in personalisation strategies between different communicative acts.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

As mentioned above, how publicly accessible a communicative act is cannot be expressed in this model, although there is an underlying assumption that personalisation is a feature usually observed for communicative acts with a certain degree of public accessibility.

2.4 Modelling personalisation

As the previous sections have shown, personalisation has many facets. It plays an important role in the actual or simulated overcoming of distance and anonymity in mass media, it manifests itself in the use of linguistic immediacy, and it is often perceived as an increase of private topics in the public sphere. In the following, I want to give an overview of some distinctions that are necessary in order to investigate personalisation in a systematic manner. Many of these aspects have already been mentioned above, and what follows can to some extent be understood as a systematic summary of what has been discussed so far. Table 2.1 provides an overview of the various dimensions of personalisation.

To start with, there are several entities that can be personalised. In a simplified model of mass media communication, three main entities are involved, all of which can be subject to personalisation. The first entity is the message, or, in the case of news media, the reported news event. Personalising the news event means, for instance, that the news article focuses on people, rather than abstract or generalised processes. Moreover, the personalisation effect is stronger in the case of private individuals than in the case of official actors. To give an example, a news report about an economic crisis would have a low degree of personalisation if it only reported on the latest statistics without mentioning any actors. A slightly more personalised news report on the same topic might present the professional opinions of experts, while an even more personalised version would, for instance, focus on the effects of the crisis on the everyday life of a specific family. I will use the term news actor to refer to any kind of person who appears in a news article. This includes official actors like politicians and experts, as well as celebrities and private individuals who are directly effected by an event or who are interviewed as witnesses. Personalising a news event, therefore, means introducing and personalising news actors in a news article.

The second entity that can be personalised is the audience. In this case, personalisation means that the audience is not conceived of as an anonymous and silent mass, but as a large group of individual readers. Synthetic personalisation contributes to the personalisation of the audience with linguistic strategies such as direct address. In addition, actual interaction with or among individual readers also increases the personalisation of the audience. If readers can talk back to text producers by sending in letters to the editor, they are no longer invisible. On online news sites, discussion sections also allow readers to interact with each other, which in some cases even leads to the formation of online communities. Members of the mass audience then have the opportunity to exchange ideas about what they have read and to get to know other readers who have read the same. In this way, members of the mass audience are able to overcome the anonymity of the communicative

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

situation and to become visible as personalised readers. Audience involvement is, therefore, an important aspect of the personalisation of mass media communication.

Text producers can also be personalised. In the case of online news sites, this means that the journalist is presented as the person who wrote the news article. In the most subtle form, personalisation of the text producer can be given through bylines identifying the journalist by name. On online news sites, there is sometimes more detailed information available about the journalist's background in profiles. Moreover, the news article can also contribute to the personalisation of the text producer, in particular when it explicitly expresses the journalist's experiences and opinions. The presentation of journalists as personalised text producers is particularly noteworthy since they are hardly ever on their own in producing news texts. Most of the other parties involved in the production, e.g. proofreaders, typesetters and layouters, are, however, not presented as being part of the process. This is also the reason why I refer to this entity as "text producer", rather than "journalist". It is actually one of the effects of personalisation that the journalist is presented as the single, personified text producer.

The personalisation of the media institution can be seen as a special case of the personalisation of text producers. In this case, the focus is not on the author of an individual article, but on the institution that publishes the newspaper or news site. In contrast to the other entities, the media institution can hardly be presented as a person. However, it can still appear as an active agent. For example, news articles can report on activities of the media institution or represent an institutional standpoint (e.g. *The Times can reveal that /. . . /*, times-uk-100116). More frequent than overt self-reference are the uses of house style and corporate identity, which contribute to the personalisation of the media institution in a more indirect way.

Another distinction, apart from which entity is personalised, is the means through which personalisation is realised. I propose that features on three main levels of texts can be distinguished. First, there are features that relate to the communicative setting of texts. This includes all factors that determine how communication can take place, e.g. the options for interaction and the available modalities (text, voice, image). This set of features is obviously closely connected to the characteristics of the medium of transmission. On online news sites, the communicative setting plays a particularly important role in the personalisation of the audience. The online setting provides various means for increasing audience involvement and it allows readers to gain much more visibility compared to older forms of news media. In addition, the technical possibility of integrating material in various modalities (text, audio, static image, video) makes the communicative setting relevant for personalisation, since non-textual modalities often involve the audience through giving them the impression that they are witnessing a news event more directly.

The article content is the second level on which personalisation can be realised. Typically, personalisation on the level of content leads to the personalisation of the news event. In some cases, however, the news actor can overlap with one of the other entities, in which case the effect of personalisation is not restricted to the news event. This occurs, for in-

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

stance, when text producers report about their own experiences. In such cases, the entities of text producer and of news actor fall together and the effects of personalisation apply to both roles. Similarly, a news report can focus on the experiences of a news actor who is presented as a reader of the news site. Again, this results not only in the personalisation of the news event, but also affects the presentation of the audience.

As a third level, finally, the linguistic realisation of the text provides an important means of personalisation. The use of second person pronouns and other forms of direct address are important ways of personalising the audience. Text producers are personalised through the use of self-referring first person pronouns and the expression of subjective stance (e.g. *in my opinion*). News events are personalised through the integration of direct speech rather than indirect speech or narrative reports. The use of self-reference and involved language in such direct speech quotes can further personalise the news actors.

There are, thus, three levels on which personalisation can be realised, and three (or four, if media institutions are counted separately) entities or roles which can be personalised. Some levels of personalisation are mainly used to personalise a specific entity (e.g. personalisation on the level of content mostly serves to personalise news events), but, at least in principle, all entities can be personalised with features from any level. Table 2.1 provides some examples of personalising features for all of the constellations. Of course, these examples are not exhaustive and more features could be added to many of the cells.

While such a grid as in Table 2.1 provides useful analytical distinctions for a systematic investigation of personalisation, it also needs to be pointed out that not all features fall neatly into a single category. A case in point is user-generated content, for instance in the form of visual or textual material provided by the audience and integrated into a news article (see chapter 4). The fact that readers of an online news site are able to submit stories, images and videos to text producers in an easy and fast way is a direct result of the communicative setting. At the same time, the integration of user-generated material into news articles usually leads to more personalised content, since it contains the perspective, experiences and opinions of individual readers. If the user-generated content is included verbatim, it might even lead to an increase in linguistic immediacy. By providing a personal point of view, the news event is personalised, but simultaneously the presentation of the view of one individual reader also personalises the audience. Thus, it is not possible to say in general whether user-generated content is mainly located on the level of the communicative setting or on the level of content, nor whether it serves to personalise the news event or the audience. In most cases, all of these aspects will come into play.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Author submitted manuscript. Consult published version for correct page numbers!

Table 2.1: Dimensions and features of personalisation

	communicative setting	**content**	**linguistic realisation**
news event and news actor	• visualisations of news actors • options for interacting with directly affected individuals	• focus on personal experiences rather than abstract concepts and processes • focus on private individuals rather than official actors	• direct speech • informal and involved language in quotes • self-reference in quotes
audience	• user feedback and interaction • user profiles • customisation of news site	• stories about or involving readers	• direct address • interrogatives • imperatives
text producer	• options for interacting with text producers • bylines and pictures of text producers • profiles and access to background information	• articles containing the text producer's experiences • opinionated articles	• self-reference (e.g. *I*) • subjective stance • informal and involved language
media institution	• corporate identity	• reports on the activities of the institution • representation of an institutional standpoint	• self-reference (e.g. *The Times can reveal that . . .*) • house style

user-generated content

Author submitted manuscript

A similar problem is posed by visual material, such as images or videos. In Table 2.1, I list visualisations as a feature of the communicative setting. This is justified insofar as the online setting allows the integration of videos, which is not possible in print newspapers. Moreover, the Internet is probably the only medium used for news reporting on which time and space restrictions are of secondary importance. In print newspapers, the available space poses a serious limitation on the number of images that can be published, whereas on television time restrictions do not allow for too many or very long videos. Online news sites, in contrast, are in the habit of publishing entire image galleries and sometimes they link to several videos in order to illustrate their articles. Server space is comparatively cheap and readers can decide for themselves how much time they want to spend viewing such visual material. Thus, the combination of language, images and videos to this large extent depends a a great deal on the communicative setting.

However, images and videos do not just play a role in personalisation with respect to their modality. I will argue in chapter 5 that visual material always tends to contribute to personalisation, but of course not all images and videos personalise news to the same extent. The content of images can be as varied as information graphics, landscape shots and close-ups of facial expressions and it is obvious that these different types of pictures do not have the same personalising effect. Moreover, the graphic realisation of visual material varies. A video of a natural catastrophe taken by an amateur on a mobile phone camera leads to a very different effect than a video of the same scene taken by a professional video crew. To be precise one would therefore have to distinguish between all three levels of visual material: its communicative setting, its content and its graphic realisation.

Not all of the features in Table 2.1 can be discussed in detail in this study. Nevertheless, all dimensions of personalisation are taken into account, and even though some cells receive less attention than others, all of them will be addressed. Chapters 4 to 8 each take their starting point with one feature (or a small group of features), which is mainly associated with either the communicative setting, the content or the linguistic realisation. However, the close interdependency between the various dimensions of personalisation will be discussed at various stages. Before proceeding to the analysis, however, chapter 3 will introduce the data.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Author submitted manuscript

3. Data collection

3.1 Working with online data

One of the advantages of online data is their accessibility. Online newspapers can be accessed from around the world, without even having to leave one's desk. In contrast, the acquisition of print newspapers outside of the country in which they are published can be more difficult. Normally, only international editions are available, which sometimes differ considerably from the local edition; down-market papers are often not available at all. An additional advantage of text from electronic sources is that it can easily be copied. In contrast to printed sources, it does not need to be digitized first, either through scanning and subsequent text recognition, or through manual typing. It is likely due to these advantages that data from online news sites are increasingly used to study news(papers) (see, for instance, Alonso Belmonte 2009, 52; Brüngel-Dittrich 2006, 11-12, 83; Chovanec 2000, 7; Valdeón 2006, 408).

However, online data also pose a number of practical problems for research. Compared to older mediums, the durability of texts published online is highly unpredictable. While it can be very difficult, if not impossible, to remove unwanted content from the Internet, there is also no guarantee that content persists over time, and even less that it persists unchanged. Online archives like the Wayback Machine[1] provide the possibility to access earlier versions of websites, back to 1996. However, this archive is far from complete. At most, it saves a few snapshots per week, and not all websites are included. Some sites might have been missed by the "crawler" (the software that automatically searches for online content), while others were skipped because the site owners explicitly excluded them from being archived. Moreover, not all pages can be archived in their original form; dynamic pages, which load data from a host each time they are accessed, provide problems for archiving and will not be accessible in their entire functionality.[2] With respect to online news sites, the Wayback Machine is not of much help since many news media organisations explicitly exclude their online sites from the archive (e.g. the *Times Online*). If an online news site is included, what can be accessed is restricted to the specific days on which the site was archived. Many links and interactive elements will, thus, no longer be functioning and a great deal of images and additional site content are missing.

[1] www.archive.org

[2] See http://www.archive.org/about/faqs.php#The_Wayback_Machine, accessed on 24/06/2010.

Author submitted manuscript

Online texts are not only unpredictable with respect to how long they remain accessible, but they also do not have the same level of fixity as texts produced in older mediums. Printed texts do not change after their production and if copies of printed texts, manuscripts or letters are edited, then the traces of this process remain visible. The same is true for texts that are handwritten on paper or parchment, or scratched into stone (Jucker 2003, 144). Similarly, spoken texts cannot be changed after their utterance. What can be edited are copies of recordings of spoken texts, but the original spoken text remains constant, even if it might no longer be accessible. This is crucially different for online texts. What one sees when accessing a particular website can be updated at any moment and the previous version is in many cases irretrievably lost. There is usually no point when one can be sure that a text is finalised and will not change its form anymore. It is also not clear, in most cases, which of the various versions of a website should be regarded as the original version. For the analysis of data published online, this means that instead of analysing a particular text, one can only analyse one specific version of a text, or several versions in comparison.

It is obvious that these characteristics of online data pose challenges for research (for further discussions of these problems, specifically with respect to online news data, see also Bateman et al. 2007; Karlsson and Strömbäck 2010; van der Wurff 2005; Ventola 2005). These problems make it necessary to save a copy of the data locally, since chances are very high that they will not remain unchanged during the period of analysis. Indeed, two weeks after I finished the data collection for this study, *BBC News* changed their commenting function. Only two months later, the *Times* announced it was closing down its online site timesonline.co.uk and replacing it with the completely redesigned sites thetimes.co.uk and thesundaytimes.co.uk. These examples illustrate the importance of collecting stable copies of online data that are not affected by such changes.

Saving local copies is not without difficulties, however. Online news sites tend to be highly complex, containing scripts and loading data interactively from the host when a site is requested. Therefore, simply saving a page in HTML format does not preserve the site exactly as it looks online and content might even be missing. Screenshots are also not an ideal solution. First, news sites are usually much larger than an average computer screen, so that one needs special tools that are capable of taking screenshots beyond what is visible on the screen.[3] However, even then, the problem remains that one only has a static image of the site, with no possibility to copy text, reconstruct links, browse through image galleries or even play videos. A combination of web page archiving tools, screenshots and manual storing of additional images, videos and audio files is therefore necessary if one wants to be able to analyse online news articles in their entirety (see also van der Wurff 2005, 7). Section 3.4 gives an overview of the tools and techniques that were used in this study.

[3] At the time of data collection, the online tool www.superscreenshot.com was capable of doing this and was used for the purpose of this study, see section 3.4.2 below.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

The language of mass media in general and news language in particular has attracted a great deal of attention from linguists and there has been a considerable amount of research on the development of newspapers and their language over time, starting with early pamphlets up to present-day print newspapers (for an overview, see Claridge 2010). In recent years, a great deal of work has been invested in preserving older newspapers and their predecessors and making them accessible in digital form. Such efforts have resulted in corpora like the Lampeter Corpus of Early Modern English Tracts containing pamphlets published between 1640 and 1740 (see Claridge 2003), the Lancaster Newsbooks Corpus of English newsbooks from 1653-1654 (see Prentice and Hardie 2009), the Zurich English Newspaper corpus, which includes early English newspapers from 1661 to 1791 (see Fries 2001; Lehmann et al. 2006), and the Rostock newspaper corpus containing English newspapers from 1700 to 2000 (see Schneider 1999, 2000, 2002). At the same time, newspapers themselves also show an increasing interest in digitizing their archives of print editions and making them accessible to the (paying) public.

While historical newspapers are thus increasingly being preserved and made accessible for research, the latest developments in news media, online news sites, pass by largely undocumented. What is to some extent preserved in the archives of news publishers and in online archives are the texts (or rather: one version of the texts) of published news articles. Their layout, however, together with the specific features of the online setting, can most often not be reconstructed. When the website of the *Times Online* was closing down on 15 June 2010, the texts of the articles published on the old site were still available from the article archive, but their layout, the interactive elements, videos, images and other additional content were lost overnight, together with all user comments that were published in response to the articles.

Manovich (2001, 6-7) argues that researchers on new media should focus less on predicting the future and more on describing and documenting the changes which we are presently witnessing. By describing and analysing five British online news sites at the beginning of the second decade of the twenty-first century, this study is contributing to such a documentation.

Author submitted manuscript

3.2 Characteristics of the online news sites

For this study, articles from five British online news sites were collected, namely the *Times Online*, the *Guardian*, *BBC News*, the *Mail Online* and the *Sun*. Except for *BBC News*, these news sites are the online equivalents of major national print newspapers. According to the National Readership Survey (NRS)[4], 87 percent of the readers of the printed *Times* and 89 percent of the readers of the printed *Guardian* in 2010 belong to the socio-economic classes A, B or C1. For the *Daily Mail*, the print version of *Mail Online*, this figure is considerably lower at 66 percent, while it is at only 36 percent for the printed *Sun*. Following Jucker (1992, 51), the print newspapers can thus be classified into up-market

[4] http://www.nrs.co.uk/toplinereadership.html, accessed on 19/5/2011.

Author submitted manuscript. Consult published version for correct page numbers!

(the *Times* and the *Guardian*), mid-market (the *Mail Online*) and down-market papers (the *Sun*).

Unfortunately similar data about the readership of online news sites are not available. For the purpose of this study it is therefore assumed that the online sites address an audience that is comparable to that of the print versions with respect to its socio-economic status. This does not mean that the actual readership of the print and the online version of a specific newspaper is necessarily the same. On the contrary, it is quite likely that there are systematic differences between the two readership groups. Nevertheless, media organisations do not appear to address a readership from a different socio-economic group in their online versions, compared to their print versions. The *Sun* does not address an up-market audience online, nor does the *Guardian* address a down-market audience in its online format. In the absence of more precise data, it seems a plausible approximation to take the socio-economic status of the targeted readership of print newspapers as an indication of the targeted audience of their online equivalents.

For the online news site of the *BBC*, estimations of the readership profile are more difficult. Again, demographic data of the visitors of the online news site are not available and parallels with the offline audience of the *BBC* are more difficult to draw than for the newspapers. Given that the *BBC* is a publicly funded organisation, it needs to address an audience that is as large as possible, which is achieved through a large range of TV and radio stations (Scannell 2005, 139-140). The TV channel *BBC News* is probably the most promising candidate for an approximation of the target audience of the *BBC News* website. According to the Ofcom PBS report on audience impact covering the period of 2006-2009, respondents of social classes A or B are more likely to be regular viewers of the TV channel *BBC News* than respondents from other social groups.[5] The BARB quarterly reach report indicates that 53 percent of the adult viewers of the TV channel *BBC News* are from the socio-economic classes A, B, or C1.[6] While this is considerably lower than the percentage of readers from classes A, B or C1 of mid-market newspapers, this is still not to be taken as evidence that this TV channel addresses a mid- or even down-market audience. The overall audience of TV is certainly very differently structured than the audience of newspapers. Audience profiles for TV channels need to be developed by comparing the proportions with which social classes are represented across different channels, in the same way as Jucker (1992, 48-52) identified such profiles for newspapers. Such a comparative perspective, albeit on the level of specific programs, has been adopted in the audience research annex to Ofcom's second public service broadcasting review of 2008. It locates the TV news programs of *BBC News* close to the ABC1 pole of the socio-economic class

[5] http://stakeholders.ofcom.org.uk/broadcasting/reviews-investigations/public-service-broadcasting/annrep/psb10/psbimpact.pdf, accessed on 24/6/2011.

[6] Based on data for the 4th quarter of 2010. Weekly reach of *BBC News* was 5,375,000 for adults of classes A, B, or C1 and 10,170,000 for all adults.
http://www.barb.co.uk/report/quarterlyChannelReachOpen?_s=4, accessed on 24/6/2011.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

continuum. Sky News and News at Ten (ITV1) are more at the centre, and Five News and ITV Evening News close to the C2DE pole.[7]

If this can be taken as evidence of the targeted readership of the *BBC News* website, it would suggest that the online site of *BBC News* can be grouped together with the *Times Online* and the *Guardian*. As my analysis will show, the three sites do indeed share many similarities and the differences between them are in most respects much smaller than their differences to the *Mail Online* and the *Sun*. I will therefore refer to the *Times Online*, the *Guardian* and *BBC News* as up-market news sites, to the *Mail Online* as mid-market, and to the *Sun* as a down-market news site.

These five online news sites were selected due to their popularity. A Nielsen/NetRatings report indicated that the *Guardian* was the most popular online site of British print newspapers in August 2007 with 2.7 million British visitors in that month. The *Mail Online*, the online site of the *Daily Mail*, was the most popular online site of a mid-market newspaper and ranked 5th overall. The *Sun*, the most popular down-market news site, was ranked 3rd.[8] The *Guardian*, the *Mail Online* and the *Sun* were, thus, chosen for investigation because they represent the most popular online sites of print newspapers for the three main market orientations (up-, mid- and down-market). While the figures in the Nielsen/NetRatings report for 2007 were taken as a basis for the selection of these three online news sites, it is likely that these figures have changed in the meantime. Unfortunately, no updated report by Nielsen/NetRatings is available. However, the Experian Hitwise Awards for News and Media in 2010 give a somewhat different picture. According to their rating, which is based on the market share of visits, but also includes online searches for brands, *MailOnline* was the most popular online version of a print newspaper in 2010, followed by the *Telegraph*, the *Guardian* and the *Sun*.[9] The online news site of *BBC News* appears to be more popular than any of the other sites. According to Hitwise, *BBC News* was the 9th most popular website in the UK in June 2011, making it even more popular than Wikipedia, Amazon or Twitter.[10] This alone makes the site worthy of more attention. In addition, it was chosen due to its similarity to online newspapers, despite not being related to a print newspaper. Comparing the online site of *BBC News* to the other online newspapers will enable me to see to what extent different mediums (TV and radio, on the one hand, and newspapers, on the other) converge online. The *Times Online*, finally, was included in the sample due to its long history, reaching back to 1785, the accessibility of its print version from 1985, and the role it has played in previous analyses of newspaper language.

[7] http://stakeholders.ofcom.org.uk/binaries/consultations/psb2_phase2/annexes/annex8.pdf, accessed on 24/6/2011.
[8] http://www.nielsen-online.com/pr/pr_021114_uk.pdf, accessed on 24/6/2011.
[9] http://www.hitwise.com/uk/2010-annual-awards/2010-winners, accessed on 24/6/2011.
[10] http://www.hitwise.com/uk/datacentre/main/dashboard-7323.html, accessed on 24/6/2011.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

3.3 Overview of collected data sets

3.3.1 The online data

A complete list of all articles that were collected for this study can be found in appendix A. Throughout the study, I will quote examples from my data set with the ID listed in the appendix. Whenever I quote examples from sites that are not included in this list, the URL and the date on which the site was accessed will be given in a footnote.

Table 3.1 provides an overview of the number of articles that were collected from the various online sites and their sections. Table 3.2 shows the number of words in each subset of the data. The data from all online news sites were selected in January 2010. Data were collected on every other day from 4 to 30 January. Letting one day pass in between collection days reduced the risk of collecting two different versions of the same article when stories are constantly updated in the case of breaking news. On each collection day, two news articles were collected from each site, one on a world news topic and one on a UK topic. In order to identify the topic category of the articles, I relied on the topic sections provided by the news sites. For the *Sun*, the distinction between world news and UK news could not be applied. The news site provides only a general category "news" and most of the articles deal with local topics or international topics presented from a local angle. The data from this site therefore all come from this generic news category, from which two articles were collected on each of the collection days.

Table 3.1: Overview of collected online data in number of articles

	Times Online	*Guardian*	*BBC News*	*Mail Online*	*Sun*	**TOTAL**
World news	14	14	14	14		56
UK news	14	14	14	14		56
News	**28**	**28**	**28**	**28**	**28**	**140**
Soft news	**14**			**14**		**28**
News columns	14			14		28
Soft news col.	14			14		28
Columns	**28**			**28**		**56**
TOTAL	**70**	**28**	**28**	**70**	**28**	**224**

Within each category, the actual articles had to be chosen. There are several strategies for doing this. A common strategy is to collect articles on one particular news story across various news publications. However, this was not an option for my own study. On the one hand, there are considerable practical problems involved, especially when newspapers with different market orientations are compared. The *Guardian* tends to cover very different topics from the *Sun* and it is often quite difficult to identify stories that are covered on all sites. More importantly, since I am not only interested in the representation of particular news events, but also in the range of topics that the various sites focus on (see chapter 6), such a selection strategy would have posed serious problems for my investigation. I therefore decided to collect the topmost article from the overview pages of the UK news and the

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

world news sections, and in the case of the *Sun* the two topmost articles from the overview page of the news section. In online settings, a position at the top of the overview page is the most prominent position, comparable with the front page position in print newspapers (Karlsson and Strömbäck 2010, 6, 12-13).[11] In many cases, the topmost articles from the UK news and world news sections were also presented on the start page of the online news site, which is, thus, more evidence that these were the articles that were given high prominence. By using this selection strategy, I was able to compare what topics were considered to have a very high news value across the different news sites.

On the fourteen days on which I collected my data, there were three news stories which appeared as the most prominent item on all the news sites. One topic was a world news event, namely the earthquake in Haiti. The other two stories concerned the UK. One of them dealt with the heavy snow falls in this period and has already been mentioned in the introduction. The other event concerned a court case on two boys who were brought to court for torturing two other boys in the town of Edlington. These stories allow more detailed comparisons across the five news sites and will therefore be used for case studies.

In addition to news articles, I also collected soft news articles from two of the online news sites, namely from the *Times Online* and the *Mail Online*. "Soft news" is a somewhat problematic category insofar as it is difficult to provide a clear-cut distinction (see also Bell 1991, 14; Tuchman 1978, 47-48). Soft news is always defined in opposition to hard news and definitions mainly rely on the content of the texts. Whereas hard news deals with content that is of direct political relevance, soft news deals with human interest stories (e.g. Burger 2005, 211; Franklin et al. 2005, 247). This distinction is complicated by several factors, however. First, Bell (1991, 14) points out that journalists often try to present soft news as if it were hard news, emphasising the relevance of the story for society at large. Second, the opposite is also the case. Focussing on the effects of an event on the lives of particular individuals is a relatively common strategy for introducing and framing

[11] In contrast to print newspapers, the placement of individual articles is not stable, however. An article may remain at the top of on overview page for only a few hours (see Karlsson and Strömbäck 2010, 12-15), whereas the front page of print newspapers does not change.

Author submitted manuscript

Table 3.2: Overview of collected online data in number of words

	Times Online	*Guardian*	*BBC News*	*Mail Online*	*Sun*	**TOTAL**
World news	11,872	10,732	9,492	10,598		42,694
UK news	10,571	12,079	9,831	23,805		56,226
News	**22,443**	**22,751**	**19,323**	**34,403**	**16,869**	**115,789**
Soft news	**29,045**			**12,743**		**41,788**
News columns	11,394			12,969		24,363
Soft news col.	14,677			14,619		29,296
Columns	**26,071**			**27,588**		**53,659**
TOTAL	**77,559**	**22,751**	**19,323**	**74,734**	**16,869**	**211,236**

Author submitted manuscript. Consult published version for correct page numbers!

articles on hard news topics (Niblock 2008, 49). Third, stylistic characteristics and the way in which the text is presented in the context of a news publication contribute to the perception of a specific news item as either hard news or soft news. Burger (2005, 212), for instance, mentions the case of German-speaking television news, where soft news items usually occur as the last news item of a news programme (see also Luginbühl and Perrin 2011, 591). In addition to their placement, these stories are also marked through the facial expression of the news reader, who, according to Burger, typically announces the item with a relaxed smile. In print newspapers and on online news sites, soft news articles typically occur in specific sections, separated from hard news stories. However, as soon as the contextual and stylistic features do not correlate with the content of the article in the prototypical way, it becomes difficult to decide whether a specific article should be considered hard news or soft news. That the categorisation into hard news and soft news is not only a consequence of the nature of the news event can also be seen in cases in which a specific news event is presented as soft news in one publication and as hard news in another (Tuchman 1978, 48). The audience orientation seems to play an important role, too, with mid- and down-market papers frequently presenting topics as hard news which are restricted to non-news sections in up-market papers (see also chapter 6).

For the collection of soft news articles for my study, I tried to identify specific sections on online news sites which contain the most prototypical soft news articles. I restricted the soft news data to two online sites, namely the *Times Online* and the *Mail Online*. For the *Sun* and *BBC News* no comparable data was available and I therefore decided to include only data from one up-market and the mid-market site. The soft news articles from the *Times Online* came from the category "Life & Style", which appeared to best represent the prototypical soft news articles published on the site. The articles deal with health topics, in particular dieting, as well as fashion, and background reports on celebrities and individuals with interesting life stories. As was the case for the news articles, the topmost article was selected every other day from 4 to 30 January. The *Mail Online* does not have a category with the same name. Therefore, I tried to identify the category or categories that contained articles which most closely resemble those in the Life & Style category of the *Times Online* in terms of content. This was necessary to ensure that the texts I compared shared a common basis for comparison. I decided that this could be best achieved by alternating my selection from the *Mail Online* between the two categories "Femail" and "Health". The fact that I tried to collect articles with similar topics on both sites means that a comparison of the topics of these articles can hardly lead to any insightful result. I will therefore restrict my analysis of this material to selected linguistic features of personalisation.

Opinion columns were also collected from the *Times Online* and the *Mail Online*. I collected two columns per collection day from each of the two news sites, one dealing with a political or hard news topic, and one dealing with a soft news topic. On the *Times Online*, columns are listed on the overview pages of each topic area. They are made recognisable as columns through the heading "Comment" and the name of the columnist. Often a small photograph of the columnist is added and some columns are additionally marked by a yellow

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

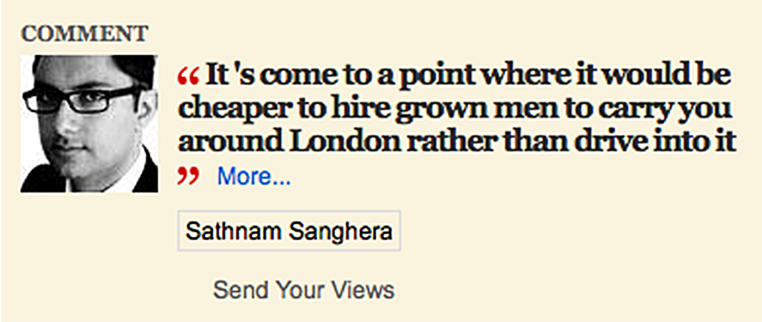


Figure 3.1: Column announcement on the Life & Style front page of the *Times Online* (5/1/2010)

Author submitted manuscript

background (see Figure 3.1). The (hard) news column was collected from the news front page and the soft news column from the Life & Style frontpage. The *Mail Online* contains a page “columnists” as part of its debate section. On this overview page, all columns are ordered according to the newspaper section with which they are associated (e.g. News, Sport, Femail, etc.). Since the columns on this overview page within each category are arranged in alphabetical order of the name of the columnist, it was not possible to always select the topmost column of a particular category, since this would have resulted in texts from a single columnist. In addition, it turned out that the columns remained on this overview page for a long time, which caused the danger of collecting the same column several times. However, the overview site also provides a section “today”, under which all columns are listed which were published on that day. In order to select columns I therefore went through this list and collected the first column related to the news section and the first column related to a soft news section. Since there was sometimes a lack of columns from the categories “Femail” and “health”, columns from the categories “TV & Showbiz” and “You mag” were also included for the soft news columns.

Whereas hard news reports are probably the best studied category of texts in newspapers, other types of articles have been investigated far less (Ljung 2000, 131-132). There is research on non-news articles, especially opinionated texts, such as editorials and columns, but linguistic studies that compare different types of texts within news publications are still not very common (but see, for instance, Grösslinger et al. 2012; Jucker 1992; Sanders 2010). This is regrettable because newspapers, whether online or print, do not only consist of hard news. In fact, there is evidence that soft news content and opinionated texts make up an increasing portion of news publications (Niblock 2008, 48). To take hard news articles as representing the language of newspapers would therefore be a gross simplification. At the same time, it is also not entirely unproblematic to include texts that are so different from each other into an overall category of “news language”. Bell (1991, 14-15) points to potential research problems for the Brown and LOB corpora, which fail to make a distinction between hard news and soft news. Furthermore, the Longman Spoken and Written English Corpus, which is the basis for the frequency information in the *Longman*

Author submitted manuscript. Consult published version for correct page numbers!

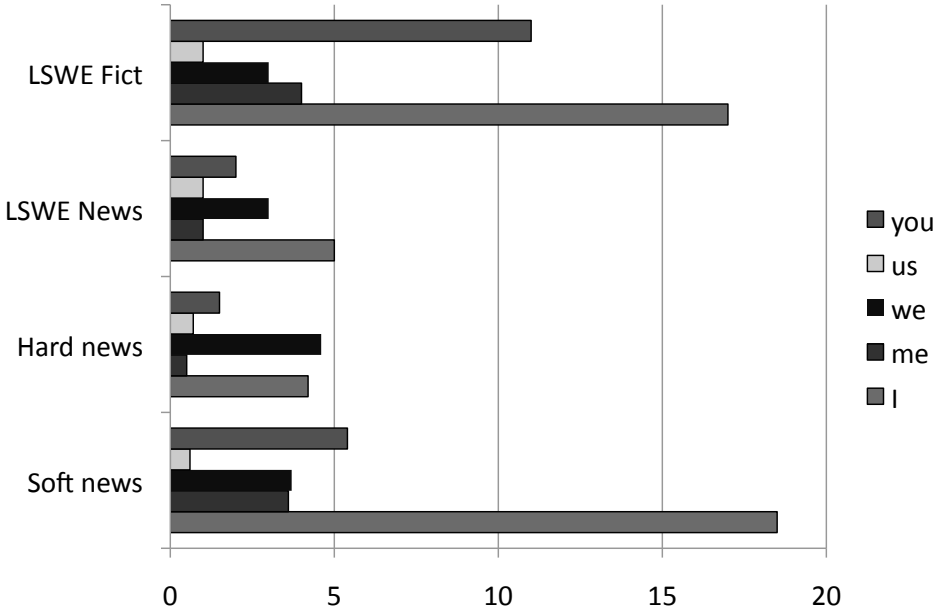


Figure 3.2: Frequency per 1,000 words of selected first and second person pronouns in soft news and hard news articles of my data, compared to the registers news and fiction in Biber et al. (1999, 334)

Author submitted manuscript

Grammar of Spoken and Written English (Biber et al. 1999), does not further distinguish between different types of articles within the news register. Biber et al. (1999, 32) provide a breakdown of the various news sections within the news data, but do not take this into account for their quantitative evaluations of linguistic features. In order to illustrate the potential problems, Figure 3.2 compares the frequency of selected first and second person pronouns in the hard news and soft news sections of my data with the two registers news and fiction in the LSWE corpus. This very rough comparison suggests that the distribution of first and second person pronouns in the soft news articles of my data has probably more similarities with the LSWE register of fiction than with the register of news. This also means that the frequency information for news given in Biber et al. (1999) must be heavily dependent on the relative proportions with which different types of news articles are represented within the register of news.

It is beyond the scope of this study to give a comprehensive account of all the various types of articles one can find on online news sites. Nor will I be able to discuss the data from soft news and columns with respect to all features of personalisation. However, by including soft news and columns in addition to hard news, I would like to draw attention to the role such articles play in online news as well as to their linguistic differences from hard news. The selective evaluations of these data carried out in this study suggest that there is a great deal of potential for further research in this direction.

Author submitted manuscript. Consult published version for correct page numbers!

Table 3.3: Overview of collected data from the *Times* from 1985 in number of articles and number of words

	number of articles	number of words
World news	18	10,172
UK news	18	13,015
TOTAL	**36**	**23,187**

3.3.2 The Times from 1985

All the articles in the data from the *Times* from 1985 were published in January, exactly 25 years previous to the online data. Since the articles in the older issues tend to be shorter than the articles in the more recent online edition, I collected slightly more articles from 1985.[12] While there are 28 news articles from the *Times Online* in the 2010 data (14 from each of the two subcategories UK news and world news), there are 36 articles from the *Times* from 1985 (18 from each of the two subcategories). They were collected in the 18 issues of the *Times* published between 7 January and 26 January.

The articles were selected with the intention of resembling the data from the *Times Online* as closely as possible. For both the UK and world news articles, the most prominent article on the front page was chosen. How prominent an article is was decided on the basis of the size of its headline and its placement. If two articles had a headline of the same size, the one placed higher on the page (including pictures presented with the article) was considered more prominent. If two articles were also placed at the same height, the one that was horizontally closer to the centre of the page was considered more prominent. And if two articles were also equal with respect to this criterion, then the article with a picture was considered more prominent. In several cases, the selected articles were only started on the front page, but continued on one of the following pages or the back page. In such cases, both parts of the article were collected. Related articles which were mentioned at the end of a complete article were not collected.

The data could be accessed via the Times Digital Archive, which – at the time of collecting my data – provided access to scanned versions of all issues of the *Times* from 1785 to 1985.[13] The database allows one either to save individual articles as image files in PNG format, or to download an entire page of the newspaper as a PDF file. Unfortunately, a text version of the articles is not available and an automatic OCR text recognition of the available formats returned results which were of very poor quality. In the end, the most effective way of transforming the articles into a textual format was manual typing. The amount of work involved in making these articles available for research is one of the reasons why not more data from earlier print newspapers were included in this study. Certainly,

[12] The news articles from the *Times* from 1985 are on average 644 words long, while the news articles from the *Times Online* are 802 words on average.

[13] In the meantime, the archive has been extended up to the year 2006.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

it would have been interesting to extend the analysis to earlier print issues of the other online news sites. In particular, the comparison with older data from down- and mid-market newspapers would have allowed important insights into the role of tabloidisation tendencies for the diachronic development of personalisation. However, it is also the case that the number of perspectives that can be integrated into any study is limited. My main focus lies on the synchronic comparison of the five online news sites, and, in particular, on the role of their market orientation. While I take an occasional glance at the variation across different categories of articles and at diachronic developments, a detailed exploration of these aspects must be left for future research.

3.4 Technical aspects

As mentioned in section 3.1 above, there are a number of problems to be solved when collecting and analysing online data. The data on the Internet is not stable; it can be modified or deleted at any time. To save a local copy is therefore an important prerequisite for any detailed analysis. However, it is by no means trivial to decide in which format online data should be saved. Every available format tends to preserve some features of the original data, while losing others (see also Karlsson and Strömbäck 2010, 7-8). The choice of the format, therefore, depends mainly on the aim and method of the subsequent analysis. To copy text from a website into a text editor may be a good solution if one is only interested in a textual analysis, whereas even a partial screenshot might be sufficient for an analysis of certain aspects of the layout.

For the present study, various parts of the analysis required the preservation of different aspects of the data. In order to be able to analyse linguistic features the text had to be available in a separate text file that could be processed further. For the analysis of interactional features, it was important to preserve the entire page on which the article was published, including the original layout and links to additional pages and resources. The analysis of visual elements made it necessary to also preserve image galleries and videos which were published together with the news article.

Unfortunately, it turned out that at the time of collecting data for this study none of the available formats and technologies was able to fulfil all these needs simultaneously. As a consequence, each article was first saved in two different formats, as a Zotero snapshot and as an extra large screenshot (see below for further explanations). Additional images, video and audio content were saved separately. From the Zotero snapshot, the text was then copied and transformed into an XML representation that allowed tagging of linguistic features.

3.4.1 Zotero snapshots

Zotero is a plugin for the Internet browser Firefox.[14] It is a tool for managing research resources and it can, in particular, be used to manage literature, similar to programs like

[14] www.zotero.org (accessed on 2/9/2011)

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

EndNote or Citavi. In addition, it also allows one to store offline versions of webpages, a function which is called "snapshot" in Zotero. To save an offline version of a webpage sounds more simple than it is. For complex websites like the online news sites I analyse, the content is distributed over various files and much of the content is generated dynamically at the moment of accessing a site. For this reason, saving the page manually with the normal browser functions did not produce any useful results and also most of the tools I tested only managed to store a small part of all relevant content. Comparatively speaking, the Zotero snapshot function provided quite good results. It did not only save all the text, but also (with few exceptions) the integrated images and it preserved a large part of the complex layout of the sites. User comments which were submitted below the article were also stored, as were links to other articles and external sites. In addition, Zotero automatically saved the title of the page, its URL and the time at which the snapshot was taken, which was of considerable help for managing the data. Since it is a tool for managing resources, Zotero could also be used to organise the data, creating collections, applying tags and adding comments. The separately collected screenshots could be added as an attachment to the respective snapshot, and the same was done for additional content such as manually saved images in galleries and videos.

Despite all these advantages, Zotero snapshots alone were not sufficient for my purpose. In many cases, the layout of the articles was distorted, with the advertisements that were integrated into the news sites posing particular problems. Side bars were also sometimes misplaced and in some cases images were missing. For this reason, the Zotero snapshots were complemented with screenshots.

3.4.2 Screenshots

Taking a screenshot of a website seems straightforward enough, but once again difficulties occurred. The main difficulty was posed by the fact that most articles of online news sites are considerably larger than a computer screen. For normal reading purposes this is, of course, no problem, since one can easily scroll down. It is a problem, however, for taking screenshots, given that most screenshot tools do what their name suggests: taking an image of the computer screen. If only part of an online news article is visible on the screen when taking the screenshot, this is also the only part visible on the resulting image. This was not good enough for my purposes, since I needed to preserve the entire news article with all surrounding content.

At the time of collecting my data, an online service provided the solution. It was called Super Screenshot[15] and generated an image file of any URL that was submitted. The tool also allowed one to select the desired resolution and to choose between two image formats, JPG and PNG. In this way, it not only allowed one to save an image of the entire web page, but the resulting image was also available in high quality, which was desirable for

[15] www.superscreenshot.com (no longer accessible)

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

my purpose. Unfortunately, the service was no longer working by the time I finished this study.

3.4.3 Inconsistencies between formats

The fact that all online data were saved in two different formats led to new complications. It was not possible to save both formats without a delay of at least a few minutes and several times it happened that an article was updated between the first and the second format, leading to two different versions of the same article. In addition, the service that was used to create the extra large screenshots did not work on two of the days on which I collected my data, so that I could only collect the screenshot format several days after collecting the Zotero snapshot. This substantially increased the probability that the article was updated between the collection of the two formats.

Several steps were taken to avoid negative effects from inconsistencies in the data. In cases in which there were major differences between the two versions, the article in question was saved again in both formats at a later point. In most cases, this later version was collected only an hour or two after the initial collection, but in some cases the collection took place several days later. To give a concrete example, the article mail-wn-100104 was published on the *Mail Online* on 4 January 2010 as the topmost article in the world news section. I collected this article on the same day, but, due to technical problems, the screenshot had to be taken again two days later, 6 January. In the meantime, the article had been updated and considerably modified on 5 January and, thus, I also collected a new Zotero snapshot. As a consequence, the version I analyse is actually the version of 5 January, even though I refer to it as the world news article of 4 January, the day on which it was selected for inclusion into my data set.

By saving a new version of articles for which there was a large discrepancy between the two formats, most problems causing inconsistencies in the analysis could be avoided. In addition, I decided that the Zotero snapshot was the relevant format for analysing features of content and language. The screenshot format was only consulted in complementation to the Zotero snapshot when the exact layout of the site was of importance. Assigning different functions to the two formats further reduces the risk of inconsistencies in the analysis.

3.4.4 XML format

After saving a screenshot and a Zotero snapshot of each online article and typing the older printed articles, the text of the articles was copied into an XML file. For the purpose of this study, an XML schema was developed that allowed one to save context information such as the date of collection, the news site and section from which it was collected and the names of authors. The texts were then annotated with additional tags for the analysis of the linguistic realisation of the articles. Tagging included direct speech, first and second person pronouns, and headlines. Tagging was done semi-automatically, using regular expressions

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

to recognise the desired strings and patterns. The results were then corrected manually (see also chapter 8.3.1).

One of the main advantages of the XML format is that it is well-suited for both quantitative evaluations of the frequency of tags and detailed qualitative analyses of features in their larger context. Using the query languages XPath and XQuery, the frequency of tags in specific sections of the data can easily be determined. Such a query can, for instance, return the number of all first person pronouns in news articles, or only those instances that occur either within direct speech or outside of direct speech. Furthermore, restricting the occurrences to specific case forms or to specific news sites or dates can be done without any additional manual work, provided the required information has been encoded in the XML tags. At the same time, the context of each occurrence of a tag is also easily accessible. Many XML editors, such as EditiX and Oxygen, make it possible to click through a list of hits in a sidebar window and thus view each occurrence in its context in the main window.

Only the texts of the main articles, together with the respective headlines, subheadlines, bylines and leads were included in the XML format. Boxes with additional information, links, visual elements, captions of images, navigational elements, and so on were not included. However, it was not clear in all cases whether a specific text segment should be regarded as belonging to the main article, or as constituting an additional element, similar to an information box. For instance, there are articles and columns of the soft news section which end with giving information about recommended books or restaurants, or with announcements of television shows on the topic of the article. In cases in which this information is presented directly following the previous text and without a distinctive layout – such as a differently coloured background, a box, or a horizontal line between article text and this information – it was considered to be part of the article and included in the XML data. One reason for this decision is that in some cases this information is very closely integrated into the previous text, so that a clear separation between article text and further information could not be made. Figure 3.3 illustrates this case with an article from the *Sun*, where there is a fluent transition from the main news text reporting on a girl bleeding from her eyes to the announcement of a documentary on the same girl.

There is one type of textual element that sometimes appears at the end of articles and which is always excluded, namely requests for the audience to submit information, personal experiences, comments, pictures or videos to the site. Such requests are in most cases clearly set apart from the article text by layout elements, so that they can be excluded on the basis of the criteria mentioned above (see Figure 3.4). These requests are also excluded in cases in which the layout is not clearly distinct, as long as they are placed at the end of an article. The reason for this is that they differ considerably from the text of the article, both with respect to their function and their (linguistic) form. This second point is of particular relevance, since it affects my analysis. One of the features I analyse based on the XML version of my data is the use of personal pronouns. Requests for comments and eye witness accounts regularly contain instances of *you* which refer to the reader, a feature

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

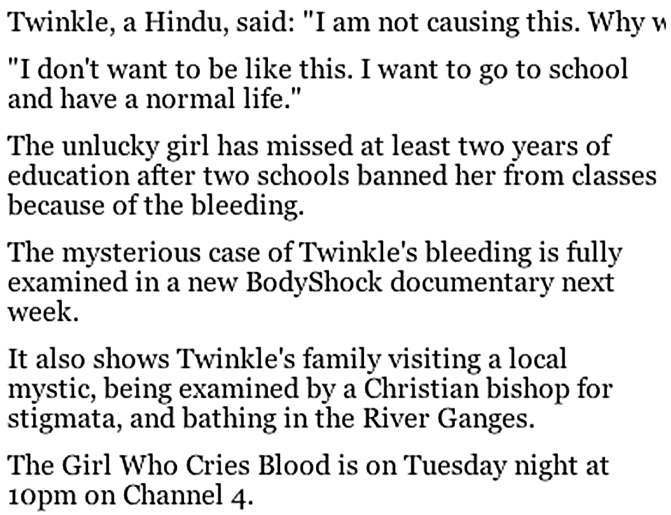


Figure 3.3: Example of a fluent transition from article text to television announcement at the end of article sun-100110b

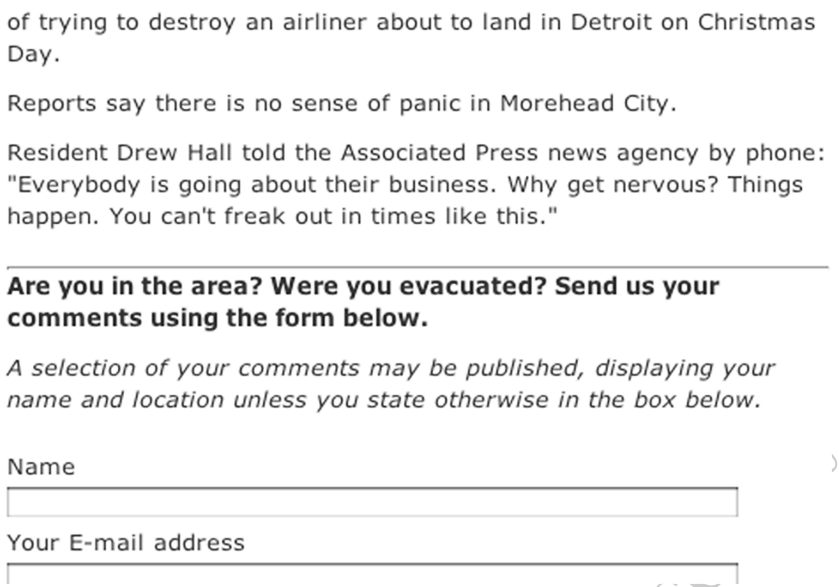
Author submitted manuscript

which is otherwise quite infrequent in news articles. Including these requests into my XML data would therefore have skewed my results.

Another issue that poses problems for deciding what to include in the XML data is the fact that sometimes several articles are combined into one. It is not in all cases obvious to decide whether two items are subsections of a single article or two separate, subsequent articles. Again, the layout provides the criterion for distinguishing between these two cases. If some text is presented in a separate box or on a distinct background from the normal article background, it is considered not to be part of the main article and therefore excluded from the XML data. This is frequently the case for articles on the *Sun*, where additional text is sometimes presented on a grey background in a separate column to the right of the main article. It can be argued that this additional text has a similar relation to the main article as an independent article, to which it is linked. The parallel to teasers of related independent texts is further increased through the fact that for some articles on the *Sun* only the first part of the additional text is displayed by default with a "more" button at the end, which expands the grey area on which the additional text is displayed, so that it can then be read completely. In contrast, text that is introduced by a second (sub-)headline, but which immediately follows the main article and is displayed in the same format as the text above is considered to be part of the main article and is included in the XML data.

One of the most difficult cases for classification is an article from the *Mail Online* (mail-uk-100124), dealing with the Edlington court case. Following the first part of the article, which provides a report of the case, is a second part, introduced by a separate headline and byline, but formatted like the first part of the article. This second part presents an account of the events by the parents of the victims. In addition, there is an inset box with

Author submitted manuscript. Consult published version for correct page numbers!


of trying to destroy an airliner about to land in Detroit on Christmas Day.

Reports say there is no sense of panic in Morehead City.

Resident Drew Hall told the Associated Press news agency by phone: "Everybody is going about their business. Why get nervous? Things happen. You can't freak out in times like this."

Are you in the area? Were you evacuated? Send us your comments using the form below.

A selection of your comments may be published, displaying your name and location unless you state otherwise in the box below.

Name

Your E-mail address

Figure 3.4: Request for audience comments at the end of article bbc-wn-100112, separated from the article text through a horizontal line and other layout elements

a grey background in the first part of the article, which contains a brief account by one of the victims. According to the criteria established above, the text in the grey inset box is not considered to be part of main article, whereas the parents' account is considered to be part of the main article. In the analysis, this decision proved to be of great importance, because the second part of the article differs considerably from other news articles with respect to its linguistic characteristics. As a consequence, the article mail-uk-100124 will receive special treatment in (parts of) chapters 7 and 8.

3.5 A comment on statistics

In appendix B, I provide results for significance tests for some central quantitative analyses of my data. The tests are mostly calculated for different pairs of data sets (e.g. comparing the *Times Online* with the *Mail Online*) and the results are given in the form of p-values, which can also be used as a measure of similarity between the data sets. However, the data on which these statistical tests are based are problematic in some respects and the results of the tests should therefore be interpreted with care.

The chi-square test, like most statistics, is "based on a random sample model, which assumes that the observed data (i.e. the corpus) were selected randomly from the language or sublanguage of interest" (Evert 2006, 177). One serious problem for the random sample model is the tendency for linguistic features to occur in clusters. If larger sections of texts are included in a corpus, a specific feature can be seriously over-represented if it happens to occur repeatedly in this section. In Evert's words:

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

> A [...] problematic cause of non-randomness is the discrepancy between the unit of measurement (typically words or sentences) and the unit of sampling (entire documents or long connected fragments). Since word types and many other linguistic phenomena tend to cluster together in the same document, observed frequencies are inflated when such documents are included in the corpus [...] (Evert 2006, 189)

The tendency for clustering together is, indeed, quite high for the features I investigate. As I will discuss in detail in chapter 8.4, first and second person pronouns have a strong tendency to occur in clusters. Sometimes only one or two sentences contain all occurrences of these pronouns in an article. For direct speech, clustering is a problem, too. Given that direct quotes usually contain more than one word, the presence of a word within direct speech makes it very likely that neighbouring words are also within direct speech.

In order to achieve complete randomness in a corpus, the unit of measurement and the unit of sampling would need to be identical (Evert 2006, 185). In this way, the negative effects of clustering could be avoided and the chi-square test would produce reliable results. This means that, from a purely statistical point of view, each word from my corpus would need to be randomly selected from a different article (Evert 2006, 185). Obviously, such a corpus does not make much sense from the point of view of a linguist, because interpretations of the use of linguistic features are impossible without context. Consequently, Evert argues that non-randomness is a problem for all corpus linguistic studies, which, in the long run, can only be solved by developing new statistical methods:

> There can be no doubt, however, that non-randomness is a ubiquitous problem in corpus linguistics. In a longer perspective, it will thus be necessary to develop new statistical methods for the analysis of corpus frequency data. These methods should be able to deal with non-random data and correct for the larger amount of variation in some way. First tentative steps have been taken – either by adjusting the unit of sampling [...] or by modifying the random sampling model itself [...] – but much work still needs to be done. (Evert 2006, 189)

In the absence of new statistical methods, linguistic corpora usually try to address the problem by limiting the size of the samples; they neither select words randomly, nor do they use a small number of complete texts. Instead, they include samples from a large number and variety of texts, and they try to ensure that the texts from which the samples are taken are independent from each other. For this study, this procedure was not an option. Given that I was not only interested in linguistic features, but also in the content of articles, it was important to include complete texts. News articles are known for their inverted pyramid structure. Information is not spread evenly through the text, but it tends to be given in condensed form at the beginning of an article and in increasing detail towards the end. This means that cutting articles after a specific number of words would have caused serious problems, because it is likely that more detailed background on events and individual actors would have been missed systematically. Indeed, it is reasonable to assume that this would not only have affected the interpretation of the content of articles, but also of their linguistic features. Direct speech quotes are often used to provide

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Table 3.4: Frequency of direct speech in two subsets of my news articles from the *Times Online*

	Words in ds	**Words outside of ds**	**Total words**
mult-4	2,561	8,993	11.554
not-mult-4	2,618	8,271	10,889
Total	5,179	17,264	22,443

additional detail and personal views on information that was previously given, and thus they can be expected to appear more frequently in later parts of an article.

As a consequence of including entire articles – which seriously violates the random sample model – the application of statistical tests to my data is problematic. The results of such tests are likely to indicate a statistically significant difference between subcorpora, which may be entirely due to the lack of independence of my samples. A short example can illustrate this point. I divided the news articles from the *Times Online* into two subsets, mult-4 and not-mult-4. The first contains the 14 UK and world news articles from my data which were published on dates which are a multiple of 4 (e.g. 4, 8, 12 January); the second contains the other half. I then calculated the number of words inside and outside of direct speech in each set (see Table 3.4).

All articles come from the same news site and the two groups are balanced with respect to the two subsections UK news and world news. Assuming that there are no underlying factors resulting in different frequencies of direct speech on particular days, one would therefore expect that there are no statistically significant differences between the two groups. However, a chi-square test (with Yates' continuity correction) results in a p-value of 0.0009. There is thus a very high probability for a significant difference between the two groups ($p < 0.001$). If statistically significant differences can be observed so easily within the data from one news site, this means, of course, that similar differences across subcorpora need to be treated with great care.

A different problem for statistical testing concerns sparse data. This is in particular the case for the evaluation of the general topic areas of news articles in chapter 6. Several of the cells do not have any occurrences (e.g. there is no celebrity news in my data from *BBC News*) and other cells contain very low counts. This means that the chi-square test is not applicable. One possible solution would be to include more data, leading to generally higher counts. However, the problem could not entirely be solved in this way, because the differences across the sites are simply too marked. Even a much larger sample of top-listed news articles from *BBC News* or the *Times* from 1985 would not result in a sufficient count for categories such as celebrity news. The only way to apply a chi-square test to such a table would be to radically merge categories. However, by doing this, one would lose analytically important distinctions, for example the distinction of a separate category of celebrity news. In other words, while it is evident that the *Sun* contains more articles

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

on celebrities in its news section than *BBC News* does in its, it is not trivial to show this difference between the two sites with a test for statistical significance.

It is not the main aim of this study to provide evidence of statistically significant differences between the subcorpora of my data. I want to illustrate, first and foremost, how specific features – independently and in combination – can contribute to personalisation. In addition, I want to point out some very general differences between the various sections of my data. These differences can often be clearly illustrated with examples, and quantitative data provide just one additional form of illustration. This is not to say that a statistically sound analysis of these features would not provide an interesting research topic. However, given that such an analysis would require different data, I leave this open for future research.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

4. Feedback and interaction

One of the most frequent and pervasive ways of personalising online news is through inviting the audience to interact directly with text producers and with each other. The following examples illustrate some of the forms such invitations can take.

(4.1) Would you like to post a comment? Please register or log in[1]

(4.2) What do you want to talk about? Post your suggestions for subjects you'd like us to cover on Comment is free[2]

(4.3) Have you seen or been involved in a news event? Is something significant, bizarre or unusual happening where you live? Have you got a story to tell or is there something you think we should follow up?[3]

(4.4) TODAY'S POLL: Who would be a better Labour leader? VOTE[4]

(4.5) Secret war: Can scanners stop bomb attacks? Discuss here[5]

Online news sites thus ask their readers to submit comments at the end of articles (example 4.1), to suggest topics for articles and comment sections (examples 4.2 and 4.3), and to participate in online polls and forum discussions (examples 4.4 and 4.5). These means of direct interaction allow readers to leave the anonymity of the mass audience and to become visible as individuals with names, opinions, and sometimes even pictures of themselves. Moreover, the fact that readers turn into active participants who contribute content changes the relationship between news and the audience with respect to the news value of relevance. Traditionally, news is often presented in a way that emphasises its relevance for the readers' own lives (Bell 1991, 157-158). In contrast, asking readers to submit their own experiences suggests that the readers' own lives are (potentially) relevant news.

The personalising effects of interactional elements are not restricted to the audience. By providing viewpoints of directly affected individuals, readers submitting content help personalise news events. Furthermore, in cases in which individual text producers are presented as interlocutors, they undergo personalisation, too. Finally, the language used both in audience content and in requests for it contain linguistic personalisation features, such as the use of the pronoun *you* to address the audience in examples 4.1 to 4.3. In other words, reader interaction is one of the places where various facets of personalisation intersect.

[1] http://www.timesonline.co.uk/tol/news/politics/article6975099.ece, accessed on 4/1/2010
[2] http://www.guardian.co.uk/commentisfree/2010/jan/21/you-tell-us, accessed on 24/1/2010
[3] http://news.bbc.co.uk/2/hi/talking_point/your_news/7593687.stm, accessed on 22/6/2010
[4] http://www.dailymail.co.uk/home/index.html, accessed on 8/1/2010
[5] http://www.thesun.co.uk/sol/homepage, accessed on 18/1/2010

Author submitted manuscript

Features promoting reader interaction are quite popular and even appear to be increasing, both in their overall use as well as in their variety of forms. For readers, the attraction may lie in seeing their own contributions published online and in interacting and debating with other readers or even text producers. For text producers, such interaction can provide them with feedback on their articles, information about the background of some of their readers, and maybe even with new sources and story ideas.[6] Not least, economic considerations play a role, insofar as reader involvement helps strengthen the tie between readers and "their" news site. Enli (2008) even suggests that audience participation can be used as a strategy to increase the legitimacy of public service providers like the *BBC*, who are required to appeal to a broad audience.

In this chapter, I am going to investigate such interactional features in the online news sites of my data. I will start by introducing a model of interactive mass communication that allows one to describe such forms of interaction without giving up the distinction between mass communication and interpersonal communication. Based on previous research of interactional elements on online news sites, I will then provide a very brief discussion of how such features have developed since the early days of online news sites. Section 4.2 gives an overview of some of the most prominent forms of audience interaction that can be found in my data, including some first observations concerning differences between the various news sites. These differences are the main focus of section 4.3, which discusses how the news sites in my sample integrate two forms of user-generated content into their articles.

4.1 Audience interaction in mass media communication

4.1.1 Modelling audience interaction

Interaction with members of the mass audience is not a new phenomenon. Communicative forms that are not purely unidirectional were used and studied before the widespread use of the Internet, in particular radio phone-ins and television talk shows (e.g. Armstrong and Rubin 1989; Bierig and Dimmick 1979; Brinker 1988; Burger 1989, 1991; Crittenden 1971; Garaventa 1993; Leitner 1983; Linke 1985; Mininni 1990; Selting 1985; Simonelli and Taggi 1985; Troesser 1983; Turow 1974). In radio phone-ins, individual members of the audience interact by phone with the moderator of the radio show and this interaction is broadcast. The boundaries between the passive audience and the active text producers are thereby transgressed by those listeners who participate in the phone-in. The phone-in, thus, "creates the illusion of radio as a two-way medium" (Crisell 1986, 181), which poses a challenge for a purely unidirectional model of mass media communication. Similarly, a model of unidirectional mass media communication is not capable of representing the situation of TV talk shows. In such shows, a host interacts with a number of guests in the presence of a studio audience, and this interaction is broadcast to a mass audience.

[6] See, however, Hermida and Thurman (2008), who find that text producers often remain sceptical about the benefit of audience participation.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

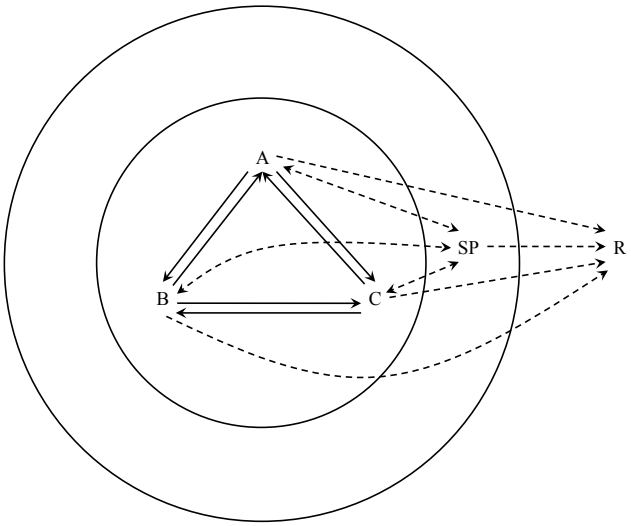


Figure 4.1: Communicative circles in a TV talk show according to Burger (2005, 21)

Author submitted manusc. ipt

Thus, the audience is divided into two segments, the first of which is present when the interaction takes place, while the second is displaced in space and often (in the case of pre-recorded shows) also in time. In addition, members of the studio audience might be actively involved in the interaction with comments or questions addressed to the guests. In the case of talk shows which are broadcast live, such comments and questions can even be contributed by members of the audience outside of the studio, either through phone-ins or more recently via online responses.

Burger (2005, 19-23) proposes a model involving different audience circles to account for such complex communicative situations. Figure 4.1 represents the communicative framework of a TV talk show. A, B and C stand for the host and two guests, who communicate in the inner circle. SP in the second circle stands for the audience present in the studio. The studio audience is addressed by the members of the inner circle and responds non-verbally (e.g. by applauding or laughing) and maybe also verbally with comments and questions. In the third circle, the mass audience R receives the communication and can be directly addressed by various members of the first and second circle. In this example, the mass audience does not respond, but, in the case of live broadcasting, it would be possible to include this option too.

In news reporting, such complex communicative situations are rare, even though they can occur in broadcast news (see Montgomery 2007, 29-30). Text producers regularly communicate with a variety of different news actors and sources, but only a fraction of this interaction is visible in the published report in the form of speech representation. The large majority of the accessible communication takes place from text producers to the mass audience. Unlike in the above example of the TV talk show, this mass audience cannot

Author submitted manuscript. Consult published version for correct page numbers!

be differentiated into various circles depending on whether or not they are present at the moment of message production. Nevertheless, the audience is not uniformly addressed. Bell (1991, 90-95) points out that mass media always address a specific segment of the general population, but that they are often consumed by individuals who are not members of this target audience. The audience can, thus, be segmented into various circles, ranging from addressees in the target audience to eavesdroppers. As an example of eavesdropping, Bell mentions media products which are only allowed to be consumed by an adult audience, but which may still end up being read, viewed or listened to by children. For online news sites, the target audience can be defined in terms of its national affiliation, political orientation, social class, linguistic background (i.e. whether or not someone speaks the language in which the news site is published) and so on. Depending on these criteria, a particular reader might be part of the addressed target audience or rather an overhearer (cf. Bell 1991, 91).

Even though the communicative situation is very different from that of television talk shows, it is still possible for the news audience to respond. In print newspapers, this takes place in the form of letters to the editor which can subsequently be published and, thus, are integrated into the main communication. In this way, responses to letters from other readers become possible, so that there is even a certain degree of interaction among members of the audience (Luginbühl et al. 2002, 78). The options for the audience to respond and interact with each other have obviously increased dramatically on the Internet. Lüders (2008) discusses the fact that individuals have gained the possibility to address a mass audience through online tools such as blogs. She observes that mass communication is no longer restricted to institutionalised mass media and that, in addition, institutionalised mass media increasingly integrate tools for interpersonal communication in order to increase audience contact (Lüders 2008, 691-692). In her view, these developments lead to a blurring of boundaries between interpersonal communication and mass communication, which she describes as two endpoints on a scale (Lüders 2008, 684, 691-692). While I agree with Lüders's observations on the increasing presence of interpersonal communication on mass media platforms, such a classification of the resulting communicative situation as a blurring between the boundaries of the two communicative forms does not seem precise enough to capture the characteristics of this setting.

Janoschka (2004, 96-101), in contrast, develops a model of interactive mass communication on the Internet which is able to account for the various combinations of mass communication and interpersonal communication without giving up the analytical distinction between the two (see Figure 4.2). The grey areas stand for elements of mass communication and the dark arrows visualise the unidirectional flow of communication from (institutional) text producers to a mass audience. In contrast, the white areas represent elements of interpersonal communication and the small white arrows stand for the multidirectional flow of information between and among members of the audience and text producers. Applied to online news sites, this means that text producers act as senders who submit content in the form of news articles, one type of "online messages" appearing on online news sites.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

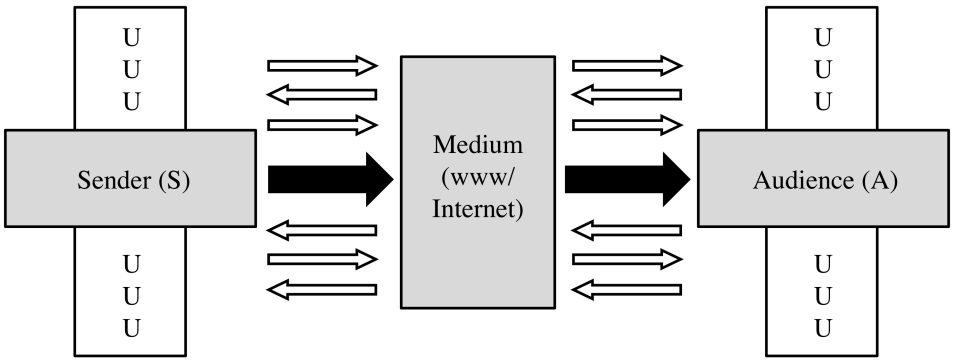


Figure 4.2: Janoschka's model of interactive mass communication on the Internet (Janoschka 2004, 98), slightly simplified

Author submitted m-

These articles are read by members of the audience, some of whom may then respond in the form of user comments, e-mails to journalists or through other interactive means. In cases in which these responses in turn are published they also become an online message of the news site. This can happen directly, for instance in the case of unmoderated user comments. Alternatively, they may first go from the audience to the text producers, who then upload them or integrate them into their own messages, for instance in the case of audience content (see section 4.2.5).

It is perhaps necessary to point out here that the term "interactive" is ambiguous in the context of online communication. On the one hand, it can refer to interaction between human participants in communication. In this sense, it stands for the characteristics of interaction, i.e. communication that is two-directional and dialogic rather than uni-directional and monologic. On the other hand, the term is often used to refer to technology that reacts to user input. In this sense, it stands for processes between a human and a non-human agent, which is known as interactivity (for a more detailed discussion of the distinction between interaction and interactivity, see, for instance, Habscheid 2005, 59-60).[7] In computer-mediated communication, the two uses can fall together. If two

[7] In addition, Wagner (2002, 29-31) distinguishes between two different levels of interactivity. In the narrower and more technical sense, users interact with a computer system when using software, for instance when selecting options and clicking buttons. In the wider sense, users interact with computer systems to access data, for instances when accessing information in a database or on the Internet.

Author submitted manuscript. Consult published version for correct page numbers!

human participants interact with each other in an online setting, the communication can be interactive in the first sense because of its two-directional and dialogic character, while it is also interactive in the second sense, due to the computer's role in responding to the users' input when mediating a message. Similarly, both senses are present in Janoschka's model of interactive mass communication (2004). In the present study, I am mainly concerned with the first sense; I focus on how online news sites promote interaction, rather than investigating the interactivity of the websites. Thus, the term "interactive" in this study is used to refer to two-directional communication. Consequently, my discussion of interactive elements on online news sites is restricted to elements that enable participants to interact with text producers and/or with each other, rather than elements that simply react to user input (e.g. hyperlinks, buttons).

4.1.2 Interaction on online news sites – a brief history

The online setting is doubtlessly an important facilitator of new forms of audience interaction in mass media. However, a look at early research on online news sites shows that these interactive forms have only developed over time. In his analysis of characteristic features of early online newspapers from around 1997, Rademann (1999) does not even mention the possibility for readers to respond. If such options were available at all, this suggests that they were not very prominent. This view is further supported by Schultz (1999, 2000), who concludes on the basis of research conducted in 1997 and 1998 that "[t]he majority of online newspapers do not even offer interactive options other than email to the newsroom" (Schultz 2000, 217).

Schmitt (1998) investigates several forms of audience interaction in her study of the online news site MSNBC Interactive in 1996. Readers were able to send e-mails, which according to Schmitt (1998, 321) resulted in several thousand e-mails each day. Only a few outstanding instances of these e-mails were selected and published on a dedicated opinion page (Schmitt 1998, 322). This illustrates how much the early online news sites followed the practices of letters to the editor in print newspapers. Schmitt also discusses two other forms of audience interaction that were already used on MSNBC Interactive at that time, opinion polls and expert chats. These two features explore the possibilities of the online setting but the detailed description by Schmitt (1998, 322-324) of how these features were used also shows that their implementation was quite different from today's polls and chats. The results of the polls were not visible immediately, but were only published in a summarised form at a later point. Similarly, in the expert chats the readers usually did not chat directly with the invited guests, but instead a chat coordinator read out the questions to an expert on the phone and then typed their answers into the chat client.

There are several reasons why early online news sites were slow to adopt features facilitating interaction with and among the audience. One of them can be found in technical restrictions. The fact that online chats required the installation of software, which was not entirely trivial, might have been an important motivation for conducting expert chats on

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

the phone (Schmitt 1998, 322). Another factor can certainly be found in the strong reliance on older forms of news publications, for instance when the well-established methods for selecting letters to the editor were applied to user feedback in the online setting. Accordingly, Boczkowski (2004, 173) observes that newspapers were slower and more conservative in adopting new technological features than organisations which did not originate in traditional mass media forms. The impact of older forms becomes particularly apparent when looking at the publication frequency of early online news sites. According to Rademann (1999, 258-259), continuous publication and updating of news was hardly done around 1997. He concludes: "There are no obvious reasons behind this artificial restraint, except for the traditional notions of printed newspapers which still tend to very much influence our understanding and treatment of their electronic counterparts" (Rademann 1999, 259).

However, even when online news sites were already well-established and the early technical obstacles had been overcome, interactional features were not always fully embraced. Comparing online news sites from five countries in 2005, Quandt (2008, 735) comes to the conclusion that the sites still did not make use of the Internet's potential for interacting with users. In a study based on 22 interviews with managers and editors of US online news sites conducted in 2003/2004, Chung (2007) identified the reasons why some online news sites were reluctant to invite their audience to interact. The most important reason appears to be the increased workload for text producers, who have to deal with a large number of e-mails (caused through sincere attempts of interaction, but also through spamming) and who also need to moderate discussion boards in order to avoid the publication of offensive messages (Chung 2007, 56-57). A similar study based on interviews with editors from British online news sites in 2004/2005 was carried out by Thurman (2008). He, too, found a wide-spread concern for the workload – and, as a direct consequence, the cost – of moderating large numbers of contributions by readers (Thurman 2008, 147-148). An additional worry was the perceived lack of quality of audience contributions. They were said to fall short of the required standards for two reasons: on the one hand, with respect to their newsworthiness, and, on the other hand, with respect to grammar and spelling (Thurman 2008, 144). The demands for cost and workload can be illustrated with some data from *BBC News*. According to Allan and Thorsen (2011, 32), *BBC News* employs "twenty-three people to handle what on an average day typically amounts to 12,000 e-mails and about 200 photographs and videos".

Despite these reservations, readers have increasingly gained new opportunities for interaction and content contribution on online news sites (see, for instance, Hermida and Thurman 2008). In some cases, the interaction with the audience was promoted in a particularly active way and with substantial resources. This is, for instance, documented in Wardle and Williams's report on a large-scale research project carried out at the *BBC* in 2007 and 2008, which is, in their own words, "the most comprehensive research to be completed on the subject of User Generated Content (UGC) in news" (Wardle and Williams 2008, 1). The project, involving among other elements various newsroom observations, interviews with journalists, senior managers and executives, surveys and focus groups, was

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

only possible through the close collaboration and support of the *BBC*, which in itself shows how much importance the organisation gives to audience interaction and contribution. The report gives an overview of the range of options for the audience to contribute content to the *BBC*, describes the practices of *BBC* journalists dealing with these contributions, provides suggestions for further developments, and indicates problems that need to be overcome, e.g. the fact that most audience contributions come from a rather small part of the general readership (Wardle and Williams 2008, 32). In their conclusion, Wardle and Williams recommend that the future development at the *BBC* should go towards a closer collaboration between journalists and the audience. Rather than integrating pieces of audience material into their own reports, journalists should develop articles together with members of the audience, and rely on their skills and expertise in the form of an extended expert network consisting of their readers (Wardle and Williams 2008, 42-43). It remains to be seen to what extent such suggestions will be implemented in the future. Given that news media organisations, especially those in the private sector, are under a great deal of pressure to work at minimal costs, the resources for such time-intensive collaborative projects might in many cases not be available.

4.2 Forms of feedback and interaction

In this section I will present and discuss the most prominent features for direct and indirect audience feedback that can be found in my data. This presentation is selective concerning the variety of forms that are offered by the sites in question or online news sites in general. One major restriction comes from the fact that I collected data directly relating to specific articles. Interactive elements appearing in special areas of the website such as blog and community sections cannot be covered in this way. Therefore, I will only touch upon one or two features of such areas, without providing a systematic treatment of these. Moreover, while some forms of interaction are quite established throughout the observed sites, others occur only once in my data, which suggests that even more forms might be found if more data had been taken into account. There is also evidence that different forms of audience interaction are still being developed and tested. In particular, with respect to the integration of news sites with external social media software, new features are continuously being added. Rather than providing an exhaustive list of ways of interaction between audience and text producers, the following analysis aims, instead, to show general tendencies and trends that could be observed on online news sites at the time the data was collected.

4.2.1 Indirect feedback

Before analysing the various forms of direct feedback and interaction between the audience and the text producers on the online news sites in my data, I would like to briefly look at the indirect feedback readers give. Online news sites do not need to ask their audience what content they particularly like. The online setting provides a range of ways in which indirect

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

feedback can be gained from the behaviour of the audience (see also MacGregor 2007). For print newspapers, circulation figures provide an indirect form of audience response, and, indeed, this has been seen as an even more influential way for the audience to provide feedback than through letters to the editor (Bell 1991, 87-88). However, circulation figures are rather limited in the range of information they provide. They indicate how many copies of a particular issue are distributed, but they do not say how many people actually read the newspaper, which articles they read and how much time they spend on individual articles. Establishing such information for print newspapers is only possible through surveys, which are time and cost intensive.

The online equivalent to circulation figures is the number of (unique) visitors to an online news site. This number is established by media monitoring organisations such as ABC[8] and provides the basis for online advertising rates. Besides such official figures that allow comparisons across various news sites, individual news sites can, of course, also determine themselves how many users access their content. Moreover, the online setting allows them to collect information on the reading habits of the audience which goes far beyond the overall frequency of access to a news site. This includes what articles are accessed, the time that is spent on specific pages, what additional material is accessed (e.g. images and videos) and the navigation paths users follow. All this information is of great interest to news sites when it comes to making their content as attractive as possible.

This information is not only used internally, but, at least in part, is also fed back to the audience. All the online news sites in my data integrate lists of "most read" or "most popular" articles, suggesting to their readers that what is read by many might also be of interest to them. These lists are created automatically and they are constantly updated. Moreover, given that the arrangement of articles on online news sites can easily be adapted throughout the day, the popularity of articles can also influence their placement on overview sites. As a consequence, the way in which articles are presented on online news sites can reflect the aggregated choices of the audience. This can be seen as a first step towards a customised presentation of news.

Indeed, it is quite surprising how little use is made of the possibility to present users with a selection of articles that are specifically tailored to their own interests, based for instance on what articles they have read in the past. Such practices are well-known from online shopping sites and would not be difficult to apply to online news sites. Still, apart from news alerts, which notify readers about new articles in selected sections, customisation of the news presentation can hardly be found in the analysed online news sites from 2010.[9]

[8] See for instance http://www.abc.org.uk (accessed on 28/2/2010)

[9] The redesigned news site of the *Times* launched in the summer of 2010 contains a function that allows one to "follow" stories of interest. Whenever a new article on a related topic is published, the reader is notified. Such functions may well become more widely used on online news sites over the coming years. Other customisation functions have disappeared, however. At the beginning of 2009, *BBC News* allowed their audience to customise the start page of their website. Readers could not only configure which topic areas they wanted to appear on the front page, but they could also choose the colour in which the site was presented. This function has disappeared in the meantime. The selection of topic areas and their order is, however, sometimes possible for smart phone applications, for instance for the iPhone app of *BBC News*.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Adaption to reader preferences is, thus, mostly restricted to mirroring the preferences of the collective audience through the prominent placement of the most popular articles.

There is another way in which readers can access articles based on personal preferences, namely through social software. All news sites integrate tools to easily post articles on social media sites, such as *Digg*, *Reddit*, and *Del.icio.us*, and on social networking sites, such as *Facebook*. Notably, *Twitter* was not included in the list of social media sites to which news articles could easily be forwarded when I collected my data in January 2010, but in the meantime it has been added to most of the news sites. Through such social media services readers can follow the suggestions of users who share similar interests and thus access lists of news articles that are (relatively) customised to their needs. For the news sites, the forwarding of articles to social media sites provides valuable information, since it complements the access rates with indications of whether or not the article was considered worth recommending.

4.2.2 User comments

User comments are the most prominent form of direct feedback, and they are central to the personalisation of the audience. Here, readers appear as actual individuals, represented by name and often also their location. While in no way being representative, user comments still provide the most direct access to the opinion of (one part of) the readership. From the point of view of readers, user comments provide an opportunity to become involved in discussions about what they read. Readers not only address editors, journalists and the general audience, but they also interact with each other. As a consequence, communities can establish themselves in the discussion sections of online newspapers, similar to online communities on message boards and online forums. This is illustrated with the following quote by reader and commentator Jay Reilly, who contributed to a panel debate on moderation practices on the *Guardian*.

> I was banned in May 2009, and the experience was surprisingly distressing. A moniker [nickname] might just be an irritation to a moderator, but it is someone's link to a community they may have spent years debating with.[10]

Reilly here clearly identifies with a community of active readers who comment and debate on the online news site regularly. The identification with a community also means a close association with the news site in question. If a community with such a high degree of involvement develops among readers, this constitutes one of the most powerful ways in which they develop a personal connection with an online news site.

However, it is not clear how common such a strong identification with a community of readers is, or whether one can even speak of an actual community for all news sites. In her study of impoliteness in the comment sections of five British online newspapers, Neurauter-Kessels (2013) found large differences in the degree of interaction among commentators on

[10] http://www.guardian.co.uk/commentisfree/2010/oct/25/panel-debate-web-moderation, accessed on 15/11/2011. I thank Manuela Neurauter-Kessels for pointing this quote out to me.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

the different sites. For instance, when analysing the number of commentators that had contributed to the sample of 600 comments taken from each site, she found that this number was much smaller for the *Express Online* (67) than for the *Mail Online* (491). The other three sites – including the *Sun* and the *Guardian* – were found in between, but clearly closer to the *Mail Online* than to the *Express Online* (Neurauter-Kessels 2013, 228). In addition, personal attacks were far more likely to receive a response on the *Express Online* than on the other sites (2013, 226). Neurauter-Kessels (2013, 305) concludes that among the news sites she analysed, the *Express Online* "stood out with more interactive patterns", which she attributed to the site's small community. Moreover, interactional patterns do not only vary across sites, but also across articles and topics. Langlotz and Locher (2012) analysed the participation framework of a total of 120 comments posted on six different articles on the *Mail Online*. The number of comments with explicit reference to a previous post varied between 0 and 8 out of 20 comments (Langlotz and Locher 2012, 1599). These findings suggest that the interaction patterns vary across news sites and articles, and that the potential for interactional exchanges among readers is not fully realised in all cases.

Nevertheless, user comments clearly are the most widespread way in which interaction with and among readers takes place. They are found on all the news sites in my data, yet some of their characteristics vary across the sites (see Table 4.1). Most news sites offer the possibility to comment directly on individual articles.[11] The most recent or most recommended comments are usually displayed directly below the article and the remaining comments can be accessed through expanding the comment section, browsing through the list, or navigating to a separate page containing all comments. All news sites allow one to sort comments according to different criteria (e.g. chronology, most recent first, most recommended first, worst rated first). Alternatively, user comments can often be submitted in special discussion or debate sections, which may or may not be linked from specific articles. In addition, some news sites ask their audience to send in comments via e-mail.

Of the news sites in my data only *BBC News* does not support comments directly on the article page.[12] Comments can instead be made in a separate "Have your Say" section, where discussion topics are launched by the *BBC News* staff. Frequently, these topics correspond to the topics of featured news articles to which they are then linked. Some of these articles also contain links to the respective debate in the "Have your say" section and, in some cases, selected comments are published on the article page. In addition, *BBC News* frequently asks its readers to submit comments via e-mail, some of which are again published online, either directly with the article or on an additional page to which there is a link on the article page.

[11] This has not always been the case. According to Thurman (2008, 141), the *Times Online* was the only one of ten British online news sites (including all five sites from my study) which integrated this form of audience interaction in April 2005. Already by November 2006, however, five additional sites from this set had added the same function, including the *Guardian* and the *Mail Online* (Hermida and Thurman 2008, 346). By the time the data for this study were collected in January 2010, the *Sun* also offered its readers the possibility to submit comments on individual articles.

[12] This had changed by 2011 and comments can now also be submitted directly below selected articles.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Table 4.1: Characteristics of the commenting function across news sites

	Times Online	*Guardian*	*BBC News*	*Mail Online*	*Sun*
Commenting on articles					
directly	67 / 70	9 / 28	n.a.	64 / 70	22 / 28
in debate section	n.a.	not linked	6 / 28	generic link	not linked
Avg. no. of comments after 1 month	71	191	1,688	142	53
Maximum size (in characters)	unspecified	5,000	500	1,000	2,000
Moderation	post publ.	post publ.	varies	varies	post publ.
Login	compulsory	compulsory	optional	n.a.	compulsory
Displayed user info	name	user name	name, location	name, location	user name
Link to user profile	n.a.	yes	yes	no	yes

None of the news sites allows comments for all articles. On the *Guardian*, the majority of articles (19 out of 28) do not support comments, and also on *BBC News* only selected articles (6 out of 28) are directly linked to a discussion section. For the *Times Online*, the *Mail Online* and the *Sun*, the default is to allow comments. The sensitivity of the article topic appears to be a decisive factor in activating or disabling the commenting function. For instance, of the 14 articles in my data which are concerned with the heavy snow fall in Britain, a rather unproblematic topic, only 4 do not allow comments. These 4 articles are published in the *Guardian* and on *BBC News*, both of which allow comments only for selected articles, and both news sites contain articles on the same topic for which the commenting function is enabled. In contrast, of the 9 articles on the Edlington court case, in which two boys were sentenced for torturing two other boys, a majority of 5 do not allow comments, 3 of which were published on the *Mail Online* and the *Sun*, which by default enable the commenting function. In the case of Fiona Donnison, a mother who admitted to having killed her children, 4 articles in my collection were published, one each on the *Times Online*, *BBC News*, the *Mail Online* and the *Sun*, none of which allows comments. Thus, while most of the articles of a rather uncontroversial topic (heavy snow fall) have the commenting function enabled, it is disabled for most (Edlington case) or all (Donnison) articles on two topics that involve violence against children.

The news sites differ with respect to how they moderate comments. Except for the *Times Online*, all sites allow readers to report offensive comments to a moderator by clicking on a special icon for this purpose. On the *Times Online*, the *Guardian* and the *Sun*, comments are published immediately and only those that violate the terms and

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Table 4.2: Rejection rates of comments on 13 discussions rated "most popular now" on *BBC News*

Jan.	Topic	Total	Processed	Rejected abs.	Rejected in %
4	Can the UK and US prevent extremism in Yemen?	2,170	1,716	489	28%
8	Should 50p be minimum price for a unit of alcohol?	684	270	1	0%
10	How are you coping with the adverse weather?	3,465	3,441	371	11%
12	Should Islam4UK be banned?	1,334	416	69	17%
14	What is your favourite Simpsons moment?	824	573	55	10%
16	Are ethnic minorities still disadvantaged?	2,828	1,577	214	14%
18	Should butter be banned?	329	151	4	3%
20	How should families be supported?	507	373	32	9%
22	Should there be a change of heart on 'have-a-go' heroes?	489	301	21	7%
24	Will UK companies face a 'bumpy recovery'?	75	68	0	0%
26	Do you agree with France's Muslim face veil ban?	1,271	280	22	8%
28	Is Tesco right to stop people shopping in their pyjamas?	769	255	21	8%
30	Your views on Blair's inquiry appearance	3,877	3,404	458	13%
Total		**18,622**	**12,825**	**1,757**	**14%**

Author submitted manuscript

conditions, for instance by being offensive, are deleted by moderators. On the *Mail Online*, the mode of moderation varies according to topic; sometimes comments are moderated prior to publication, and sometimes only offensive comments are deleted after publication. *BBC News* also distinguishes between two moderation levels. For "fully moderated" discussions, all comments are screened by a moderator prior to publication. For "reactively moderated" discussions, a further distinction is made between registered members, whose comments are directly published, and non-registered users, whose comments are screened by a moderator prior to publication.

BBC News publishes information about the number of comments that are rejected for each discussion. Table 4.2 presents the rejection rates for the most popular discussion on each day on which I collected data.[13] The rejected comments are given in absolute frequency, as well as in percent of all processed comments, where the number of processed

[13] The most popular discussion on 6 January was also the most popular discussion on 10 January. The information for this discussion is only listed once and, therefore, there is no discussion listed for 6 January.

Author submitted manuscript. Consult published version for correct page numbers!

comments consists of all comments which are no longer in the moderation queue. As these frequencies show, a considerable number of comments are regularly rejected by moderators on *BBC News*. 14 percent of all processed comments on these thirteen topics were rejected, with rejection rates ranging from 0 percent to 28 percent for individual topics (the average rejection rate across the discussions is 9.8 percent). It is also noteworthy that only one of these thirteen topics was reactively moderated, namely the discussion on 14 January, "What is your favourite Simpsons moment?". All the other topics were fully moderated. This indicates that the moderators on *BBC News* considered even topics such as "Should butter be banned?" or "Is Tesco right to stop people shopping in their pyjamas?" likely to attract offensive comments.

Two final ways in which the comment function differs between the news sites are size and user information. The maximum size for user comments varies between 500 characters on *BBC News* and 5,000 characters on the *Guardian*. The *Times Online* is the only news site which did not specify any restrictions of comment size at the time of data collection. Comments of up to 2,000 characters can be found, but comments with more than 1,000 characters are quite rare. On the new website that was launched in June 2010, the length of user comments is restricted to 3,000 characters. The *Times Online*, the *Guardian* and the *Sun* require their users to log into their user accounts before they are able to submit comments. On *BBC News*, logging in is not compulsory, but comments by registered users are published without pre-moderation for "reactively moderated" discussions. On the *Mail Online*, the commenting function is not linked to the user account, but commentators are asked to provide their name and location, which are displayed below the comment. For the sites that require commentators to log into their accounts, the user name is displayed with the comment. On the *Guardian* and the *Sun*, the user name is linked to the respective profile page, which also includes an overview of previous comments made by the user.

Their high degree of visibility and frequency of use make user comments by far the most important feature for representing readers' opinions on online news sites. Through user comments, readers can voice their opinion in front of an audience that is at least potentially as large as the audience of the article to which they respond. As mentioned before, this is not in itself entirely new. Letters to the editor have always provided readers with the option to comment on articles and news events. However, there are important differences between letters to the editor and user comments with respect to the selection process and editorial interventions. Letters to the editor undergo a selection process which for most newspapers results in only a small proportion of all letters being published. In addition, the letters that are published are edited and often shortened before publication (see for instance Stewart 2005, 626-629 for some insights into the editorial practices for letters to the editor in the *Times* in the 1980s). User comments, in contrast, are not edited and the selection process is inverse to letters to the editor; It is not the best submissions which are selected for publication, but rather unsuitable comments which are selected for deletion. As long as user comments do not violate the editorial guidelines, for instance by containing harassment, illegal content or spam, they are published exactly as they were

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

formulated, which can explain why they frequently contain aggressive and impolite content (see Neurauter-Kessels 2011). In short, user comments open up an entirely new dimension of audience feedback. Ordinary readers can make their opinion public in a forum that is promoted by large, commercial news media organisations and that is visited not only by a restricted online community but by a more general mass audience.

4.2.3 Opinion polls

Opinion polls are a second way in which users can make their opinion known to text producers and other readers. In contrast to user comments, the feedback is restricted to selecting one of several predefined answers, usually expressing agreement or disagreement with a statement. With respect to personalisation, an even more important difference to user comments is that a poll only shows an aggregated result of all responses (percentage of users agreeing / disagreeing with the statement). It is not clear who selected which answer. Still, opinion polls contribute to personalisation, not least because they ask readers to interact with the news site, in most cases including forms of direct address of the audience. Moreover they provide a way of integrating the audience in the news report, presenting the "voice of the audience" as one component of the news article.

Opinion polls are much less frequent than user comments. In my data, only four articles from two online sites integrate opinion polls.[14] Two of these articles were published on the *Mail Online*, one with an article reporting on heavy snow fall in Great Britain, the other accompanying an article on the recession there. The polls are displayed on the main article page, inset in the text like a picture. After selecting one of the answers, the results of the poll are displayed. Figure 4.3 illustrates the placement of opinion polls on the *Mail Online*.

Two articles on the *Times Online* provide a link to opinion polls in a box to the left of the main text of the article with the heading "Multimedia". The polls themselves are located on a separate page, which is part of the "Money Central" weblog. Thus, the polls and their results can only be seen when the user navigates away from the main article. In both cases there is more than one poll on the same topic. In the case of the two polls relating to a financial reform speech by president Obama (examples 4.6 and 4.7), the first poll asks for the reader's position with respect to the suggested reform, while the second asks for the readers' opinion about future actions. The second article on the *Times Online* which has related polls is about the recession in Britain. Three questions are asked in the polls (examples 4.8 to 4.10); the first asking for a general assessment of the situation, with a focus on the location of the reader, the second asking for an assessment of the personal situation of the reader, and the third asking for a prediction of the future development of the personal situation of the reader.

[14] This only includes articles which either display the poll on the page of the article or where the article site provides a link to a thematically related opinion poll. It does not include polls that are located on separate discussion or debate sections of the site (unless there is a direct link from the news article).

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

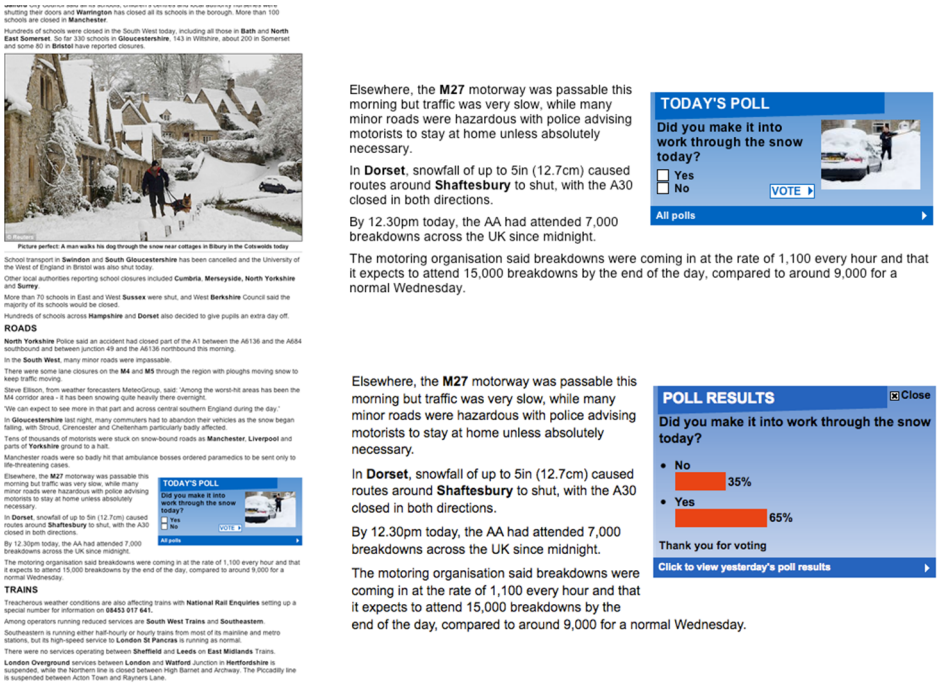


Figure 4.3: Poll in article published on the *Mail Online*, on the left in context of the article, on the right in blow-ups before and after voting (mail-uk-100106)

(4.6) Do you support Barack Obama's planned crackdown on banks? (times-wn-100122)

(4.7) Should the British Government copy Barack Obama's proposed crackdown on the banks? (times-wn-100122)

(4.8) Do you think that the recession is over where you live? (times-uk-100126)

(4.9) Has your personal financial position worsened or improved since the recession began? (times-uk-100126)

(4.10) Do you expect your personal financial position to improve or worsen over the next 12 months? (times-uk-100126)

Examples 4.8 to 4.10 illustrate very well how polls can be used to personalise news events by emphasising how individuals are personally affected by a specific event. Asking readers to assess the situation focussing on their "personal financial position" and the location where they live presents the recession from the point of view of ordinary citizens and their experiences, rather than as an abstract topic of national politics. It increases the readers' personal involvement, not only by activating them to participate in a poll, but also by letting them frame the news event in their own experience. The role of opinion polls for the representation of personal experience is also expressed in the short text that introduces the polls (example 4.11).

(4.11) Official statistics published today show that the UK has finally emerged from recession. But while the economy may be marginally bigger now than it was three months ago, the many people and businesses still struggling to cope

Author submitted ma

Author submitted manuscript. Consult published version for correct page numbers!

> might find the recovery hard to believe. And with tax rises, higher interest rates and growing unemployment still to come, many people's finances will get worse before they get better.
> We want to know what you think and to hear your stories about how the recession has affected you and your area. Vote in our e-Poll below and leave your stories in the field below that. (times-uk-100126)

This text starts by mentioning the official statistics and points out that such statistics do not always correspond with how individuals experience a situation personally. The poll is seen as a way to capture and represent this personal experience, as well as to provide a starting point for submitting even more extended personal accounts. Polls, thus, let readers participate and, at the same time, increase the news value of a story by emphasising its direct effects on the lives of the readers and introducing a personal angle.

4.2.4 Contact details and profiles

While comment sections and user polls provide platforms of interaction between readers and text producers and/or among readers, the publication of contact details enables interaction through different forms of communication. Contact details can mostly be found for the journalists who are identified as the authors of an article, but in some cases they are also available for readers who submit user comments. E-mail addresses are most common, but links to Twitter feeds and blogs are also sometimes given, which provide further ways in which interaction can take place.

The *Guardian* provides the most detailed information about journalists. Here the names in the bylines of the articles are formatted as hyperlinks which lead to the profile of the journalist. These profiles usually contain a small portrait photograph (sometimes called a "mug shot"), a list of links to previously published articles, and some further information about the journalist. This information can be quite brief and it mostly concerns his or her position within the newspaper (see example 4.12), but sometimes it includes previously held positions, fields of expertise and awards. In some cases, this information also includes personal interests and links to the Twitter page of the journalist (example 4.13).

(4.12) Andrew Culf is a deputy news editor for the Guardian (guardian-uk-100110)

(4.13) James Sturcke has been a news reporter for The Guardian since 2005. He previously freelanced for several national newspapers and, before that, was a reporter at Newcastle's Evening Chronicle. He likes photography, biking and pootling around in sparsely populated places. He's now a photographer. His work can be seen at James Sturcke Fotografía and you can follow him on Twitter @jsturcke (guardian-wn-100114)

The other sites in my collection do not have journalist profiles, but they use some similar elements. Bylines on the *Mail Online* also contain hyperlinks which lead to a site with articles previously published by the journalist. However, these sites do not contain any further information on the journalist. The *Sun* provides the e-mail address of journalists at the end of their articles. On the *Times Online*, additional information about text producers is only given in the case of columnists. There, it is sometimes given in the form

Author submitted manuscript

of e-mail addresses at the end of the text. For some regular columnists, a box is presented next to their column containing a mug shot, links to the main site of the columnist, and personal information. Example 4.14 shows the text provided in the information box on the columnist Libby Purves. For guest contributors, such information is sometimes given at the end of the column text, for instance in example 4.15.

(4.14) Libby Purves worked for some years for BBC Radio 4, as a reporter and a presenter on the Today programme and, since 1983, has presented Midweek. She joined The Times as a columnist in 1990. She received an OBE in 1999 for her services to journalism and was Columnist of the Year in the same year. In her spare time she writes bestselling novels. Her opinion column appears in the The Times on Mondays. (times-sc-100126)

(4.15) Malcolm Turnbull is a former leader of the Liberal Party and was chairman of the Australian Republican Movement from 1993-2000 (times-nc-100122)

BBC News is the only site that does not usually indicate the author of articles. Several of the articles, however, contain boxes labelled "analysis", which start with the name, affiliation, and sometimes a mug shot of the analyst.

Providing contact details is once again not an entirely new practice and it is also not a practice restricted to the online setting. Printed newspapers have always published contact addresses to which advertisements and letters to the editor can be directed. Providing contact details for individual journalists is a more recent practice, but still not necessarily restricted to online sites. In her analysis of historical newspapers from 1700 to 2000, Schneider (2002) identifies various stages of reference to the authors of news texts. Names are not given until around 1900 and they first occur only in popular newspapers (i.e. down- and mid-market papers) (2002, 141). Until the end of the period she investigates, popular newspapers give the names of text producers more frequently than up-market newspapers, which, according to Schneider, makes "news reports in popular papers appear much more personalised" (2002, 142). However, among the 36 news articles that I collected from the up-market *Times* from 1985, there is only one article which does not give the name of the journalist. Mug shots by authors are another feature that Schneider (2002, 143) finds in popular newspapers in the second half of the 20th century. These are almost completely absent in my data from the *Times* from 1985 and can only be found for columns on soft news topics, e.g. Penny Perrick's column on the "Monday page", which was intended to be of particular interest to the female readership (see Stewart 2005, 201). Thus, while background information and contact details for journalists were given even before the Internet age, the information that is found on online news sites is much more varied and, at least in some cases, also much more rich.

Such additional information on journalists relates to personalisation in two ways. First, as mentioned above, contact information enables the interaction between readers and text producers. Given that clicking on an e-mail address is sufficient to open the user's e-mail program, the interaction takes far less effort than sending a physical letter to a journalist by post. Second, providing background information on journalists emphasises that an actual person with his or her specific field of expertise, experience and interests is the author of

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

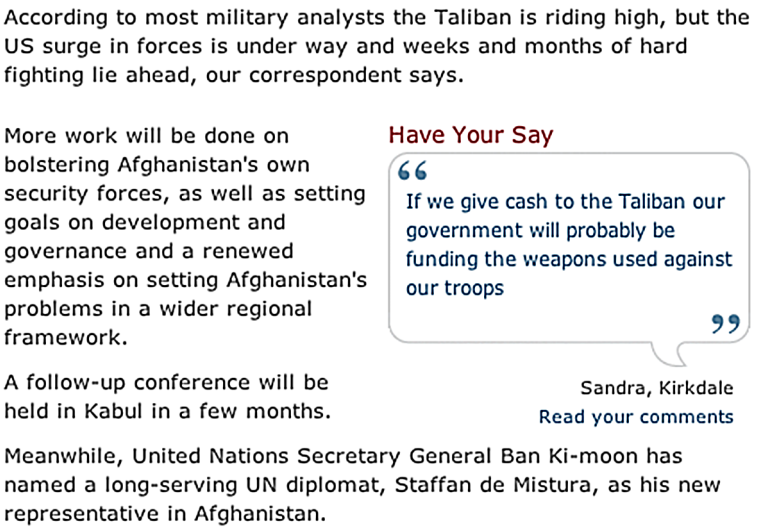


Figure 4.4: Example of a selected user comment integrated into an article as speech balloon (bbc-wn-100128)

the text. This lets text producers appear far less abstract in comparison to providing only their name or no name at all. Indeed, such background information and mug shots are the main ways of representing the text producer as a person, at least in news articles where overt self-reference to journalists usually cannot be found (but see chapter 8.5 for some notable exceptions).

4.2.5 Audience content

The term "audience content" is used by Wardle and Williams (2008, 14) to refer to all types of materials that are produced by the audience and that are integrated into online news articles, e.g. personal experiences of a news event, eye-witness accounts, pictures and videos. In contrast to user comments, which can also be considered a form of user-generated content, this material is published by journalists and is presented as part of the editorial content. The two categories are not always strictly distinct, however, since selected user comments can be integrated into articles. An example of this can be found on *BBC News*, where a speech balloon containing an opinion that was expressed in a user comment is used in an article as an illustration (see Figure 4.4).

Another way of transforming user comments into editorial content is the creation of pages consisting of selected user comments. Two days after a feature story which triggered a high number of comments, the *Times Online* published 14 of these comments on a separate page with the heading "Changing from gay to straight: what you thought". The comments were introduced with the following text:

> Patrick Muirhead came out for the second time, only this time in reverse; in his article for the Times he described his decision to turn straight after being gay for 20 years.

Author submitted manuscript. Consult published version for correct page numbers!

> His explanation met with an interesting response from Times readers. The reaction ranged from encouraging to hostile to, at times, philosophical. Here's what you had to say. (http://www.timesonline.co.uk/tol/life_and_style/men/article6994861.ece)

Here, the "interesting response" by the readership apparently provided grounds for the publication of a separate article. Such an article with selected user comments might serve as an additional entry point to the original article (which can be reached through clicking on the phrase *turn straight after being gay for 20 years*), thus increasing its readership. In addition, it foregrounds the function of online comments to provide a space for controversial discussion and it invites the audience to become involved (additional comments can also be submitted on this second page). It is remarkable that the comments that were selected for this new article include several comments that are not only highly provocative, but which also challenge the news site's terms and conditions of use by containing offensive and homophobic content:

(4.16) **Olde Jannner [sic] wrote:** I think (well I am hetero) you will enjoy yourself a whole lot more putting it where it was designed to go.

(4.17) **David Braham wrote:** Bisexual and homosexual people have no control and lead undisciplined lives

The selection of comments was criticised for being "one sided" by readers leaving comments on this page and commentators advised the audience to read the comments on the original article instead.

Besides the integration of user comments into articles, the contribution of pictures and videos is one of the most important forms of audience content. The fact that nowadays most mobile phones have a camera makes it is easy for eye-witnesses of potentially newsworthy events to capture these visually and send them to news media. The two events for which by far the most user-generated content was presented were the heavy snow fall in Great Britain and the earthquake in Haiti. In the first case, the material ranges from amateur videos of traffic problems (bbc-uk-100106) to pictures of snowy landscapes and the readers' own snowmen (e.g. mail-uk-100108, times-uk-100110, sun-100106b). In the case of the Haiti earthquake, the material includes amateur videos of the moment the earthquake started (e.g. mail-uk-100114) as well as pictures and videos of the destruction, victims and rescue attempts.

Audience material can also be found in the form of eyewitness accounts. A *BBC News* report (bbc-wn-100118) on a Taliban attack in Kabul, for instance, links to a separate page with four eyewitness accounts, one of which is reproduced in example 4.18.

(4.18) **Abdul, travel agent, Kabul**

> The travel agency where I work is about 1km (0.6 miles) from where the battle started. I was walking to work when I heard the huge explosion right in front of the shopping centre.
> I wasn't sure if it was a bomb or a rocket, then hundreds of people started running out.
> I gave a taxi driver $2 to drive me away as fast as he could. We are now in our travel agency, waiting for the police to clear the area.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

> We can hear a helicopter right above us. There are a lot of people still running away.
> I've never seen anything as close as this before. I am from Khost, north of Kabul, it's a mostly Taliban area!
> I have only been in Kabul two months but I will stay because of the good salary.
> Interviewed at 0910 GMT (1340 local time)
> (http://news.bbc.co.uk/2/hi/south_asia/8465085.stm)

The person who provides this information is not a spokesperson or an expert speaking in a professional function. He is identified only by first name and profession and is (quite literally) presented as "the man in the street" who is affected by the event on his way to work. The function of this eyewitness account is to provide personal experience rather than information. Much more detailed information about the incident can be found in the main article. Montgomery (2010) finds very similar functions in data from broadcast news interviews with victims and witnesses of the London bombings in 2005. The patterns he observes clearly indicate that also in his data, the eye-witness accounts by these private individuals did not mainly serve the purpose of providing information, but of "[telling] us what it felt like as it happened" (Montgomery 2010, 205).

Eyewitness accounts by private individuals (i.e. not regular news actors, professional experts or spokespersons, but people speaking in a non-professional capacity) have established a relatively strong presence in online news reports. This has two effects on the personalisation of news. On the one hand, the news event is presented from the point of view of ordinary people "like you and me" who are directly affected. This invites readers to identify with their perspective and personalises the news event. On the other hand, it also implies that the news site is interested in such personal accounts and, thus, the mere publication of audience content can already constitute an implicit request for readers to submit their stories and images.

Eyewitness accounts, images and videos are also explicitly searched for and requested by the online news sites. The article which linked to the eyewitness account in example 4.18, for instance, ends with a request to share relevant information, pictures and videos with the *BBC* (see example 4.19). At the end of an article on the heavy snow fall in Great Britain, the *Mail Online* even announces a special e-mail address to which "weather-related photographs" can be sent (see example 4.20). Similar requests can be found at the end of many articles published on *BBC News* and the *Sun*, sometimes asking for very specific information. The request in example 4.21 is asking people who know Sindy, a woman with a butane gas addiction who appeared on a TV show, to contact the *Sun*.

(4.19) **Are you in the area? Have you been affected? If you have any information you wish to share with the BBC you can do so using the form below:**
You can also send your pictures or videos to yourpics@bbc.co.uk, text them to +44 7725 100 100 or if you have a large file you can upload here.
(bbc-wn-100118)

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

(4.20) Have you taken any weather-related photographs you'd like to share with us? Email them to weatherpics@dailymail.co.uk and we'll publish the best (mail-uk-100108)

(4.21) **DO** you know Sindy? **CALL** The Sun newsdesk on 0207 782 4104, **TEXT** 63000 or **EMAIL** exclusive@the-sun.co.uk (sun-100122b)

In addition to specific requests at the end of articles, several of the online news sites contain more general requests for user-generated content. The *Sun* includes a line at the top of each article, asking readers who "got a story" to contact them and, in some versions, even offering to pay for information. *BBC News* has a special page in the "Have your say" section which explains how to submit content and suggestions for stories.[15] The *Guardian* also asks for user-generated content on the start page of its "Comment is free" section (see Figure 4.5). For the *Times Online*, no requests for audience content were found in my data. Nevertheless, the site contains picture galleries containing user images of the heavy snow fall in Great Britain.

Figure 4.5: Request for user-generated content on the *Guardian* (http://www.guardian.co.uk/commentisfree, accessed on 17/8/2010)

Explicit requests for audience content are particularly interesting since they invite the audience to enter into an interaction with the news site. The audience is directly addressed and it is suggested that individual readers are in possession of information that is of interest to the news site and its audience. The fact that readers can also see content that is marked as audience material substantiates this claim. To some extent, this can be seen as an inversion of the news value of relevance. The implication is not that the news event is relevant for the readers' own lives, but rather that their own experiences may be relevant for the news.

4.3 Strategies for integrating user-generated content

The previous section has given a brief overview of some of the most common ways in which the interaction between text producers and the audience is facilitated on online news sites. As mentioned before, the features that could be investigated are numerous, varied and their exact forms, functions and uses keep developing. Most of the news sites in my sample experienced some redesign in the year following the data collection, and features relating

[15] http://news.bbc.co.uk/2/hi/talking_point/your_news/7593687.stm (accessed on 17/8/2010).

Author submitted manuscript. Consult published version for correct page numbers!

to the interaction with the audience belonged to the aspects that underwent the most profound changes. Within the data I selected, the use of individual features is also not always consistent within each site and some features are used quite infrequently. Rather than comparing the sites with respect to individual features, I would, therefore, like to describe some more general strategies for how the news sites deal with the possibility to interact with their audience and to integrate content provided by them.

As the discussion above has shown, all the news sites integrate features that let their readers interact with text producers and with each other. User comments, for instance, are used on all the sites I analysed, and audience content is also present on all of them. Explicit requests for audience content were found on all the sites except for the *Times Online* and it occurred in various realisations. Since user comments and audience content, thus, seem to play an important role for the news sites in my sample, I will focus on the ways in which these two types of user-generated content are presented. In particular, I am interested in the extent to which user-generated content is separated from editorial content. I will suggest that three main strategies can be distinguished: 1) separation of user-generated content from editorial content; 2) integration of user-generated content into editorial content; and 3) blurring between user-generated content and editorial content. In order to illustrate these strategies, I will look at articles which cover the heavy snow fall in Great Britain in January 2010. This is a topic which was covered by all the news sites and to which readers are quite likely to contribute content, describing their own experiences and sending in pictures.

Before presenting these three strategies in more detail, I would like to point out some of the methodological limitations of my approach. Given that I do not have any inside information from the text producers of the data I analyse, I cannot know which parts of a news article and its surrounding material were actually contributed by the audience. Of course, this would deserve further investigation in the form of a separate study, making use of ethnographic methodologies to analyse newsroom practices. The importance of combining product-based approaches to news media (i.e. analysing newspapers) with process-based approaches (i.e. analysing practices in the production of newspapers) has repeatedly been pointed out in recent years (e.g. Catenaccio et al. 2011; Cotter 2010; Czarniawska 2011; Paterson and Domingo 2008; Van Hout and Jacobs 2008; Wardle and Williams 2008). More work is certainly needed in this area. However, such projects are very time intensive (not only for the researcher,. but also for the text producers who are investigated and whose co-operation needs to be secured) and for the present study such a perspective could therefore not be integrated.

Therefore, if I refer to audience content in the following, I do not so much mean that the content in question is known to have been submitted by a reader, but rather that content is presented in a way that suggests it has the status of audience content. By doing so, I adopt the perspective of the general audience, rather than the perspective of the text producers working for the news sites. The general audience in fact has as little access to background information about how the material was collected as I do. As far as

Author submitted manuscript.

Author submitted manuscript. Consult published version for correct page numbers!

personalisation effects are concerned, what counts is the appearance of the news article, not its production process. An article that appears to integrate content submitted by members of the audience implicitly (or sometimes even explicitly) claims to represent the experiences of actual readers. This simultaneously leads to a personalisation of the news event, and also gives a face, or at least a voice, to the mass audience. Whether or not the content was actually submitted by a reader is, from this point of view, of secondary importance.

4.3.1 Separation: The Times Online and the Guardian

One way to deal with user-generated content is to separate it from editorial content. Both the *Times Online* and the *Guardian* can be said to apply such a strategy. The online setting allows users to submit comments and send e-mails containing story ideas or images, but the main text of news articles does not appear to bear any traces of such audience content.

From the *Times Online*, only one article on the heavy snow fall was included in my corpus. The article covers weather forecasts, the lack of salt stocks for gritting, casualties caused by the weather conditions, consequences on premier league football games and a gas supply alert. The information is attributed to a range of official sources and organisations: the Met Office, the Highways Agency, the Police, the power company National Grid, the UK's biggest salt mine and charities. The nature of the information as well as the way in which it is presented do not suggest that any parts of it were contributed by readers of the news site. Readers are also not directly addressed, neither within the article text nor in boxes and other elements around the text that relate to the snow fall, except for the heading "Your comments" introducing the user comments at the bottom of the page.

Nevertheless, user-generated content is present. On the one hand, it appears in the comment section at the bottom of the page, where, at the time of collecting, 48 comments had been posted. On the other hand, the side bar on the right contains a link to a slide show, which contains 16 images that were submitted by readers. Both these features are not prominently advertised on the site. The fact that comments are available for the article is indicated by a small speech balloon and the number of comments in a light grey font at the beginning of the article. The announcement of the reader pictures is only the sixth item in the side bar and two other picture galleries with pictures taken by professional photographers are advertised before. It is also noteworthy that there is no invitation for readers to submit further images.

The *Times Online*, thus, does not seem to promote user-generated content. It provides a space for submitting comments and it uploads some reader images in a picture gallery, but these elements appear to be rather peripheral to the news article that covers the event. While a single article is, of course, too small a basis for making generalising claims, the other news articles in my sample confirm this impression. User-generated content plays a marginal role on the *Times Online* and, when it occurs, it is presented in clear separation from the editorial content.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

A similar situation can be found in the three *Guardian* articles that deal with the heavy snow. Like on the *Times Online*, the information in the articles is attributed to official sources, with the exception of a quote in one of the articles, which is attributed to an Internet discussion forum, albeit not one related to the *Guardian*. There are also no links from the article pages to audience material, in particular no reader images, which can be found on all the other news sites. One of the articles links to a gallery, but the pictures there all come from news agencies. Two of the three articles do not even allow users to submit comments, while the one article for which comments are enabled did not contain any at the time of data collection. One month later, 26 comments had been submitted, which makes this the least frequently commented *Guardian* article in my sample for which commenting was active. There are several possible explanations for this. On the one hand, it is possible that commentators on the *Guardian* prefer different topics for their debate. On the other hand, it is possible that comments were instead submitted to other articles on a related topic. The article which received most comments on the entire *Guardian* site between 8 and 14 January was a blog post with the title "Britain's cold snap does not prove climate science wrong". It was posted on 6 January and received more than 2,000 comments within eight days. Thus, this blog post relating the news event to scientific and political debates received much more attention from commentators than the more factually oriented news report.

The lack of user-generated content may appear more surprising for the *Guardian* than for the *Times Online*, given that the former seems to promote its community of readers much more actively. In contrast to the *Times Online*, users on the *Guardian* can create personal profiles with information on themselves and there are special sections on the website with information and resources for the "community". However, a closer look at the site suggests that audience material tends to be restricted to certain parts of the website. On the "comment is free" page, for instance, users are asked to suggest stories and a link to stories that were commissioned after such suggestions is also provided (see also Figure 4.5 above). In addition, this page provides a range of links to particularly popular and frequently commented texts on the site, inviting readers to participate in the discussions. It can also be observed that the most frequently commented texts in January 2010 were all either part of the "comment is free" section or blog posts. Thus, discussions tend to take place in these two sections of the *Guardian*, in which direct address of the audience and requests for content can also frequently be found.

Thus, while the *Guardian* tries to involve its readers more actively than the *Times Online*, this does not have a large impact on the sections of the site in which news articles are published. News articles are largely free from audience content and while comments are possible for about one third of all news articles (9 out of 28, see Table 4.1 above), they are not particularly promoted.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

4.3.2 Integration: The BBC News

BBC News is much more active in eliciting user-generated content than the *Times Online* and the *Guardian* and rather than separating audience material and comments from editorial content, the former are integrated into the latter. My sample contains three articles from *BBC News* that deal with the snow fall. All of them end with a request to submit pictures, videos, comments and experiences. Moreover, the requests are differently formulated in each case and relate to the topics that are covered in the article. On 8 January, the article deals with the effects of the weather on the national football league and horse racing and correspondingly, the request for audience material asks for effects on local sports clubs (see example 4.22).

(4.22) Are you affected by the "big freeze"? Send us a picture or a video showing how cold the temperature is where you live. Are you a sports fan? Has the cold weather affected your local club? [...] (bbc-uk-100108)

Six days later, the article discusses the risk of floods posed through the thawing of snow and, at the end of the text, readers are asked to send their *pictures of ice and thawing snow* (bbc-uk-100114).

The text of the main news articles is very similar to the texts on the *Times Online* and the *Guardian*, insofar as it seems to rely on information that is provided by official sources such as the Met Office, the Police, the Highways Agency, EDF Energy, spokespersons for schools, train companies and airports, etc. In one of the articles, there is a brief account of a TV presenter who was stuck in the snow with her car, but, given that she is not an ordinary reader of the site, this cannot be considered audience material. However, in contrast to the other two up-market news sites, *BBC News* places links to audience material more prominently and it addresses its readers in side bar links, image captions, and boxes that are embedded into the text. All three articles, for instance, contain an information box that is placed in the text as an inset. It is entitled *Weather and travel info* and it contains links to several websites, advertised with directives (see example 4.23). The right-hand sidebar contains links to features, many of which are announced in the form of questions (e.g. example 4.24). They are directed at the audience and suggest that these are topics about which readers may want to know more, or about which they may even have enquired at the site.

(4.23) Check if snow is forecast in your area at BBC Weather (bbc-uk-100106)

(4.24) Why is snow so bad for potholes? (bbc-uk-100114)

The implication of an exchange with readers is even more salient in the case of the links provided further down in the side bar, under the heading "Guides and Info". The items in this section include links entitled *Q&A: How to cope with the cold* and *Your tips on beating the freeze*. The first of these sites contains a series of questions, such as *What should I take on car journeys in case I get stranded?*, followed by practical information. This can be seen as a case of synthetic personalisation, where mass media simulate an interaction

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

with the audience which has never taken place. The second site, containing tips on beating the freeze, is structured into sections like *What to wear* and *How to keep warm at home*. In contrast to the Q&A site, the answers here are provided by readers. Each heading is followed by a number of recommendations, the author of which is indicated by name and location (see examples 4.25 and 4.26).

(4.25) If you have to go out and walk about in the icy conditions, put an old pair of woolly socks over your shoes. This should help you stop slipping about. Claire T, Enniskillen, Northern Ireland[16]

(4.26) Go to Australia until April. Les Sutcliffe, Catherine de Barnes, West Midlands[17]

These contributions were most probably selected from comments in the "Have your say" section of the site. There, a discussion on *How are you coping with the adverse weather*, accessed via a link from the articles, elicited more than 3,000 comments. All three articles also link to image galleries with 13-16 pictures submitted by readers. The pictures are different for each of the articles, so that the total number of reader pictures is considerably higher than for any of the other news sites.

Thus, *BBC News* gives much more presence to user-generated content than the *Times Online* and the *Guardian*. It places links to audience material and comments prominently in sidebars and insets and it also uses very active strategies for eliciting user-generated content. This is not restricted to a dedicated area of the website, but combined with the editorial articles on hard news topics. The prominent integration of user-generated content increases its personalising potential, not only by giving it more visibility, but also by signalling to the audience that their contributions are valued. Nevertheless, user-generated content is clearly marked through layout and explicit declarations. It is, therefore, used to enrich and complement editorial content.

4.3.3 Blurring: The Mail Online and the Sun

This is different on the *Mail Online* and the *Sun*, where the boundaries between editorial content and user-generated content can become blurred. There are two main aspects that lead to this blurring. On the one hand, images by readers are used to illustrate articles, alongside images from news agencies. On the other hand, the articles combine information from official sources with personal stories of snow incidents of a less serious nature. In both cases, it is usually not possible to determine by looking at the final article alone whether a particular piece of information or a picture was sent to the online news site by a reader, whether it was obtained through a news agency, or whether a reporter of the news site tracked down the story by different means. However, by reporting the stories of "ordinary" people in (what is presented as) their own words, the news site at least stages an interaction with potential readers. Moreover, demonstrating that the news site is interested

[16] http://news.bbc.co.uk/2/hi/uk_news/8447680.stm, last accessed on 14/1/2011
[17] http://news.bbc.co.uk/2/hi/uk_news/8447680.stm, last accessed on 14/1/2011

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

in publishing this kind of information also reinforces the requests for stories and pictures, which are particularly prominent on the *Sun*.

The *Mail Online* contains a story about a pregnant woman who was stuck in the snow for twelve hours, together with her daughter (example 4.27). The story is presented in direct speech, suggesting an interaction between her and the *Mail Online*. This segment occurs between more factual information about traffic problems and meteorological information. It is, thus, part of the main news article and there is no indication in the text that it should have a different status than the information provided by official sources.

(4.27) Carla Holt said she and her 13-month-old daughter Lily-May were stuck for 12 hours in the freezing conditions. She said she received no support from the police overnight and was only able to leave the road when it was partially cleared at 6.30am today.
The 23-year-old said: "We went through hell. I am eight months pregnant, I couldn't go to the toilet all night, I couldn't warm the bottle up for my baby daughter. It was very frightening.
"There were loads of cars parked up, just on the motorway. No-one knew what was going on – there was no-one to help.
"We didn't see any police, we've heard that the Army is out but we didn't see anyone – it's not very good really." (mail-uk-110106)

The *Sun* also includes such personal stories in its news reports. One of the articles published as the main news article of the day is even mostly devoted to such a story (sun-110104b). It describes the experiences of a group of people who were locked in a pub for three days, including only very little information about the general weather situation towards the end.

Perhaps even more interesting is the use of readers' images on the two sites. Whereas the *Times Online* and *BBC News* restrict readers' images to dedicated image galleries, the *Mail Online* and the *Sun* also include them in their main articles. On the *Mail Online*, readers' images seem to be used rather as an exception and most pictures that are used in the article are marked as coming from press and photo agencies such as Reuters, Getty Images and PA. Two articles, however, mail-uk-110108 and mail-uk-1101, contain pictures of an igloo built by students and of snow sculptures which give the impression of readers' images. They are not attributed to any well-known photo agency, but carry names which seem to be personal names, and both the subject of the photos as well as their realisation do not point to professional photography. These pictures are otherwise presented in exactly the same way as the pictures by agencies and they are used to illustrate the main article.

On the *Sun*, pictures are less regularly marked for copyright and the boundaries between professional photography and reader images becomes even more blurred. The pictures illustrating the snow articles include satellite images of Britain from NASA, images that are marked as coming from press agencies, unmarked images that seem to be professionally produced (e.g. an image of Manchester airport in article sun-110106b), and images that are unmarked and might have been sent in by readers (e.g. the image of an overturned car in the same article). In one case, the image of a person sitting on a snow dragon contains a caption saying that it was sent in by "a Sun reader" (sun-110106b). Again, all these

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

pictures serve to illustrate the main text of the article and they are all formatted in the same way.

The overall impression this creates is that of user-generated content as one type of raw material for news articles, alongside more traditional sources. This is particularly striking in the case of the *Sun*, which prominently advertises paying readers for interesting content. The resulting personalising effects are quite high for the news events, which are frequently framed in terms of the personal experiences of ordinary individuals. What is less pronounced than on *BBC News* is the presence of a community of readers that builds a complementing counterpart to the editorial team and a possible partner for interaction.

4.4 Summary and conclusion

In this chapter, I discussed the role of feedback and interaction for the personalisation of online news. After introducing some theoretical background and previous research, I gave a brief overview of the most prominent features in my data. These can be grouped into indirect feedback (e.g. number of unique visitors, sharing of articles on social networks), and direct feedback and interaction (e.g. user comments, opinion polls, submitting audience material). I limited my discussion to only a selected set of such features instead of attempting to give a comprehensive account, but, nevertheless, it became clear that the forms of feedback and interaction are diverse on these sites, with new features continuously being added. All these features are specific to the online setting of the news sites, even though precursors of such interactive features can be found in older mediums.

A comparison of the five news sites in my corpus showed many parallels between them. First and foremost, all news sites make use of the options for interacting with their audience online. User comments seem to be the most important feature on all the sites, but all of them also allow (or even encourage) their readers to contact them directly and to send in content. All news sites also list which articles are most frequently read, indicating that indirect feedback is collected on all of them.

Apart from these parallels, there are also some differences between the sites. On the one hand, not all the sites promote interaction with their audience to the same extent. *BBC News* is particularly active in inviting its audience to contribute content, and requests for stories are also placed very prominently on the *Sun*. The *Mail Online* is somewhat less active in this respect. While the *Guardian* promotes a visible community of readers, their forums for interaction are largely separated from the news content. The *Times Online*, finally, appears to be the news site which is most reluctant to invite its readers to make contributions.

Independent from how actively user-generated content is requested, the news sites also apply different strategies when it comes to presenting this content. In this respect, the market-orientation of the sites seems to play an important role. On the down-market *Sun*, and, to a somewhat lesser degree, on the mid-market *Mail Online*, user-generated content is not always separated from editorial content. Without clear declarations of

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

user-generated content, the boundaries between the two types become blurred and it is often unclear whether, for instance, a particular image was received from a reader or obtained through an agency. This is very different on the three up-market sites, which all contain user-generated content that is clearly marked through layout, placement and explicit declarations. *BBC News* is noteworthy for the high degree of integration of user-generated content and editorial content, while still maintaining a clear boundary between the two.

Effects of personalisation were identified for the audience, text producers and news events. All interactive features invite readers to become actively involved on the news site and, thus, lead to the personalisation of the audience. The involvement of the audience is further increased when readers personally relate to a news site through the identification with its online community. Moreover, the fact that readers become visible with their names, opinions and experiences contributes in a different way to the personalisation of the audience. Similarly, text producers can also be personalised through profiles providing additional background information. These profiles are connected to interactive elements, insofar as they often contain contact details, allowing readers to interact directly with text producers. News events, finally, are personalised through the fact that the integration of audience content leads to an increase in personal viewpoints and experiences which help to frame news events in personal rather than abstract terms. Thus, while interactive features are an aspect of the communicative setting of online news sites, they also have an impact on the content of articles and on their language. Similarly, their personalising potential is not restricted to one particular group of actors. Instead, interactive features have manifold and often closely interrelated effects on the personalisation of online news.

Author submitted manuscript.

Author submitted manuscript. Consult published version for correct page numbers!

5. Visual elements

When seven-year old Kiki was saved after being captured under the rubble of the Haiti earthquake for eight days, the picture of his rescue (Figure 5.1) was called "the most iconic image of the Haiti earthquake" by the *Mail Online* (mail-wn-100122). It shows the joy

Figure 5.1: Photograph showing the moment Kiki was rescued (mail-wn-100122)

of the boy and his rescuers and, thus, provides a positive emotional counterpoint to the prevailing reactions of shock and despair in the aftermath of the disaster. The image also gives a face to the people who are directly affected by the earthquake, namely its victims and the rescue teams, and it lets the viewer adopt the perspective of a member in the crowd witnessing the rescue efforts and their outcomes.

Similar images of successful rescue operations had a strong presence in the news coverage of the earthquake, as had images of death and injuries, shock and despair. Photographs shot on location are an absolute necessity for present-day news coverage of most types of news events. Increasingly, still images are also complemented with video reports, often providing an even more vivid account. Consequently, visual elements play a central role in news articles. This alone would make them an object of interest for the study of personalisation. However, they are made even more relevant through the fact that news images tend

Author Manuscript

to focus on people and that images in general are closely associated with the evocation of emotions.

In this chapter I am going to investigate the role of visual elements for the personalisation of news articles. While the focus will lie on still images which are either embedded within articles or presented in image galleries, I will also include the discussion of some selected videos. I will start with a quantitative overview of all visual elements in my news data, providing support for the claim that visual elements are presented for the vast majority of the most prominent news articles, and pointing out some first differences between the news sites. In section 5.2, I will discuss in detail some of the main characteristics and functions of images that are relevant for the personalisation of news. The chapter will end with a case study on the use of images in news reports on a single topic, the Edlington case.

Author submitted manuscript

5.1 Overview of visual elements in news articles

Table 5.1 presents an overview of the number of visual elements contained in the news articles of my data.[1] The numbers are based on all 28 UK news and world news articles for each of the five online sites, and on the 36 news articles from the printed *Times* from 1985. All images and videos that are directly visible on the article page are counted as embedded, whereas visual elements that are only accessible via a hyperlink are listed as linked. The numbers only include images and videos which are thematically related to the article in question and which serve as illustrations for the text. Furthermore, I have listed only the number of image galleries, not the number of images included in these galleries. This is due to the fact that for 16 galleries on the *Times Online* the entire galleries had not been saved locally, so that the exact number of images cannot be established.[2] For those galleries for which the images are available, the numbers range from four images up to 33, with an average of 14 images per gallery.

Concerning linked elements, the comparatively high number of image galleries on the *Times Online* is almost entirely due to only four articles which contain a total of 17 links to image galleries (called "slide shows" on the *Times Online*). These four articles all report on the Haiti earthquake and some of the galleries appear with several articles, so that the total number of galleries to which these articles link is actually lower than 17. The fact that the remaining 24 articles from the *Times Online* overall contain only links to six galleries (one of which is also about Haiti) shows that the earthquake has an extraordinary status with respect to image galleries on the *Times Online*. Of the other news sites, *BBC News* is the site that most frequently links to image galleries (not statistically significant),

[1] See appendix B for statistical tests.

[2] When collecting data from the *Times Online*, it seemed that links to thematically related image galleries were always presented in insets with the heading "Multimedia". Only when the articles were studied more closely did it turn out that for four of the articles on the earthquake in Haiti additional image galleries were provided in the right-hand sidebar. By the time this was discovered, the image galleries were no longer accessible. Another gallery on a news article about the Togolese football team was likewise missed.

Author submitted manuscript. Consult published version for correct page numbers!

Table 5.1: Visual elements in news articles from five online news sites and the printed *Times* from 1985

		embedded			linked	
news site	**articles**	**images**	**videos**	**per article**	**galleries**	**videos**
Times Online	28	17	12	1.0	23	6
Guardian	28	24	6	1.1	7	3
BBC News	28	51	28	2.8	11	71
Mail Online	28	221	7	8.1	3	0
Sun	28	104	13	4.2	2	15
Times 1985	36	15	–	0.4	–	–

Author submitted manuscript

and it is also the site which has by far the most links to videos (significant at $p < 0.001$, see appendix B). This might in part be a consequence of the fact that the affiliation with *BBC* television stations means that video material is easily available.[3] Again, the number of links to videos varies greatly depending on the topic. Three articles on the earthquake in Haiti contain a total of 15 links, as do three articles on the heavy snowfall in the UK. In contrast, eleven of the 28 articles do not contain any links to videos at all. The few links to videos and image galleries on the other news sites show similar patterns, so that it can be said that linked visual elements generally tend to occur with specific topics, while being largely absent from most articles.

Embedded visual elements tend to be more equally distributed over the various articles of a news site. The *Times Online* usually has one visual element embedded at the top of the article, between the headline and the main text. This is either an image or a video, except for one article without any illustrations and two articles with two images each, presented as a slide show to click through. The *Guardian* has a very similar layout, which usually includes one image or video between the byline and the main text. Two of the articles on the *Guardian* do not contain any illustrations, while three articles present two portrait images, side by side, in the place where usually a single image can be found. Furthermore, a map is embedded in the text of one article in addition to the illustration at the top. While these two up-market sites, thus, tend to contain one embedded visual element per article, the other three online news sites include considerably more illustrations on average. On *BBC News* the articles have between one and nine embedded visuals, although only three articles contain more than four. The layout is generally more varied, with images and videos not only occurring at the beginning of the article, but also as insets within the text. The difference of embedded visual elements between *BBC News* and the two other up-market news sites is significant at $p < 0.01$.

The *Mail Online* has most embedded images by far, with individual articles including up to 21 embedded visuals (significant at $p < 0.001$, except for the *Sun*). While videos

[3] The *Times Online* and the *Sun* similarly tend to integrate videos produced by *Sky News*.

Author submitted manuscript. Consult published version for correct page numbers!

are usually placed at the end of the text, images appear throughout. Often sequences of two to four images interrupt the flow of the text and two of the articles even consist mainly of sequences of images (twelve each) with only a very brief introductory text. On the *Mail Online*, the boundaries between heavily illustrated articles and image galleries are fluent; these two articles are presented on the news overview page in the same way as other articles, but the fact that they consist almost exclusively of images and captions is a more typical feature of image galleries. Indeed, the three image galleries I identified on the *Mail Online* were pages that were formatted exactly like these articles, with a brief introductory text followed by a sequence of images and captions. There were no pop-up galleries available, such as on the other sites, but the links to these collections of images shared characteristics with the links to image galleries on other news sites and the fact that the pages consist almost entirely of images was taken as justification to regard them as galleries. On the *Sun*, galleries appear in pop-up windows and can, thus, be differentiated more clearly from articles with many images. The number of embedded visuals on the *Sun* varies between one and ten. Images and videos are spread across the articles, but, unlike in the *Mail Online*, they do not always span the entire width of the article.

The printed *Times* from 1985 does not contain any videos or hyperlinks to online content.[4] In 1985, the only option for illustrating news reports on the *Times* were images in black and white, either from photos, information graphs or hand-drawn caricatures. Such images, mainly photos, are used in my data from the *Times* but less frequently than in the data from the online news sites (not significant for the *Times Online* and the *Guardian*). Of the 36 news articles included in the study, seven articles contain one illustration each and four articles contain two illustrations, while 25 (almost 70 percent) do not contain any illustrations at all.

5.2 Images and personalisation

The most apparent way in which images contribute to personalisation is through focussing the reader's attention on the individuals involved in news events. The majority of images in news reporting tend to show people, as my analysis in section 5.2.2 below will demonstrate. Mostly, these are either the named key actors involved in the news event or unnamed individuals who, in one way or another, are affected by it. By depicting them, images foreground these people and direct attention to their roles, their perspectives, and to evaluations of them. In many cases, images also provide visual indications of the identity of the actors that go beyond the textual descriptions. Therefore, images are instrumental in the personalisation of news events and actors.

[4] URLs pointing readers to online resources can be used similarly to hyperlinks in printed newspapers, though they were not yet used in 1985. The only form of hyperlinks (understood in a broad sense) that can be found in 1985 are references to other articles, either within the same issue or in earlier issues of the newspaper. Nowadays printed newspapers regularly contain URLs to additional online resources. On page 6 of the printed *Times* of 6 January 2010, readers are, for instance, directed to a webpage of the *Times Online* containing "Pictures, live updates and breaking news on the freezing weather".

Author submitted manuscript.

Author submitted manuscript. Consult published version for correct page numbers!

Moreover, images are relevant for personalisation through their ability to create involvement. While interactional elements create involvement by letting the audience participate actively in communication, visual elements can create involvement in the form of an emotional response on the part of the reader. Scheufele (2001, 146-147) is one of the many researchers who expresses the view that visuals create more emotional involvement than texts. He attributes this to the ability of images to create the impression that viewers can experience the reported events more directly than through textual accounts. Stöckl (2004a, 98, 247, 381-382) mentions the elicitation of emotions and their quality of providing documentary evidence as the main function of images in contrast to written or spoken language. More specifically related to news reporting, Schröder (2010, 177) argues in a study on the (changing) roles of images in German newspapers that images carry the main emotional load of news reports. According to him, emotions play a marginal role in news texts, but they are frequently expressed in the accompanying images. Thus, he sees emotionalisation and the expression of emotions as one of four main functions of images in present-day print newspapers, together with text-related functions, conveying information and influencing the reception process, e.g. through attracting attention (Schröder 2010, 169). The relation between images and emotions in news reporting has also been investigated empirically. Knieper (2006), for instance, analyses the emotional reaction of people in Germany after the Tsunami in the Indian Ocean in December 2004. His findings suggest that people who informed themselves regularly through television news showed more reactions indicating stress and fear than other participants in the study. This might indicate that the visual representation of the catastrophe in the news coverage on television was responsible for a more intensive emotional response.

Despite a considerable and increasing research interest in the role of both emotions and of images in mass communication, there are still open questions about the interrelations between them (see Müller and Kappas 2011, 310). While there seems to be general agreement that there is a close relation between images and emotions, the theoretical foundation of this relation has not yet been fully explored. For Kappas and Müller (2006, 3), the common denominator between images and emotions lies in the fact that both function on the basis of associations. They see this as a contrast to texts, which they say is based on rational argumentation (Kappas and Müller 2006, 3). However, it seems obvious that not all images create the same degree of involvement and, consequently, they do not all have the same effect for the personalisation of mass media communication. The factors that influence the degree of involvement elicited by an image are manifold. They include what is depicted, how it is depicted, and the context in which the image is presented. Other factors are not related to the presentation of the image, but rather to its reception. Müller and Kappas (2011) point out that the response to images changes depending on the context in which they are received, so that the same image can have a different effect when viewed in a scientific context than when viewed in a private context. Moreover, the identity of the recipient and his or her general attitude towards what is depicted influence

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

the degree of involvement that an image creates (see Müller and Kappas 2011, 319; Wolf 2006, 58).

Not all of these factors can be discussed in the following. In particular, it is beyond the scope of this study to investigate effects that relate to different reception contexts and recipients. Instead, I will focus on three groups of features related to the presentation of visual elements in news articles. These concern 1) the functional relations between visual and textual elements; 2) the content of visual elements; and 3) the composition of images. I will argue that all three aspects influence the degree and the way in which visual elements contribute to personalisation.

5.2.1 Functional relations between visual and textual elements

There are two basic ways in which the question of the functional relations between visual and textual elements can be approached. On the one hand, one can ask which function images fulfil for the corresponding texts; on the other hand, one can ask which function texts fulfil for the corresponding images (Burger 2005, 414). Burger argues that due to interdependencies between texts and images, the two questions can often not be clearly separated. Moreover, giving preference to one of the two perspectives would mean that either text or image is seen as central. For instance, in asking for the functions that an image fulfils for a particular text, the text is presented as central, with the image fulfilling secondary functions (see also Schröder 2010, 176). The opposite is the case if one reduces the question to the function that a text fulfils for an image.

A primacy of text over image has been claimed for text types that are related to online news. For television news, Burger (2005, 414) argues that the text is intended as primary and images as secondary information, even though this might be perceived differently by the audience and despite the fact that in producing television news it is usually the image that is produced first, with text added at a later point.[5] Moreover, in print news, the text is usually seen as the primary element. According to Langton (2009, 52-53, 96), news stories are mostly developed by writers and word editors and while good images help promote a news story, they are often treated as illustrations, rather than alternative ways of communicating content.

Assuming that similar processes are at work in the production of online news, this would suggest that the functions images fulfil for texts are more relevant in this context than the functions texts fulfil for images. However, such a view would ignore the fact that news articles do not only contain one text, but rather a variety of textual elements (and, it might be added, also different types of visual elements). Captions are one type of textual element that, in particular, seems to be subjugated to the image with which it occurs. Similarly, headlines and leads serve to establish the general context not only for the main text of the article, but also for the accompanying images. Irrespective of whether

[5] However, there is research indicating that the relation between text and image in television news is more complex, and that there are various patterns of interaction between the verbal and the visual, see for instance Lauerbach (2010).

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

images appear as the primary element, e.g. in galleries, or whether they play a secondary role, there is, thus, always quite literally a (con-)text whose function for the image can be investigated. Consequently, both the functions of texts for images and the functions of images for texts are relevant in online news, and the same can be argued for print news.

Starting with the functions of visual elements for texts, three of the functions mentioned by Burger (2005, 418-421) appear to be of particular interest for personalisation, namely arousing interest, creating authenticity, and emphasising that an event is ongoing.[6] That images are able to attract the attention and interest of the news audience is well-known (e.g. Donsbach 1991, 137, 145). Eye-tracking experiments have, for instance, shown that large images serve as the main entry point into a news story for readers of print newspapers (Garcia and Stark 1991). This is a finding of which journalists are aware (see, for instance, Blum and Bucher 1998, 66) and it is certainly one of the main reasons why good images help promote news stories to more prominent parts of the news publication (Langton 2009, 100). On the investigated online news sites, the importance given to visual elements can be seen from the fact that only three out of 140 news articles (2 percent) did not contain embedded images or videos (see section 5.1 above). In the large majority of articles, images or videos were already embedded at the beginning of articles, between the headline and the main text, thus serving as an eye-catching element and potential entry point into the article. Furthermore, on the overview pages, images are often used as teasers for articles.

The fact that images attract attention and can raise interest in a story can be attributed to several factors. One of these is the use of colour. The large majority of images in online news (but also in present-day print news) is in colour. As the results of eye-tracking experiments suggest, the use of colour generally attracts the attention of readers (Stark Adam et al. 2007). However, the fact that images create interest in a story is also closely related to the other two functions of visual elements mentioned above, creating authenticity and emphasising that an event is ongoing. Images and even more so videos let readers see the reported events for themselves and, thus, provide visual evidence of what is described in the text (Bednarek and Caple 2012, 115; Schröder 2010, 174). Creating authenticity in this way is one of the main functions of images in news reporting and it applies to the large majority of visual elements in this context (Burger 2005, 420). A typical example from my data involves images and a video of the opening of the world's tallest skyscraper in Dubai (bbc-wn-100104, mail-wn-100104), which illustrate not only the opening ceremony itself, but also the size of the building. Similarly, the large number of images and videos of the earthquake in Haiti serve to provide evidence for the extent of the catastrophe.

While the function of creating authenticity in this most general sense can be found on all the news sites in my data, there are also variants which are specific to some news

[6] Burger (2005, 418-420) mentions two additional functions of visual elements for texts, which are however not directly relevant to this study. These are "establishing referentiality" and "providing illustrations". With the first, he refers to the fact that in television news (or video reports more generally) images can help determine the reference of referring expressions (e.g. *these*, *she*) which might be ambiguous from the (spoken) text alone. With the second, he refers to illustrations, in particular information graphics, that help the audience understand the text.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

sites. For instance, several articles on the mid- and downmarket sites *Mail Online* and *Sun* provide images with visual evidence of condemnable behaviour. One article on the *Mail Online* reports about a man sentenced for having fabricated a heroic military career in order to impress his wife (mail-uk-100112). As part of this fabrication, he participated in a parade, displaying 17 military medals which he had bought rather than earned. An image of his participation in the parade is embedded in the report, accompanied by an information box identifying each medal. Readers are, thus, invited to check for themselves that he indeed displayed a large range of medals unlikely to have been acquired by a single individual. Another example of illustrating condemnable behaviour comes from the *Sun*, in an article about a woman addicted to inhaling lighter fuel (sun-100122b). The article contains a video, taken from a television talk show on which she and her family appeared, which shows the woman repeatedly inhaling the gas during an interview. In both cases, the possibility to compare textual information with visual evidence provides authenticity and involves the reader.

In the news data from the upmarket sites, comparable examples of visual elements illustrating condemnable behaviour cannot be found. This is not very surprising, given that such topics are hardly treated as prominent hard news items on these sites (see also chapter 6.2). However, even when similar topics are treated in the up-market sites, e.g. in the health section, images are not used in the same way. The soft news data from the *Times Online*, for instance, contain an article entitled *"Our mother's drinking was wrecking us"*, which deals with the problems of children of alcoholics (times-sn-100126) and, thus, has certain similarities in terms of topic to the article from the *Sun*. The text strongly relies on interviews with two affected teenagers, but the image used as an illustration does not display the alcoholism of their parents. Instead, the image shows a glass of wine being held by a person, presumably a woman, whose face is not in the picture. This image gives the distinct impression of being staged and thus does not fulfil the role of providing visual evidence for the text. It serves a more iconic function, visually representing the generalised (rather than personal) problem of alcoholism.

Instead of evidence of condemnable behaviour, a different type of visual evidence can frequently be found in news articles on up-market sites, namely videos from political speeches and press conferences. These videos appear with news articles that incorporate information released as part of the displayed event. A typical example is an article in the *Times Online*, based on a press conference by Chancellor Alistair Darling. The embedded video contains a section of Darling's speech, including some of the quotes that are integrated as direct speech into the text of the article. Readers, thus, have the option to watch and listen to his speech and verify that the quotes represent his statements accurately.[7] Similar examples can also be found in articles covering the Chilcot inquiry, an inquiry about the role of the UK in the Iraq war. My data contains articles on this topic from all three up-market news sites and all of the sites make use of videos showing politicians at the hearing. One article on *BBC News* even embeds a live video stream from

[7] See also chapter 7.2 for how such videos affect the functions of direct speech.

Author submitted manuscript.

Author submitted manuscript. Consult published version for correct page numbers!

the hearing (bbc-uk-100118). Being marked as *LIVE* and with the note *One minute delay* in the upper-right corner, the video provides direct access to what is happening. Thus, it fulfils the third function of visual elements mentioned by Burger, namely emphasising that the event treated in the news article is currently ongoing.[8] Such live videos are particularly powerful in creating immediacy and letting readers perceive themselves as "being in touch" with the news (Lorenzo-Dus 2009, 79).

Arousing interest, creating authenticity, and emphasising that an event is ongoing are, thus, three functions of visual elements which are very closely related. Emphasising that an event is ongoing can be seen as a special case of creating authenticity, since it not only provides evidence for the reported fact, but also for the reported fact being currently in progress. Both of these functions can contribute to arousing the interest of readers, who – in the context of news reporting – are first and foremost interested in real and recent events. Making readers interested in a story can be seen as a first step in creating involvement and, as such, it is a possible contribution to personalisation. If images, furthermore, provide evidence that can be checked and evaluated, this allows readers to engage with a news story through these images, thereby further increasing their personal involvement.

According to Burger (2005, 414), the main function that texts fulfil for images is to reduce their range of possible interpretations. Captions of images in news articles or image galleries often fulfil this function by providing a textual description of the image. The description only rarely represents in words what can be seen on the image with the same range of possible interpretations. Instead, images on their own usually allow for more interpretations than the text and the caption, therefore, suggests to the reader what he or she is supposed to see (see also Schröder 2010, 175). Hall (1973, 177-178) makes this point in relation to how five British newspapers cover the story of the resignation of Home Secretary Reginald Maudling. All of them use a portrait photo of Maudling, and while the photos are very similar, the interpretations assigned to each of them through the image caption range from anger through tragedy to "a look of resignation".

This function of image captions to reduce the range of interpretations frequently affects the extent to which people who are visible in an image are perceived as central to what the image expresses. Figure 5.2 illustrates how the interpretation inscribed by an image caption can even turn a potentially impersonal image into an image with a person focus. The image was published in an image gallery presented with an article on the *Guardian* (guardian-uk-10010). The centre of the image shows a river with trees on both sides, with Fountains Abbey in Ripon, North Yorkshire, visible in the background. Only after reading the caption will one probably notice the two small silhouettes walking towards the abbey. The walkers are so small that it was even deemed necessary to specify their position in the image – *(left)* – since they might otherwise have been difficult to spot. Nevertheless, by

[8] It is noteworthy, however, that the fact that the video depicts an ongoing event is only established through the use of textual elements (*LIVE*, *One minute delay*). It could therefore be argued that "emphasising that an event is ongoing" is not a function of the visual elements in the video, but rather of the textual elements appearing with the video.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

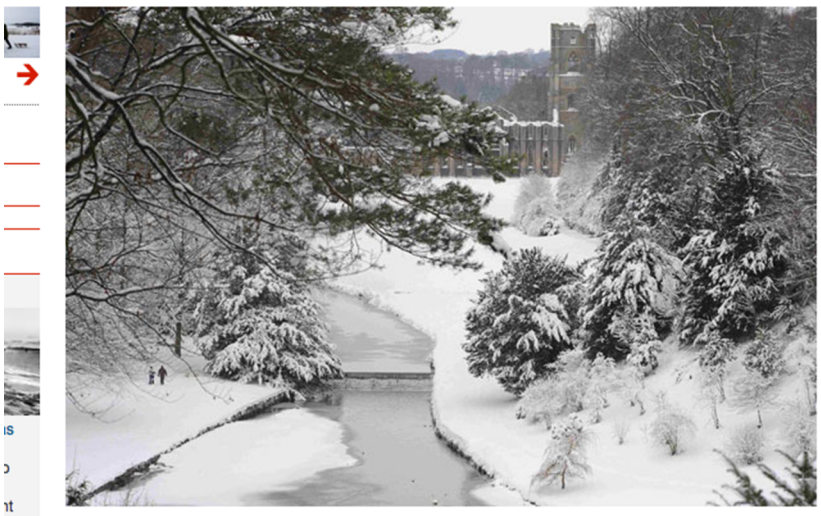


Figure 5.2: Image from image gallery in guardian-uk-100110. Original caption: "9 January, Ripon, North Yorkshire: Walkers (left) approach Fountains Abbey."

providing a caption that focuses on them, they become central to the interpretation of the image.

A similar effect can be observed with image captions for another photograph of a snowy scene. The image shows a street at night in a residential area, with cars parked left and right. Both the street and the cars are covered in snow and more snow is still falling. The exact same photograph appeared in image galleries on the *Sun* and *BBC News*, but the captions are quite different.

(5.1) Blanket. . . scene in Crystal Palace, south London last night (gallery linked from sun-100106b)

(5.2) People in the south-east of England were waking up to snowy scenes for the first time this year (gallery linked from bbc-uk-100106)

There are several differences between the captions in 5.1 and 5.2 that could be discussed. Most interesting from the point of view of personalisation is the fact that the caption from *BBC News* focuses on people and their experiences, despite the fact that they are not even visible in the image. The link between what is visible and the personal experiences of individuals waking up in the morning can be created through the fact that the image depicts a residential area at night and, thus, one can assume that the residents are asleep and will be *waking up to snowy scenes*. Like this, image captions can create a person-oriented interpretation of images, even when no people are depicted.

Related to their role in specifying the interpretation of images is the function of captions to provide background information. This often consists of information about where and when images were taken.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

(5.3) An empty grit box on Queensdown Gardens, Brislington, Bristol (gallery linked from times-uk-100110)

Example 5.3 is the caption for an image showing an empty grit box. The caption describes what can be seen in the image and identifies its location. Specifying the location of the grit box reinforces the function of the image to provide authenticity for the news report. Without the caption, readers could see the image of an empty grit box, but it would not necessarily be clear whether the image could be interpreted as evidence of the lack of grit. It could also be an archived image, used iconically to represent the reported shortage. By declaring the box's location it is made clear that this is not a generic grit box, but that this particular box can be found in the specified location. Despite the fact that the caption does not indicate the time when the image was taken, it is also implied that it is a recent image, directly relevant to the reported events. Therefore, only the caption allows the audience to read the image as evidence of the reported shortage of grit and it is only through the caption that the image can provide authenticity to the news report. This example illustrates well how interwoven the functions of textual and visual elements often are in news reporting (and not only there; see, for instance, Stöckl 2004a, 95).

Captions providing additional information become even more relevant for personalisation when they contain background information on individuals. In particular, this is the case for images that show private individuals who are unknown to the general audience. Example 5.4 is the caption for an image showing a snow-covered street with a man on a dog sledge.

(5.4) Weather-proof: Shane Wilkinson takes his Siberian Huskies, Molly, Zia, Nikita and Ash out on a training run in the snow today at Wilton in Wiltshire (mail-uk-100106)

On the one hand, the caption identifies the time and place at which the image was taken, thus allowing readers to draw their own conclusions about the snow situation and, therefore, potentially increasing their involvement. On the other hand, the caption also introduces an otherwise unknown individual and his dogs by name. This increases personalisation in a different way, by framing a news event (the heavy snow fall) in terms of the experiences of an ordinary individual.

5.2.2 The content of visual elements

Despite the fact that captions have a strong influence on how images are perceived and how they contribute to personalisation, the actual content of images is probably the more important factor. For instance, it seems intuitively plausible that the image of a crying victim of a natural catastrophe is perceived as more personal than the image of a snow-covered street devoid of people. A first approach to the personalising potential of image content could therefore be to classify images according to whether or not they depict people.

However, analysing the content of images is far from trivial. According to Petersen (2006, 39-40), the content analysis of images is faced with the problem that so far no easily

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

identifiable categories for the quantitative analysis of images have been established. The identification of such categories appears to be considerably more difficult than for texts, which may be related to the above-mentioned observation that images tend to be more open to interpretation than texts.

Deciding whether an image depicts people seems to be a relatively simple classification. Nevertheless, while trying to apply this classification to the images in my data, I was soon faced with a multitude of problems. There is, for instance, no obvious boundary between images depicting people and images without them. While there are cases that clearly belong in one or the other category, there are also pictures in between, for instance pictures on which people play a very minor role, on which they are not clearly recognisable or on which they are depicted small or in marginal positions. From the point of view of personalisation, these images would not necessarily be grouped with images on which people are very central. Applying discreet categories to a feature that is realised as a continuum will always result in borderline cases that are difficult to classify. The fact that person orientation in images is a matter of degree, thus, poses the first problem for analysis. The second and related problem concerns the question of how much and what part of a person needs to be visible in order for the image to qualify as one depicting people. There is an image in my data which only shows the legs of a person shovelling snow (guardian-uk-100110). It needs to be decided whether this should be treated in the same way as images showing someone's face. A third problem is posed by the type of image. While the majority of images consist of photographs, there are also other types of visual elements. A *BBC News* article dealing with the debate about banning Muslim face veils in France includes visualisations of various types of headscarves. The images are created with graphical software and the women wearing the scarves are central to the images, but not depicted realistically. Again, it is not immediately clear how such a case should be treated. A related problem is posed by images depicting images. One image in a gallery, for instance, shows the hand of a person holding a photo of his wife up to the camera (times-wn-100118). A final problem relates to image captions and their effects on the perception of images. The image discussed above in Figure 5.2 does not at first appear to have a person orientation. In fact, the people are so small that the casual viewer might very well miss them. After reading the caption, however, they become more important for the image's interpretation. This problem can also extend to other textual elements of the news article, such as the headline or the main text, which may affect the interpretation of images. Consequently, any classification of images needs to take such effects into account.

A very rough analysis suggests that about 70 percent of all embedded images in my data have a strong focus on people. Included in this are images which have people as a central and clearly visible element, no matter whether the image is a photograph or not. The image caption was only consulted in borderline cases, e.g. when a person was central but his or her face was not visible, or when people were visible in more peripheral parts of the image. These results should, however, be treated with great care, due to the problems outlined above. The categories are not stable enough to survive inter-coder

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

reliability testing and, therefore, the results can at best be taken as a very rough indication of certain trends. They suggest, for instance, that the large majority of all images have a rather clear person orientation. There are not many clear differences across the various news publications, given the problematic nature of the classification. The news site with the lowest proportion of person oriented images is *BBC News*, with about 45 percent. This seems to be mainly due to the site's more frequent use of maps and information graphics, compared to the other sites. The absolute number of pictures with person orientation is not lower for *BBC News* than for the other up-market sites. The *Times* from 1985 is noteworthy too. Of the 15 embedded images in news articles, there is only one image without person orientation, namely an information graphic. With the exception of a small caricature, all the other images show photographs with a very clear person focus, mostly depicting politicians.

If even such a basic classification produces results of questionable reliability, it does not seem promising to attempt a quantitative evaluation of the content of these images that goes into more detail, e.g. dealing with the roles of the depicted individuals or their actions. In order to be successful, such an analysis would need to specify in a very detailed manner how images need to be classified, taking into account the various phenomena that can be found. Given that my data covers a large range of topics and a correspondingly large range of types of images such an endeavour would hardly result in reliable and insightful generalisations. Instead, the content analysis of images requires a certain homogeneity of data. This homogeneity can be given in various ways. Bell (2001) demonstrates the applicability of content analysis to images with the example of an analysis of 40 covers of a women's magazine, 20 of which were published between 1972 and 1974 and 20 between 1996 and 1997. The data in this case is very homogeneous; with only one exception, all the covers show a single woman. This enables the analysis of more subtle differences, e.g. of the age of the models, their physical attributes, compositional elements (e.g. gaze and distance) and the arrangement with other elements on the cover.

A different example can be found in a study by Wardle (2007). Wardle uses a less restricted data set of images from news articles from three decades (1930s, 1960s, 1990s), two countries (UK and US) and six different newspapers (three up-market and three tabloids). A certain homogeneity is still achieved through the common topic of the articles, all of them dealing with child murders. Wardle classifies the content of the images according to three main groups of actors or entities who are regularly involved in this type of news event. The first group, called "personal", includes images of individuals who are personally involved as offender, victim, or relatives. The second group, "institutional", contains images of the police investigation and court proceedings. "Societal" images, finally, show the effects on local communities and society more generally (Wardle 2007, 267). These categories are stable enough to score well in inter-coder reliability testing and they provide meaningful results. Wardle (2007, 268-270), for instance, is able to show that societal images, and in particular images focussing on community grief and community anger, make up a much larger proportion of visuals in the 1990s, compared to the 1930s.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

These two examples give a first impression of the degree of homogeneity that is required for the content analysis of images. Generally, the more the investigation focuses on details, the more homogeneity needs to be present in the data. Without at least a common topic, it is very difficult to compare the content of images in various news publications in a meaningful way. As a consequence, a more detailed analysis of image content will be postponed to the case study in section 5.3.

5.2.3 Interactive functions of images: how the content is depicted

One of the most influential attempts at characterising images in a systematic way has been undertaken by Kress and van Leeuwen (1996). Their social semiotics approach to images results in "inventories of the major compositional structures" of visual design (Kress and van Leeuwen 1996, 1). Of the various aspects of images they treat, the one most interesting for this study concerns the interactive meanings of images, the ways in which images address the viewer. Kress and van Leeuwen's aim is to characterise, for instance, in which ways images can create involvement and detachment, how some images are perceived as intimate and others as impersonal, and whether the actors in the image are represented as powerful or not. For this, they identify a number of visual features for which they suggest that they carry out interactive functions, such as expressing power relations and creating involvement. The model has not been without criticism (see, for instance, Bateman 2008, 40-53; Forceville 1999; for an overview, see Machin 2009), and, indeed, when trying to apply their categories to my data, it soon becomes apparent that they cannot be used for a straightforward classification of the interactive meanings of these images. In what follows, I will first briefly introduce the main categories that are relevant for this study, and then go on to illustrate the limitations for their application.

The first categorisation of interactive elements in images concerns the fact whether or not the represented actors in an image gaze at the viewer. Kress and van Leeuwen (1996, 122) argue that a gaze at the viewer "creates a visual form of direct address". They call this type of image "demand" images, and regard them as "image acts", in parallel to speech acts. Their counterparts are "offer" images, in which the represented actors do not gaze at the viewer. The two types of images are associated with different roles for the viewer and the represented actors. In offer images, the represented actor is the object and the viewer is the subject of the look; in demand pictures, the opposite is the case (Kress and van Leeuwen 1996, 124). On the one hand, the roles of subject and object are related to the respective grammatical functions. On the other hand, they are also seen as expressing social power relations. Discussing images from Australian primary school text books, Kress and van Leeuwen (1996, 124-126) interpret the absence of gaze from Aboriginal people as an act of objectification, while the existence of gaze from white people is taken to indicate their status as subjects with whom the viewer is invited to engage as equals.

The second categorisation addresses the size of frame of the image, which is associated with various degrees of social distance (Kress and van Leeuwen 1996, 130-135). Three

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

main categories are distinguished, defined with respect to the represented actor. A close shot shows the head and shoulders of the actor and is taken to express an intimate or personal distance. Frames that cut off the actor at the knees are classified as medium shots and stand for social distance. In long shots, the actor occupies not more than half of the height of the frame and this is taken to indicate impersonal distance. Given that distance and size of frame are continua, additional categories such as extreme close shot or medium long shot can be distinguished. Kress and van Leeuwen (1996, 130-131) argue that the association of the size of frame with different degrees of social distance is based on social relations in everyday interactions, which means, for instance, that strangers tend to remain at a certain distance from each other and, thus, have a view of their interactant which is more similar to a medium shot than to a close shot.

The angle at which an image is taken provides the basis for the remaining two categorisations to be discussed here. The horizontal angle is divided into frontal and oblique, which are associated with involvement and detachment, respectively (Kress and van Leeuwen 1996, 140-146). A frontal angle is given when the frontal plane of the image producer (and likewise the viewer) and the frontal plane of the represented actors are parallel. Formally, this is defined by the vanishing points in the image which need to fall within the vertical boundaries of the frame (Kress and van Leeuwen 1996, 142). For most images, a frontal angle is given when the actors face the viewer (i.e. their shoulders are parallel to the horizontal boundaries of the image). If the frontal plane of the actors in the image is not parallel to the horizontal boundaries of the image, the angle is oblique. According to Kress and van Leeuwen, a frontal angle expresses involvement with the represented actors, whereas an oblique angle expresses detachment.

> The horizontal angle encodes whether or not the image-producer (and hence, willy-nilly, the viewer) is "involved" with the represented participants or not. The frontal angle says, as it were: "what you see here is part of our world, something we are involved with." The oblique angle says: "what you see here is *not* part of our world; it is *their* world, something *we* are not involved with." (Kress and van Leeuwen 1996, 143)

This categorisation seems somewhat less intuitive than the previous two. Kress and van Leeuwen discuss some examples of images in which the frontal and oblique angles appear to be congruent with their interpretation, but it does not become clear why the meanings of involvement and detachment should be necessarily associated with different horizontal angles. An explanation relating to face-to-face interactions, like the one given for the size of frame and social distance, is not provided and also does not seem immediately obvious.

The last categorisation concerns the vertical angle, which is related to power (Kress and van Leeuwen 1996, 146-148). A low angle in which the represented actor is depicted from below, thus towering over the viewer, is taken to express the power of the actor. A high angle, in contrast, stands for viewer power and an eye level angle for equality. Compared to the interpretation of the horizontal angle this seems more intuitive and can more easily be related to other situations in which a higher position is associated with power.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

While applications both by Kress and van Leeuwen themselves as well as by other researchers have shown that their framework can lead to interesting insights (e.g. Bell 2001; Jewitt and Oyama 2001; Kress and van Leeuwen 1996), applying their model is not entirely unproblematic. Kress and van Leeuwen (1996) point to some problems themselves. They mention, for instance, that size of frame and angle are continua rather than discreet categories, a fact which leads to similar problems as those discussed with respect to the presence and absence of people in section 5.2.2 above (see also Forceville 1999, 167-169). As long as individual images are analysed in detail, this might not cause major difficulties, since the degree to which a feature is given can easily be taken into account. However, as soon as the model is applied on a larger scale, the categories would need to be defined more clearly.

A far more serious problem is posed by the fact that it is questionable whether the interpretations they propose can be generalised (see also Forceville 1999, 172-173). A few examples will help illustrate some of the ways in which various factors can overrule the meanings that, according to their model, are expressed by compositional structures. Figure 5.3 shows two images with different vertical angles. The image on the left shows victims of the Haiti earthquake, photographed from a low angle. The image on the right shows the US president Barack Obama during his first State of the Union address, photographed from a high angle. According to the classification proposed by Kress and van Leeuwen (1996) the actors in the left image, the victims, should appear to have power over the interactive participant, thus over both the producer of the image and the viewer. The opposite would be the case for the right image, where the represented actor, Obama, is predicted to appear to be in a less powerful position than the image producer and the viewer. Both these interpretations seem highly unlikely and stand in contradiction to the way in which the actors are presented in the article text. Apart from the fact that Obama, as the president of the United States, is clearly a person in power, the image caption in no way suggests that the depicted scene is to be understood as a moment of defeat (see example 5.5). Instead, *Positive* is reflected in his positive facial expression and the visible applause from those standing behind him.

(5.5) Positive: President Barack Obama delivers his first State of the Union address and makes a commitment to job growth (mail-wn-100128)

In contrast, the victims in the image on the left are presented as powerless and in need of external help. This interpretation is again supported by the image caption:

(5.6) The Red Cross said that the three million inhabitants of Port-au-Prince would need food, water and shelter for months to come (times-wn-100114)

The caption points out that the need for help is not only immediate, but will continue for months, and it attributes this assessment to an international humanitarian organisation. These two aspects mean that the image can be read as an appeal for help to the reader, who likely is in a position to donate money and therefore in a position of power over the

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Figure 5.3: Victims of the earthquake in Haiti and US president Obama during his first State of the Union address (times-wn-100114, mail-wn-100128)

depicted actors. While I refer to the captions as evidence supporting my interpretations, I would still argue that these are also the likely interpretations of these images in isolation. The contradiction to the classification by Kress and van Leeuwen (1996) is, therefore, not caused by the caption, but by the images themselves. It is rooted in the ability of the viewer to visually distinguish between people in need and people in power, based, for instance, on their appearance and facial expression, and it is further supported by the viewer's knowledge about the world in general and the context of the images in particular (e.g. knowing that Obama is the president of the United States). In the case of these two pictures, these factors clearly overrule the effect of the vertical angle.

The images in Figures 5.4 to 5.6 illustrate problems of equating distance of shot with social distance. Figure 5.4 shows a large group of people in the Haitian capital Port-au-

Figure 5.4: Scene after the earthquake in Haiti (guardian-wn-100114)

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Figure 5.5: Image appearing in a slide show related to an article on a Taliban attack in Kabul (bbc-wn-100118)

Author submitted manuscript

Prince after the earthquake. They can be seen sitting and lying on blankets on the ground and the image caption states that they have lost their houses. Given that the individual people in the picture each do not occupy more than half of the height of frame, this is a long shot. According to Kress and van Leeuwen (1996), this image should therefore be classified as expressing an impersonal distance. However, such an interpretation seems to contradict the content of the image, which rather suggests a social distance. Arguing this point is not trivial, given that Kress and van Leeuwen's definitions of the various degrees of distance rely on judgements of whether "[P]eople are portrayed *as though* they are friends, or *as though* they are strangers" (1996, 132, emphasis in the original). Such judgements are likely to be subjective and hard to substantiate. For the image in Figure 5.4, one could argue, however, that it is possible to imagine that it was taken from a position within the group, or at least from a position close enough to allow for interaction with some of the group's members. If this point is accepted, then the image would express social distance, rather than impersonal distance (see Kress and van Leeuwen 1996, 130-131). Similarly, the image in Figure 5.5 can be argued to express a closer distance than what would be predicted by the model. Given that the image is a medium shot, it would be expected to represent social distance, the distance at which impersonal business interaction occurs (Kress and van Leeuwen 1996, 130). If one imagines oneself to be part of this situation, an ongoing Taliban attack in Kabul, it does not seem likely that the interaction with the depicted actors would have the character of a formal business situation. Not only does the visible tension of the situation suggest otherwise, but also the fact that the image allows a backstage view, a view behind the scenes, showing how journalists report on the attacks.

I would even argue that image 5.5 expresses a closer social distance than the close shot in Figure 5.6. This image shows a teenager who was convicted for rape. Images similar to this can frequently be found with news reports on crimes. They are often

Author submitted manuscript. Consult published version for correct page numbers!

Figure 5.6: Image of a teenager convicted for committing rape (sun-100126a)

Author submitted manuscript

released by the authorities, as is the case here, where the source of the image is indicated as *Staffordshire Police*. The image gives the distinct impression of having been taken for official purposes, probably being a police mug shot, or possibly a passport photo. This impression is created through the well-known conventions of such pictures, which require a clearly defined size of frame, a neutral background, a straight view into the camera and a neutral facial expression. The fact that passport photos are easily recognised as such is also noted by Kress (2010, 115), who further argues that passport photos "document a specific, regulated social relation of subject to authority" and that passport photos show individuals "as imagined by the purposes of authority". This already makes it clear that they certainly do not represent an intimate relation. Despite being a close shot showing only the head of the actor, the image in Figure 5.6 expresses a high degree of formality. The size of shot and the degree of distance, thus, do not in all cases correspond in a direct way.

Likewise, one can also argue that the fact that the depicted actor in Figure 5.6 directs his gaze towards the camera does not necessarily present him as a subject endowed with agency, with whom the viewer is invited to engage as an equal. Given the recognisable conventions of the picture and the context in which it occurs, such an interpretation is hardly plausible. Instead, the actor is (publicly) represented as a convicted criminal, which can be seen as part of his punishment. He is the object of the look of the readers, satisfying their curiosity. Thus, problems can be encountered in the application of all the categories: gaze, size of frame and angle.

In recent years there has been a large interest in images, image-text relations and multimodality (e.g. Baldry and Thibault 2006; Bateman 2008; Bucher 2010; Diekmannshenke et al. 2011; Dürscheid 2007b; Jewitt 2009; Kress and van Leeuwen 2001; Kress 2010; Royce and Bowcher 2009; Schmitz 2007; Stöckl 2004a; Stöckl and Grösslinger 2010; Ventola et al. 2004; Wolf 2006). Nevertheless, there have so far been no comparable classifications of image components that provide valid generalisations of their interpretation and that could

Author submitted manuscript. Consult published version for correct page numbers!

be applied on a large scale. This is perhaps not all too surprising. Like language use, communication through and with images is a complex process, depending on the interplay of different layers of structure and meaning, the world knowledge of participants and the context. Accordingly, it is implausible that one specific meaning or interpretation can be assigned to individual presentational and compositional features of images, a point that has been argued before (see, for instance, Bateman 2008; Bucher 2007, 2010, 2011; Forceville 1999). Drawing a parallel with linguistic analyses, this would maybe be akin to claiming that a certain word or utterance always means the same, in all contexts and uses. Instead, as Bucher (2010, 58-65) states, multimodality needs to be studied from the point of view of how the various resources are employed in actual communicative acts. In other words, one could say that the inventory that Kress and van Leeuwen (1996) develop as part of their "grammar of visual design" needs to be complemented by a pragmatics of multimodality. This is not a simple endeavour and thus the fact that structural components of visuals cannot directly be translated into specific functional effects certainly makes the study of images and multimodal documents challenging – but also quite fascinating.

5.3 Case study: Visual elements in reports on the Edlington hearing

I have argued above that the way in which images involve the reader and how they contribute to personalisation is dependent on many factors, including their interaction with related textual elements, their content and composition, and shared knowledge about conventions of image presentation. Generalisations for each of these aspects are problematic, and the fact that they interact with each other makes generalisations even more difficult. The following section will therefore be devoted to a case study, comparing the use of images across the various news sites in articles relating to a single topic, the Edlington court hearing. This hearing concerned the case of two brothers, aged 10 and 11, who were sentenced to indefinite detention with a minimum of five years for torturing and seriously injuring two other boys, aged 9 and 10, in the town of Edlington.

The news reports on the Edlington case are particularly interesting for studying the role images play in the personalisation of news. Given that it is common practice in crime news reporting to include images of offenders and victims (see, for instance, Wardle 2007, 268), the text producers reporting on this case were faced with a serious problem – the fact that the identities of both offenders and victims are protected, due to their young age. Consequently, no photographs were released by the authorities, and journalists were forbidden to publish any information, including images, that would make it possible to identify any of the parties involved. One could assume that this would lead to an absence or, at least, a lower frequency of visual material in these articles. This is not the case, however. Nor is it the case that the visual material used only showed less personal subjects, such as photographs of the crime scene, the police investigation and the subsequent hearing. Such images and videos are used, sometimes even with a personalising effect, but, as I will

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Table 5.2: Visual elements in news articles reporting on the Edlington case

		embedded		
news site	**date**	**images**	**videos**	**linked elements**
Times Online	22/1/2010	1		1 gallery (9 images)
Guardian	22/1/2010	1		profile page (1 image)
BBC News	22/1/2010		2	5 videos
Mail Online	22/1/2010	8		
	24/1/2010	9		
Sun	22/1/2010	6	1	4 videos
	24/1/2010	7	1	4 videos
Total		32	4	10 images, 13 videos

Author submitted manuscript

show, some of the news sites also employ a great deal of creativity in order to present images of the children without revealing their identity.

The Edlington hearing was the top listed UK news story on all the news sites on 22 January, the day on which the sentence was passed. Two days later, the following collection date for my data, the story still provided the top listed news item for the *Mail Online* and the *Sun*. Consequently, there are seven articles on the topic in my corpus, one each from the three up-market sites and two each from the mid- and down-market sites. All of the articles have embedded visual elements and some contain links to additional material (see Table 5.2). The article on the *Times Online* contains an image gallery with 15 images and the *Guardian* contains a link labelled *Profiles: Edlington attackers and victims*, which includes another image. Furthermore, the articles on *BBC News* and the *Sun* contain links to videos. The links to videos in all cases also include the same videos that were already embedded in the article, and two of the four linked videos in the second *Sun* article were already included in the first article. All in all, this amounts to 42 images and 17 videos, although due to repetition the number of unique videos is only 11 and some of the images also occur repeatedly, both across sites and for the *Mail Online* and the *Sun* between the two articles.

A comparison with the figures in Table 5.1 shows that the number of embedded visuals is not generally lower than the average number of embedded elements per article. *BBC News*, with an average of 2.8 embedded visuals per article, has only 2 embedded videos in its Edlington article; in contrast, the *Sun* with an average of 4.1 has 7 and 8 embedded visuals in its two articles, thus, considerably more than average. This already provides a first indication that visual elements were not necessarily lacking as a consequence of the protected identities of the victims and the offenders.

The two sites with articles containing video content, *BBC News* and the *Sun*, are the two sites also generally containing most videos (see Table 5.1 above). The first embedded video on *BBC News* shows statements by the police, the prosecutor and the victims' families being read out to the media after the sentence was passed. The statement by the victims'

Author submitted manuscript. Consult published version for correct page numbers!

families is not read out by themselves, but by a police sergeant. The same sergeant also appears in the second video, in which he recounts how he discovered one of the victims at the crime scene. The third video shows a press conference reacting to the criticism against Doncaster Council and the fourth video is an interview with the Children's Secretary, Ed Balls, on the same topic. The final video does not deal with the Edlington case directly. It is a report from inside East Moore, a secure children's home for young offenders. Except for this last video, which is only loosely related to the topic of the article, all videos on *BBC News* show officials giving statements as part of their professional roles. Each video is taken from one setting only and there are no anchors and no voice overs. They thus create the impression of providing the audience with direct access to the respective press conferences and professional actors involved in the case.

The videos in the *Sun* articles tend to show shorter segments of similar and sometimes identical press conferences and interviews. In some cases, these segments are combined into longer *Sky News* video reports of the crime. These start with an anchor in a studio introducing the topic and then continue with segments showing a reporter at the crime scene, passages from interviews and extracts from press conferences. In contrast to the videos on *BBC News*, the embedded video in the first article also includes an interview with a private witness, a woman from Edlington who saw one of the victims searching for help. Several videos also show the police sergeant who found one of the victims at the crime scene. One of the linked videos consists of an interview with the sergeant, apparently taking place in the same setting (close to the crime scene) as the one on *BBC News*. The 30 seconds of the video are entirely devoted to the interviewee's emotions as the transcript in 5.7 shows.

(5.7) It was the most distressing thing I've ever dealt with. I've been in the job twenty-two years, and it was horrific, obviously. We get involved with numerous searches, hundreds of searches in a twenty-two-years' service, but never would you expect to come across the scene like we'd come across on that day. So obviously, as the automatic pilot kicks in, initially, but after the incident when he got in the ambulance, into the helicopter, then yes, I was upset. I've got children of my own, it's horrific to think what could have been, if we hadn't got to him in time. (fourth video linked to from sun-100122a; transcript by author)

The *BBC News* video also includes a passage with the emotional response of the sergeant and the questions by the interviewer make it clear that this topic was covered intentionally. However, the *BBC News* video is about five times as long as the video on the *Sun* and in addition to his emotions the interviewee also talks about the sequence of events on finding the victim, how the victim recovered after the attack and the fact that, according to his professional experience, the crime was exceptional. In contrast, the video on the *Sun* focuses exclusively on his emotions.

It is not only the spoken statements in the videos that reflect a stronger emphasis on emotions in the *Sun* than on *BBC News*. The settings in which interviewees and reporters are presented fulfil a similar purpose of involving the viewer. The videos on the *Sun*

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Table 5.3: Image content in news articles reporting on the Edlington case (classification based on Wardle 2007)

Category		*Times*	*Guardian*	*Mail*	*Sun*	**Total**
Personal	Victims	0	0	3	0	3
	Offenders	0	1	1	5	7
	Victims and offenders	0	0	1	1	2
Institutional	Courtroom	0	1	2	2	5
	Police	5	0	4	2	11
Other	Area shots	4	0	5	1	10
	Other	1	0	1	2	4
Total		10	2	17	13	42

Author submitted manuscript

regularly include footage from Edlington in general or the crime scene in particular. In some cases the videos show archived scenes of the police investigation after the crime, while voice overs report on the court hearing. The information that the offenders recorded part of their attack with their mobile phones is accompanied by video footage that appears to be taken by a hand camera, moving through the undergrowth. Even though it is highly unlikely that these video sequences published in the *Sun* come from the original recordings taken by the offenders, their imitation still creates the impression that one is directly witnessing the evidence and the crime. In other cases, reporters are shown at the crime scene and there are also sequences with reporters on a playground where the victims and the offenders allegedly met. These images are particularly likely to evoke an emotional response, since they show a seemingly safe place for children to play in connection with a violent attack. In comparison with the videos on *BBC News*, the videos on the *Sun* thus contain more visuals of the crime scene and a stronger focus on emotions.

To start with the analysis of the 42 images, an overview of their content is given in Table 5.3. The categorisation is based on Wardle (2007), who analysed the image content in reports on child murders. The category "victims" also includes images on which the victims' families can be seen, while images showing family members of the offenders are also included into the category "offenders". One additional category, "victims and offenders", was introduced for an image occurring on two news sites, showing both the mother of one of the victims and the two offenders. Four of Wardle's categories – "community grief", "community anger", "nostalgic images" and "politicians" – did not occur in my data. This is probably a consequence of the small number of images in my sample, the different nature of the crime and the fact that the reports in my data covered only one of the crucial moments of the unfolding events, the day on which the sentence was passed and, in the case of the *Mail Online* and the *Sun*, the immediate aftermath. Visual representations of community grief would be more likely when someone dies, which, luckily, was not the case. Images of politicians were present in later reports on the ensuing public debate about the appropriateness of the sentence length, as well as in reports on the investigation of the

Author submitted manuscript. Consult published version for correct page numbers!

social services department responsible for Edlington.[9] Due to the small number of images, no inter-coder reliability testing was carried out for the classification of image content in Table 5.3. However, most of the images will be discussed in the following.

The 10 images on the *Times Online* show area shots and photos of the police. Indeed, the two categories are not always easy to distinguish. There are two long shots of a police tent at the crime scene, guarded by a police officer. Another image shows the path leading to the crime scene, with a police officer visible in the distance. Several images show the town of Edlington. On two of them, police officers are visible, but they do not seem to be directly involved in the investigation of the crime. One image shows an aerial view of the area, with labels identifying the scenes of the crime. The image classified as "other" shows a street sign of Auburn road, the place where one of the victims was found. While people are visible on some of the images, they are either too far away to be recognisable, or they are shown from behind. The only image with people likely to be identified shows two police officers for whom it is not clear whether or not they are directly involved in the case. Similarly, the captions of the images do not identify the depicted actors. Thus, the images published with the *Times Online* article clearly focus on the scene of the crime, the town of Edlington and the police investigation. There are no images in this article which can be said to have strong personalising potential. The main effect for personalisation consists in allowing readers to see the area of the crime scene for themselves. Compared to the crime scene pictures on the *Mail Online* and the *Sun*, the area shots, however, remain more distant. There are no images of the place in the wooded area where one of the victims is said to have been found and the image of a police jacket lying on the ground after having been used to keep the victim warm is also not present in the article on the *Times Online*. Rather than providing specific details of the crime scene, the images on the *Times Online* provide the reader with a general impression of the area.

In contrast to the *Times Online* article, the article in the *Guardian* makes use of one type of image allowing news sites to depict central actors without violating the protection of their identity, namely court room sketches. The embedded image shows a sketch of the general scene, with (among others) the prosecution barrister, the judge, and the two offenders with their solicitors and social workers (see Figure 5.9 further down). While all the other individuals are drawn with clear facial features, the faces of the two boys are blurred and not recognisable. The linked profile page presents another sketch, this time only showing the two boys (Figure 5.7). Again, their faces are not shown in any detail. Still, the image is quite expressive. It shows one boy resting his head on his hands and the other, lying with his head on the table. They appear to be either tired or bored, a theme that is repeated in other sketches of the two boys, published in the *Mail Online* and the *Sun*. Thus, despite the fact that their faces are not clearly visible, the image expresses a characterisation of the two offenders.

[9] See, for instance, http://www.dailymail.co.uk/debate/article-1245758/MELANIE-PHILLIPS-Camerons-right-Broken-Britain-tax-breaks-reverse-descent-savagery.html and http://www.dailymail.co.uk/news/article-1267251/Doncaster-Council-Government-intervention-urged-Edlington-scandal.html (accessed on 9/1/2012)

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

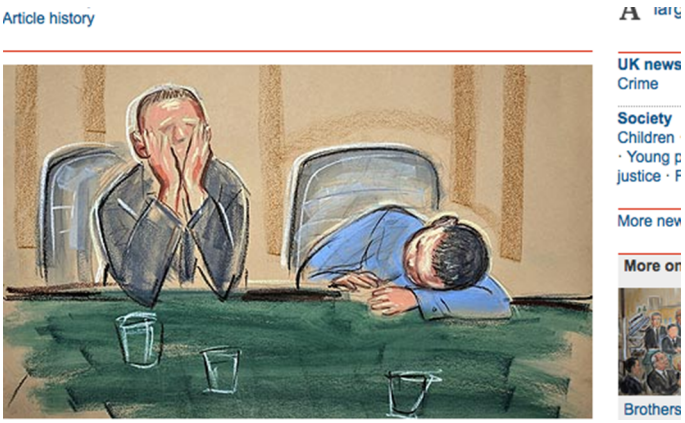


Figure 5.7: Court room sketch of the two offenders (profile page of guardian-uk-100122)

The images in the articles on the *Mail Online* are more varied. They include area shots of Edlington and the crime scene, a portrait of the judge, and two photographs of police officers reading a press statement after the sentence. There are also various court room sketches, including one showing the judge assessing the weight of a stone, and one showing the two offenders in a similar pose as in the sketch published on the profile page of the *Guardian*. In contrast to the other news sites, the articles on the *Mail Online* also include several images focussing on the victims and their families. Images of the victims must have been even more difficult to obtain than images of the offenders, for while the identities of the offenders were also protected, they were present at the hearing and consequently depicted in the courtroom sketches. This was not the case for the victims. The only image of one of the victims is created by zooming in on a section of the general courtroom scene that was published in the *Guardian* article (see Figures 5.8 and 5.9). This section shows a photograph, used as evidence and lying next to the judge, which shows the upper part of the head of one of the victims to document his injuries. Thus, given that no photographs of the victims are available, the *Mail Online* uses a detail from a court room sketch showing a photograph with one half of an unrecognisable face. Still, this image must have been considered central, since it appears in both of the articles from the news site.

Another court room sketch, which is used in the first *Mail Online* article and the second *Sun* article, shows a scene between the mother of one of the victims and the two offenders. The boys are leaving court and the mother is shown from behind, turning with raised arms towards the boys, from whom she is separated by a screen (Figure 5.10). The captions of the images (examples 5.8 and 5.9) state that the sketch shows how she *screams* (*Mail Online*) and how she is *hurling abuse* (*Sun*).

(5.8) The mother of one of the victims screams at the brothers as they leave court following sentencing (mail-uk-100122)

(5.9) Fury . . . sketch shows mum of young victim hurling abuse at yobs (sun-100124b)

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Figure 5.8: Sketch of photograph showing injured victim (mail-uk-100122)

Figure 5.9: Image showing the entire sketch, from which the detail in the lower right corner was taken (guardian-uk-100122)

Author submitted manuscript

It is noteworthy that the caption from the *Sun* uses the more informal terms *mum* and *yobs* where the *Mail Online* uses the neutral and more detached terms *mother* and *brothers* (see also chapter 6.4). The personalising effect of the image showing this highly emotional scene is thus further reinforced through the use of language of immediacy in the *Sun* caption.

Overall, the personal images in the *Mail Online* articles can be said to focus on the victims and their families. The court room sketch of the mother of one of the victims is the first image in the first *Mail Online* article, and the second article also opens with an image of one of the victims' parents. They are shown from behind, standing arm in arm and looking towards a wooded area that resembles the images from the crime scene (Figure 5.11). The image serves to introduce the main theme of the article, which is largely devoted to a long interview with the parents. They remain anonymous and can hardly be identified on the basis of the image. Nevertheless, the image provides an additional focus on the parents and their emotions.

Author submitted manuscript. Consult published version for correct page numbers!

Figure 5.10: Court room sketch showing the mother of one of the victims and the two offenders, who are leaving court (mail-uk-100122, the same image is also used in sun-100124b)

Figure 5.11: Photograph of victim's parents (mail-uk-100124)

Author submitted manuscript

In the two *Sun* articles, the sketch in Figure 5.10 is the only image that shows the victims or their families. In contrast, there are four images of the offenders and one of their mother. One image showing the offenders is a court room sketch similar to the one included on the *Guardian's* profile page, but the others are all based on photographs. The image of their mother shows her from head to waist with her face obscured. She does not seem to be looking into the direction of the camera and might not have been aware of the photograph being taken. This is different in two of the images of the offenders. They are based on portrait photographs of the boys, in one case combining the portraits of both brothers into one image that is used very prominently at the beginning of the first article (Figure 5.12). This double portrait is particularly remarkable. Given that the entire frame of the two photographs is filled with the heads of the boys, whose faces are completely blurred, all that remains visible is the fact that these photographs are portraits of two

Author submitted manuscript. Consult published version for correct page numbers!

Figure 5.12: Anonymised portraits of the offenders (sun-100122a)

individuals. One could argue that the only function this image fulfils is personalisation. It points to the fact that the two offenders are people whose identities cannot be revealed. Apart from this, it does not contain any informational value.

While the images in the *Mail Online* foreground the victims and their families, the images in the *Sun* articles focus on the offenders. Apart from the image of the offenders' mother and the double portrait, there is another similarly anonymized portrait of one of the boys, as well as a pixelated medium shot of his brother in the second *Sun* article. Moreover, this second article is introduced by an image of a sign that is said to have been installed at the offenders' home, probably intended as a joke, which reads *BEWARE OF THE KIDS*. A smaller version of the same image is also included in the first article and both occurrences were classified as "other" in Table 5.2 above. Despite the fact that the image does not depict the offenders, they are still its topic, and, thus, a total of 8 out of 13 pictures from the *Sun* directly refer to the offenders or their mother.

Summing up the results of the case study, various patterns of personalisation can be observed. On the most general level, one can distinguish between news sites focussing on institutional visuals and news sites focussing on personal visuals. The two up-market sites *Times Online* and *BBC News* clearly belong to the former category. They do not include any visuals of victims and offenders, but focus on the police investigation (*Times Online*) and interviews and press statements with actors professionally engaged in the case (*BBC News*). These visuals mainly involve the audience by providing authenticity and visual evidence for the information provided in the text. In contrast, the mid- and down-market sites *Mail Online* and *Sun* have a much stronger personal focus in their visuals. Even though these articles also contain a large number of institutional images, the emphasis clearly lies on the more prominently placed images of the key actors. However, while the *Mail Online* focuses on the victims, the *Sun* focuses on the offenders. Both sites come up with creative ways to present images of the boys and their families, despite the fact that their identities are protected. These images have a high personalising function through showing emotional scenes and through reducing the informational value of the image to

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

the fact that the key actors were people. The *Guardian* takes up an intermediate position. It only includes two images, both showing court room scenes with the offenders.

5.4 Summary and conclusion

This chapter investigated the role of visuals in the personalisation of news articles. Both still images and videos are central to personalisation. On the one hand, this is a consequence of their tendency to focus on people; on the other hand, it is due to their ability to create involvement. The precise effects, however, depend on a large range of factors, of which I discussed, in particular, the functional relations with textual elements, the content of visuals, how the content is depicted, but also the context in which visuals occur, as well as shared knowledge about image conventions. I argued that it is not possible to identify separate classifiable features of images that can be related directly to specific personalising effects, which makes it difficult to undertake a large-scale comparative analysis of visual elements in online news. Instead, a case study investigating a small number of visual elements pointed out different personalisation strategies employed by the various news sites in their coverage of one particular topic, the Edlington case. The results of this case study can be seen as a starting point for further research into the use of visual elements for personalisation in online news. They also provide one, albeit limited, indication of more general differences in personalisation strategies across the various news sites.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Author submitted manuscript

6. News actors

The degree to which content on online news sites is personalised by focussing on individual actors depends on several factors. First, it is a question of the type of topic that is treated. A financial news article about the economic development of a country can generally be expected to have less of a personal focus than a health article discussing a new diet. Indeed, of the two articles whose headlines are given in examples 6.1 and 6.2, the second has a much stronger focus on individuals, their private lives, experiences and emotions.

(6.1) UK economy emerges from recession (bbc-uk-100126)

(6.2) I'll never be that chubby little girl again! How Anna Richardson overcame the childhood fear that drove her to comfort eat (mail-sn-100104)

While such correlations between the general topic of an article and its focus on (private) actors can frequently be observed, this is not a necessary relation. On the one hand, the financial news about the economic recovery could also be presented by introducing a person who is directly affected by it, for instance by making profits. On the other hand, the article on dieting could be presented without introducing personal angles, relying entirely on new results from medical research. The specific aspects of a story that are included and the number and type of actors that are presented, thus, deserve further attention. This point can be illustrated with two more headlines, one from a news article (example 6.3) and one from an opinion column (example 6.4).

(6.3) France MPs' report backs Muslim face veil ban (bbc-wn-100126)

(6.4) In a burka you're cutting me off as well as you (times-cs-100128)

Both articles deal with the question of whether face veils, such as burkas and niqabs, should be legally banned. In the first article, individual news actors have a low presence and the actors that do occur are all represented with respect to their professional functions, for instance being members of a parliamentary commission in charge of compiling a report on this issue. In the second article, in contrast, several actors are represented with respect to their personal experiences. The columnist bases her argument mainly on episodes from her past in which a burka provided obstacles for interaction. In one case, she was herself required to wear a burka in her professional role as a journalist. In another case, she met a woman wearing a face veil while visiting a park with her son and, thus, she had the experience in her private role as a mother.

Author submitted manuscript

It is usually the case that articles with such a strong degree of personalisation appear in sections devoted to soft news or in opinion columns, but experiences of private individuals can also be integrated into news articles. Moreover, even when articles include the same actors, they can differ with respect to how these actors are represented. The following three headlines all report on the same news event and they all refer to the same actor, albeit with different effects for personalisation.

(6.5) Appeal court frees man jailed for attacking burglars (guardian-uk-100120)

(6.6) "Have-a-go hero" businessman Munir Hussain freed by Court of Appeal (times-uk-100120)

(6.7) Dad who fought burglar is freed (sun-100120a)

The personalisation effect depends on the amount of information given about the actor, the choice of referring expressions, and whether official or private roles of the actor are emphasised.

I will start this chapter by discussing the various types of actors that can be found in news articles and their effects regarding personalisation. In section 6.2, I will investigate the relation between the general topic areas of news articles and the types of actors that can typically be found in them. This section also compares the news sites in my data with respect to the presence of both topic areas and actors. Section 6.3 provides a closer look at the various options for representing news actors in more or less personalised ways through the choice of referring expressions. I will end this chapter with a case study comparing the presence of different types of actors and their representations in five news articles on the Edlington hearing, thus continuing the case study of the previous chapter.

Author submitted manuscript

6.1 Official, private, and celebrity actors

At the core of personalising content in news media lies the news value of personalisation (see also chapter 2.1). In the words of Bell (1991, 158), this news value "indicates that something which can be pictured in personal terms is more newsworthy than a concept, a process, the generalized or the mass". The same news value can also be found in Galtung and Ruge (1973), where it is referred to as "personification". Galtung and Ruge define it as follows:

> The thesis is that news has a tendency to present events as sentences where there is a subject, a named person or collectivity consisting of a few persons, and the event is then seen as a consequence of the actions of this person or these persons. The alternative would be to present events as the outcome of 'social forces', as structural more than idiosyncractic outcomes of the society which produced them. In a structural presentation the names of the actors would disappear [...]. (Galtung and Ruge 1973, 66)

Both definitions point to an opposition between named individuals and their actions, on the one hand, and more abstract, generalised processes, on the other. Looking at news articles, it soon becomes clear that this opposition has to be understood as a continuum,

Author submitted manuscript. Consult published version for correct page numbers!

rather than as two discreet categories. Most news articles contain references to general processes as well as to identifiable (if not named) actors. Indeed, all of the 140 news articles from the online sites in my data contain references to named actors, and the same is true of the 36 news articles from the *Times* from 1985. The distinction is, therefore, a matter of degree, depending on the relative focus on people and their actions.

What is missing from the above definition is the observation that different types of actors lead to different personalisation effects. Schneider (2002, 146) states that a distinction needs to be made between official actors (which she calls "public figures") and private actors. Official actors are individuals who appear as news actors as part of their professional role and in an official capacity. Politicians are the prototypical example, but spokespersons, representatives of interest groups, leading executives of companies, and experts also fall into this category. This is the category that Galtung and Ruge seem to have had in mind in the passage quoted above. It evokes scenarios of actors such as politicians appearing on the news due to their responsibility for important events and developments that most likely affect the public at large. The second group, private actors, consists of "ordinary" people who happen to be in the news because they are personally affected by a news event. Talking about the criteria that make private individuals appear in the news, Bell (1991, 194) argues that "the surest way is to be a victim of crime, accident or disaster – that is, to have negative or unexpected things happen to you" (see also Hodge 1979, 161). In a study on interviews in broadcast news, Montgomery (2010) includes two quite different groups of private actors, namely victims of the London bombings in 2005 and successful contestants at the Beijing Olympics. Schneider (2002) lists a number of additional roles in which private individuals can typically be found. Instead of being the victim, private actors can also be the perpetrator of a crime, or they can appear as heroes, e.g. rescuing other individuals, or as witnesses (see also Schneider 2000, 327-328). A final group discussed by Schneider are "curiosities", people who appear in the news due to possessing unusual properties (2002, 147, 166).

I mentioned above that official actors are sometimes referred to as "public actors" or "public figures". My preference for the term "official" rather than "public" is caused by the terminological problems of "public" and "private" that I discussed in chapter 2.3. I suggested adopting the terminological distinction between public and non-public, on the one hand, and private and non-private, on the other, as proposed by Dürscheid (2007a). The public/non-public distinction refers to accessibility and the private/non-private distinction is used to describe whether something affects the interest of an individual or a collective. Applied to actors, one can say that the main difference between the two groups, as described above, is that "ordinary" people represent their own interests, whereas actors appearing in a professional role represent the interest of a larger group or a more abstract collective, e.g. the government, their party, the law, the scientific community, an interest group, the executive board of a company, etc. The difference between the two groups can, therefore, be understood as a difference between private and non-private actors, which

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

would be the most consistent choice of terms. However, given that "non-private actors" is not a very intuitive term, I prefer to use "official actors".

Of course, actors could also be classified with respect to the public/non-public distinction. In this case, the criterion would be how accessible or visible actors are to the general audience, which would result in a distinction between "knowns" and "unknowns", a categorisation proposed by Gans (1979, 8-15). There are two problems with this categorisation, however. First, the distinction between individuals who are known and individuals who are unknown is not clear-cut and more difficult to make than the distinction between private and official actors. Gans (1979, 11), for instance, argues that murderers of well-known people become knowns themselves. While he reserves the possibility for a transition from unknown to known for murderers of known individuals, one could also argue that whoever appears prominently in the news is no longer entirely unknown. This would, however, raise problems for the application of the distinction to news actors, since not many unknowns would remain. Second, Gans's discussion of knowns in the news suggests that the classification is to a large extent based on the profession or official function of the person. His knowns include the president, presidential candidates, leading federal officials, state and local officials, astronauts, and professionals (1979, 10). Gans (1979, 9) himself points out that some of the knowns "are not necessarily known by name but occupy well-known positions, like governor of a large state or mayor of a troubled city". If this is the case, it seems that the underlying property of individuals classified as knowns is not so much whether or not they are known, but whether or not they represent an official function. Consequently, a classification relying upon the criterion of whose interests actors represents seems more suitable.

The distinction between official and private actors is important for personalisation in several ways. For official actors, the main personalising effect caused through their presence is the personalisation of the collective they represent. Presenting a representative of a group or organisation is more personal than referring to this group or organisation in abstract terms, for instance by quoting business reports or press releases. In contrast, private individuals stand for themselves only. They do not serve the personalisation of a collective, but rather they personalise the news event more generally by focussing on how their private lives are related to it.

Moreover, there is a general tendency towards a stronger person focus and a greater identification potential for private rather than for official actors, which means that private actors also tend to have a stronger personalising effect overall. The audience is generally much more likely to identify with private actors, presented as ordinary people "like you and me", than with official actors. Readers can be expected to relate to official actors mostly through how their actions affect them directly or indirectly. For instance, if the union negotiator Len McCluskey is quoted with respect to a possible strike of the cabin crew of British Airways (guardian-uk-100118), readers might think about direct effects on themselves, such as possible complications for their holiday plans or how this would affect the value of their shares in the company. They might also think about more indirect effects,

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

such as general effects on the economy and the politics of the country in which they live or where they have relatives or business associates. It seems far less likely for them to identify with the negotiator and think about how they would feel if they were in his place (although this might in fact be the response of a small number of readers). In contrast, this is exactly the response triggered by many private individuals in news reports. If a victim of the earthquake in Haiti tells of his or her experiences, this will in most cases not affect readers personally. Rather, they are likely to identify with the victim and assess the severity of the event by their own emotional response. Likewise the severity of a crime can be perceived in (inverse) relation to the potential to identify with the offender. The less comprehensible the actions of the offender – in other words, the less one can see oneself act like the offender in the same circumstances – the more condemnable the crime. Thus, while official actors are present in the news because their actions have an effect on the general public, private individuals tend to be present because their experiences provide the audience with a personal entry point into the event.

However, the distinction between private and official actors is not always unproblematic either and while there might be strong tendencies with respect to how different types of actors are typically represented, these should not be mistaken for universally valid rules. For all types of actors, one can distinguish between presentations that foreground private aspects and presentations that foreground non-private aspects. In prototypical constellations, private actors are presented with a strong focus on their private lives and official actors with a strong focus on non-private aspects, but different constellations can also occur. On the one hand, private actors can appear in news reports without providing much insight into their private lives. On the other hand, official actors are shown in a more private way when their personal opinions, experiences and emotions are presented. In a study of the presentation of expert sources in a health debate, Boyce (2006, 897-898) found that scientists were often quoted with respect to their personal opinions, based on their own experience:

(6.8) As a mother of two young boys I understand how concerned and frightened some parents must be feeling about the MMR vaccine. No parent wants to take any risk with their child. Both my boys, aged six and nine, have had two doses of MMR and if I had to take them to be immunised tomorrow I would. (The Sun, 7 February 2002, quoted from Boyce 2006, 897)

The source of the quote is a child health expert, Dr Helen Bedford (Boyce 2006, 897). The fact that she is quoted at all is a consequence of her professional expertise and, thus, she is an official actor, representing the scientific community. At the same time, however, the way in which she is presented foregrounds her personal experience in a private context, i.e a context which does not affect anyone apart from her immediate family. Another example of presenting private aspects of official actors is the emphasis on their emotions.[1] Gripsrud argues that this form of personalisation can, in particular, be found

[1] See also chapter 8.4, where I investigate how first person singular pronouns are used in direct speech by official actors to express judgement, commitment, personal experiences and emotions.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

in down-market newspapers: "If the prime minister gives a speech about something and gets emotional (tears fill her or his eyes), the popular press will typically focus more on the speaker's emotional state than on what was actually said in the speech about a political issue" (Gripsrud 1992, 85). Emotions are one of the most private and non-public aspects that can be discussed. They are private in the sense that only the individual having them is directly affected by them, and they are also non-public in the sense that apart from this individual no one has direct access to them.

As a final example of a focus on private aspects of official actors, one can think of stories about the love lives of well-known people, which often appear in the form of scandals. Again, their love lives are one of the most private aspects of news actors, normally not affecting anyone apart from the individuals directly involved. As soon as certain norms are transgressed, however, these private matters are perceived to affect the interests of the larger public, in particular in the case of official actors with a strong representative function like politicians. To the extent that the love lives of such actors can actually have an effect on the general public, one can argue that their love lives are less private than that of "ordinary" people. At the same time, the sexual escapades of a politician are still a more private topic for a news report than his or her latest political speech.

A special role is played by so-called celebrities. They are not official actors in the same way as politicians or experts, since they do not represent a larger group. Nevertheless, they are not private actors comparable to "ordinary" people who happen to be personally affected by a newsworthy event. They are famous, which means that they appear in the media on a regular basis. Often the media coverage concerns specific skills or properties they posses, for instance being a successful sportsperson, singer or actor, or having extremely wealthy or famous parents. In some cases, individuals are simply (momentarily) famous for being famous, for instance participants appearing on reality television shows. When the fame is based on skills and activities that involve public performances, non-private topics have a particularly strong media presence. The concert of a rock singer, for instance, is clearly a non-private event, given that it affects many. It affects those who attend it, but also the fans who are unable to attend, people living close to the venue who are disturbed by the music, and so on. Celebrities are, however, also frequently presented in relation to private topics: engagements, divorces, pregnancies, holiday destinations, hair cuts, diets, addictions and depressions are all common topics for news articles on celebrities (see also Conboy 2006, 187). All of these do not immediately affect anyone except for the protagonist and his or her closest relatives, although fans may, of course, feel emotionally affected by the latest developments in the life of their star. Again, one can argue that the private lives of celebrities are less private than those of "ordinary" people. Part of the role of being a celebrity involves the fact that certain aspects of their private lives are publicly discussed and celebrities are even known for staging aspects of their private lives for the media in order to profit from the resulting attention.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

6.2 Main topics and actors in top-listed news articles

What counts as news is a question that is not answered in the same way by different news sites. As a consequence, there are large differences in the frequency with which topic categories can be found in their news sections – and, indeed, whether certain topics can be found at all. A topic may be treated as news on the *Sun* while appearing in the section "Life & Style" or "Arts & Ents" on the *Times Online*, and being absent completely from *BBC News*. Moreover, different priorities are given to topics even if they are treated as news. A topic appearing at the top of the news front page of the *Guardian* may well appear much further down on the *Mail Online*. This is important to keep in mind, given that my analysis is only based on the topmost articles.

Market orientation plays an important role for the selection and placement of topics. It is, for instance, well-known that mid- and down-market publications tend to have a stronger focus on human interest stories than their up-market counterparts (Conboy 2006, 12; Harcup and O'Neill 2001, 273; Hodge 1979, 166). However, news sites with the same market orientation also differ with respect to which stories they emphasise. On 26 January 2010, for instance, the *Times Online*, *BBC News* and the *Mail Online* all contained an article about the end of the recession as the top-listed article in their UK news sections. This topic also featured prominently on the UK news front page of the *Guardian*, but only in second position. The article appearing first treated the Iraq enquiry, a topic on which the *Guardian* reported in great detail in January 2010 and which appeared as the topmost UK news article on 4 of the 28 days on which I collected data. On the same collection days, *BBC News* presented this topic only twice at the top of its UK news section, the *Times Online* once and the mid- and down-market sites never. Both topics, the Iraq enquiry and the end of the recession also appeared on the news front page of the *Sun*, but not as the most prominent item. Instead, the two top-listed articles dealt with the rape conviction of a 13-year-old and with the extra-marital affair of the golfer Tiger Woods. Both topics were absent from the front pages of the up-market sites and they were less prominently placed on the *Mail Online*.

In the following, I would like to provide a bird's eye view of the presence of different topics and types of actors in the news articles from my corpus. In order to do so, I establish a systematic relation between general and re-occuring topic categories in news (e.g. crime, politics, and celebrity news) and typical actor constellations. This point seems almost trivial for some of the categories. Celebrity news is by definition about celebrities, and the claim that official actors have a strong presence in articles about politics will hardly be contested. For crime news, the main actor groups of victims, offenders, and official actors as part of the police or legal system have been identified before (Wardle 2006, 2007). As I will argue in the following, other topic categories show systematic patterns for the presence of different types of actors too.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

6.2.1 Topic categories in news articles

In a first step, the news articles in my corpus were categorised into the nine general topic areas "politics and finance", "war and terror", "catastrophe", "crime", "public protests", "weather", "celebrity news", "bizarre news", and "other". "Politics and finance" covers all articles which focus on decision-making processes by the government, political debates, diplomatic interaction between countries, press statements by politicians, political speeches, strike announcements, as well as news from the financial markets. "War and terror" is used to classify articles that directly report on terror attacks and events from war zones. Articles dealing mainly with the political reactions to such events are classified as "politics and finance". This is, for instance, the case for articles about Obama's announcement of new security measures after the attempted bombing of an air plane heading to Detroit, and articles reporting from the Iraq war enquiry. The online articles from the "catastrophe" category all deal with the earthquake in Haiti. Three articles from the *Times* from 1985 were also included in this category. Two of them deal with a cholera outbreak in an Ethiopian refugee camp and the third with the rescue operation of Ethiopian Jews endangered by famine. The "weather" articles all deal with the heavy snow fall in the UK in January 2010. "Crime" articles either report directly about recently committed crimes or, more frequently, about court cases, in particular the sentences passed to alleged perpetrators of murder, rape or assault. The category titled "public protest" is used for a small number of articles which report on various forms of protest by interest groups or members of the public. The "celebrity" category was used for articles dealing with celebrities such as music stars and sports persons, as defined in the previous section. Most of the articles from this category focus on the private lives of such individuals, their extra-marital affairs, and criminal or otherwise problematic behaviour involving them or their close relatives. The category also includes two articles related to the Haiti earthquake, one of which focuses on the singers involved in a charity single launched by the news site and the other reporting on the personal experiences of a singer who went to Haiti to help the victims. Since both these articles focus much more on the celebrities than on the earthquake, they were included in the category of "celebrity". The category labelled "bizarre news" refers to what is sometimes also known as the "other news" (Conboy 2006, 33-34; Langer 1998). It includes only three articles dealing with private individuals who made the news for unusual behaviour or properties. The category "other", finally, was used for articles which did not clearly fall into any of the other categories.

I identified these topic categories in my data in a bottom-up approach, based on similarities across the articles. I decided to form rather broad categories in order to increase the stability of the classification and to reduce ambiguous examples. In particular, the separation between finance news and politics is often very hard to make, given the close connection between political issues and the developments at financial markets during a financial crisis, which was one of the reasons for not distinguishing between the two.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

The reliability of the classification was tested by a second coder, who independently classified a subset of 56 out of 176 articles (32 percent). The second coder classified the articles according to the definitions above, without further instructions. The relative agreement between the two codings was 87.5 percent (agreement in 49 of the 56 cases). The kappa value, a measure correcting for chance agreement, was 0.84. Given that the kappa value is a conservative measure, a value of more than 0.80 is usually considered very high (see Altman 1991, 404; Landis and Koch 1977, 165). The classification of articles according to topic category, thus, has a very high reliability.

6.2.2 Actor constellations in different topic categories

After having established the topic categories, the second step was analysing the presence of different types of actors in the articles from each category in more detail. As mentioned above, the presence of actors can be understood as a continuum, ranging from complete absence to a very strong focus as a main actor. This makes the quantification of actors difficult. Therefore, rather than attempting an exact quantification, I decided to apply a classification system that distinguishes between only four levels. Level $+++$ is used if an actor type occurs as one of the central actors in articles of a particular topic category with hardly any exceptions. Level $++$ indicates that an actor type is sometimes present as a central actor in this topic category. Level $+$ is used if an actor type is hardly ever present in a topic category, even though at least one instance was found in my data. And level $-$, finally, is used when an actor type was never observed in articles of this topic category in my data. The resulting patterns are presented in Table 6.1.

Unsurprisingly, the major type of actors appearing in articles in the category "politics and finance" are politicians. Other common actors in this category are professionals interviewed as experts and interest groups directly involved in the events. An example of the former can be found in an article on the end of the recession in which several economists present their outlook on future developments of the economy in Britain (times-uk-100126). An example of the latter is the representative of a children's charity, who is interviewed about a proposed change of laws regarding sex offenders (bbc-uk-100124). Private individuals are quite rare. If they appear, they do so mostly in connection with crimes or terror attacks that become the object of political debates and decisions. The private individuals are then the offenders, who tend not to be given much space and presence compared to the political actors. One of the few "politics and finance" articles that give considerable presence to private individuals is an article dealing with increasing hate crimes after victories of the British National Party (guardian-uk-100116). The private individuals in this case are victims reporting on how they experienced the attacks.

Official actors are mostly presented in a non-private way in "politics and finance" articles; the reader does not learn details of their private lives and their personal experiences. One of the exceptions is a news article on the *Times Online*, which deals with the political consequences of the extra-marital affair of Iris Robinson, MP for Strangford and wife

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Table 6.1: Presence of different types of actors by topic category, based on 176 news articles from five online news sites and the *Times* from 1985

topic category	no. of articles	actor type: official	actor type: private	actor type: celebrity
Politics and finance	70	+++	+	−
War and terror	22	+++	++	+
Catastrophe	24	+++	++	+[(1)]
Weather	14	+++	++	+
Crime	23	+++	+++	−[(2)]
Public protests	3	+++	+++	−
Celebrity news	7	++	+	+++
Bizarre news	3	+	+++	−
Other	10	n/a	n/a	n/a

key

+++	regularly present as (one of the) main actor(s)
++	sometimes present as main/important actor
+	rarely present / typically absent
−	not observed in my data

[(1)] only on the *Sun*
[(2)] crime news stories about celebrities are classified as "celebrity news"

Author submitted manuscript

of Peter Robinson, the Northern Ireland First Minister (times-uk-100108). The affair, as such, is part of the politicians' private life. However, according to the article, Iris Robinson provided money to her lover to set up a café without declaring her financial interest, despite a possible conflict of interest with her position in the local council. This meant that her private love life conflicted with her non-private role as MP and affected the interests of the public she represented. Thus, even when the private lives of official actors are discussed in political articles, this has a close connection to their official roles.

At the opposite end, professional actors have least presence in "bizarre news" and "celebrity news". The category "bizarre news" consists, by definition, of articles focussing on private individuals who are presented for unusual behaviour or properties. The three articles in my data deal with an Indian girl who "cries tears of blood" (sun-100110b), a man with a face tattoo that looks like a pair of glasses (sun-100116b) and a family who appeared on a television talk show because the mother is addicted to sniffing large amounts of lighter fuel (sun-100122b). In all these cases, the main actors do not appear in the context of their profession or in an official role, but as private actors. They are always presented with a strong focus on their private lives, in particular those aspects that made them appear in the news. If official actors are present, they appear in the role of experts. Similarly, the articles in the category "celebrity" also tend to focus on the private lives of the main actors, in this case celebrities. Celebrities were sometimes presented with respect to non-private

Author submitted manuscript. Consult published version for correct page numbers!

topics, for instance in a *Sun* article presenting the contributors to the charity single for the victims of the Haiti earthquake. Official actors were found more frequently in "celebrity" articles than in the "bizarre news" category, mainly due to judges and legal representatives who appear when celebrities are involved in legal matters.

The categories "war and terror", "crime", and "public protests" take on an intermediate position with respect to the presence of private actors. Articles in these categories tend to include a mixture of private individuals and official actors, in more or less predefined roles. For "war and terror" articles, private individuals appear either as victims or as terrorists. Details about the private lives of the former tend to emphasise the tragedy of the events, whereas details about terrorists try to shed light on their motives and the reasons for their radicalisation. The same is true for the victims and offenders in "crime" news. Both article categories also include professional actors, such as politicians, the police and representatives of the legal system. The articles of the category "public protests" typically feature interest groups and their representatives, alongside politicians commenting on the protests and private individuals who present their own views. In all three categories, there is some degree of variation concerning the presence of private individuals and personal view points. However, this variation does not very clearly point towards differences between news sites and their market orientation. The differences that exist in how the sites present these topics are more subtle than the simple presence or absence of particular types of actors. For "crime" articles, some of these differences will be investigated with a case study in section 6.4.

The two remaining categories, "weather" and "catastrophe", are special because – at least for the data from the online articles – they all deal with the same two topics in my sample. The "weather" articles all report on the heavy snow fall in the UK and the articles of the category "catastrophe" deal with the earthquake in Haiti. All the news sites reported on these events and most even contained several top-listed articles on each of these topics in the period in which I collected data. The presence of different types of actors for these two categories changes throughout the different stages of the news event. On the first day of reporting about the Haiti earthquake, for instance, official actors had by far the strongest presence. In later stages, private actors gained more prominence, first as victims and witnesses, then as miracle survivors, and later in connection with riots.

6.2.3 Topic categories across news sites

In a third step, the data was evaluated with respect to the frequency with which topic categories appear in the various news publications (see Table 6.2 and Figure 6.1). In addition, I compared the articles of each category across the sites to see whether there are systematic differences in the presentation of actors within each category. It proved very helpful to work with groups of articles based on topic categories. Without this restriction, the articles would have differed so considerably that a comparison of actor constellations would hardly have been possible.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

On the up-market sites, 43 percent of the top-listed articles belong to the category "politics and finance", which includes mostly official actors. "Crime" is the only category in which a strong presence of private individuals can be found at all on the up-market sites, but none of the up-market sites contain more than 11 percent of crime articles. The remaining categories with a strong focus on private individuals ("public protests" and "bizarre news") cannot be found at all. This relation is almost inverted on the *Sun*, where only a single article from my data (4 percent) was classified as "politics and finance", and where 50 percent of the articles belong to categories with a strong presence of private individuals, namely, "crime", "public protests" and "bizarre news". In addition, the *Sun* is the only news site that more or less regularly contains news articles dealing with celebrities. This corresponds to previous studies, which have named the strong presence of news about celebrities and bizarre events as a characteristic feature of tabloid journalism (Bingham 2009, 229; Conboy 2006, 33-34,185-206). The articles from the printed *Times* from 1985, finally, show an even starker contrast to the online articles from the down-market *Sun*. The only categories that were found are "politics and finance", "war and terror" and "catastrophe", the three categories which were least frequent on the *Sun* and most frequent on the up-market sites. 75 percent of all articles fall into the category "politics and finance", making this category 1.7 times more frequent than on the up-market online news sites from 2010.

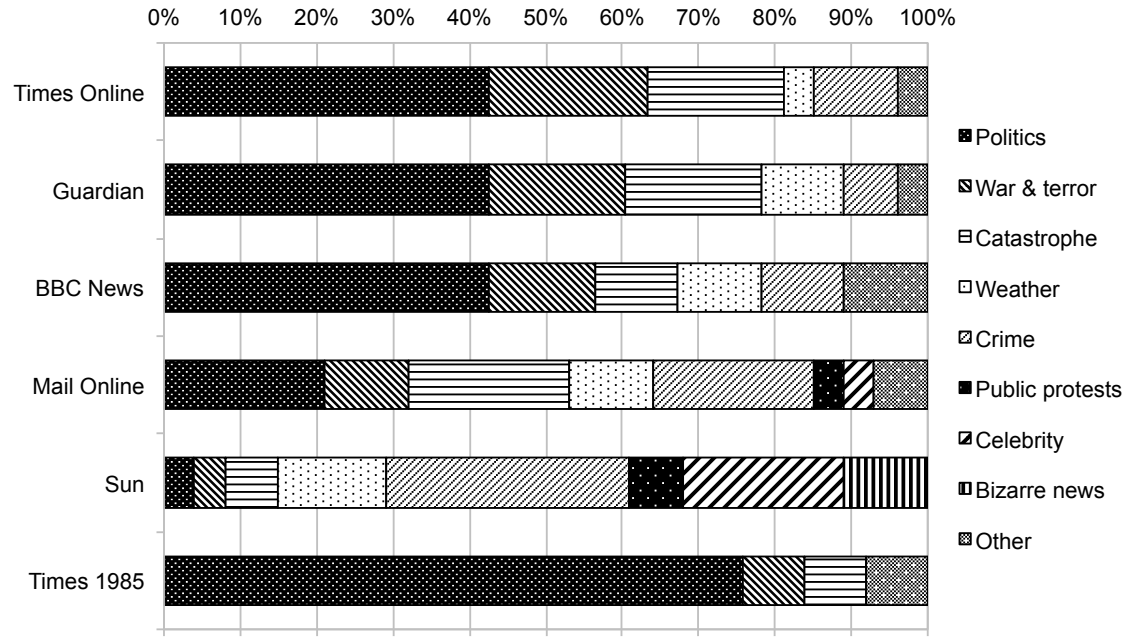


Figure 6.1: Content categorisation of news articles from five online news sites and the *Times* from 1985

Author submitted manuscript. Consult published version for correct page numbers!

Table 6.2: Content categorisation of news articles from five online news sites and the *Times* from 1985

	Times Online		*Guardian*		*BBC News*		*Mail Online*		*Sun*		*Times 1985*	
Politics and finance	43%	(12)	43%	(12)	43%	(12)	21%	(6)	4%	(1)	75%	(27)
War and terror	21%	(6)	18%	(5)	14%	(4)	11%	(3)	4%	(1)	8%	(3)
Catastrophe	18%	(5)	18%	(5)	11%	(3)	21%	(6)	7%	(2)	8%	(3)
Weather	4%	(1)	11%	(3)	11%	(3)	11%	(3)	14%	(4)	0%	(0)
Crime	11%	(3)	7%	(2)	11%	(3)	21%	(6)	32%	(9)	0%	(0)
Public protests	0%	(0)	0%	(0)	0%	(0)	4%	(1)	7%	(2)	0%	(0)
Celebrity news	0%	(0)	0%	(0)	0%	(0)	4%	(1)	21%	(6)	0%	(0)
Bizarre news	0%	(0)	0%	(0)	0%	(0)	0%	(0)	11%	(3)	0%	(0)
Other	4%	(1)	4%	(1)	11%	(3)	7%	(2)	0%	(0)	8%	(3)
Total	100%	(28)	100%	(28)	100%	(28)	100%	(28)	100%	(28)	100%	(36)

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

It is interesting to note that the articles within each category do not show major differences across the various sites regarding the presence of particular types of actors. Thus, fewer of the articles I collected from the mid- and down-market sites fall into the category "politics and finance" compared to the data from the up-market sites. However, the few politics and finance articles from these sites do not include more private actors than the politics and finance articles from the up-market sites. It is, for instance, not the case that the politics and finance articles from the mid- and down-market sites include more personal viewpoints on political events by "ordinary" citizens than the up-market articles. Thus, at least on the basis of this small sample of articles, it seems that the main difference across the sites is related to what type of topics they present as prominent news.

6.3 Representing the actors

Saying that the same types of actors tend to be present in articles on the same topic does not mean that there are no differences with respect to how actors are represented on the various news sites. Examples 6.9 to 6.12 present the headlines of four articles dealing with the same news actor, a mother who killed two of her children.

(6.9) Mother held over Sussex car boot children's murders (bbc-uk-100128)

(6.10) Car boot mother Fiona Donnison charged with children's murders (times-uk-100130)

(6.11) Mother who 'killed her two young children and left their bodies in car boot' had lost baby to cot death (mail-uk-100128)

(6.12) Mum confesses: I've killed my children (sun-100128a)

While the main actor is always the same person, the headlines differ considerably with respect to how she is represented. In 6.9, the least personalised version, she is not named. She appears in the thematic role of patient rather than agent (*Mother held*), and it is not explicitly stated that she is suspected of having killed the children, even though this is clearly implied. This is different in headline 6.10, which states the charge and also identifies her by name. Headline 6.11 does not give her name, but provides information about her private life, namely that she *had lost [a] baby to cot death.* Finally, headline 6.12, the most personalised version, not only refers to her with the informal term *Mum*, but also contains a statement in the first person, presented as her own words: *I've killed my children.*

van Leeuwen (1996) investigates the "representational choices" that are available when representing actors. His detailed framework contains several categories that are relevant to personalisation, among them the distinction between nomination and categorisation (van Leeuwen 1996, 52-59). Nomination represents individuals with their personal identity (e.g. *Fiona Donnison*), whereas categorisation represents individuals in relation to collective identities and roles they share with others (e.g. *Mother*). van Leeuwen (1996, 47, 53) suggests that nominated actors (i.e. actors identified by name) tend to serve as points of identification for the audience. This suggests that named references contribute to personalisation. Similarly, Chovanec (2000, 223) argues that referring to actors by name can be

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

a way of presenting them in a more personal and more sympathetic way, while at the same time establishing them as news actors with an elite status. In his longitudinal study of the news coverage of the court case of Louise Woodward, an English au pair charged in the United States over the death of a child, Chovanec is able to identify the moment at which article headlines in the Electronic Telegraph refer to Louise Woodward by name, rather than by a category label such as *English nanny* or *British girl*. According to him, this marks the moment at which she acquires an elite status, becoming a newsworthy actor in her own right. At the same time, the reporting about her also becomes more favourable (Chovanec 2000, 220-223).

However, examples 6.9 to 6.12 suggest that the tendency for proper name reference to represent actors in a more personal and sympathetic way than category labels needs to be treated with care. The label *Mum* appears to focus on her as a person at least as much as *Fiona Donnison*, and one of the most sympathetic headlines seems to be the one in 6.11, as it provides background information about her loss of a baby to sudden infant death syndrome that may to some extent serve as a possible explanation for her act. Thus, proper name reference alone is an unreliable indicator for the personalisation of news actors.

One important factor influencing the effect of proper name reference versus descriptive labels is the choice of terms. Based on the referring expressions in the headlines of his data, Chovanec (2000, 225) argues that using only the actor's surname *Woodward* is a more detached reference than *Louise Woodward* and that the use of her first name *Louise* is more intimate and casual than the use of her full name. Likewise, referring to the actor in examples 6.9 to 6.12 as *Mrs Donnison* has a different effect than referring to her as *Fiona* or with the nickname *Fi*. Jucker (1996, 378) provides a list of all types of labels that are used to refer to news actors. These are names, descriptive labels, pronouns and appositions consisting of a descriptive label and a name (e.g. *Car boot mother Fiona Donnison*). He further distinguishes between different forms of naming actors, as listed in Table 6.3.

Table 6.3: Name formats referring to news actors (based on Jucker 1996, 378)

Name	**Example**
Title + surname	*Mrs Donnison*
Title + full name	*Mrs Fiona Donnison*
Surname	*Donnison*
Full name	*Fiona Donnison*
Given name	*Fiona*
Nickname	*Fi*

The only format that does not occur in any of the articles on this news event is *Mrs Fiona Donnison*. Indeed, in the entire online data, there is only one occurrence of the title + full name pattern. In contrast, this type is not uncommon in the data from the *Times* from 1985, where it is often used in first named references to actors. All news sites

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

introduce Fiona Donnison with her full name early in the main text, sometimes after a descriptive label representing her as a mother. Later references to her usually take the form of title + surname, which is realised as *Ms Donnison* on *BBC News* and as *Mrs Donnison* on the *Mail Online* and the *Sun*. The *Times Online* mostly refers to her with her surname only in later references, but includes one instance of *Mrs Donnison* in a direct speech quote. Another quote contains one occurrence each of her given name and her nickname, as can be seen in example 6.13.

(6.13) Afterwards, as she was ushered away, two people shouted, "We love you, Fiona," and, "We love you, Fi." (times-uk-100130)

The fact that the use of her given name and her nickname occurs only in quotes by people feeling close to her underscores the personal character of these forms. Given names and, even more so, nicknames are normally used among people who are well acquainted with each other and consequently the use of these forms in news articles provides a more private perspective on these actors than the other forms.

Table 6.4: Selection of descriptive labels referring to Fiona Donnison

	Label	**Site**
1	*the woman*	sun-100128a
2	*a former City insurance manager*	sun-100128a
3	*a 43-year-old former City worker*	mail-uk-100128
4	*a former Lloyds syndicate worker who also has two teenage sons*	bbc-uk-100128
5	*a mother*	mail-uk-100128
6	*The mother of two toddlers found dead in the boot of her car*	times-uk-100130
7	*The mum*	sun-100128a

In a similar way, descriptive labels can also be arranged on a scale with respect to their level of personalisation. Although their positioning is perhaps more open to interpretation than that of different name formats, it still seems clear that referring to an actor as *a former City insurance manager* has a different effect than referring to her as *Mother* or *Mum*. The descriptive labels in Table 6.4 are a selection of the terms used to refer to Fiona Donnison, some of them appearing as appositions together with names. The labels are again ordered according to their perceived level of personalisation and help point out some of the criteria that contribute to more personalised representations. The first label, *the woman*, represents the actor as a gendered person, but it is otherwise very general. Less personalised labels are still imaginable, for instance *someone*, but they are not used to refer to this actor in the four articles. The second label provides more information about her person, but presents her exclusively in her professional, non-private role. Labels 3 and 4 also focus on her professional role, but include more personal information in the form of her age and the fact that she has two sons. Label 4 is thus the first label in the list that refers to her private role as a mother, which is also the focus of labels 5 to 7. Label 6 is considered more personalised than 5 due to the larger amount of information about the actor, i.e.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

the fact that she had two toddlers and that she owns a car. Label 7, finally, contains the term *mum* rather than *mother*, a term that is associated with children addressing their own mother and, thus, evokes a more private perspective on the actor. In sum, the two main factors influencing the degree of personalisation realised through a descriptive label are the amount of information about the actor and the degree to which this information is private. In addition, informal terms can increase personalisation by evoking a private context, similar to the use of nicknames.

There are at least two factors that complicate the analysis of the role of referring expressions for personalisation. The first of these is related to the position in which referring expressions occur. Jucker (1996) demonstrates that the format of expressions referring to news actors varies depending on where in an article they appear. In his data, surnames and descriptive labels are the most common format in titles, while full names account for more than two thirds of all references in the lead paragraph, the place in which the most explicit references tend to occur. In the body of the article, the reference formats are more varied (Jucker 1996, 385). Chovanec (2000) also discusses the relation between the format of referring expressions and their position within the article. Furthermore, he shows that the format of expressions referring to a particular news actor – in his case Louise Woodward – changes over time. Early reports tend to use indefinite descriptive labels in the headline and the lead paragraph and introduce her name only in the body of the article. Throughout the period in which the case is covered by the newspaper, articles become more likely to contain a named reference first in the lead paragraph and then later already in the headline (Chovanec 2000, 236-240). He explains this by observing that named reference in the headline tends to be reserved for actors who are known to the audience (Chovanec 2000, 232, 240). Indeed, in examples 6.9 to 6.12 above, 6.10 was the only headline referring to Fiona Donnison by name, and this was also the headline of the one article that appeared on 30 January, two days after the publication of the other articles. On 28 January, the UK news section of the *Times Online* listed an earlier article on this event with the headline *Woman arrested after children found dead in a car*. This indicates that the headline with the named reference was not used with the first article published about this event on the *Times Online*. These findings suggest that a comparison of different types of referring expressions across various news corpora needs to take into account both where in an article the expression occurs, and also whether or not a person is already established as a known news actor through previous articles.

The second factor complicating the analysis of the representation of actors is that referring expressions are only one of the ways in which actors are personalised. Apart from the fact that features such as visual elements and direct speech play an important role (see chapters 5 and 7), personalising information about actors is, of course, also presented in articles outside of referring expressions. I have argued above that the information about Fiona Donnison's loss of a baby to sudden infant death syndrome adds to the personalisation of headline 6.11. While this is not part of a referring expression, it provides personal background to the main actor and represents her in a more sympathetic way. Such

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

information can be provided in many forms throughout articles. In the present example, the articles on the *Mail Online* and the *Sun* provide more private information than the articles on the two up-market sites. They include quotes from neighbours (example 6.14), information from Fiona Donnison's social networking profile (examples 6.15 and 6.16), but also information for which no source is given (example 6.17).

(6.14) Jill Radford, 61, who lived across from the Donnisons, said: "I have never met the couple in the house or even seen them even though they are just opposite me.
"There were two young children and there were always toys in the garden. I heard that the couple separated just before Christmas and put the house up for sale. It is desperately sad." (mail-uk-100128, two consecutive paragraphs)

(6.15) In her Friends Reunited listing, she describes her children as "wonderful" and says: "Now a lady of leisure – although hopefully that won't be for too long." (sun-100128a)

(6.16) In her profile, she also revealed that she drove a BMW 5 Series and that her hobbies included cookery, socialising with friends, tennis, football, golf and skiing. (mail-uk-100128)

(6.17) Mrs Donnison lived in the six-bedroom property with the kids and husband Paul until her marriage hit the rocks. The house was put on the market before Christmas and a "sold" sign was posted outside last night. The mum moved to a rented three-bedroom house in Lightwater, Surrey. (sun-100128a)

Developing a classification scheme that accounts for private information in its various degrees and realisations and that allows a comparison across large amounts of news data is beyond the scope of this study. Instead, I will, therefore, turn to a case study that compares the representation of actors and the presence of private content on the five online news sites with the example of one news event.

6.4 Case study revisited: Personalising content in reports on the Edlington hearing

In chapter 5.3, I presented a case study investigating the use of visual elements in reports on the Edlington case. In this section, I want to continue this case study by looking at the representation of the actors and the content of the news reports in detail. The articles covering the Edlington court hearing fall into the content category of "crime news". They include all types of actors that can commonly be found in crime news reporting: offenders, victims and representatives of the police and the legal system. As a consequence of the extraordinary nature of the crime having been committed by young children, the story received a great deal of media attention and the background of the offenders was discussed in detail. In addition, the case also had political consequences when it became known that social services had already been aware of problems in the family of the offenders, but had failed to intervene in time. In my case study, I will first give a rather general overview of how much attention the story received on the various news sites of my data. I will then analyse the content of five articles in detail, all of which were published on the same day, one from each of the five news sites. In this step, I will be particularly interested in

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

looking at the main angles from which the story is presented and in finding out how much prominence each of the story lines and the involved (groups of) actors is given. In a last step, I will then investigate how the various actors are represented, in particular by looking at referring expressions for the victims and the offenders.

On 22 January, the day that the Edlington hearing ended and on which the sentence was passed, this was the top-listed UK news story on all online news sites included in my study. However, not all news sites gave the story the same degree of prominence, which can be seen when looking at how long it remained on the front page. On *BBC News*, the Edlington hearing had already disappeared entirely from the UK news front page by the next day. The *Times Online* and the *Guardian*, in contrast, still both presented an article on the Edlington court case as the top-listed UK news topic on 23 January, in both cases dealing with the demand of child welfare campaigners to increase the sentence of the two offenders to a minimum of 10 years. Since this was one of the days on which I did not collect data, these articles were, however, not included in my corpus. One day later on 24 January, the story had moved further down on the page, with a news article on the inspection of the role of Doncaster council being listed on position four (*Guardian*) and five (*Times Online*). After that, the story did not appear again on the UK front page of the two sites, at least not before the end of January, by which time I had stopped collecting data. In contrast, the *Mail Online* and the *Sun* had articles on the story on the top position on 23 and 24 January, and still included articles lower down on the UK frontpage several days later. The last front page reference to the story on the *Mail Online* was found on 26 January and on the *Sun* one day later, on 27 January. Thus, it appears that the story had a greater appeal for the mid- and down-market sites than for the up-market sites. This corresponds to the finding above that crime articles account for 21 percent of all top-listed news articles on the *Mail Online* and even 32 percent on the *Sun*, while only 7-11 percent of the top-listed articles on the up-market sites fall into this category.

For the following detailed analysis, I am only looking at the articles published on 22 January. Before I start this analysis, however, one problem related to the data collection needs to be mentioned. I saved snapshots of the articles in the afternoon of 22 January, shortly after the sentence was passed. The saving of all data in its various formats takes time and consequently not all articles were saved at exactly the same moment. In this case, the *Times Online* article was saved a bit earlier than the articles from the other news sites. When analysing the articles more closely, it turned out that the *Times Online* article was much shorter than the other articles. The fact that it ended with *More soon...* further indicated that this might have been a first version, soon to be updated. Comparing the saved version with the version that was still online several months later indeed confirmed that the article was updated later that day and that the final version, last edited around midnight of 22 January, was 3.4 times as long as the version I had collected. By contrast, some of the articles from the other news sites had later undergone some minor changes and additions, but not in any way comparable to those of the *Times Online* article.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

My approach to collecting online news articles explicitly regards these data as fluid and under construction (see chapter 3.1). While I could have saved articles only a few days after their publication in order to avoid major updates, I decided against this on purpose. On the one hand, there is no point in time at which one can be certain that articles are no longer updated. In some cases, decisions by the Press Complaints Commission prompt modifications months after the publication date.[2] Even at a later date, stability of the text could therefore not be guaranteed. On the other hand, such an approach would have precluded the possibility for observing characteristics of these data that are a direct consequence of their fluid character. For instance, the fact that the *Times Online* article ends with the expression *More soon...* is in itself noteworthy. It expresses a temporal immediacy that points to the ongoing nature not only of text production, but also of the events themselves. As another example, features for interacting with the audience tend to change over time. User comment sections are often closed at some point and requests for submitting eye-witness accounts and additional first-hand information may be removed when news is no longer breaking. The risk of missing such characteristics was the main reason for deciding against collecting data at a later stage.

While it is not generally a problem for my analysis if the versions I include in my corpus are later changed on the online sites, it is a problem in this case. If I want to compare in detail how one particular story is presented on various online news sites, I need to ensure that the versions I analyse are indeed comparable. For the *Times Online* article version saved in my corpus, this is not the case. For this reason, I will in the following include two versions of the *Times Online* article in my analysis. The earlier version is part of my corpus, and this is the version that elsewhere in this study I refer to as times-uk-100122. The updated version is only used for this case study and provides a better basis for comparison with the articles from the other sites. Including two versions of the same article has the interesting side-effect that it provides a glimpse at how articles are updated. It will show, if only for this single case, which parts of the story were already present in an early version, and which aspects were added only later.

As a first general observation, one can say that there are many similarities between all the articles. They mostly report on the same aspects of the story and they are clearly based predominantly on the same sources and material. This can be seen from the fact that many of the quotes appear almost identically in all articles. The differences across the sites relate, on the one hand, to the space and prominence that is given to each of the thematic aspects, and, on the other hand, to how the various actors are described. The headlines of the articles in 6.18 to 6.23 give a first indication of these differences.

(6.18) Edlington child torturers jailed for at least five years (times-uk-100122)

[2] Such a complaint was for instance filed for article mail-wn-100112, see http://www.pcc.org.uk/news/index.html?article=NzA3OA== (accessed on 9/3/2012). Due to delays in getting access to supporting evidence for the complaint, the decision by the Press Complaints Commission was only posted on 21/4/2011, more than 15 months after the publication of the article. In this case, the decision did not call for any further changes to the article, but the case shows that changes can occur long after the publication date.

Author submitted manuscript.

Author submitted manuscript. Consult published version for correct page numbers!

(6.19) 'Sadistic' Edlington child torturers jailed for at least five years (times, updated version)

(6.20) Edlington brothers jailed for torture of two boys (guardian-uk-100122)

(6.21) Edlington torture attack brothers detained (bbc-uk-100122)

(6.22) Torture victim's mother vents her rage at evil brothers as parents face child neglect charges (mail-uk-100122)

(6.23) Devil brothers are locked away (sun-100122a)

Except for the *Mail Online*, all the headlines focus on the fact that the offenders were sentenced. On the *Times Online* and the *Guardian*, they are said to be *jailed*; on *BBC News*, they are *detained*; and on the *Sun*, they are said to be *locked away* – a less formal expression. The *Times Online* additionally gives the duration of the sentence, which becomes important in later demands by interest groups to increase the sentence. The *Guardian* merely provides the offence, *for torture of two boys*. On the *Times Online* and *BBC News*, this information is integrated into the referring expressions *Edlington child torturers* and *Edlington torture attack brothers*. The *Sun* headline gives the least amount of detail. It only refers to the offenders in an emotionally charged way as *Devil brothers*. Similarly negative expressions are used to refer to the offenders in the headlines on the *Mail Online* (*evil brothers*) and in the updated version of the *Times Online*. In the latter case, the attribute *'Sadistic'*, given in inverted commas, is based on a quote by the judge, who described their actions as *prolonged, sadistic violence*. Finally, the headline on the *Mail Online* is the only one that does not focus on the sentence. Instead, the two topics mentioned are the emotional reaction of the mother of one of the victims, and the possible consequences for the parents of the offenders.

Summarising these observations, one can conclude that the focus of the headlines differs in a way that is quite consistent with the findings of the analysis of visual elements in the coverage of this news event (see section 5.3). The up-market news sites, in particular the *Times Online* and *BBC News*, were found to include mainly institutional visuals with a focus on actors professionally involved in the case. Similarly, the headlines of the up-market news sites focus on the legal outcome of the hearing. The visuals on the *Mail Online* were found to have a strong focus on the victims and their families, which is also true for the first part of its headline. The *Sun*, with many images of the offenders, also foregrounds them in the headline. Moreover, through avoiding more technical legal terms like *jailed* or *detained*, the context of a court hearing is much less evoked here than on the up-market sites.

I classified the content of the articles according to the nine categories "court", "social services", "social context", "police emotion", "attack", "offenders", "charges against parents", "victims", and "other" (see Table 6.5 below). "Court" was used for all passages describing the events in court, in particular the pronouncement and justification of the sentence by the judge and the argumentation of the defendants and their lawyers. Descriptions of the episode of the mother of one of the victims insulting the offenders were also included in this category, as were descriptions of the reaction of the offenders on hearing the sentence.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

The category of "social services" contains all passages discussing the failure of the services and the police in Doncaster to intervene before the attack, given that they knew about the problematic family life of the offenders. "Social context" was used to classify statements that relate the attack to other cases of violence committed by children, in particular the murder of James Bulger. This category also contains statements by politicians who see the case as a manifestation of more profound social problems in Great Britain. In one case, the amount of public money spent on institutions for child offenders is discussed in relation to their perceived effectiveness. The category "police emotion" was created to compare the content of the articles with the findings of the case study on visual elements. It only contains two small passages in which police officers describe their emotional reactions to the case. Detailed descriptions of the attack were classified as "attack", while background information about the offenders, their family life and their motivations for the attack were categorised as "offenders". The category "charges against parents" was used for passages that spell out possible legal consequences for the parents of the offenders. "Victims", meanwhile, contains passages focussing on the aftermath of the attack for the two victims and their parents, while "other", finally, contains the two words *more soon...* at the end of the *Times Online* article, as well as two section headings from the *BBC News* article which did not clearly fall into any of the other categories.

The category "court" could have been used for most of the content, given that most information in the articles was based on evidence and statements made during the hearing. It was, however, reserved for passages that did not fall into any of the other categories and in which the context of court events was very clearly evoked. Longer passages describing the attack or the background of the offenders frequently start with reported speech, introduced by matrix clauses like *The court was told* or *Sheffield crown court heard* (see example 6.24).

(6.24) The court was told how the boys were approached by the brothers on 4 April and lured to an isolated wood ravine where they were strangled, stripped and forced to sexually abuse each other.
The elder victim was seriously injured when pieces of ceramic sink were dropped on his head.
After their tormentors left the scene the injured boy told the younger victim: "You go and I'll just die here." (bbc-uk-100122)

In such cases, the matrix clause was classified as "court", while the remaining passage was classified independently, in this case as "attack". Otherwise, not many problems for the classification were encountered. Care was taken that comparable information was always classified in the same way across the various articles. Given that the information was very similar, mainly differing with respect to the amount of detail provided, this did not prove very difficult.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Author submitted manuscript. Consult published version for correct page numbers!

Author submitted manuscript

Table 6.5: Content analysis of news articles reporting on the Edlington case, published on 22 January 2010. Words by category in percent (and absolute)

	Times Online		*Times Online*, updated		*Guardian*		*BBC News*		*Mail Online*		*Sun*	
Court	69%	(214)	52%	(557)	31%	(262)	48%	(454)	21%	(620)	33%	(326)
Social services	16%	(49)	18%	(190)	45%	(371)	14%	(129)	23%	(658)	10%	(101)
Social context	0%	(0)	14%	(149)	2%	(16)	0%	(0)	5%	(148)	0%	(0)
Police emotion	0%	(0)	0%	(0)	0%	(0)	4%	(34)	0%	(0)	2%	(19)
Attack	15%	(46)	6%	(60)	13%	(107)	15%	(140)	18%	(523)	36%	(352)
Offenders	0%	(0)	11%	(113)	4%	(37)	8%	(73)	19%	(565)	17%	(167)
Charges against parents	0%	(0)	0%	(0)	5%	(41)	0%	(0)	5%	(131)	0%	(0)
Victims	0%	(0)	0%	(0)	0%	(0)	11%	(105)	9%	(254)	2%	(16)
Other	1%	(2)	0%	(0)	0%	(0)	1%	(5)	0%	(0)	0%	(0)
Total	100%	(311)	100%	(1069)	100%	(833)	100%	(940)	100%	(2900)	100%	(980)

A very general look at the quantitative data in Table 6.5 already leads to some interesting observations. Except for the first version of the *Times Online* article and the article on the *Mail Online*, all articles are quite similar in length, ranging from 833 words in the *Guardian* to 1,069 words in the updated version on the *Times Online*. The *Mail Online* version, in contrast, is 2,900 words, and therefore about three times as long. It provides much more detail and background and it is also the article that contains most words for six of the nine categories.

Looking at the relative figures indicating the proportion of each category within the entire article, one can see that the largest category on the *Times Online* and *BBC News* is "court", and on the *Guardian* it is "social services". Together with "social context", these are the categories with least focus on individual actors, as I will show in a more detailed discussion of each category below. These three categories make up more than half of the articles on all the up-market sites. Their sum is highest on the earlier *Times Online* article with 85 percent and lowest on *BBC News* with 62 percent. With 49 percent, they account for about half of the text on the *Mail Online* and for slightly less, namely 43 percent, on the *Sun*. In contrast, the categories "attack", "offenders", "charges against parents" and "victims" are the categories which have a strong focus on the private actors directly involved in the event. They are more prominent on the *Sun* and the *Mail Online* than on the up-market sites. While the up-market sites, thus, have a stronger focus on the institutional settings and the social context (although with a considerable degree of variation between them), the mid- and down-market sites give roughly equal space to topics dealing with the institutional setting and to topics dealing with private individuals who were directly involved. Finally, the comparison of the two versions of the *Times Online* article shows that text was added for several of the categories. All the categories already present in the first version were expanded and information about the social context and the offenders was added.

The category "court" is one of the largest categories for all the articles and the information provided in this category is very similar across the sites. Example 6.25 presents a passage that is typical of the type of information that can be found in this category.

(6.25) The brothers pleaded guilty to causing their victims grievous bodily harm with intent, and intentionally causing a child to engage in sexual activity. They had denied a more serious charge of attempted murder and the prosecution accepted their pleas so there was no trial. (sun-100122a)

All articles include information about the sentence that the two offenders received, and all articles contain one particular quote from the judge. Example 6.26 presents the quote as it was included in the article from the *Mail Online*, but the other articles also contain it with only minor variations.

(6.26) The judge said: "The fact is this was prolonged, sadistic violence for no reason other than that you got a real kick out of hurting and humiliating them". He said: "The bottom line for the two of you is that I'm sure you both pose a very high risk of serious harm to others." (mail-uk-100122)

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

From the point of view of personalisation, this quote is noteworthy because the judge is directly addressing the offenders. It also contains the rather informal expression *you got a real kick out of* and the second part of the quote presents the high risk posed by the offenders as the opinion of the judge (*I'm sure*), rather than as a conclusion based on factual evidence and assessments by other experts. Despite the fact that the source of the quote is an official actor, the quote, thus, contains various elements of linguistic immediacy that contribute to personalisation.

Another episode covered by all articles describes how the mother of one of the victims was screaming at the offenders when they were led out of court. The differences across the sites are more marked in this case (see examples 6.27 to 6.30).

(6.27) One of the mothers yelled abuse as the brothers were led out of court, banging a glass partition with her fist and sobbing. (guardian-uk-100122)

(6.28) As they were led out of court the mother of their younger victim swore at them and called them "evil". (bbc-uk-100122)

(6.29) As they were led out of court, the mother of their younger victim banged on a dividing screen in the courtroom and shouted: "You evil little b******ds. "You f****** a***holes. I hope someone does that to you." (sun-100122a)

(6.30) As they were led out of court, the mother of their younger victim banged on a dividing screen in the courtroom. "You evil little b******s," she shouted. "I hope someone does that to you." The woman, who had been staring intently at the boys during the hearing, was led away from court by a family member as the two boys were bundled away. (times-uk-100122, updated version)

While *BBC News* and the *Guardian* include a narrative report of her swearing, the other sites represent it as a direct quote, censoring the offensive language by using asterisks. The direct speech representation creates a much more vivid and dramatic effect, compared to the more detached account on *BBC News* and the *Guardian*.[3] However, this difference does not clearly correlate with market orientation. Both versions of the article on the *Times Online* contain her insults in direct speech and it is in fact the updated version of the article on the *Times Online* which provides the most detailed account, adding information about her behaviour during the hearing and how she was *led away from court by a family member* afterwards.

This episode is by far the most personalising passage within the "court" category. It focuses on a private individual, the mother of one of the victims, and her emotional reaction to the offenders. The remaining text in this category summarises the sentence and presents the argumentation of the judge and in some cases the legal representatives of the offenders, as well as their pleas. It focuses much more on the institutional setting than on the private actors and it does not contain detailed information about their personal lives.

The category labelled "social services" does not contain any strongly personalised text. The articles mainly quote from official reports and represent statements by officials, such as in example 6.31. The only exception to this pattern is example 6.32 from the *Mail Online*, presenting a quote from an anonymous neighbour criticising social services.

[3] For a more detailed discussion of the functions of direct speech for dramatisation and personalisation, see chapter 7.2.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

(6.31) A Safeguarding Children Board report, seen by BBC Newsnight, found the attack could have been prevented. Chair of the board, Roger Thompson, said: "No-one could have predicted the severity of the attack. However, the review has concluded there were serious failings in local services." (bbc-uk-100122)

(6.32) One woman, who lives on the family's estate, said she had repeatedly pleaded with the local council and police for something to be done after witnessing the behaviour of the family. "If social services had acted sooner, maybe those kids wouldn't have been fighting for their lives and whatever," she said. (mail-uk-100122)

Overall, the two categories of "court" and "social services" do not contain a strong focus on private individuals and emotions. The content presented also tends to be public in the sense that it is mainly based on press statements and statements made during the sentence hearing, intended for a public audience.

Three of the articles establish the relevance of the events in a broader social context, namely the updated article on the *Times Online* and the articles from the *Guardian* and the *Mail Online*. In the *Guardian* article, this occurs in a very brief form, referring to the case as *the most notorious crime committed by British children since James Bulger was murdered in 1993* (guardian-uk-100122). The other two articles contain longer passages, both also comparing the attack with the murder of James Bulger. In addition, they include statements by David Cameron, who referred to the case as an example of more general social problems in British society. The article on the *Mail Online* contains the most personalised passage from this category. It discusses the costs of running institutions for young offenders and includes a quote by the mother of one of the victims, complaining about having to help finance such institutions through her taxes.

(6.33) Although huge amounts of money are poured into the centres, 75.3 per cent of children released from custody in secure units, secure training centres or young offenders' institutions reoffend within 12 months. The mother of one of their victims, from Edlington, near Doncaster, said she was disgusted that her taxes would be used to give "these nasty boys the best education". (mail-uk-100122)

This is the only passage in this category in which the perspective of a private actor is presented.

The category titled "police emotion" contains only 53 words in total, divided across the two articles from *BBC News* and the *Sun*. In chapter 5.3, I discussed videos embedded in the articles from *BBC News* and the *Sun*, showing an interview with the police officer who found one of the victims. In both articles, the videos contain passages describing his emotions. The same two articles also contain descriptions of police emotions in the text. In the *Sun* article, the same police officer is quoted indirectly and on *BBC News* a Detective Superintendent is quoted directly with a very similar statement.

(6.34) Speaking about the case, Det Supt Mick Mason of South Yorkshire Police said: "This is clearly the most traumatic investigation I have ever come across. Not the most difficult – but the most traumatic." (bbc-uk-100122)

Such accounts of police emotion are absent from all the other articles.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

"Attack" is one of the two categories, together with "offenders", which clearly take more space on the mid- and down-market sites than on the up-market sites, both in terms of proportion of the overall article as well as in number of words. The *Times Online* only includes a rather factual account of the attack, including the date, the age of the offenders and the victims, the duration of the attack, as well as the acts carried out by the two offenders (example 6.35). Such information is included in all of the articles, but in the mid- and down-market articles it tends to be more detailed and dramatic. Example 6.36 contains a short excerpt of the description of the attack as it appeared on the *Sun*. It shows how the information that *the boys were strangled* is expanded by providing more details. Moreover, both the quote from the offender as well as the capitalisation and bold face for the word *strangle* dramatise the account.

(6.35) The brothers, aged just 10 and 11 at the time of the attack, lured two boys, aged 9 and 11, away from a park in Edlington, near Doncaster, on April 4 last year and subjected them to 90 minutes of terror. Sheffield Crown Court heard how the boys were strangled, hit with bricks, made to eat nettles, stripped and forced to sexually abuse each other. (times-uk-100122, updated version)

(6.36) One brother tried to **STRANGLE** one of the lads with a clothes line. He wound it round his neck three times, asking: "Are you dying now?" (sun-100122a)

The description of the attack on the *Guardian* is quite similar to example 6.35, adding the information that *[t]he older boy almost died from his injuries*. This information is also included in the article on *BBC News*, here in the form of an indirect quote from a witness who *told the court he feared the youngster had died*. In addition, *BBC News* is the only of the up-market sites that includes direct quotes of the victims addressing each other.

(6.37) After their tormentors left the scene the injured boy told the younger victim: "You go and I'll just die here." (bbc-uk-100122)

Similar and even longer passages of such direct quotes from the victims during the attack also occur in the articles from the *Mail Online* and the *Sun*. They have a highly personalising function, not only by giving a voice to the victims, but also by creating a very vivid and dramatic scene of the events (see also chapter 7.2). The article from the *Mail Online* also contains a similar dialogue between the offenders.

(6.38) When the younger brother said he wanted to go, the elder allegedly told him: "Wait, I ain't killed 'em." Minutes later he caught up with his accomplice and said he thought one boy was dead, "cos he's just laid there saying nowt". (mail-uk-100122)

In this case, the situation is, once again, dramatised by providing direct quotes. However, whereas example 6.37 emphasises the suffering of the victims, example 6.38 emphasises the cold-bloodedness of the offenders. In addition, it is noteworthy that the quotes of the offenders' speech are represented in non-standard English. Such marked deviations from standard English are quite rare, even in direct quotes. Clark and Gerrig (1990, 785), for instance, state that newspapers conventionally represent direct speech quotes in formal

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

register and standard language and they add: "To do otherwise is to mark it for some purpose." It seems that in this case the purpose of using non-standard features is to contribute to the negative characterisation of the offenders. There is another quote with a non-standard expression being attributed to one of the offenders. This quote occurs on *BBC News*, the *Mail Online* and the *Sun* in passages classified as "offenders" and it represents the reason one of the offenders gave for carrying out the attack.

(6.39) The elder brother told police he attacked the boys because he had been bored and "there were nowt to do". (bbc-uk-100122)

Naming boredom as the reason for their attack is in keeping with the court room images of the two offenders in which they were presented as being bored or tired by the court events (see chapter 5.3). Such images were present on two of the three sites including the quote, namely the *Mail Online* and the *Sun*. As was the case with the expression of police emotion, there are, thus, strong parallels between the content of the article text and its visual elements.

The characterisation of the offenders and their family background is the main topic of the passages in the "offenders" category. Except for the first version of the *Times Online* article, all articles contain information in this category. In all accounts, the family of the offenders is described with the terms *violent/violence* and *chaotic/chaos*. Except for the *Guardian*, which presents the briefest version, access to drugs, horror films and pornography are mentioned as well (see example 6.40).

(6.40) The court earlier heard the brothers, who were placed in foster care a few months before the attack, had a chaotic upbringing with their violent father, mother and five brothers. They were shown "extreme" horror films and the younger brother had access to pornographic DVDs and smoked cannabis grown on his father's allotment. (bbc-uk-100122)

The description on the *Sun* again adds some additional information, in particular that one of the offenders had already before the attack been described by a psychiatrist as a risk to the community. The *Mail Online* contains by far the most detailed account of the offenders' background. The article lists a range of different offences the two boys had committed prior to the attacks on their two victims. Information about the family's general lifestyle is attributed to neighbours who are reported to have been *terroris[ed]* by *senseless vandalism and sporadic violence*. The article also gives more room to accounts describing the father of the offenders as committing violence against his children and their mother.

(6.41) The father, who was a heavy drinker, would rain blows on the boys if they attempted to step in to protect their mother, the court heard. The boys would imitate what they had seen at home, assaulting adults and children they came across in their daily life. (mail-uk-100122)

While most information about the offenders' background presents them in a very negative way, passages describing their violent upbringing also present them as victims. The passage in example 6.41 quite explicitly shifts responsibility for their violent actions to

Author submitted manuscript

their father, by stating that they imitated his behaviour. The source of this information is not explicitly named, but it appears within a longer passage which starts and ends with statements attributed to the legal representative of one of the offenders, suggesting that this information was part of his defence. The responsibility of the offenders' parents is expressed even more explicitly in a small passage on the *Guardian* and in a more detailed version on the *Mail Online*, stating the possibility that the parents will be charged for child neglect. As the analysis of the headlines earlier in this section has shown, the *Mail Online* even includes this information in its headline, thus presenting it as very central to the case.

The most noteworthy observation about text classified as "victims" is that there is hardly any. The victims are present in the passages classified as attack, but further background information about them and their families is hardly provided. Wardle's results of a diachronic study of the representation of child murders would suggest a stronger presence of the victims. According to her findings, offenders had a very strong presence in data from the 1930s and 1960s, while in the 1990s "the story of the offender had almost disappeared, replaced by the emotive drama surrounding the grieving family" (Wardle 2006, 523). There are several possible explanations for this difference between her findings and the data from the Edlington case study. First, the Edlington case is not a murder case. This is particularly important because the identities of the surviving victims are protected, thus making them less accessible for text producers. Second, the offenders are not adults, but children. This clearly sets the attack apart from other cases of violence against children and explains the stronger focus on the offenders. Finally, the data were collected and evaluated according to different criteria. Wardle (2006, 517, 520) evaluates a large number of articles on four cases of child murder in the 1990s, two in the US and two in the UK. Each of these articles is assigned to one category, based on its predominant focus. In contrast, I compare the relative weight of topics within a single article from each news site. It is possible that the sites published additional articles with a stronger focus on the victims at a different date. Indeed, the second *Mail Online* article in my data, published on 24 January, has a very strong focus on the victims and their families.

Returning to the articles in my case study, the *Sun* contains a short description of the immediate reaction of the victims' families to the hearing. In addition, *BBC News* and the *Mail Online* include information from a statement from the victims' families that was read out by a police officer after the hearing. In it, the families describe their experiences and emotions, as well as those of their children. Example 6.42 presents an extract of this statement taken from the *BBC News* article. Except for punctuation marks, the passage is identically included on the *Mail Online*.

(6.42) "Their physical and emotional scars will live with them for a long time to come.
"Their relationship with each other has been seriously affected and their parents have been left with a strong sense of guilt, which they didn't begin to deserve, about whether they could have done more to protect their boys."
(bbc-uk-100122, two consecutive paragraphs)

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

The *Mail Online* contains two additional passages, one representing a statement from the father of one of the victims expressing his anger, and one providing information about the victims' recoveries.

As a last step in the analysis of the six articles reporting on the Edlington hearing, I am going to briefly present the main referring expressions that are used for the victims and the offenders. Given that the names are not revealed for any of these actors, the main interest lies in the descriptive labels that can be found. On all the news sites, the most frequent expressions used to refer to the offenders are *the [two] brothers*, followed by *the [two] boys*. The terms *brother[s]* and *boy[s]* are also often used to refer to one of the offenders by using expressions like *the elder/older/younger brother* and *one of the boys*. All these expressions are rather neutral, but more emotionally charged expressions can also be found, in particular on the *Mail Online* and the *Sun*. The mid-market *Mail Online* repeatedly refers to the offenders as *evil brothers*. It also includes one instance each of *hell boys*, *young monsters*, and *sadistic pair*. In addition, a subheadline reading *Young sadistic killed is now "likeable and well-behaved"* presumably contains a typographic error and must have been intended to include the referring expression *young sadistic killer*. The expressions *devil brothers*, *evil brothers* and *evil pair* are also used as referring expressions on the *Sun*. These terms describe the offenders as inhuman, or even as non-human (*monsters*, *devil*), and they all occur outside of direct quotes, thus representing these evaluations as facts. In contrast, such evaluative referring expressions are mostly absent from the up-market sites. There are no comparable expressions on the *Guardian*. The evaluative expressions on *BBC News* are *attackers*, *tormentors*, *well-known troublemakers* and *Edlington torture attack brothers*, the last of which is used in the headline. Compared to the expressions on the mid- and down-market sites, these are less emotionally charged and they do not present the offenders as non-human beings. The updated *Times Online* article similarly refers to the offenders as *perpetrators* and *attackers*. It also includes two expressions which otherwise can only be found on the mid- and down-market sites, namely *evil little b******s* and the attribute *sadistic* in *"sadistic" Edlington child torturers*. The first case appears in the quote from the mother of one of the victims that was discussed above. In the second case, *sadistic* is put in inverted commas, signalling that this term is quoted and, thus, does not represent the evaluative stance of the news site.

The two victims are mostly referred to with the terms *victims* and *boys*. Emotionally charged expressions again occur mainly on the mid- and down-market sites. The *Mail Online* twice refers to them as *innocent schoolboys* and one reference to *the younger victim* contains the apposition *barefoot and covered in blood*. Likewise, the *Sun* refers to them as *a pair of innocent young boys* and *terrified chums* and it also contains the expressions *the frightened lad* and *the blood-soaked lad*. Thus, both sites include references to the victims that emphasise their innocence and their injuries, while on the *Sun* their emotions also become part of referring expressions. The only comparable referring expressions on up-market news sites occur on *BBC News* with one instance each of *the blood-soaked boy* and *the injured boy*. In sum, the emotionally charged expressions found in the Edlington

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

articles quite literally reflect how victims and offenders are generally represented in crime news: "[v]ictims and offenders are the lynchpin of the coverage, with the 'innocence' of the former juxtaposed with the 'evil' of the latter" (Wardle 2008, 146).

The detailed analysis of the article content allows three main observations. First, it shows that – with few exceptions, such as the emotional outburst of the mother of one of the victims – the categories "court", "social services" and "social context" focus on the institutional setting and on official actors. These categories have a strong presence on the up-market sites, accounting for more than 60 percent of the text from each article. They contain less personalised content than the categories "attack", "offenders", "charges against parents" and "victims", which have a stronger presence on the mid- and down-market sites. Second, the *Mail Online* and the *Sun* repeatedly contain the most detailed, most dramatic and generally the most personalised passages within each category. The difference between the up-market news sites, on the one hand, and the mid- and down-market sites, on the other, is, thus, not only a difference in the amount of space devoted to each category, but also in the way in which information is presented. Third, the discussion of several of the examples shows the interrelation between personalisation features related to the content of the articles, on the one hand, and linguistic features and visual elements, on the other.

6.5 Summary and conclusion

This chapter investigated various ways in which the presence and representation of news actors can contribute to personalisation. I argued that it is, in particular, the presence of private actors that leads to personalisation. On the one hand, this is due to the tendency of presenting private actors with respect to their personal experiences; on the other hand, private actors provide more possibilities for identification. The analysis of actor constellations across news topics showed that while official actors can regularly be found in most articles, private actors tend to be more restricted to certain topics, especially news about crime, public protests, and bizarre news. These topics were found to be more frequent in the data from the mid- and down-market news sites than on the up-market news sites, leading to a stronger presence of private actors in the former. Within each topic category, in contrast, there were no clear differences across the sites with respect to whether private actors are present. However, the personalisation effects can still vary due to differences in the representation of actors, for instance through the choice of referring expressions.

A case study on news articles about the Edlington hearing illustrated some differences across the five online news sites. It showed that they differ not only with respect to how much space they give to individual aspects of a news story, but also with respect to how actors are represented. In particular, the close connection between the content, visual elements and linguistic features became apparent. In several cases, news sites which included information in textual form had been found in chapter 5.3 to include the same information in visual form. In addition, the personalisation of the actors was often closely

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

related to linguistic features, such as the use of direct speech, emotionally charged referring expressions, and non-standard language.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

7. Direct speech

On 7 January 2010, the US president Barack Obama gave a press statement in which he announced measures that were taken as a consequence of an attempted suicide bombing of a plane on 25 December 2009. One passage of this thirteen-minute statement received particular attention from the news media: the passage in which Obama took over the ultimate responsibility for the event. All three articles in my data reporting on this press statement represented this part in Obama's own words, thus in the form of direct speech:

(7.1) President Obama said last night that ultimately it was his responsibility to keep Americans safe from terrorists as he announced a major shake-up of US security in the wake of the Christmas Day airline bomb plot.
"The buck stops with me," Mr Obama declared during a long statement at the White House, in which he decried the security failures that allowed a Nigerian man with known links to extremists, and with plastic explosives sewn into his underwear, to board a US-bound flight. (times-wn-100108)

(7.2) President Obama took responsibility for the security failures, saying: "Ultimately the buck stops with me." (mail-wn-100108)

(7.3) Mr Obama criticised "systemic" intelligence failings over the plot, but said: "The buck stops with me." (bbc-wn-100108)

What are the functional differences between representing this statement as *Obama took responsibility* or as *"the buck stops with me"*? In other words, what are the effects of representing this statement in the form of direct speech, rather than in indirect speech or in a narrative representation of the speech act? The relevant information, namely the speech act value, could be represented in all of these forms. Indeed, in two of the reports the direct speech account does not add any information that was not already given previously in indirect speech (7.1) or in a narrative representation of the speech act (7.2).

The reason for choosing a direct speech representation in addition to (or instead of) an indirect representation may well be found in the different functions that the various forms of speech representation fulfil. Using direct speech suggests that the journalist had access to the exact wording of the statement. If not, he or she would be unable to use direct speech without violating its faithfulness claims. Direct speech signals to the reader: this is what Obama actually said, thus highlighting the factual status of the quote. This is even reinforced in the case of the *Times Online*, where a video at the top of the article shows Obama delivering this passage of his speech. In addition, a link is provided to the website of the White House where one can find the entire transcript, as well as watch a

Author submitted manuscript

video with the complete statement. All this enables readers to "see for themselves" that this is, indeed, what Obama said.

However, this is not the only function that might motivate the use of direct speech in this case. The direct speech representation also adds flavour to the report by including the colloquial expression *the buck stops with me*, rather than the more formal and detached formulation *taking over responsibility.* Moreover, the formulation *the buck stops with me* adds intertextual references, referring to the phrase *the buck stops here*, which is strongly associated with former US president Truman (McCullough 1992, 577).[1] It would probably not be appropriate for a news article published on one of these three news sites to include the colloquial formulation in an indirect quote, especially if it had not been used by Obama in his statement. Likewise, implied meanings and intertextual references are less likely to be expressed in indirect speech and narrative representations of speech acts. Direct speech, therefore, allows the integration of more flavoured expressions through colloquial formulations and intertextual references.

By enriching the article with the flavour of Obama's own words, he is also given a voice in the text. Whereas indirect speech and narrative representations contain only the voice of the reporter and give little presence to the news actor, direct speech brings the news actor's own voice into the text.[2] Thus, direct speech not only differs from other forms of speech representation with respect to the faithfulness claims about the wording of a statement, but it also gives a different degree of presence to the person of the news actor. The use of direct speech is, thus, an important way in which news articles are personalised (see also Bös 2010a,b; Schneider 2002).

In this chapter, I am going to look at how news actors are personalised through the use of direct speech. I will start with some terminological distinctions and classifications of forms of speech representation. This will be followed by an outline of the functions of direct speech in news reporting in general and the shifts that can be observed with respect to these functions in my online data. After presenting some quantitative data about the use of direct speech in different parts of my corpus, I am going to investigate the sources whose speech is quoted in more detail. On the one hand, I will identify tendencies across the different news publications with respect to whether or not sources are introduced by name. On the other hand, the proportion of private and official sources in the different news publications will be of interest. This investigation of the sources of direct speech reveals different degrees as well as different characteristics of the personalisation of news in the publications included in my corpus.

[1] I thank Herbert H. Clark for bringing this point to my attention.

[2] It should be noted, however, that the choice of reporting verb in direct *and* indirect speech can express the stance of the journalist (see, for instance, Floyd 2000). This is, therefore, a feature by which text producers are personalised and which would deserve further attention.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

7.1 Forms of speech representation

Speech representation refers to the reproduction of statements that were made previously, in what is usually called the "anterior discourse", in a new context, the "posterior discourse" (see, for instance, Short et al. 2002). Various forms of speech representation can be distinguished, depending on how closely the representation corresponds to the original. In direct speech, the actual words and structures are reproduced, in indirect speech the propositional content of a statement is represented without necessarily using the same expressions, while in narrative representations of speech acts only the speech act value of the original utterance is reported. Further distinctions can be made with respect to the mode of the anterior discourse, which leads to the differentiation between speech, thought and writing representation.

Before outlining these different forms of speech representation, a few comments on terminology need to be made. A number of terms have been used in the literature to refer to "reporting what someone said". These include "speech reporting" (e.g. Caldas-Coulthard 1994; Collins 2001; Coulmas 1986), "speech presentation" (e.g. Bell 1991; Brüngel-Dittrich 2006; Jucker 2006; Leech and Short 1981), "speech (and thought) representation" (e.g. Vandelanotte 2009), "discourse presentation" (e.g. Short et al. 1999, 2002), "discourse representation" (e.g. Fairclough 1988, 1992, 1995) and "citation"(e.g. Calsamiglia and López Ferrero 2003). In addition, it is also frequently referred to as "(direct) quotation" (e.g. Bell 1991; Bös 2010a; Brendel et al. 2011b; Cappelen and Lepore 1997; Clark and Gerrig 1990; Mayes 1990) and the segments of direct speech in a text are called "quotes" or "quotations", with "quotes" being the more commonly used expression in news contexts (see Bell 1991, 209).

Which terminology is used depends mainly on the research tradition in which speech representation is studied (see Short et al. 2002, 333-336). Speech reporting would maybe be the most obvious choice for the study of newspaper data, due to its close connection with reporting as the central activity of journalists. My own preference for "representation" is motivated by my aim to stress its performative aspect. My central focus is not on the accurateness of the repetition of what sources said, but on how voices are created in a news article (see also Fairclough 1988, 1992, 118-120). The use of "speech representation" therefore alludes to a concept of representation as proposed by Derrida (1982), which incorporates both the aspects of citation and performativity (Thomassen 2007, 117). With this, I would like to emphasise that direct speech in news articles does not only have the status of a (more or less accurate) copy of a previously uttered statement. These statements – and, through metonymic extension, those who made them – are also represented through direct speech in the sense of 'making them present' or 'bringing them into presence'. In Waugh's words, "the reported speaker is, so to speak, on the scene" (Waugh 1995, 137-138). It is this second, performative aspect of speech representation that is most relevant for the personalisation of news articles.

In order to avoid misunderstandings caused by switching between different terminology, I will in the following always use "speech representation" as a cover term for the different

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

ways of "reporting what someone said". This is also the case when I discuss literature on the topic which uses different terminology. Thus, Leech and Short's categories of speech representation below are in fact categories of speech *presentation*, and Short et al. variously use discourse presentation, discourse (re)presentation, and discourse report/(re)presentation in addition to speech, thought and writing (re)presentation in their publications. In addition, both "quotation" (for the action of quoting) and "quote" (for the segment that is quoted) will sometimes be used in the discussion of direct speech in my data.

7.1.1 Leech and Short's classification

One of the best-known models for the classification of different forms of speech representation was developed by Leech and Short (1981, 318-351) as part of their analysis of style in fictional writing. Figure 7.1 lists the five categories they distinguish, together with the examples they use to illustrate the categories. Their reference point is direct speech and they describe the other four forms of speech representation in relation to this category. This is also the reason why this is the only category for which only one possible realisation is given. Variation would also be possible for the reporting clause in direct speech. In their example, the reporting verb could, for instance, be *he promised* instead of *he said*, and the reporting clause could also stand at the end of the reported speech.

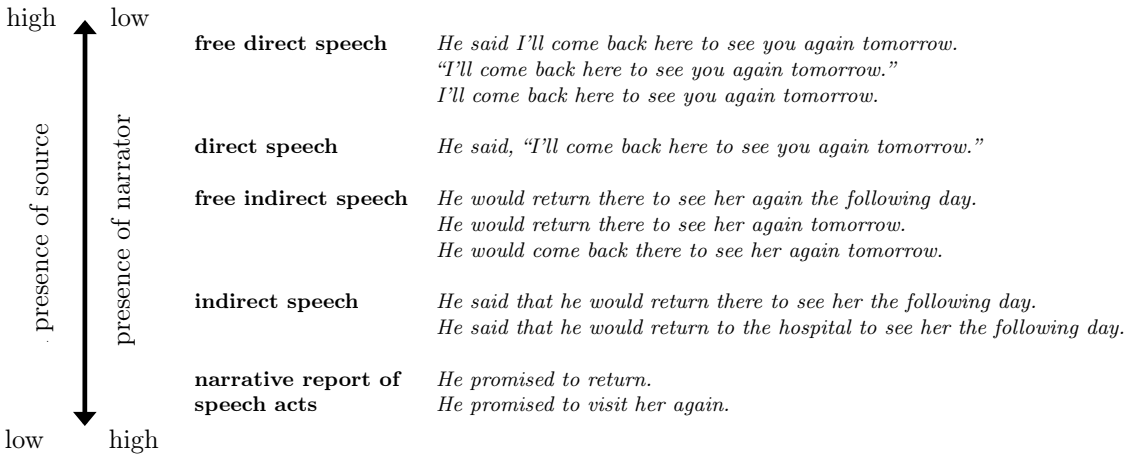


Figure 7.1: Overview of categories of speech representation with examples in Leech and Short (1981, 318-351)

These five categories build a cline with respect to the degree of "interference" in the report through the narrator (Leech and Short 1981, 324), with least interference in the case of free direct speech and most interference for narrative reports of speech acts. In both forms of indirect speech, there is more interference from the narrator than in forms of direct speech, since the content of what was said is expressed from the point of view of the narrator, leading to a change in deictic terms. Both forms of direct speech, in contrast, imply that they represent what the source said in his/her original words. Interference is least in the case of free direct speech, a variant of direct speech in which the quotation

Author submitted manuscript. Consult published version for correct page numbers!

marks or the reporting clause or both are missing. Quotation marks and reporting clauses are "features which show evidence of the narrator's presence" (Leech and Short 1981, 322) and their removal, thus, reduces the presence of the narrator in the text. Narrative reports of speech acts are at the other end of the cline. In this form the narrator reports a speech act without giving more details concerning the content or wording of what was said (Leech and Short 1981, 323). Here the narrator is more present than in any of the other forms of speech representations and the source of the reported speech is least present.

For the purpose of my own analysis, the narrator's presence can be seen as being inversely proportional to the presence of the source whose speech is reported. Thus, the reported source is most present in (free) direct speech and least present in narrative reports of speech acts. Since I am interested in the personalisation of news actors in this chapter, I will, therefore, only focus on the use of direct speech in my data.

Not all of these categories are equally common and their use varies considerably across text types and genres. Free indirect speech, for instance, while rare in many texts, is a common device in modernist fictional writing (Abrams and Greenblatt 2000, 1906). In present-day newspapers, by contrast, it can hardly be found, in particular not in news articles. Moreover, free direct speech, which was a stylistic feature of news reporting until the early 20th century, is not used outside of headlines in present-day news articles (Bös 2010b, 83). Apart from headlines, the only free form that is regularly found in present-day news articles is direct speech with missing reporting clause (see also Slembrouck 1986, 87-90). Even this only occurs in cases of continued speech representation, when the reporting clause is given for the first part of direct speech, but missing for later parts. Direct speech without quotation marks is very rare. Garretson and Ädel for instance state that in their corpus of US newspaper data some cases of direct speech without quotation marks can be found, but that they are "infrequent enough to be negligible" (2008, 168).

Thus, direct speech with missing reporting clause is the only free form of speech representation that regularly occurs outside of headlines in news articles. Since in this case the difference between free and non-free forms is not important for my argument, I will not apply this terminological distinction in what follows. Instead, I will only distinguish between direct speech, indirect speech and narrative reports of speech acts.

7.1.2 Faithfulness claims

The three basic types of speech representation, direct speech, indirect speech and narrative report of speech act, differ with respect to their faithfulness claims. A schematic overview of the different faithfulness claims is given by Jucker (2006, 109) (see Table 7.1). According to this, all of the different forms claim to represent the speech act value accurately. For narrative representations of speech acts this is the only claim being made, whereas direct speech and indirect speech also claim to faithfully represent the propositional content of the anterior speech event. Faithfulness with respect to the actual wording of the anterior speech event is only claimed by direct speech.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Table 7.1: Faithfulness claims in different forms of speech representation according to Jucker (2006, 109)

Form of reported speech	*Speech act value*	*Propositional content*	*Words and structure*
Direct speech	Yes	Yes	Yes
Indirect speech	Yes	Yes	No
Narrative representation of speech act	Yes	No	No

Author submitted manuscript

The fact that speech representations claim to be faithful does not mean that they are. The question of the veracity of speech reporting has received attention by researchers, especially in contexts like news reporting, where misrepresentation of statements can have political and legal consequences. It has been demonstrated that direct speech in news reporting sometimes contains formulations that were unlikely to have been used in the original statement (see, for instance, Caldas-Coulthard 1994). Moreover, it has been argued that even if the verbatim reporting of a statement is accurate, the speech act value of the anterior discourse can change and new implicatures can be generated if it is contextualised differently (Ikeo 2009). In this case, the claim to represent the wording of the original statement is fulfilled, but the claim to represent the speech act value accurately is violated. In the context of spoken conversation and storytelling, Tannen (1986) observes that speech representation is most often not accurate. She finds that speakers in conversation frequently use direct speech as a stylistic device to present a dialogue that has never been spoken in anterior discourse (1986, 314). Even when direct speech represents what has been spoken before, this representation is often not accurate. Tannen argues that due to contextual limitations, which force the speaker to rely on his or her memory of the anterior discourse, direct speech in conversation cannot be expected to represent the words and structures accurately (1986, 314). Therefore, she proposes referring to speech reporting in this context as "constructed dialogue" (1986, 311). This position is shared by Mayes (1990), whose analysis of informal spoken interaction indicates that many instances of direct speech are not faithful representations of the wording of anterior speech. She concludes that, at least for her type of data, "direct quotes fall along a continuum from plausible to impossible" (Mayes 1990, 326).

The regular lack of faithfulness of direct speech in some text types raises the question whether the strong faithfulness claims outlined above actually hold in these cases. If, as Tannen argues, it cannot be expected that speakers in casual conversation remember the exact wording of the anterior discourse, this might lower the expectations of the audience with respect to the accuracy of the words and structure of direct speech. If there is a conventional agreement that direct speech in this context is usually not entirely accurate,

Author submitted manuscript. Consult published version for correct page numbers!

then one could argue that the faithfulness claims that are evoked are less strong. A similar problem is posed by fictional writing, in which there is no anterior discourse that could be quoted accurately. Observations like these have led some researchers to reject the idea of associating specific faithfulness claims with different forms of speech representation (e.g. Collins 2001, 50-58; Vandelanotte 2009, ch. 4; see also Short et al. 2002 for an overview of the arguments against the notion of faithfulness).

Short et al. (2002) address this problem by arguing for a contextualised view of faithfulness claims. Depending on the context in which direct speech is used, the faithfulness claims can be weakened or even cancelled. They list eight factors which they suggest affect the salience of faithfulness claims of direct speech. They see this not as an exhaustive list, but call for further research to identify additional factors. The eight factors are (Short et al. 2002, 349-353):

1. **Anterior discourse accessibility:** The more accessible the anterior discourse, the more salient the faithfulness claim. From this follows, for instance, that quoting written sources usually evokes stronger faithfulness claims than quoting spoken sources, since the latter are generally less accessible.

2. **Posterior discourse accessibility:** If a text is predictably inacessible, e.g. due to not being persistent, then the faithfulness claims of speech representations in this text are less strong. Consequently, speech representations in informal conversations evoke weaker faithfulness claims than in a recorded interview.

3. **Importance of (wording of) what is being reported:** If the content and wording of what is being reported is particularly important, then the faithfulness claims that are evoked are particularly strong. Short et al. mention statements by heads of state as one example where the exact wording can be of great importance, and where therefore the expectations with respect to the faithfulness of quotes of such statements are particularly strong.

4. **Memorability of the original:** For particularly memorable utterances the faithfulness claims are also particularly strong, e.g. in the case of famous quotes.

5. **Status, social role and personality of the producer of the original discourse:** Short et al. name high social status and persons whose feelings are easily hurt as two examples of sources which are usually quoted with a high degree of faithfulness.

6. **Social role, personality and attitude of the reporter:** How strong the faithfulness claims are can also be influenced by the role of the reporter and his or her relation to the quoted source.

7. **Text-type or speech context:** As already mentioned above, the text type plays an important role in deciding how strong the faithfulness claims of speech representations are. In informal conversations, they tend to be less strong than in public speeches, while in legal texts they are stronger than in personal letters.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

8. **Part of text in which reporting occurs:** Short et al. suggest that some elements of texts can affect faithfulness claims of speech representations that occur in them. They mention as an example newspaper headlines, which sometimes contain direct quotes in which the wording of the original utterance is altered.

I would like to suggest that not all of these factors can work equally well in both directions, leading either to a weakening or a reinforcement of faithfulness claims. If one assumes that speech representations by default evoke the faithfulness claims listed in Table 7.1, and that these faithfulness claims can be weakened or cancelled, then factors that can result in weakening and cancellation are particularly interesting. From the above list, this applies in particular to the accessibility of anterior and posterior discourse (factors 1 and 2), and the text type and part of text in which the speech representation occurs (factors 7 and 8). The importance of the original and its wording (factor 3) will typically work towards a reinforcement of faithfulness claims in cases of particularly important statements, but can also work in the direction of weakening and cancellation. In contrast, it is not conceivable how the (low) memorability of the original could lead to a cancellation of faithfulness claims. For the social roles and personalities of source and reporter (factors 5 and 6), reinforcement would also be the typical effect, but the opposite is possible.

In addition, it is noteworthy that faithfulness claims can never be stronger than described in Table 7.1. Even when several reinforcing factors apply, indirect speech will never evoke claims to represent the actual words and structures of the original sentence, and narrative representations of speech acts will also not evoke claims about the propositional content of the original. Thus, reinforcing factors only become effective when they counter weakening factors.

For my own data, one case in which speech representation is affected by factors that weaken faithfulness claims is when direct speech appears in headlines (factor 8). In this case, the claim that direct speech represents the actual wording is not fully effective (see also Short 1988, 67-69). Due to the specific functions and constraints of news headlines, which need to maximise the eye-catching effect while at the same time being very brief, parts of quotes can appear slightly altered. An example of this can be found in the headline of the *Times Online* article reporting on Obama's speech which was introduced at the beginning of this chapter.

(7.4) Obama: Our security failed and the buck stops with me (times-wn-100108)

The reporting clause is here reduced to an indication of the speaker and the reporting verb and quotation marks are missing. Still, the personal pronouns *Our* and *me* suggest that the headline represents what Obama said in (free) direct speech, which would create the claim that the words and the structure are represented faithfully. In this case, however, the first part of the quote, *Our security failed*, had not been included in these words in Obama's speech. Instead, the headline represents a much more condensed form of some of his statements. Such modifications of direct quotes in headlines do not appear to be a rare exception in my data. It seems reasonable to assume that the audience, who are familiar

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

with the functions of headlines, will not raise the same expectations about the faithful rendering of words and structures of direct quotes in headlines as when direct quotes occur in the body of a news article.

An interesting point is raised by Short et al. (2002, 352), who "suspect" that down-market newspapers are less careful about accurate quoting than up-market newspapers. They see this as an instance of factor 7, which means that the type of newspaper in which direct speech occurs can have a weakening influence on the faithfulness claims that are evoked. A similar observation is made by Burger (2005) about German and Swiss down-market newspapers. He goes as far as to suggest that there is probably only one case in which it can be assumed that direct speech in these publications is quoted verbatim, namely when they quote other media (Burger 2005, 100).

7.1.3 Modifications to Leech and Short's classification

Leech and Short's classification of speech representation was originally developed for the stylistic analysis of fictional writing. It has since been applied in very different domains, such as for the study of praise quotations on paperbacks (Ikeo 2009) and Early Modern English courtroom discourse (Włodarczyk 2007). In particular, the model has been used for the study of newspaper data from various sources, both present-day and historical, British English, as well as American English and German (Bös 2010b,a; Brüngel-Dittrich 2006; Garretson and Ädel 2008; Jucker 2006; Short 1988; Short et al. 1999, 2002; Slembrouck 1986). These applications have resulted in a range of modifications and extensions, some of which will be discussed in the following.

One factor that complicates the classification above is that in news reporting indirect and direct speech are often combined within a single quotation. The most common form of this is what Garretson and Ädel call "partial direct reported speech", which is also known as "slipping" (Collins 2001, 13-14; Fairclough 1988, 126; Page 1988, 36-37; Schuelke 1958; Slembrouck 1986, 54), "mixed quotation" (Cappelen and Lepore 1997) and "embedded quotation" (Short et al. 1999, 43). This refers to indirect speech which incorporates segments of direct speech. The direct speech element can be very brief, as in examples 7.5 and 7.6. Example 7.7 shows an indirect quote with three noun phrases in direct speech. In example 7.8, finally, almost the entire quotation is given in direct speech.[3]

(7.5) To the amazement of friends and colleagues he declared that a life in Westminster was "incompatible" with being a father. (mail-sn-100110)

(7.6) Experts said that Day would have been a "world famous Rambo character" if he had earned the medals in battles. (mail-uk-100112)

(7.7) The Supreme Court had criticised the growth of "anonymity orders" granted by courts in recent years, which had become an "ingrained habit" among court officials without "the slightest justification". (sun-100130a)

[3] Burger (2005, 92) further distinguishes between "slipping" and "partial quotation" ("Teil-Zitat"). He uses "slipping" to refer to cases like 7.8, where a single change from indirect to direct speech (or vice versa) occurs. "Partial quotation" is used to refer to short, syntactically dependent expressions in direct speech which occur in indirect speech or in narrative passages, as in 7.5 to 7.7.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

(7.8) Hamas, which Israel has blockaded in Gaza for three years and which it tried to topple in a brief war last year, said that it would "retaliate for this Zionist crime at the appropriate time and place". (times-wn-100130)

This last example is particularly interesting from the point of view of personalisation. The indirect part *that it would* here affects only the subject of the quoted sentence, so that what would presumably have been *we* in the source statement becomes *it* in indirect quotation. Thus, the first person plural pronoun *we* – which would have contributed to a personalised representation of Hamas and which can also be used with vague meaning – was avoided in favour of the impersonal third person pronoun *it* by representing the subject in indirect speech (see also chapter 8.4). At the same time representing the main part of the statement in direct speech clearly distances the journalist (and the newspaper) from the content of the statement. Thus, this example illustrates well how combining direct and indirect speech can be employed to create or avoid personalised representations of news actors and how this can express stance-like qualities.

A special case of embedded quotation is so-called "nested speech" (Garretson and Ädel 2008, 177-179). Nested speech occurs when the quoted material integrates further quotes. Garretson and Ädel (2008, 166) found that more than 5 percent of all the tokens of speech reporting in their data from US newspapers from 2004 contained nested speech. Nested speech can be found with direct and indirect speech in both positions. Thus, direct quotes can include direct speech, such as *Thank you* in example 7.9, or indirect speech as in example 7.10. Similarly, indirect speech can incorporate reported statements in indirect speech, as in example 7.11, or in direct speech, as in example 7.12.

(7.9) Carmen Michalska, a rescuer with the Greek team, said: "He was holding the light to help us saw. He just said 'Thank you' when we pulled him out." (guardian-wn-100124)

(7.10) He said: "One of my friends has a shop there. He told me two men entered - insurgents, yes - and were yelling at people to get out of the building. He said he left his shop open and ran away. Police were coming in as he ran out." (bbc-wn-100118)

(7.11) Foreign consultants inside the ministry building said that after initial blasts and gunfire shortly before 10am (6am GMT) they were moved into a different building and told to stay in the corridors. [...] They were told that shots outside the building were from a gunfight between Afghan security forces and the militants. (guardian-wn-100118)

(7.12) She said that Mr Obama had called her to commiserate, telling her: "You can't win them all." (times-wn-100120)

Garretson and Ädel (2008) point out that such nested speech, which is not uncommon in news reporting, poses particular challenges for journalists, who need to judge the risk of misrepresentation of the nested quote. For my own analysis of the role of direct speech for personalisation, nested speech does not pose a serious problem. Direct speech, whether integrated into quotes or not, can always fulfil the same range of functions with respect to personalisation. However, in the case of nested direct speech, such as in examples 7.9 and 7.12, the personalising effect of direct speech does not affect the news actor presented

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

in the article, but a fourth party to whom the news actor refers. Nevertheless, this can also contribute to the overall personalisation of an article.

One of the most substantial modifications to Leech and Short's classification has been proposed by Short et al. (1999, 2002). They observe that thought and writing can also be reported in forms parallel to those for speech representation. Short et al. therefore develop a classification scheme that consistently distinguishes between speech, thought and writing representation (Short et al. 1999, 2002). They argue that thought and writing representation, while often subsumed under the same heading, differ from speech representation in important ways, for instance with respect to the accessibility of the anterior discourse, i.e. the original speech, thought or writing that is quoted (Short et al. 2002, 349-350). While the anterior discourse typically remains accessible, at least in principle, in the case of writing representation, this is less frequently the case for speech representation and it is outright impossible for thought representation. This has consequences for the faithfulness claims that are made in each case, and Short et al. point out that a lack of differentiation between these forms "sometimes results in confusion" in the literature on faithfulness claims of speech representation (Short et al. 2002, 342).

While it is true that different conditions apply in speech, thought, and writing representation, which calls for an analytic distinction between the three types, it is also the case that this distinction poses a number of problems for practical application. At least for data from news sites, it is often not clear at all whether a specific instance is a case of speech or writing representation. Usually, the only indications that can be used to determine whether the original statement was produced in spoken or in written form are the context and the speech act verb. However, the speech act verb is a very unreliable indicator. As Short et al. themselves state, the verb *say* is often used to introduce reported writing (Short et al. 2002, 335). Similarly, *it was thought* is a common way to report statements without naming the source. An example of this can be found in a news report on two British MPs' attempt to challenge the position of Prime Minister Gordon Brown.

(7.13) All eyes were on the Cabinet as the news emerged. It was thought that if one or two supported the call, then Mr Brown's position could be in jeopardy. (times-uk-100106)

In this example, *It was thought* is used to introduce the prevalent opinion on the issue. In order for the journalist to be able to access such a general opinion it must have been stated, either in spoken or in written form. Consequently, this is a case of speech (or writing) representation. To consider this an instance of thought representation is taking the meaning of the reporting verb too literally.

An additional problem in news reporting is posed by press conferences, where written statements are often read out. This raises the question whether the reporting of these statements should be considered speech or writing representations, even more so in cases in which the manuscript that is read out is made accessible to journalists or the general public. For instance, Obama's press statement that I introduced at the beginning of this chapter was accessible on the website of the White House in the form of an electronic

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

manuscript and as a video file. In this case, it is impossible to determine whether the author of a news article attended the speech, whether he or she saw the video recording or read the manuscript, and from which of these sources he or she quoted.

Finally, it should also be noted that spoken language that builds the basis for news reporting is usually recorded, unlike informal everyday conversations. This means that even for clear cases of speech representation in news reports, the expectations about the accuracy of the representation are not fundamentally different from those in writing representations. In other words, readers expect journalists to represent direct speech accurately, no matter whether the anterior discourse was spoken or written.

It can thus be concluded that the distinction between speech, thought, and writing representation is valuable in pointing out that not everything that is referred to as speech representation has actually been spoken in the anterior context. This distinction may also be usefully applied where the modality of the anterior context is not in question, for example in fictional writing. For present-day news data, however, the categorisation does not seem to be applicable in a meaningful way. Therefore, in the following, I will not distinguish between speech, thought and writing representation, but refer to all three types as speech representation.

Author submitted manuscript

7.2 Functions of direct speech

Journalists do not always present their sources' statements in the form of direct speech. They make use of all the available forms of speech representation in order to achieve different effects in their texts. Bell (1991, 209) even argues that direct speech is used only as an exception in news media. Normally, journalists present the information they receive from their sources in more indirect forms. I will later investigate to what extent this is true in my data, and whether there are indications that this might have changed over the last twenty-five years. Before looking at these quantitative findings, I want to explore the various functions that direct speech fulfils in news texts and assess to what extent and in which respects these contribute to the personalisation of news. Previous research has investigated the functions of direct speech in a large range of different contexts. In the following, I will restrict the discussion to functions that are directly relevant for news articles.

A classic, but still relevant, account of the functions of quotes is provided in Bell (1991, 207-209), who argues that direct speech is used in news media to fulfil three main purposes. The first of these is related to the role of quotes as reportable facts. Journalists are expected to report facts which they are able to prove, if need be, in court. They need to be able to provide evidence that what they report is actually true. If a statement is recorded (in spoken or written form), then presenting parts of it in the form of direct speech is a convenient way of fulfiling these requirements. In Bell's words, quotes in the form of direct speech function as "particularly incontrovertible fact[s]" in journalistic reporting (1991, 207). This function is especially salient on online sites, which allow text producers to

Author submitted manuscript. Consult published version for correct page numbers!

present the recording of the original statement together with the reported speech. There are two main ways in which this can be done. On the one hand, a video or audio recording of the statement can be presented together with the article in which the direct speech occurs. On the other hand, it is possible to link to other online sites which provide records (in written form or as video/audio recordings) of the anterior speech. In both cases, the audience is able to check the reported speech against the original statement. Even though most readers will probably not follow links to external sites to check the veracity and accuracy of the speech representation, the mere possibility to do so will serve to emphasise the factual status of direct speech.

The second function Bell names is "to distance and disown" (1991, 208). Direct speech is used by journalists to mark the content of the quote explicitly as not being their own opinion, nor that of their news media organisation. Positions of extremist interest groups are regularly reported in this way, which was already illustrated with example (7.8) above, reporting a statement by Hamas. Similarly, in the following example, all the extremist statements by a suicide bomber were represented in direct speech.

(7.14) In an interview on September 26 last year al-Balawi said he had "been moulded on the love for jihad since my childhood". He vowed to "take up arms, and to wear an explosive belt, to avenge the killing of children and women in the Gaza War".
He also said that he decided to leave his pro-jihadist writings in favour of "real jihad on the ground, because I came to realise that preaching about jihad is not enough ... You have to carry out jihad in practice." (times-wn-100106)

As in example (7.8) above, there is repeated switching between direct and indirect representations. The pattern is the same as in the previous example too. The direct speech sections contain the main parts of al-Balawi's statement. The use of direct speech marks the content as the personal opinion of al-Balawi and makes it clear that this position is not shared by the newspaper. On the one hand, this presents al-Balawi in a more personalised way; on the other hand, it emphasises the aim of the journalist to adopt an objective stance. The indirect reporting includes the subject of the sentence (e.g. *he had, he vowed to*), which means that, except for one instance, al-Balawi is always referred to with the third person pronoun *he*, rather than with *I*. The exclusion of the subject from the direct quote can be seen as a way of de-personalising al-Balawi and of presenting him in a less immediate way, and, thus, it limits the personalising effect of the direct speech account (see also chapter 8.4).

The distancing function of direct speech is most clearly visible in the case of so-called scare quotes, which "alert readers that a term is used in a nonstandard (or slang), ironic, or other special sense" (Chicago Manual of Style, 365). This illustrates that direct speech can not only be used to distance the author from the content of the quote, but also from its style. Colloquialisms, technical expressions or debatable word choices are often put into quotation marks.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

(7.15) The President's sensational intervention could mean that the door is slammed on the "fat cat" banking which brought the world to the precipice. (mail-cn-100122)

(7.16) Pregnant with my little brother Ben, she was suffering with "placenta previa" and would have to spend the next two months in bed. (mail-sn-100104)

(7.17) With Hypnodiet, you'll "reprogramme" your subconscious mind. (times-sn-100104)

(7.18) Even if financial inducements did persuade some couples to marry who currently live "in sin" (as we used to say), or live apart, I worry about the potential undesirable consequences for the status of marriage. (mail-cn-100110)

As can be seen from these examples, scare quotes are not always direct speech in the strict sense. They do not always report a statement that was made previously (for a more detailed discussion, see, for instance, Brendel et al. 2011a; Gutzmann and Stei 2011). Therefore, it could be argued that such instances should be excluded from the discussion of direct speech. However, scare quotes fulfil very similar functions to direct speech in mixed quotes and to differentiate between the two proves to be difficult (Cappelen and Lepore 1997, 430; Garretson and Ädel 2008, 168). Based on the available context, it is often not possible to decide whether a short quote is repeating anterior discourse or not. Short et al. (1999, 44) discuss an example which at first sight appears as a scare quote without anterior discourse, but which turns out to be a quote when more context is taken into account. Similar examples can also be found in my data. In 7.19, for instance, "toxic home life" could easily be interpreted as a scare quote, marking an unconventional word choice by the journalist. However, when reading reports on the court hearing on other news sites, such as in 7.20, it becomes clear that the expression was used by a lawyer in court and that its use in 7.19 is a case of direct speech.

(7.19) The boys also watched ultra-violent movies, including Saw, and porn DVDs as part of their "toxic home life". (sun-100124b)

(7.20) Peter Kelson QC, who is representing the older brother, said: "These were gruesome movies in the extreme – that is an influence in this toxic home life." (mail-uk-100122)

In addition, examples like 7.15–7.17 above can be interpreted as quoting an undefined or hypothetical source. In spoken language, when no quotation marks are available, scare quotes are often marked by the phrase "so called" (Nacey 2012: 118-119; see also Bell 1991, 208), which implies that the expression is used by undefined sources.

The third function of direct speech, according to Bell (1991, 208-209), is to "add [...] the flavour of the newsmaker's own words" to a text. Typical examples of this function are found in short quotes that contain concise and colourful formulations. The focus here lies on the language that is used, not on the content. Often what is presented in direct speech does not add any new information, but presents an alternative and original way of phrasing.

These three functions of direct speech are not always present to the same extent. Depending on topic, context and the formulation of a statement, one function may be

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

particularly emphasised. The source of a statement seems to be an important factor in this. In example 7.14, for instance, the function of distancing and disowning is most salient, which is typical of quotes by interest groups and individuals expressing extremist positions. For professional news actors like politicians, in contrast, the function of providing reportable facts seems to be of special importance. On the one hand, this is a group of sources for whom it is particularly valuable and necessary to ensure that what is reported is correct, because consequences in the case of incorrect or imprecise reporting are very likely. On the other hand, the more important a news actor, the higher the news value of his or her statement. In other words, some statements are "reportable facts" only when they are uttered by news actors like spokespersons and politicians in their professional capacity.

Although one function of direct speech can be particularly salient in a specific quote, it is also true that the functions are not mutually exclusive. In fact, they very often co-occur. In the example that I presented at the beginning of this chapter, Barack Obama's statement "the buck stops with me" was represented as direct speech in various news media. As I have argued above, the use of direct speech stresses that this was, indeed, Obama's statement. The factual status of the statement becomes even more incontrovertible in those reports which integrate video sequences showing Obama making the statement, or when the link to the entire speech on the website of the White House is provided. The quote also serves to make the report more vivid and colourful by presenting Obama's own words, which in this case include the use of a colloquial expression and intertextual references. At the same time, the use of quotation marks also signals that this is not a formulation by the journalist, thus distancing the newspaper from the informal choice of expression, which would not be appropriate for the reporting style of the up- and mid-market news sites in which the report appeared. This illustrates well how the different functions of direct speech are often present simultaneously, complementing each other.

As the examples above show, all of Bell's functions can be found in my data. However, there is also an additional function of direct speech in my data which is not explicitly named by Bell (1991), but which is mentioned by other researchers (Burger 2005, 97-101; Stenvall 2008). This concerns the use of direct speech for reporting emotions. Stenvall (2008, 1570) points out that reporting emotions challenges the journalistic ideals of objectivity and factuality of reporting. Direct speech provides a way to report emotions without violating these ideals, since the emotions are clearly marked as those of the news actor. As Stenvall puts it:

> [T]he 'objective' style of [hard news] reports presupposes that the journalist's voice is backgrounded, which means that authorial emotions ought to be totally excluded. Strictly speaking, the only way to report *other* people's individual feelings 'objectively' is to resort to direct quotes, that is, to let the person in question describe how s/he feels. (Stenvall 2008, 1572, emphasis in the original)

The same point is also made by Tuomarla (2000, 70-72), who argues that direct speech in French newspapers is used to dramatise by adding content that would violate the norm

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

of objectivity if it were not marked as direct quote. Thus, reporting emotions through direct speech relates to its function to distance and disown what is said. Still, this use of direct speech is frequent and specific enough to distinguish it from the more general distancing function that can be found, for instance, in the reporting of extremist positions. Direct speech that is used to report emotions typically does not present any new factual information or opinions of the news actors. It mostly occurs after the main events have been stated and lets the sources – often witnesses, victims, or parties who have been directly affected by a news event in other ways – tell and relive their emotions in retrospect. This use of quotes is not restricted to newspapers, as Montgomery (2010) shows in his study of the same feature in broadcast news. Example 7.21 is a typical instance of this use in my data.

(7.21) A pregnant woman today told how she trudged through snow in just her pyjamas to reach hospital to give birth to twins after her car became stuck in blizzard conditions. [...] "I was so scared, I was really panicking. I think I was in shock," she said. "I just wanted to get there for the twins." (mail-uk-100108)

This passage appears at the end of a news report on heavy snowfall in Great Britain, accompanied by a photo of the parents with one of their newborns. Five sentences, which were omitted in the example above, give additional details about the journey and the twins before the emotions of the mother are reported. It seems clear that such retellings of emotions are not included in news reports for their informational value. Schneider (2002, 155-157) argues that "first person quotations", as in example 7.21, are used to dramatise the reporting and to personalise those who are affected by a news event. In addition, Luginbühl (2012, 258-259) points out that quoting directly affected sources in this way creates immediacy and emotional involvement.[4]

This use of direct speech has many parallels with what Tannen (1986) calls "constructed dialogue" in spoken interaction. This refers to the retelling of personal experiences in the form of reported dialogue, which – despite being presented as direct speech – was often not actually spoken in those words before. According to Mayes (1990, 348), the ability of direct speech to dramatise narration stems from the fact that it "allow[s] the speaker to tell about [an event] as if it were presently occurring". A similar point is made by Bublitz and Hoffmann (2011, 441-442), who argue that direct speech increases narrative immediacy by allowing authors to "propel their audience into the ongoing action of their verbal accounts". Bös calls constructed dialogues in news reporting "dialogic first-person quotations" (Bös 2010a,b). Her findings, based on the Rostock Newspaper Corpus, show that the use of dialogic first-person quotation was most frequent in the 19th century and has constantly decreased since the beginning of the 20th century (Bös 2010a, 230-231). Even though it can no longer be found very frequently in newspapers at the end of the 20th century, Bös

[4] Direct speech has also in other contexts been identified as a feature for creating involvement, such as in courtroom discourse (Collins 2001, 68-69; Włodarczyk 2007, 156) and in classroom interaction (Baynham 1996). According to Collins (2001, 68-69), this is largely due to the fact that direct speech demands more interpretative work on the part of the audience, who need to adopt a more active role in evaluating the text.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

(2010a, 239) argues that its "dramatizing function" still plays an important role in soft news. In my data, constructed dialogue can be found in news reports on the Edlington court hearing, in particular in one article from the *Sun*. In this example, the interaction between the two victims and later between one victim and the villagers is presented in the form of direct speech (see also chapter 6.4).

> (7.22) After the evil brothers fled the scene, the younger boy knelt next to the older lad and asked him if he was OK.
> He replied: "No. I can't see and I can't move my body."
> He went on: "You go and I'll just die here."
> The younger lad raised the alarm when he was found wandering, dazed and covered in blood. He told shocked villagers: "I don't know where my friend is." (sun-100122a)

Throughout the report, direct speech occurs at the most dramatic moments and clearly fulfils dramatic effects. It does not become clear from the text what the source of these quotes is. One might suspect that the victims were asked at the hearing to repeat what had happened and that this involved repeating what they had said at the time of the event. However, according to a news report in the *Mail Online*, the families of the victims accepted a deal which meant that their sons did not have to testify in court (mail-uk-100124). The dialogue might have been reported by another witness, who had heard it from one of the victims. It is also unclear whether the news report repeats a constructed dialogue that was presented during the hearing, or whether individual statements that were made during the hearing were arranged in this way by the reporter. Thus, it is a typical case of constructed dialogue in which the factual status and veracity of what is reported is far less important than its dramatic effect. Tannen argues that:

> constructed dialogue in conversation and in fiction is a means by which experience surpasses story to become drama. Moreover, the creation of drama from personal experience and hearsay is made possible by and simultaneously creates interpersonal involvement among speaker or writer and audience. (Tannen 1986, 312)

The reporting of emotions as in example 7.21 above is in a very similar way used for the "creation of drama from personal experience" with the aim of involving the audience. By shifting the focus from factual information to emotions and personal experience, it, thus, fulfils an important function for the personalisation of news.

In sum, it is possible to identify four main functions of direct speech in my data: presenting a statement as a reportable fact, distancing the journalist and the newspaper from the content of a statement, colouring the text with the flavour of the news actor's own words, and reporting emotions. These functions often co-occur and cannot always be easily separated from each other. I have argued that the first of these four functions takes on a new dimension online, where (links to) the source's statement can be presented together with the direct speech account. The other three functions all directly contribute to a more personalised presentation of news texts. In the case of direct speech serving to distance the journalist from the content of a statement, this content is presented as

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

the personal opinion of the source, whose position is not generally shared. When direct speech is used to bring colour into a text, the style of the text is marked as the personal style of the source, which is not appropriate for the objective and distanced reporting style of the news media. Finally, direct speech that reports emotions focuses on the personal experience of individual news actors, rather than on impersonal facts.

7.3 Frequency of direct speech

7.3.1 Quantifying direct speech

In order to compare the frequency of use of direct speech across different news sites, article types, or other dimensions, one first needs to define the unit of comparison. What unit is suitable for measuring direct speech? This question is less straightforward to answer than it might seem. The following example contains stretches of direct speech from a single source, introduced with two reporting clauses, spread over three paragraphs and marked by four pairs of quotation marks. Depending on which of these criteria is used, one could (more or less convincingly) argue for counting this example as containing one, two, three or four direct quotes.

(7.23) Foreign Ministry spokesman Ma Zhaoxu said the US should "respect the facts" and stop making "groundless accusations against China".
"The US has criticised China's policies to administer the internet, and insinuated that China restricts internet freedom," he said in a statement posted on the foreign ministry website.
"This runs contrary to the facts and is harmful to China-US relations."
(bbc-wn-100122)

In previous research, different forms of speech representation have usually been measured and compared to each other in "instances" of use. Short et al. (1999, 2002), for example, manually tagged a corpus containing texts from fiction, newspapers, and biographies. They classified each stretch of words according to 20 different types of speech, thought or writing representation and narration (plus some "portmanteau" tags for ambiguous cases) (Short et al. 2002, 238-239). They then analysed the frequency with which different tags had been assigned. They compared, for instance, the frequency of use of direct speech, indirect speech and narrative representation of speech acts (Short et al. 1999, 48), and the frequency of use of direct speech across different sections of their corpus (Short et al. 2002, 342). Similarly, Garretson and Ädel (2008, 169) use the number of tags as a basis for comparison to study the use of direct speech and indirect speech across 11 US newspapers. As far as can be deduced from the methodological descriptions provided in the respective studies, the number of instances of direct speech in Short et al. (1999, 2002) and Garretson and Ädel (2008) corresponds more or less to the number of pairs of quotation marks.

However, a disadvantage of counting pairs of quotation marks is that this measure is very dependent on conventions of typography and layout. If a direct quote continues over a paragraph break, the second paragraph usually starts with opening quotation marks, even if there were no closing quotation marks before the paragraph break.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

(7.24) In a podcast from No10, the Prime Minister said: "Because of the cold snap we are also facing record levels of demand for gas.
"National Grid has confirmed that it expects supplies to meet demand. I can assure you: supplies are not running out." (mail-uk-100110)

Taking the number of pairs of quotation marks as a measure, one would need to decide whether to count this as a single instance or as two instances. This is further complicated by the fact that this convention is not followed in exactly the same way across different publications, nor is it applied entirely consistently within each publication. In addition, the length of paragraphs also seems to vary across the different news sites. This means that news sites with shorter paragraphs would end up having more instances of direct quotes if cases like 7.24 were to be counted as two instances.

A further disadvantage of this measure is that the length of the quote is not taken into account. This means that single word quotes, e.g. in scare quotes, count as much as direct speech that spans several sentences. Mixed quotes, consisting of indirect speech with individual expressions in direct speech, even end up counting more than a longer statement in direct speech.

(7.25) In late 2002, he [Goldsmith] was expressing concern about "optimism" in the government that the UN security council had provided legal cover for an invasion. He warned Jack Straw, the foreign secretary, not to assume that it would be "all right on the night". But on his return from Washington he now took the view, he said, that while a fresh UN decision would have been a "safer" course, a "reasonable case" could be made for war without going back to the UN. (guardian-uk-100128)

(7.26) In a statement to the inquiry, released ahead of her appearance this afternoon, the former civil servant [Wilmshurst] said the invasion was not only illegal but would damage the reputation of the UK as a law-abiding nation. "Collective security, as opposed to unilateral military action, is a central purpose of the charter of the United Nations. Acting contrary to the charter, as I perceived the government to be doing, would have the consequence of damaging the United Kingdom's reputation as a state committed to the rule of law in international relations and to the United Nations." (guardian-uk-100126)

Example 7.25 is a mixed quote, containing four short segments of direct speech which include a total of nine words. The longest segment, *all right on the night*, includes five words, whereas the others are only one or two words long. Example 7.26, in contrast, contains only a single direct quote which is 59 words long. If the number of pairs of quotation marks were to be taken as a measure for quantifying direct speech, example 7.26 would count as only one instance, whereas example 7.25 would count four times as much. This does not correspond to the intuition that the long quote in 7.26 contains more direct speech than the short quotes in example 7.25.

To avoid these problems, I use the percentage of words that appear within quotation marks as a measure of comparison. This measure is robust with respect to typographic conventions. Whether or not longer stretches of direct speech are segmented by several pairs of quotation marks does not affect the frequencies I calculate. It has the additional

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

advantage that scare quotes, which have a questionable status in the discussion of direct speech, do not contribute much to my frequencies, since they usually contain no more than one or two words. The number of words has also been used as a secondary measure of comparison in Short et al. (1999, 48). Since the calculation of this figure poses fewer methodological problems, it also provides a better basis for comparing my own results with theirs.

7.3.2 Direct speech across subcorpora

Table 7.2 shows the percentage of words that occur within direct speech in all news articles in my data. For the online news sites, the percentage of words in direct speech varies between 22.9 percent in the data from *BBC News* and 29.4 percent in the data from the *Guardian*. With 13.9 percent the *Times* from 1985 clearly has a lower frequency than all the online news sites (statistically significant at $p < 0.001$; see appendix B for statistical tests).

What becomes clear from these figures is that direct speech is quite frequent in these data. Thus, Bell's claim that journalists present information from their sources in indirect speech and use direct speech only as an exception does not apply (Bell 1991, 209). Since I have not determined the amount of indirect speech in my data, I do not know whether direct or indirect speech is used more frequently. It seems clear, however, that direct speech cannot be considered an exception in online news articles, if it contributes roughly one fourth of all words. Even the frequency of direct speech in the printed *Times* from 1985 is too high to regard it as an exception.

If it were not for the data from the *Guardian*, it would appear that up-market news sites, with 23.1 percent (*Times Online*) and 22.9 percent (*BBC News*), contain less direct speech than the mid- and down-market news sites with 28.0 percent (*Mail Online*) and 29.0 percent (*Sun*). However, the *Guardian* contains more direct speech than any of the other news sites. This high amount of direct speech in the *Guardian* is partly caused by one article with an extraordinarily long quote. The article reports on two British MPs, Patricia

Author submitted manuscript

Table 7.2: Number of words within direct speech in news articles from five online news sites and the *Times* from 1985

News site	Words in ds	Total words	Words in ds in %
Times Online	5,179	22,443	23.1%
Guardian	6,686	22,751	29.4%
BBC News	4,425	19,323	22.9%
Mail Online	9,648	34,403	28.0%
Sun	4,886	16,869	29.0%
Times 1985	3,220	23,188	13.9%

Author submitted manuscript. Consult published version for correct page numbers!

Hewitt and Geoff Hoon, who sent an e-mail to labour MPs, calling for a secret ballot about Gordon Brown's leadership (guardian-uk-news-100106). This e-mail is reproduced in the article and was classified as direct speech in my data, which means that 50 percent of the words in this article occur in direct speech (905 of 1,806). If this article is excluded from the sample, the percentage of words within direct speech in the news articles from the *Guardian* is reduced to 27.6 percent (5,781 of 20,945). Nevertheless, this is still much closer to the figure for the *Mail Online* than to the figure of the other up-market news sites.

One interpretation with respect to market-orientation can still be made if these findings are compared to previous research. Short et al. (1999, 48) found in their analysis of British newspapers from December 1994 that the percentage of words occurring in (free) direct speech was about 16 percent in broadsheet newspapers and about 27 percent in tabloids. I have not collected any data from mid- and down-market sites from 1985, but their figure for broadsheet newspapers fits quite well with my own findings. If direct speech in the *Times* was at around 14 percent in 1985, and if the data from the *Times Online* can be taken to indicate an increase in direct speech in news articles published by the *Times*, then a slight increase by 1994 could be expected. If this is accepted, it would mean that Short et al.'s data was calculated in a sufficiently similar way to my own data to allow comparisons.

Comparing the findings of Short et al. (1999) to my own results leads to two observations. First, the percentage of words in direct speech is higher in the more recent online data; and second, the rise is much more marked for up-market papers, so that the difference across the different news sites in 2010 is much smaller than in Short et al.'s study. Conclusions from this observation need to be drawn with caution. The data vary both with respect to the time of publication and the mode of publication, and it is also not entirely clear whether differences in the method of calculation may affect the comparability of the findings. Still, the fact that the difference across the sites is so much smaller in my data could be read as an indication that up-market news reporting may have followed the trend towards an increased use of direct speech that was initiated by down-market news reporting. This would correspond with previous findings on changes in British newspapers between 1700 and 2000 (see Schneider 2002, 155-157).

For four of the five online news sites, the figures for UK news and world news can be compared (Table 7.3). The *Sun* is not included here, since the distinction between world news and UK news could not be applied to the articles on this site (see chapter 3.3.1). For all of the four online sites, the UK news articles contained more words in direct speech than the world news articles. This difference is most marked for the *Guardian*, with 23.8 percent of direct speech in world news and 34.4 percent in UK news. However, the figures for the UK news data from the *Guardian* again contain the one article with a particularly high amount of direct speech. Without this article, the percentage of direct speech in UK news would be 31.6 percent (3,232 of 10,213). It can also be observed that the variation across the sites is smaller for the world news articles than for the UK news articles. Moreover, for the

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Table 7.3: Number of words within direct speech in world news and UK news articles from four online news sites and the *Times* from 1985

	World news			UK news		
News site	**In ds**	**Total**	**In %**	**In ds**	**Total**	**In %**
Times Online	2,283	11,872	19.2%	2,896	10,571	27.4%
Guardian	2,549	10,732	23.8%	4,137	12,019	34.4%
BBC News	2,140	9,492	22.5%	2,285	9,831	23.2%
Mail Online	2,445	10,598	23.1%	7,203	23,805	30.3%
Times 1985	1,368	10,172	13.4%	1,852	13,016	14.2%

printed *Times* from 1985 a difference between world news and UK news can be observed, but it is smaller than for most of the online news sites. Statistical testing indicates that the difference between world news and UK news is significant at $p < 0.001$ for three of the sites, namely the *Times Online*, the *Guardian* and the *Mail Online* (see appendix B).

One possible explanation for the higher percentage of direct speech in UK news is that it might be easier to gain access to statements that can be quoted in the case of national topics than for international news. Many of the quotes in the UK news articles come from British news actors such as politicians and official spokespersons. For international topics, the journalists of the British news sites might have fewer opportunities to interview the relevant news actors themselves and might therefore have to rely more on news agencies.

The comparison of different types of articles on the *Times Online* and the *Mail Online* also shows some interesting results (Table 7.4). There does not seem to be a clear tendency with respect to whether soft news articles contain more or less direct speech than news articles. For the *Times Online* the figures are somewhat lower for soft news articles (18.7 percent) than for news articles (23.1 percent), but comparable to the subsection of world news articles (19.2 percent, see Table 7.3). For the *Mail Online*, the percentage of direct speech in soft news articles (28.7 percent) is about the same as that in news articles (28.0 percent). For both news sites, the amount of direct speech in opinion columns is much lower than in news and soft news, and, with 6.2 percent for the *Times Online* and 5.8 percent for the *Mail Online*, the amount of direct speech on the two sites is remarkably similar.[5]

It is perhaps not surprising that columns contain less direct speech than other types of articles. Given that they represent the opinion of the writer, there is less need to substantiate claims by using quotes in their function as "particularly incontrovertible facts". Moreover, authors of columns are not under the same restrictions of using formal and detached language as authors of news articles. While quotes in news articles enable jour-

[5] For statistical tests of these differences, see appendix B. The difference between news and soft news is not significant for the *Mail Online* and the difference between columns on the *Mail Online* and the *Times Online* is also not significant. All the other differences are significant at $p < 0.001$.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Table 7.4: Number of words within direct speech in news articles, soft news articles and columns from two online news sites

	News			Soft news			Columns		
News site	**Ds**	**Total**	**In %**	**Ds**	**Total**	**In %**	**Ds**	**Total**	**In %**
Times Online	5,179	22,443	23.1%	5,421	29,045	18.7%	1,614	26,071	6.2%
Mail Online	9,648	34,403	28.0%	3,652	12,743	28.7%	1,592	27,588	5.8%
Total	14,827	56,846	26.1%	9,073	41,788	21.7%	3,206	53,659	6.0%

nalists to "spice up" the text by including formulations that would be inappropriate outside of direct speech, this is not necessary for columnists. They are able – and may even be expected – to use vivid, subjective and emotional language throughout their texts (see also Sanders 2010, 243).

7.4 Source types

The quotes that are represented as direct speech in news articles come from a variety of different sources. They include news actors for whom the appearance of their quotes in newspapers is part of their profession or official role. Spokespersons are a case in point, but also politicians, experts, and representatives of interest groups, such as charities and trade unions, are regularly quoted in their official capacity. Private individuals are quite a different type of source. They are ordinary citizens who are quoted in the news because they were directly affected by an event as a victim or witness, or because they committed an offence that made the news. As a third group, quotes by celebrities can sometimes be found. In other words, sources fall into the same categories that were established for news actors in chapter 6: official, private, and celebrity.

The source is not indicated by name for all quotes, and, in some cases, direct speech is not attributed to any source at all. From the point of view of personalisation, these distinctions are, of course, relevant. If a source is referred to by name, this is a more personalised presentation than if the source remains anonymous. The presence of private sources in a text is a particularly effective way of personalisation, given that they tend to have a large potential for identification for the audience. Moreover, I have argued above that specific types of sources tend to co-occur with specific functions of quotes: speech representations of an official news actor, such as the Prime Minister, very often serve as reportable facts, whereas private sources, such as victims, very often express personal experiences and emotions. The stronger degree of personalisation for private sources, thus, results from the combined factors that the source is a private individual with whom the audience can identify relatively easily, and that the content of the quote is more personal and emotional. Therefore, it seems worthwhile to take a closer look at the presence of different types of sources in my data.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

7.4.1 Categorisation of source types

Since it was not possible to analyse all the quotes in my data in detail, I focussed on a randomly generated set of 600 quotes from different sections of my corpus: 100 quotes from each of the online news sites and 100 from the data from the *Times* in 1985. Except for the data from the *Sun*, I collected 50 quotes each from the UK news section and the world news section. The statistical software R was used to generate 50 random numbers between 1 and the number of quotes in the respective section, and the quotes at the position of each random number were then selected for the analysis of their sources. Direct speech in headlines differs from direct speech in the main body of the text in several respects, one of which is that the source is often not clearly indicated. Therefore, I decided to include only quotes which appeared in the main body of the text and exclude all cases of headlines, subheadlines and leads for this step of the analysis.

As I have argued in section 7.3.1, counting instances of direct speech is not a straightforward task and, consequently, this procedure for the selection of quotes suffers from some weaknesses. It is possible, for instance, that several segments of a longer quote are selected for the analysis of their source, in which case the same source appears more than once in my evaluation, even though one could argue that the individual segments all belong to a single quote. The method is also not sensitive to the length of quotes: short quotes have the same probability of being selected for analysis as longer quotes, even though they have less presence in the text. This means that if the length of quotes differs systematically for the various source types (e.g. if quotes from official sources tend to be shorter than quotes from private sources) then the source type with shorter quotes will be over-represented in the evaluation. These points need to be kept in mind when interpreting the results. Furthermore, randomly generated numbers are not necessarily distributed evenly. In the case of the quotes from the UK news articles from the *Mail Online*, for instance, 16 out of 50 quotes were selected from a single article (mail-uk-100124). This is an article that is particularly long and which also contains more quotes than the other articles. In addition, the random number generation also over-represented these quotes, compared to the quotes in the other articles. I will discuss the impact of this in my interpretation of the results below.

For the classification of my random quotes, I distinguished two levels: the naming of the source and the type of source. In a first step, I classified whether the source of the quote was indicated by name or whether he or she was anonymous. The source was classified as "named" as soon as his or her name was given anywhere in the article in which the quote in question occurred. This is an important point, since it is very common in journalistic writing to vary the way in which sources are referred to throughout the article. As a consequence, sources are often introduced by name when they first appear, but are later referred to by their function only. Anonymous sources are sometimes more or less explicitly introduced as such, e.g *an employee at the nearby Unity Dearne Club, who did not wish to be named* (mail-uk-100110) or *[a] diplomatic source* (sun-100108a). In other cases, they are

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

collective sources or spokespersons who are referred to only by their function, as in *British Rail said* (times-uk-850109). Impersonal sources are also classified as anonymous, e.g. in the case of scare quotes. The third category, "identified", was used for cases in which the name of the source was not given, but the source was specified in a way that he or she was identifiable beyond the context of the article. There are two different ways in which this can occur. On the one hand, sources can be identified through their function. If only one person fulfils this function, the identification is not ambiguous and the source can therefore be recognised in the real world. This occurred only once in my random quotes, namely in an article from the *Times* from 1985, where a source was referred to as *[the] NUM general secretary*. On the other hand, it is possible that the name of a source cannot be given for legal reasons. This occurs, for instance, in articles on the Edlington court hearing. Since the identities of both offenders and victims are protected, the reports never use their real names, nor those of their parents, when quoting them. Instead, they are referred to as *the younger brother*, *the older victim*, *the mother of the younger victim*, etc.[6] This way of referring to the actors does not allow one to identify them in the real world, but it enables readers to establish the correct co-references across different reports and media. For this reason, the category "identified" was assigned.

In a second step, I classified the type of source. I distinguished four main groups of sources: official actors, celebrities, private actors, and impersonal quotes. The category "undecided" was used for a few cases which proved to be difficult to assign to either of these groups. The classification into these general source types followed the definitions of official, private and celebrity actors provided in chapter 6.1. Thus, official sources are actors who represent the interests of a larger collective and who are quoted as part of their profession or official function. Actors who represent their own interests and who appear in the news because they are personally affected by an event are classified as private sources. Celebrities are persons of interest like music stars and actors, who frequently appear in the news, and who are often quoted on matters of their private lives. In addition to these three main categories, there are a few instances of "impersonal" quotes, for which no relation to any specified source could be established, e.g. in the case of scare quotes and quotes from contracts and statutes.

For official sources and for private sources, several subcategories were further distinguished. The category "government" was assigned to all sources that represented the state (e.g. ministers, police officers, judges), to politicians (from all officially recognised political parties, also of the opposition) as well as to representatives of the UN and its organisations (e.g. the WHO). Not included were sources who were explicitly presented as spokespersons. These were assigned to the category "spokespersons", together with cases in which a government organisation was presented as the source, such as in *the Iranian Foreign Ministry claimed* (times-wn-100112). The "spokespersons" category only contained spokespersons of government organisations. Spokespersons of private organisations, as well as other in-

[6] In one article, mail-uk-100124, the victims are referred to as *James* and *Michael*. Even though it is not stated explicitly, it can be assumed that these are not the real names of the two boys.

Author submitted manuscript.

Author submitted manuscript. Consult published version for correct page numbers!

dividuals that appear in their role as a member of a private organisation, are assigned to the category "interest group". This includes companies, non-government organisations, unions, sports teams, but also political interest groups and extremist groups like Hamas and Al Qaeda. The category "media" is reserved for cases in which one news site quotes another published or broadcast news report. The category "expert", finally, is assigned to sources who appear due to their professional expertise, but who are not directly involved in the events (since this would make them part of the "interest group" category). Frequent examples of experts are sources who used to serve official functions in government or in an interest group, e.g. *former CIA officials* (times-wn-100106) or *a former UN humanitarian chief* (guardian-wn-100116), but also university professors or medical doctors who are consulted for their professional opinion. However, if a former politician was questioned at a hearing about events that took place while he or she was in office, this was classified as a governmental source, not as an expert.

Private actors appeared in three possible roles. The most common role was that of "victim", which included victims of natural disasters, crimes, and extreme weather situations. In one case, a girl who reportedly "cries tears of blood" (sun-100110b) was also classified as a victim. Furthermore, close relatives of victims (e.g. parents and partners) were also included in this category. "Witnesses" are those sources who report on such events without being directly affected. Finally, in the case of crimes, the perpetrators of the crime are sometimes quoted, in which case the source is tagged as "offender".

It probably goes without saying that the classification of source types is not always clear-cut. Often the information provided in a news article was not sufficient to decide which category should be assigned to a source. Additional background information on individuals and organisations was then consulted. There were also some cases of multiple sources, e.g. of a government institution and an interest group who presented a statement together. In some cases it was also quite difficult to decide whether a short segment of text in quotation marks should be treated as a scare quote ("impersonal") or whether it was part of a mixed quote which could be attributed to a source.

Most difficulties occurred in deciding which of the categories within one of the three main groups should be applied. For instance, members of consultant agencies who are interviewed about the economy and recent government decisions were classified as experts, but it could easily be argued that they are representatives of interest groups. The president of the association of chief police officers was classified as a member of an interest group, but, as a police officer, he would also qualify for the category "government". Similarly, it is not always easy to distinguish between victims and witnesses, especially in the case of major catastrophes like the Haiti earthquake, where local witnesses must also be directly affected by the event.

There are also some difficulties for assigning categories across the three main groups of sources. In the case of the Togolese football team, who were interviewed about their decision not to participate in the Africa Cup of Nations, the team members were classified as members of an interest group (the football team). Given that the quoted team mem-

Author submitted manuscript.

Author submitted manuscript. Consult published version for correct page numbers!

bers are very successful football players, one could also have argued for the classification "celebrity". Moreover, since the reason for their departure was that they had been the target of an attack, resulting in the death of three people, one could also have classified them as victims. However, the articles appeared a few days after the attack and their main focus is not the attack itself, but the decision of the team to leave. The decision to classify them as official sources rather than celebrities was based on the immediate context in which the quotes occur. The quoted sources clearly speak on behalf of their team, thus representing not primarily their personal interests, but the interests of a group. This is emphasised by the fact that the two sources who are mainly quoted are the coach, Hubert Velud, and the captain, Emmanuel Adebayor. They also use plural forms and explicitly refer to the decisions of the collective they represent, as in the following quote.

(7.27) Adebayor told a French radio station: "We all decided to do something good for the country and play to honour those who died. Unfortunately, the head of state and the country's authorities have decided otherwise. We will pack up and go home." (times-wn-100110)

According to the definitions for private, official and celebrity actors presented above, "official source" seemed to be the most fitting category.

The distinction between official news actors and witnesses in the case of the Haiti earthquake proved to be challenging too. Members of charities and governmental rescue operations who travelled to Haiti to help were naturally also witnesses of the disaster and reported on what they saw. Since they would not have been present, were it not for their official affiliation, they were classified as official sources (members of interest groups or governmental organisations). Conversely, the local residents who were interviewed as witnesses were sometimes also introduced with their professional affiliation. It was often difficult to distinguish between professional and private news actors in these cases and the decision was usually based on the context in which the quote appeared. Thus, a priest who reported on the difficulties of providing medical help and food to victims of the earthquake was classified as an official news actor, while a man who was running a local radio station and who was quoted on how he experienced the earthquake was classified as a witness. Again, arguments for different classifications could be made.

The reliability of the classification was tested with the help of a second coder, who independently classified a subset of 120 out of the 600 quotes (20 percent). Three different classifications were tested: the classification into named, identified and anonymous sources; the classification of source types into official, celebrity, private and impersonal; and the sub-classification of source types into finer-grained distinctions like government, spokesperson, media, etc. The second coder relied exclusively on the definitions presented above, without further instruction. The results are given in Table 7.5, both in relative agreement between the two codings and with the kappa value, a more conservative measure that corrects for chance agreement. Kappa values of more than 0.80 are usually considered very high, and even values between 0.61 and 0.80 are still good (see Altman 1991, 404; Landis and Koch 1977, 165). All categorisations are therefore sufficiently reliable for drawing conclusions.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Table 7.5: Results of reliability testing for the classification of sources based on 120 quotes

Classification	Absolute agreement	Relative agreement	Kappa value
named / identified / anonymous	113	94.2 %	0.87
official / celebrity / private / impersonal	110	91.7 %	0.80
subcategories	95	79.2 %	0.73

However, given that some of the source type subcategories have only few occurrences and that the reliability of this classification is lower than 0.80, I will base my interpretations mainly on the results for the first two classifications.

7.4.2 Named, identified and anonymous sources

Table 7.6 presents the quantitative results for the naming of sources in the news articles from the five online news sites and the *Times* of January 1985.[7] As mentioned above, the randomly selected quotes from the *Mail Online* have a strong bias, with 16 quotes coming from a single article. In all of these quotes, the source is "identified" without giving the source's name. The article reports on the Edlington hearing and the names of the quoted sources could not be given. Therefore, the data for the *Mail Online* is given in two ways. First, the total number of occurrences in each category is given, then, in brackets, the absolute and relative frequencies of all quotes in the sample which did not appear in this article are indicated. In this way, it is possible to compare the data of the *Mail Online* to the other media, both with the given bias and without.

Table 7.6: Named, identified and anonymous sources in randomly selected direct speech quotes from the news sections of five online news sites and the *Times* from 1985

	Times Online	***Guardian***	***BBC News***	***Mail Online***	***Sun***	***Times 1985***
named	74	86	87	65 (65, 77%)	72	64
identified	0	0	1	22 (6, 7%)	8	1
anonymous	26	14	12	13 (13, 15%)	20	35
Total	100	100	100	100 (84, 100%)	100	100

As the numbers show, naming of the source is the most frequent case in all news media, while the identification without name is quite rare. Named sources are most frequent for the *Guardian* and *BBC News* and least frequent in the data from the *Times* from 1985. The identified source in the case of the *Times* from 1985 is a quote attributed to the *NUM general secretary* (times-uk-850121), and all the identified sources in the data from the online sites refer to parties involved in the Edlington case. The anonymous sources do

[7] See appendix B for statistical tests.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

not seem to show a clear pattern at first sight. They are most frequently found in the *Times* from 1985, followed by the *Times Online* and the *Sun*. However, if one looks at some instances more closely, differences become apparent. For the *Times* from 1985, the anonymous sources fall into three categories. First, there are cases of scare quotes, as in example 7.28, for which no source can be named because the quote is not attributed to anyone in particular. Second, there are quotes that come from spokespersons or groups of speakers, of whom none is named as an individual source, such as in example 7.29. Third, there are cases in which newspapers and radio broadcasts are quoted, as in example 7.30, and where again it is not an individual person who is identified as a source.

(7.28) It is understood the refugees on the so-called "long march" had heard by word of mouth in their primitive villages that they would find new homes and food in Israel. (times-wn-8501)

(7.29) In a statement on the situation in the industry, the NUM said last night "The board and union also agreed that if negotiations are resumed it should be in the earlier part of the week commencing January 28 [...]" (times-uk-850125)

(7.30) "The Soviet-American statement has gone beyond expectations," said the Warsaw daily Zycie Warszawy [...] (times-wn-850110)

Examples like these can also be found in the *Sun*, but in addition there are also cases of private sources who are quoted as anonymous witnesses, as in example 7.31. Such cases do not occur in the *Times* from 1985, and they are rare in up-market online news sites. When anonymous witnesses appear on these sites, this usually concerns political matters and the sources are quoted because they are experts or insiders due to their profession, for instance in example 7.32. The sources are presented anonymously to protect their identity, which at the same time, creates the impression that the text producer has privileged access to confidential political information (see Tuomarla 2000, 169). The only cases of anonymous private witnesses who were quoted in the data from the up-market news sites were the two brief quotes in example 7.33. Here, it does not so much seem to be the case that the sources did not want to be named, but rather that their names are not known to the reporter.

(7.31) The woman, who did not want to be named, added: "How can somebody do this? It's terrible." (sun-100128a)

(7.32) "He had provided information that checked out, about people in al-Qaeda whom he had access to," a senior intelligence official, told the newspaper. (times-wn-100106, third comma in the original)

(7.33) Afterwards, as she was ushered away, two people shouted, "We love you, Fiona," and, "We love you, Fi." (times-uk-100130)

In their analysis of sources in the reporting on the US presidential campaign of 2004, Garretson and Ädel (2008) distinguish between unnamed and anonymous sources. Examples like 7.28 to 7.30 would be classified as unnamed, whereas cases in which sources explicitly wish to remain anonymous, as in example 7.31, are classified as anonymous. Applying this distinction to my data, one can say that explicitly anonymous sources, especially in the role of private witnesses, are a feature that seems to be typical of the *Sun* and to a somewhat lesser extent of the *Mail Online*. In contrast, unnamed sources in the

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Table 7.7: Official and private sources in randomly selected direct speech quotes from the news sections of five online news sites and the *Times* from 1985

	Times Online	***Guardian***	***BBC News***	***Mail Online***		***Sun***	***Times 1985***
government	52	52	54	31	(31, 37%)	30	58
spokesperson	7	12	19	4	(4, 5%)	6	6
media	1	1	0	0	(0, 0%)	3	6
interest group	22	20	12	17	(17, 20%)	8	20
expert	6	8	3	2	(2, 2%)	7	0
official	**88**	**93**	**88**	**54**	**(54, 64%)**	**54**	**90**
celebrity	**0**	**0**	**1**	**0**	**(0, 0%)**	**9**	**0**
witness	3	0	2	4	(3, 4%)	8	1
victim	2	7	3	31	(16, 19%)	21	0
offender	2	0	4	8	(8, 10%)	3	0
private	**7**	**7**	**9**	**43**	**(27, 32%)**	**32**	**1**
impersonal	**5**	**0**	**2**	**3**	**(3, 4%)**	**3**	**7**
undecided	**0**	**0**	**0**	**0**	**(0, 0%)**	**2**	**2**
Total	100	100	100	100	(84, 100%)	100	100

Author submitted manuscript

up-market papers tend to be official sources, such as spokespersons and representatives of interest groups, who are not named because they speak for a group rather than for themselves, such as in example 7.34.

(7.34) Some aid agencies had already started trying to help those in affected areas but the spokesman [of the Disasters Emergency Committee] said: "The aid effort isn't on the scale that's required ... we need people's help to scale it up urgently." (guardian-wn-100114)

In addition, the higher number of anonymous sources in the *Times Online* compared to the other up-market sites is also due to a slightly higher number of impersonal quotes (see also Table 7.7).

7.4.3 Private and official sources

A first, striking observation of the distribution of different types of sources in Table 7.7 concerns the differences between the data from the *Times* from 1985 and the online sites. With the exception of one witness, all the sources are official, if the quote is not impersonal.[8] Indeed, even this single occurrence of a private witness (example 7.35) is a disputable case.

(7.35) A member of the Jewish community who met the escapers when they landed up to 18 months after fleeing Lithuania, said: "I think Sugihara-San saved 10,000 people. When they landed in Japan I would stand next to the immigration officer, helping with translations," Mr Alex Triguboff said.

[8] The two undecided cases concern quotes that were co-attributed to a governmental organisation and an interest group. Both, therefore, clearly have official sources.

Author submitted manuscript. Consult published version for correct page numbers!

> "Sometimes they'd just photographed one person's visa. Again and again I saw the same name and the same date on the visa." (times-wn-850119, second quote included in random selection)

This quote appears in a report about a former Japanese consul, honoured for helping Jews escape from Lithuania in 1940. The witness describes how he encountered those who had escaped in Japan. From the report, it does not become entirely clear in what capacity he *would stand next to the immigration officer, helping with translations.* Since he is not introduced with any (former) official affiliation, this was classified as a private source, even though it is still somewhat less private than the typical case of a private witness in the data from the online sites, as, for instance, in the following examples (see also example 7.31 above).

(7.36) Reports say there is no sense of panic in Morehead City.
Resident Drew Hall told the Associated Press news agency by phone:
"Everybody is going about their business. Why get nervous? Things happen. You can't freak out in times like this." (bbc-wn-100112)

(7.37) Another neighbour, 61-year-old Jill Radford, said: "There were two young children and there were always toys in the garden."
"I heard the couple separated just before Christmas and put the house up for sale."
"I didn't know anything about this until I got home at 4pm today and saw all the police activity. It is desperately sad." (sun-100128a, first and third part of quote in random selection)

In these examples, the sources are presented as a *resident* and a *neighbour*, respectively, which clearly emphasises that they speak in a private capacity. Furthermore, they not only report what they saw, but they also express their personal attitude and their emotions. This results in a much more personal presentation than in example 7.35 from the *Times* from 1985. There, the direct speech seems to have a stronger function of presenting reportable facts, e.g. the number of people who were possibly saved and the details of the procedure, which could presumably not be verified by any other evidence than the source's statement. Indeed, previous to this quote, the report names two much lower figures which were suggested as the number of visas written by the consul (4,500 and 6,000). Even if this example is counted as a case of a private source, it fulfils different functions than the direct speech quotes from private sources in the more recent online data.

The official sources correlate quite well with the level of market orientation across the different media. All the randomly selected quotes from the up-market news media, including the older *Times* data, have between 88 percent and 93 percent of official sources. With 54 out of 100 random quotes the *Mail Online* and the *Sun* have the same number of official sources. However, if one takes the bias of the data from the *Mail Online* into account and excludes all quotes from the over-represented article, then the percentage of quotes with official sources is 64 percent, which is slightly higher than for the *Sun*, but still considerably lower than for the up-market news sites. A chi-square test comparing the frequency of official sources and private sources showed no statistically significant differences across the three up-market online sites. Moreover, the difference between the

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Mail Online and the *Sun* was not significant. All the other differences were significant at $p < 0.001$ (see appendix B).[9]

Thus, one can say that private sources are more frequent in the online data from 2010 than in the print data from 1985, and that also within the online data, private sources are more frequent on mid- and downmarket sites than on up-market sites. This corresponds to previous findings on the presence of private sources in the news. In her analysis of newspapers from 1700 to 2000, Bös (2010b, 85) finds an increasing tendency to quote private sources, which accelerates quite drastically towards the end of the 20th century. My findings suggest that this trend is still ongoing. Given that down-market newspapers often initiate changes in news reporting that are later adopted by up-market papers (see, for instance, Schneider 2002), it is possible that even more private sources will be quoted in up-market papers in the future.

The *Sun* seems to be the only news site that more or less regularly contains direct speech quotes from celebrities in its news data. In my random section, the only quote from a celebrity that appears on one of the other sites is example 7.38

(7.38) TV presenter Annabel Giles was among the hundreds of drivers stranded there. She was stuck in her car on the A3 for more than 22 hours en route to Brighton.
She told the BBC News website: "You can't sleep because the traffic keeps moving on a couple of inches and everyone starts beeping their horns and you have to wake up and get into gear and move on". (bbc-uk-100106)

This quote by Annabel Giles was one of only two direct speech quotes in a long report on problems caused by the heavy snow fall in the UK. In contrast to the celebrity quotes on the *Sun*, the source is not the main topic of the article. Instead, she is quoted in the context of a news report of immediate national relevance and the fact that she is a person of interest is of secondary importance.

The quotes with celebrity sources in the data from the *Sun* come from three different articles. One reports on the alleged love affair of England's football captain John Terry with the girlfriend of one of his team-mates; one on the planned release of a charity single by the *Sun* for the victims of the Haiti earthquake, to which bands and singers like Take That, REM and Susan Boyle were to contribute; and one on the musician Wyclef Jean and his charity organisation for Haiti. There are three more articles from the *Sun* in my data which deal with celebrities (see also chapter 6.2): Tiger Woods and his love affair; a report on the father of Ronnie O'Sullivan, a famous snooker player; and one on the drug death of the ex-girlfriend of rock star Pete Doherty. Thus, the presence of direct speech quotes with celebrity sources is a direct consequence of the fact that the *Sun* regularly reports on celebrities in its news section. In the news articles from the other news sites, in contrast, only one article with a similar celebrity focus can be found, namely an article from the *Mail Online* which also reports on the alleged love affair of John Terry, the football captain.

[9] Due to sparse data, the difference between the *Times* from 1985 and the up-market online sites could not be tested.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

7.5 Summary and conclusion

This chapter investigated the role of direct speech in the personalisation of news. I argued that, by giving a voice to news actors, direct speech plays an important role for their personalisation. When direct speech is used to express emotions and dramatise descriptions, its personalising potential is particularly strong. However, the functions of distancing the journalist from the content of the style of the quote can also contribute to more personalised presentations of news events and actors.

The distribution of direct speech across the different sections of my data showed that direct speech is clearly more frequent in the online data than in the printed *Times* from 1985. I suggested that this might be one effect of up-market news media following a trend set by mid- and down-market publications, although more research would be needed to test this hypothesis. A clear difference between the online news sites and the printed *Times* from 1985 could also be found in the analysis of different types of sources in a randomly selected sample of quotes. With one possible exception, private sources were only found in the data from the online sites. This again seems to point towards an increase in personalised news presentation in the data from 2010, compared to the data from the *Times* from 1985. Moreover, the larger number of named sources in the online data (especially from the *Guardian* and *BBC News*) compared to the data from the *Times* from 1985 can be seen as reflecting a stronger focus on the person of the source. Impersonal quotes are less frequent and sources are often presented by name. This tendency towards personalisation is even stronger than it appears from the evaluation of the 100 random quotes alone since the online data contains considerably more quotes overall than the data from the *Times* from 1985.

The analysis of the sources in randomly selected quotes also showed clear differences across market orientation. Both the *Mail Online* and the *Sun* have considerably more quotes from private sources than any of the up-market news sites. In addition, the *Sun* appears to be the only site which more or less regularly contains quotes by celebrities. This corresponds with the finding that celebrity news on the *Sun* has a similar prominence as topics like war and terror do on the up-market sites, as I have shown in chapter 6. This is again related to personalisation strategies that turn the lives of certain individuals into newsworthy stories. The *Sun* also takes a special position with respect to the naming of sources. The anonymous sources in the *Sun* are to a large extent the result of private sources – in particular in the role of witnesses – who do not wish to be named. They, too, are related to personalisation tendencies, in this case the trend of this down-market news site to emphasise the experiences of private individuals.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Author submitted manuscript

8. Personal pronouns

First and second person pronouns play an important role in the personalisation of news. Some of the ways in which they contribute to personalisation are illustrated in the following headlines.

(8.1) Now *we*'re running out of grit: Councils forced to ration dwindling supplies as Britain faces weekend of chaos (mail-uk-100108)

(8.2) Does eating less make *you* healthier? (times-sn-100130)

(8.3) Mum confesses: "*I*'ve killed *my* children" (sun-100128a)

All three headlines use first and second person pronouns to attract the attention of the reader. In all three cases, personalisation plays an important role, but each headline differs with respect to the precise effects. The headline in 8.1 contains a first person plural pronoun that includes the target audience. An alternative version of the headline without a first person pronoun would be *Now Britain is running out of grit*. In contrast to this version, the headline including *we* creates a collective identity that includes the news site as well as the audience, thus aligning the site with the position of the reader. It also implies that actual individuals are affected, and that these individuals include the readership. In this way, the direct relevance of the news event for the audience is emphasised. Similar effects are created with the use of *you* in 8.2. Here, too, the headline suggests that the information that is provided in the article is directly relevant for the reader. However, while the reader in 8.1 is part of a larger group, defined through a common nationality, he or she is directly addressed as an individual in 8.2. This direct address of the reader by the text producer introduces an interactional element into mass communication. In 8.3, a headline already discussed in chapter 6.3, the pronoun does not refer to the reader or the text producer. It occurs within direct speech, representing a statement by a news actor. This actor is presented in informal terms as *Mum* and representing her quote as a first person statement in direct speech further increases the immediacy and personalised character of the headline. The use of the first person singular pronouns *I* and *my* presents her as an individual who acted and who is aware of her actions.

Such uses of first and second person pronouns can not only be found in headlines, but also in the body of articles. The examples above illustrate that a variety of different uses of first and second person pronouns can be found in my data. The effects for personalisation differ depending on which pronoun is used, whether or not it occurs in direct speech, who it refers to and what the specific context of its occurrence is.

Author submitted manuscript

In this chapter, I am going to investigate these different dimensions. My aim is, first and foremost, to describe the large spectrum of effects that first and second person pronouns can have for personalisation. I will demonstrate that first and second person pronouns are used to personalise any of the three main entities in the mass communication situation: the audience, the news actors and the text producers. For this purpose, it is necessary to interpret individual instances of the various pronouns in detail, differentiating between different forms and their place and context of occurrence. In addition, I will look at differences in the use of first and second person pronouns across different news sites and different types of articles. For this comparison, I will look, on the one hand, at the frequency with which the pronouns occur inside and outside of direct speech in different news publications and articles. On the other hand, I will also try to characterise qualitative differences in their use across the various sections of my data, identifying typical uses for each news site and article type.

I will start with some general characteristics of first and second person pronouns, in particular with a systematic overview of the possible types of referents each pronoun can have and with a brief summary of how they have been discussed as a feature of linguistic immediacy and involvement in previous research. This will be followed by some methodological remarks and a first quantitative overview in section 8.3. The actual analysis is then structured into three sections. Section 8.4 focuses on personal pronouns within direct speech quotes in news articles from five online sites and the printed *Times* from 1985. In section 8.5, I will look in detail at the infrequent occurrence of first and second person pronouns outside of direct speech in the same data. Section 8.6 will compare these results to the use of first and second person pronouns outside of direct speech in columns and soft news, focussing in particular on the use of first person singular pronouns.

Author submitted manuscript

8.1 Reference of first and second person pronouns

Typically, first and second person pronouns are used with specific reference, referring to speaker/writer and the addressee, respectively (Biber et al. 1999, 328). However, there are a number of additional uses of first and second person pronouns. Following Kitagawa and Lehrer (1990, 742), one can distinguish between three groups of uses: referential, vague, and impersonal or generic. These different uses will be briefly described in the following.[1]

Referential use is not only the most typical use of first and second person pronouns, but also the use for which the identification of the referent seems more or less straightforward. First person singular pronouns with referential meaning always refer to the speaker/writer. Second person and first person plural pronouns are ambiguous; there are several possibilities to which participant(s) in a communicative situation they can refer. A first problem of identifying the referent of second person pronouns is posed by the fact that there are no distinct forms for second person singular and plural pronouns in standard English. Thus,

[1] For more information about the differences between referential, vague and generic reference, see Kitagawa and Lehrer (1990). For a much more detailed discussion of pronoun reference in general, see Wales (1996).

Author submitted manuscript. Consult published version for correct page numbers!

you can refer to a single addressee, a plural addressee, or an addressee together with a third party.[2] The distinction between singular and plural referents can be made explicit through expressions such as *you all*, *you two*, *you guys*, or by dialectal forms like *yous* (Biber et al. 1999, 330). Wales (1996, 16-17) lists a whole range of non-standard second person pronouns in different varieties of English which allow this distinction. However, distinct forms for singular and plural are not available in standard English and whether *you* has a singular or plural referent needs to be derived from the context.

A similar situation is given for *we*, which can have either inclusive or exclusive meaning. In English, the same forms of the first person plural pronoun are used to refer to either the speaker and the addressee, the speaker and a third party, or the speaker, the addressee and a third party. In the following two examples from Biber et al. (1999, 329), the ambiguity is resolved by specifying the referent of *we* explicitly. Such an explicit specification is the exception, however, and, as with *you*, the referent is usually derived from the context.

(8.4) *I and Vicki* we don't really like caramels.
(8.5) We've got a bond in common, *you and I.*

In addition, Wales (1996, ch. 3) names several uses of *we*, restricted to specific situations, in which its meaning does not correspond to 'you and I' or 'I and another'. These include "royal *we*" and "authorial *we*", which refer to a singular speaker, and "medical *we*" and *we* in baby-talk, where *we* refers to the addressee only.

Bredel (2002, 171) argues for German that second person singular pronouns can be used with speaker reference in so-called inner dialogues, in which the speaker addresses him- or herself. A relevant example of English singular *you* occurs in a *BBC News* article on an attack against the Togolese football team at the Africa Cup of Nations.

(8.6) "[...] I told myself: 'If *you*'re still there on the ground in Angola, why not (play)?'" (bbc-wn-100110)

The Togolese player Emmanuel Adebayor here reports his thoughts on the question of whether or not to participate in the competition after the attack. The thoughts take the form of an inner dialogue, in which he addresses himself with *you*. In inner dialogues, speaker and addressee are the same person, which allows *you* to refer to the speaker.

Vague uses can be observed with *we* and *you* and describe cases in which the pronoun refers to individuals through their membership of a group, such as a nationality, religion or political party. Thus, vague *we* is identical with what Quirk et al. (1985, 350) call "rhetorical *we*", namely *we* "used in the collective sense of 'the nation', 'the party'". Particularly interesting in the context of media discourse is the vague use of *we* by politicians who simultaneously speak on behalf of themselves and their party or the government. Using *we* to refer to the politician and party when addressing a larger audience would be an exclusive

[2] The same is, of course, true for the forms *your* and *yours* – although not for the reflexive form, where *yourself* can only refer to a single addressee and *yourselves* can refer to a plural addressee or an addressee together with a third party. In the following, the subject case forms *I*, *you* and *we* will be used to refer to all pronominal forms of the respective person, i.e. subject case, object case, possessive pronoun, possessive determiner and reflexive pronoun (see Biber et al. 1999, 328).

Author submitted manuscript.

Author submitted manuscript. Consult published version for correct page numbers!

use of *we*, but, as Wales (1996, 62) argues, the boundaries between inclusive and exclusive uses are often blurred in political discourse. Speakers often use a potentially inclusive *we* to align themselves with the audience and construct a common identity, for instance by sharing a common party affiliation or nationality. This use is illustrated by the following example from my data from an article that appeared in the *Times* from 1985, reporting on negotiations between the US and Russia about nuclear arms.

(8.7) Although clearly delighted at the outcome, however, Mr Reagan warned Americans last night that negotiations to reduce the nuclear threat would be difficult. He told a televised press conference: "But *we* will persevere. While *we* must continue to resist actions of the Soviet Union that threaten *our* freedom and vital interests and those of other nations, *we* must also be prepared to work togheter where ever possible to strengthen the peace". (times-uk-850110)

While the first use of *we* seems to refer to the negotiators directly, the following uses of *we* and, in particular, *our* in *our freedom* includes a much larger group, defined through a common nationality. Wales (1996, 62) calls uses of *we* in which the implied referent is defined through a common nationality "patriotic *we*". In news articles, patriotic *we* can occur in the direct speech of sources like politicians. However, patriotic *we* is also used by text producers, especially in opinionated texts. Westin (2002), for instance, in her analysis of newspaper editorials in three British up-market newspapers from 1900 to 1993 finds that rhetorical uses of *we* referring to the nation were particularly frequent around the time of World Wars I and II.

Vague uses not only occur with *we*, but also with *you*. Here, the second person pronoun projects an identity of "the other", often in contrast to the identity created through vague *we* (see also Wales 1996, 58-61 on 'them and us' attitudes). An example of this can be seen in the following extract from my data, which quotes a tape broadcast on al-Jazeera on which the speaker (purportedly Osama bin Laden) claims responsibility for an attempted terror attack.

(8.8) "America will never dream of living in peace unless we live it in Palestine. It is unfair that *you* enjoy a safe life while our brothers in Gaza suffer greatly. God willing, our raids on *you* will continue as long as *your* support to the Israelis will continue." (guardian-uk-100124)

The tape is addressed to President Obama and some uses of the second person pronoun might, therefore, refer to him, for instance in *your support to the Israelis*. However, it also becomes clear that some uses of the second person pronoun refer to a larger group. The *you* in *raids on you* cannot refer to Obama only, since the attacks that are mentioned were an attempted airplane bombing and the attacks of 9/11. Thus, this *you* refers to a vaguely specified group of individuals who identify with *America*, the subject of the first sentence.

First and second person pronouns can also be used with generic or impersonal reference. In this use, the pronoun does not refer to a vaguely defined group of individuals (e.g. those with a common political orientation or nationality), but to people in general. The personal pronouns are used to formulate general truths (example 8.9) or to generalise opinions and

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

experiences (examples 8.10 and 8.11), and they could be replaced by *one* (Kitagawa and Lehrer 1990, 742; see also Laberge and Sankoff 1979 on impersonal uses of *tu* and *vous* in French).

(8.9) When *we* start talking *we* often cease to listen. (Biber et al. 1999, 330)

(8.10) Granted, "laugh" is a perfectly decent word – but one would hope that *we*'d have more than one option when it came to the second best activity available to humans, after having a cheese sandwich. (times-cs-100118)

(8.11) The fact is, when *you*'re under the media microscope, *you* can't afford to make a mistake – or *you*'ll be on the carpet, red or otherwise. (mail-sn-100122)

In example 8.10, *we* co-occurs with *one* and the use of *we* rather than repeating *one* may well be stylistically motivated. Wales argues that impersonal *you* often co-occurs in sentences with *one* and that this is one way to avoid the repetition of *one*, which many perceive as "pretentious and affected" (Wales 1996, 81-82). In example 8.10, the use of the numerical *one* later in the same sentence might additionally discourage the repetition of *one* and favour the use of *we*. Impersonal uses occur most often with *we* and *you*, but Kitagawa and Lehrer also present an example of an impersonal use of *I*.

(8.12) We form a frame of script for this kind of situation ... Thus, in order to be able to take the subway in New York *I* simply need a 'taking a subway' script or frame, if *I* have one, and supply now relevant specific information about the situation. But at the same time, *I* may – even if *I* take the subway daily – be reminded of yesterday's trip when *I* met this strange man, or last year's when there was a fire in the subway. If *I* do not have a frame or script, *I* may well be reminded of the rather vague and remote (i.e., macro-) information from the model *I* built when some years ago *I* took the subway in New York. (Kitagawa and Lehrer 1990, 741-742)

As Kitagawa and Lehrer argue, the fact that the text in which this passage occurs is written by more than one author makes it implausible that *I* is used to refer to the writer. Rather, this example is very similar to examples 8.9 to 8.11 above; *I* has generic reference, referring to anyone, and it could be replaced by *one*, *you*, or *we* without changing the informational content of the message (Kitagawa and Lehrer 1990, 742). Still, the choice of generic *I*, *you*, or *we* rather than *one* changes the effect of the text. As Biber et al. (1999, 331) put it, generically used personal pronouns "retain a tinge of their basic meaning", i.e. the meaning they have in their prototypical referential use (see also Quirk et al. 1985, 354). Thus, one can argue that generic first and second person pronouns are a more personalised alternative to *one* or passive constructions. The term "impersonal use" is, therefore, somewhat misleading in the context of a study on personalisation and I will in the following refer to these uses as "generic uses".

The personalising potential is clearly visible in a subtype of generic uses of *you*, which Kitagawa and Lehrer call "life drama" (Kitagawa and Lehrer 1990, 749). They give the following example:

(8.13) *You* are in Egypt admiring the pyramids and feeling that *you* have really left *your* own world and time behind when suddenly *you* meet *your* next-door neighbor from home. (Kitagawa and Lehrer 1990, 749)

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

As in example 8.11 above, *you* here is used generically; it could potentially refer to anyone. This use occurs in what Kitagawa and Lehrer (1990, 749) call "the 'scene-setting' portion of a mini-tale" and they argue that this use differs from other generic uses in a number of formal respects, e.g. by not being restricted to uses with the present tense. What is more interesting in the context of the present study is the function of such life drama uses to generalise personal experience (see also Laberge and Sankoff 1979, 428-429, on similar uses of *tu* in French). The scenario that is described seems to reflect the personal experience of the speaker, or at least of someone the speaker knows. The use of *you* rather than a pronoun that refers to the person who experienced the story (*I* or *he/she*) makes it clear that this could happen to anyone, including the addressee. The choice of pronoun invites the addressee to identify with the person who experienced the episode and, thus, creates involvement.

An example of a life drama use of *you*, in which the author is clearly not the person who experienced the events first hand, occurs in a soft news article on the *Times Online*.

(8.14) Imagine that *you* are Iris Robinson. *Your* husband has dutifully, doggedly marched in Paisley's shadow for decades, always the bridesmaid and never the bride, but finally his day has come and "the Doc" has gone. *Your* husband doesn't have the charisma or the sex appeal that made Paisley such an elemental force of politics. But *you* do. (times-sn-100112)

The author initiates the life drama scenario by explicitly asking the reader (to whom the first *you* refers) to identify with Iris Robinson, whose experience is then described. This experience is very specific and does not have the same implication of a general truth as the generic use of *you* in example 8.11. *You* refers to Iris Robinson, but it generalises her position to include the addressee and potentially anyone.

Bredel (2002, 172) also identifies uses of second person pronouns that generalise personal experience. In her case, the data come from German narrative interviews. While Kitagawa and Lehrer (1990) are mostly concerned with the potential of *you* to create dramatic tension, Bredel (2002) is interested in how the second person singular pronoun is used to establish different speaker perspectives. In contrast to Kitagawa and Lehrer, she insists that the second person pronoun in examples like 8.13 and 8.14 still refers to the person who experienced the episode, despite its generalising use. Thus, the life drama use of *you* is far less impersonal than other generic uses of *you*. It refers more clearly to the experiencer of an episode and it is also more clearly directed at the addressee, who is invited to identify with the specific position that is described.

The reference of first and second person pronouns can be summarised as follows (see Table 8.1). All first and second person pronouns can occur with specific reference to speaker/writer and/or addressee. For the first person plural, this reference is ambiguous between inclusive (speaker/writer and adressee) and exclusive (speaker/writer and third party) meaning. In the case of second person pronouns, the reference is ambiguous between singular and plural referents. These ambiguities are usually resolved through the context. In addition, there are a number of special uses such as "royal *we*" and "medical *we*", where

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Table 8.1: Summary of first and second person pronoun reference

	I	*we*	*you*
specific reference			
	speaker (sg.)	speaker (pl.) speaker & addressee speaker & 3rd person(s) speaker & addressee & 3rd person(s)	addressee (sg.) addressee (pl.) addressee & 3rd person(s)
special referential uses			
"royal *we*"		speaker (sg.)	
"medical *we*"		addressee (sg./pl.)	
sg. *you* in inner dialogue			speaker = addressee (sg.)
vague reference			
e.g. "patriotic *we/you*"		speaker & others	addressee & others
generic			
	anyone	anyone	anyone
"life drama" subtype			anyone ([+experiencer], [+addressee])

we refers to a single speaker/writer or to one or more addressees, respectively. First person plural pronouns – and, to a lesser extent, second person pronouns – can also have vague reference, referring to a group that is not clearly defined. A relevant example of this use is what Wales (1996) calls patriotic *we*. Generic uses, in which the pronoun refers to people in general and could be replaced by *one* can occur with all first and second person pronouns, although they are rare for *I*. A special case of generic use of the second person pronoun is given by what Kitagawa and Lehrer (1990, 749) call the "life drama" subtype. One could argue with Bredel (2002) that this use not only has generic reference, but that it also has a strong experiencer and addressee orientation.

The categorisation into referential, vague and generic uses of first and second person pronouns is not always straightforward, however. The distinction between vague uses and referential uses that include third parties is often not clear-cut, and vague and generic uses also cannot always be clearly distinguished. Biber et al. (1999, 329-330) mention academic prose as a genre in which it is often unclear whether *we* refers to the author(s), the author(s) and the reader, or to people in general. In mass media communication, the situation can be further complicated by complex communicative situations involving different audience circles (see chapter 4.1.1). In radio or TV news interviews, for instance, the speakers can switch between addressing the interlocutor of the interview situation and addressing the mass audience (Bell 1991, 96-97). Whether or not the audience is included in the address may not always be clear. In her analysis of the pronoun use of TV news presenters reporting in the war in Iraq in 2003, Ferrarotti (2009), therefore, includes a category of potentially inclusive uses of *we* and *you*. With this category, she accounts for cases in which it is not clear whether the mass audience is included in the referent of the pronoun

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

or not. Furthermore, the identity of the mass audience is necessarily unknown, which raises problems not only for referential uses, but also for vague uses. Whether or not readers will feel included or excluded by the use of a vague pronoun depends on their identity, beliefs, political orientation, and so on (see also Wales 1996, 62). As a consequence, the intention of the speaker and the interpretation of the reader may not correspond. A speaker/writer can use a vague *we* with the intention of including the audience, but individual readers "may resist inclusion" Wales (1996, 62).

8.2 First and second person pronouns as features of immediacy and involvement

First and second person pronouns have repeatedly been identified as features of linguistic immediacy (e.g. Ágel and Hennig 2006a, 58; Koch 1999, 409), often in connection with their deictic function of referring to the speaker/writer and the addressee in a way that indicates or simulates co-presence. Ágel and Hennig, for instance, list person deixis – which they argue is mostly expressed through first and second person pronouns (2006a, 58) – as one of the parameters indicating a high agreement between the spatial and temporal situation of the interlocutors (Ágel and Hennig 2006b, 23, and appendix 382). Similarly, Puschmann (2010, 181) argues that first and second person pronouns in corporate blogs are used to "[construct] a deictic center that can be retrieved by the blog reader". He regards this as one of the parallels between spoken discourse and the language used on the blogs he analyses.

The most important aspect of linguistic immediacy for this study is the function of first and second person pronouns to indicate personal involvement. In a study on interaction on online message boards, Claridge (2007, 97) even calls the analysis of pronouns "[a] fairly straightforward approach to personal involvement and interactiveness". She argues that in her data *I* marks personal involvement and is used to express personal opinions, attitudes and feelings, whereas *you* is used to address the other forum participants and is a good indicator of the degree of interaction (2007, 97-99). Biber comes to the conclusion that first and second person pronouns are highly indicative of involved texts in general. In his multi-dimensional approach to register, he first identifies clusters of linguistic features which frequently co-occur in texts and to which he then assigns functions in a second step. The distinction between involved and informational texts is one of the dimensions resulting from this cluster analysis. Involved texts are characterised by "high interpersonal interaction or high expression of personal feelings" (Biber 1988, 106), and the presence of first and second person pronouns accounts for two of the ten most indicative features for locating texts at the involved end of the scale (Biber 1988, 89, 1995, 114). Chafe likewise names first person references as one of the involvement features that can be found in spoken conversation, and adds that "Second person reference would seem to be also a symptom of involvement, but there were too few examples in our data to demonstrate anything of

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

interest" (Chafe 1982, 46). Likewise, Katriel and Dascal (1989, 290-291) see the frequent use of first person pronouns as an indication of speaker involvement.

In mass communication, the involvement function of first and second person pronouns is particularly salient when they are used to address the text producer and/or the audience. Directly addressing members of a mass audience and, thus, creating a pseudo-interaction between text producer and addressee is an example of what Fairclough (2001, 179) calls synthetic personalisation (see chapter 2.1). He argues that synthetic personalisation is a strategy of text producers and official news actors in mass media to simulate solidarity and to "relate to members of their audience as individuals who share large areas of common ground" (Fairclough 2001, 160). Direct audience address is typical of advertising, but it can also be used by radio announcers and even in newspapers, in particular in headlines of mid- and down-market newspapers (Jucker 2000, 638-639, 2011, 255). Jucker lists direct audience address as one of the strategies for "trying to overcome the unidimensionality of [mass] media communication", which may be motivated by the wish to increase audience appeal (Jucker 2011, 254-255; see also Valdeón 2006, 411-412). Similarly, Montgomery (1986) argues that for DJ radio monologues the frequent use of second person pronouns (along with other forms of address) serves to simulate co-presence with the audience and to foreground the interpersonal dimension. O'Keeffe (2006) discusses personal pronouns as markers of pseudo-intimacy in various types of media interactions such as radio phone-ins, chat shows and political interviews on television and the radio. She also argues that the pronouns *I* and *you*, as well as inclusive uses of the pronoun *we*, are used to "create an illusion of an interpersonal relationship between strangers" (O'Keeffe 2006, 97).

Generally, first and second person pronouns are not a typical feature of news language. Bell (1991, 208) even argues that first and second person pronouns do not occur outside of direct speech in news articles. The close connection between the use of first and second person pronouns and the use of direct speech is also pointed out by Mair and Hundt (1997, 76). They suggest that the increasing use of these pronouns in British and American newspapers between 1961 and 1991/92 is probably a consequence of the increasing use of direct speech. Still, first and second person pronouns contribute to what Mair and Hundt (1997) call the increasing colloquialisation of newspapers (see also Hundt and Mair 1999). Simon-Vandenbergen (1986, 22) analyses the use of personal pronouns in British qualities and tabloids from 1983 from the point of view of readability. She finds that the frequency of first person pronouns relative to all pronouns differs considerably across the different newspapers, with fewer first person pronouns in qualities than tabloids. She, too, stresses the close relation between the frequency of first person pronouns and the frequency of direct speech in her data. However, these findings do not seem to hold equally for all types of articles. In her analysis of involvement features in British newspaper editorials across the 20th century, Westin (2002, 43-45) finds that first person plural pronouns are quite frequently used outside of direct speech. According to her results, the frequency of *we* in editorials is highly dependent on the topic. She explains the higher frequency of first person plural pronouns up to the 1940s with a frequent use of rhetorical *we* used

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

to create a common national identity in the context of war situations (Westin 2002, 45). This is supported by the findings of Xekalakis (1999, 220), who investigates foreign news reports between 1785 and 1985. In her data, too, first person plural pronouns with vague reference are frequently used in war times in constructions like *our troops*, *our losses* and *our bombing*. She also points to shifts that occurred during the 200 years covered by her study. While first person plural pronouns outside of direct speech are not unusual in the early sections of her data, they can no longer be found in the later sections (1999, 227). Xekalakis (1999) relates this finding to changes in journalistic practices during this period, which require a more objective reporting voice and, consequently, restrict the use of *we* to direct quotes (1999, 227).

Based on these previous findings some hypotheses about the distribution of first and second person pronouns in my data can be formulated. It can be expected, for instance, that the frequency of first and second person pronouns will be higher in the data from the mid- and down-market sites than in the data from the up-market sites, and that it will be higher in the data from the *Times Online* from 2010 than in the data from the *Times* from 1985. With respect to different types of articles, occurrences of first and second person pronouns in news articles can be expected to be restricted to direct speech. Columns, in contrast, are likely to contain first person pronouns outside of quotes. On the one hand, these may be occurrences of rhetorical *we*, similar to the uses Westin (2002) found in editorials. On the other hand, the more argumentative character of columns is likely to surface in a higher frequency of parenthetical constructions like *I think*. Such constructions have been shown to be frequently used in argumentative discourse where they fulfil an intersubjective function, inviting the audience to share the perspective of the speaker (Fetzer and Johansson 2010, 245). Given that columns tend to present the personal perspective of the text producer, such constructions with first person singular pronouns are likely to be found.

Author submitted manuscript

8.3 Frequency of first and second person pronouns

I want to analyse first and second person pronouns in my data with respect to their frequency and their effects for personalisation across different sections of my corpus. I also want to compare their use across different online news sites, compare it to the older data from the printed *Times* from 1985 and include different types of articles. In addition it is crucial to distinguish between occurrences in direct speech and occurrences outside of direct speech. Including all these aspects makes it necessary to adopt a selective approach. My aim, therefore, is not to give a comprehensive account of all characteristic features of all uses of first and second person pronouns in all types of articles. Instead, I will focus on some of the most salient tendencies in the various contexts.

I am first going to give a quantitative overview of all occurrences of first and second person pronouns in my data, which I will start with a discussion of the methods I used to identify these occurrences and some problems I encountered. The following sections

Author submitted manuscript. Consult published version for correct page numbers!

will then investigate first and second person pronouns that occur in direct speech in news articles, their infrequent occurrences outside of direct speech in news articles, and, finally, their use in different types of articles, focussing in particular on the use of *I* outside of direct speech.

8.3.1 Identification of first and second person pronouns

Since first and second person pronouns occur only in a limited number of forms, they are well-suited for automatic tagging. Using regular expressions, all the forms given in Table 8.2 were searched for and annotated with the respective tags. I included not only subject case and object case forms of the personal pronouns but also possessive pronouns, possessive determiners and reflexive pronouns. These latter forms can have the same referents as the personal pronouns and can therefore be assumed to fulfil equivalent functions for personalisation. As can be seen in Table 8.3 – and as could be expected – possessive and reflexive pronouns are, however, considerably less frequent than personal pronouns. Given that there are no separate forms for second person singular and plural pronouns (except for the reflexive pronoun), the second person forms could not be automatically annotated with respect to number. In the following, I will therefore only distinguish between first person singular, first person plural and second person pronouns.

Table 8.2: First and second person pronouns and related forms that were included in the analysis (table based on Biber et al. 1999, 328)

person	personal pronoun		possessive		reflexive pronoun
	subject case	object case	determiner	pronoun	
1st sg.	*I*	*me*	*my*	*mine*	*myself*
1st pl.	*we*	*us*	*our*	*ours*	*ourselves*
2nd	*you*		*your*	*yours*	*yourself / yourselves*

In most cases, the automatic tagging of the forms was unproblematic. However, there were a few cases that needed disambiguation. For the form *us*, a case-sensitive search was necessary; *us* and *Us* were tagged as pronoun, but *US* was excluded, as it is frequently used instead of *U.S.* to refer to the United States. A manual search through the occurrences of *US* ensured that no capitalised forms of the personal pronoun were missed. A few cases of *I* were identified which stood for the numeral 'one' in *World War I* and which were manually excluded. Finally, *mine* is not only a possessive pronoun, but also a noun, for instance in *salt mine*. Since both uses are not very frequent, all occurrences of *mine* could easily be checked manually.

Table 8.3 summarises the frequencies with which the different forms were found in the online data. First person singular pronouns were the most frequent group of forms, followed by first person plural forms, with second person forms being the least frequent. This is in agreement with the distribution of personal pronouns that Biber et al. (1999, 334) found in their news data. Within each category, the subject case form is the most frequent, followed by the possessive determiner and the object case form. Possessive and

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Table 8.3: Frequency of first and second person pronoun forms in the online data (all article types), normalised per 10,000 words (and absolute)

form	**norm.**	**(abs.)**
I	82.6	(1,744)
me	13.5	(285)
my	20.5	(432)
mine	0.3	(7)
myself	1.8	(37)
1st sg.	**118.6**	**(2,505)**
we	43.2	(912)
us	7.4	(157)
our	14.3	(303)
ours	0.1	(3)
ourselves	0.4	(9)
1st pl.	**65.5**	**(1,384)**
you	30.0	(634)
your	8.0	(169)
yours	0.0	(1)
yourself	0.6	(12)
yourselves	0.1	(2)
2nd	**38.7**	**(818)**
TOTAL	**222.8**	**(4,707)**

Author submitted manuscript

reflexive pronouns are quite infrequent for all persons. In the following, I will always refer to the total of forms found for each person, without discussing differences between the use of the individual forms.

A problem was posed by one of the articles in the UK news section of the *Mail Online*, mail-uk-100124, an article reporting on the Edlington court case which was already noted for its unusual properties in chapter 7.4. At 4,796 words, it is considerably longer than any other news article, the second longest containing 2,903 words and the total number of words in the 28 news articles from the *Mail Online being* 34,403. This means that this article contains 14 percent of all words in the news data from this site. In addition, the article contains some atypical features. A large part of the article reports on the experiences of the parents of one of the tortured boys. This section contains a high proportion of direct speech and an atypically high frequency of first and second person pronouns. As Table 8.4 shows, 155 of the 537 first and second person pronouns from the news section of the *Mail Online* come from this article, which amounts to about 29 percent. For first person plural pronouns, the article contains as many as 35 percent of all instances in this section of the data (89 of 257).

Including articles which have such a large impact on the feature under analysis will lead to skewed results. Therefore, I decided to exclude this article for the quantitative analysis of first and second person pronouns. This means that the quantitative results for the news articles from the *Mail Online* are based on the remaining 27 articles only. This

Author submitted manuscript. Consult published version for correct page numbers!

Table 8.4: Frequency of pronouns in news articles from the *Mail Online* with and without article mail-uk-100124, normalised per 10,000 words (and absolute)

Mail Online		**1st sg.**		**1st pl.**		**2nd**		**Total**	
	Words	norm.	(abs.)	norm.	(abs.)	norm.	(abs.)	norm.	(abs.)
All news	(34,403)	60.8	(209)	74.7	(257)	20.6	(71)	156.1	(537)
100124	(4,796)	95.9	(46)	185.6	(89)	41.7	(20)	323.2	(155)
w/o 100124	(29,607)	55.1	(163)	56.7	(168)	17.2	(51)	129.0	(382)

being said, it is, of course, still noteworthy that an article that is so atypical of the news section of online sites is published as a top-listed article of the UK news section on the *Mail Online*.

8.3.2 First and second person pronouns across subcorpora

Table 8.5 gives a first overview of the frequency with which first and second person pronouns were found in the news articles of my data. Overall, the normalised frequencies (instances per 10,000 words) indicate that all first and second person pronouns are clearly much more frequent in the online data than in the *Times* from 1985. This is further supported by statistical testing, which indicates a significant difference ($p < 0.001$) between the *Times* from 1985 and all the online sites (see appendix B for statistical tests). Within the online data, there are different patterns for the individual pronouns. First person singular pronouns show the clearest correlation with different market orientations. In news articles from the up-market sites, first person singular forms have a frequency of 43-49 instances per 10,000 words. With 82, this frequency is much higher for the down-market *Sun* (significant with $p < 0.001$). The mid-market *Mail Online* has a slightly higher frequency than the up-market sites (not significant), but still considerably lower than the *Sun* ($p < 0.001$).

A similar (albeit less marked and not statistically significant) variation can be observed in the second person forms. Again, the up-market sites have fewer occurrences than the *Sun*, and the *Mail Online* takes an intermediate position. The first person plural forms show a less clear pattern. The *Times Online* is the site on which they occur most frequently and the *Sun* is the site on which they are least frequent, but the *Guardian* and *BBC News* have frequencies which are much lower than those of the *Times Online* (significant at $p < 0.01$). Except for the *Sun*, first person plural forms are most frequent on all sites, followed by first person singular forms. For the *Sun*, first person singular forms are most frequent.

Looking at the overall frequencies of personal pronouns across different types of articles in the *Times Online* and the *Mail Online*, some even more striking patterns become apparent (see Table 8.6, for statistical tests see appendix B). First person singular forms are 5.5 (*Mail Online*) to 6.3 (*Times Online*) times more frequent in soft news articles than in news articles ($p < 0.001$). This corresponds to the fact that soft news articles often focus on the (life) stories of individuals, whose experiences are sometimes even presented

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Table 8.5: Frequency of pronouns in news articles from five online news sites and the *Times* from 1985, normalised per 10,000 words (and absolute)

		1st sg.		**1st pl.**		**2nd**		**Total**	
News site	Words	norm.	(abs.)	norm.	(abs.)	norm.	(abs.)	norm.	(abs.)
Times Online	(22,443)	43.2	(97)	80.6	(181)	15.1	(34)	139.0	(312)
Guardian	(22,751)	48.8	(111)	68.1	(155)	14.1	(32)	131.0	(298)
BBC News	(19,323)	44.5	(86)	58.5	(113)	14.5	(28)	117.5	(227)
Mail Online	(29,607)	55.1	(163)	56.7	(168)	17.2	(51)	129.0	(382)
Sun	(16,869)	82.4	(139)	53.9	(91)	23.7	(40)	160.1	(270)
Times 1985	(23,188)	9.9	(23)	30.6	(71)	2.2	(5)	42.7	(99)

Table 8.6: Frequency of pronouns in different types of articles on the *Times Online* and the *Mail Online*, normalised per 10,000 words (and absolute)

News site		**1st sg.**		**1st pl.**		**2nd**		**Total**	
Category	Words	norm.	(abs.)	norm.	(abs.)	norm.	(abs.)	norm.	(abs.)
Times Online									
News	(22,443)	43.2	(97)	80.6	(181)	15.1	(34)	139.0	(312)
Columns	(26,071)	130.4	(340)	71.3	(186)	66.7	(174)	268.5	(700)
Soft news	(29,045)	272.7	(792)	46.1	(134)	70.2	(204)	389.1	(1,130)
Mail Online									
News	(29,607)	55.1	(163)	56.7	(168)	17.2	(51)	129.0	(382)
Columns	(27,588)	124.3	(343)	68.1	(188)	54.0	(149)	246.5	(680)
Soft news	(12,743)	304.5	(388)	62.0	(79)	67.5	(86)	434.0	(553)

Author submitted manuscript

in the form of first person narratives. Columns also have a considerably higher frequency of first person singular pronouns than news articles do ($p < 0.001$). Here, the use of first person singular pronouns may be connected to the argumentative nature of the texts. For second person pronouns, a similar pattern can be observed. The difference between news and columns is significant at $p < 0.001$, but the difference between soft news and columns is smaller and not statistically significant. The first person plural pronouns show a less clear pattern.

As I have shown in section 8.1, *we* and *you* can have a range of different interpretations. In order to interpret the findings for these forms, their use needs to be investigated more closely, taking into account whether they are used with referential, vague or generic meaning. In addition, it is important to distinguish between pronoun uses inside and outside of quotes in order to interpret the findings in a meaningful way. As I mentioned in section 8.2, it has often been pointed out that first and second person pronouns occur in news articles mainly in direct speech quotes. Differences in the frequency of first and second person pronouns across different newspapers and across time have, therefore, usually been attributed to differences in the use of direct speech. As my findings of chapter 7.3.2 show, the news sites do indeed differ in their use of direct speech, with mid- and down-market sites containing more direct speech than most up-market sites (with the exception of the *Guardian*). Therefore, investigating the use of pronouns inside and outside of direct speech

Author submitted manuscript. Consult published version for correct page numbers!

separately will allow me to show whether differences in the use of direct speech are a sufficient explanation for the observed differences in the use of first and second person pronouns in news articles. Furthermore, looking at first and second person pronouns outside of direct speech might also indicate different strategies of personalisation across different types of articles.

8.4 First and second person pronouns within direct speech in news articles

Table 8.7 summarises the frequency of first and second person pronouns that occur within direct speech in the news articles of my data. The normalised frequency is calculated on the basis of the number of words that occur in direct speech only. Thus, how much direct speech is used in the articles from the various sites does not affect the normalised frequency. It is, therefore, noteworthy that the general differences that were observed with respect to the overall distribution of the pronouns (Table 8.5) can be found again in this evaluation. The pronouns are generally less frequently used in the data from the *Times* from 1985 than in the online data, although there are marked differences for the different forms. With 217 instances per 10,000 words, the frequency of first person plural forms is only slightly lower than for the *Guardian* and *BBC News*, and even higher than for the *Sun* (all these differences are not statistically significant). In contrast, first person singular and second person forms are far less frequent in direct speech from the *Times* from 1985 than in direct speech from the online sites. For first person singular pronouns, the difference is significant at $p < 0.01$ for the *Times Online* and at $p < 0.001$ for the other sites. For second person pronouns, the difference is not significant for the *Guardian*, but it is significant at $p < 0.01$ for the other up-market sites and at $p < 0.001$ for the *Mail Online* and the *Sun*. Within the online data, first person singular pronouns are again used more frequently on mid- and down-market sites than on the up-market sites (significant for the *Sun*), and a similar but less pronounced and not statistically significant tendency can also be observed for the second person pronouns. The first person plural pronouns again show a less clear pattern.

Author submitted manuscript

Table 8.7: Frequency of pronouns in news articles occurring inside of direct speech in five online news sites and the *Times* from 1985, normalised per 10,000 words (and absolute)

		1st sg.		**1st pl.**		**2nd**		**Total**	
News site	Words	norm.	(abs.)	norm.	(abs.)	norm.	(abs.)	norm.	(abs.)
Times Online	(5,179)	139.0	(72)	343.7	(178)	65.6	(34)	548.4	(284)
Guardian	(6,686)	166.0	(111)	231.8	(155)	47.9	(32)	445.7	(298)
BBC News	(4,425)	194.4	(86)	237.3	(105)	63.3	(28)	494.9	(219)
Mail Online	(7,459)	218.5	(163)	219.9	(164)	68.4	(51)	506.8	(378)
Sun	(4,886)	282.4	(138)	171.9	(84)	81.9	(40)	536.2	(262)
Times 1985	(3,220)	62.1	(20)	217.4	(70)	15.5	(5)	295.0	(95)

Author submitted manuscript. Consult published version for correct page numbers!

Table 8.8: First and second person pronouns in quotes by private and public sources

	Official sources		**Private sources**	
	abs.	in %	abs.	in %
Number of quotes	377	100.0%	98	100.0%
Containing any pronoun	165	43.7%	68	69.4%
Containing 1st sg.	57	15.1%	39	39.8%
Containing 1st pl.	108	28.6%	30	30.6%
Containing 2nd	26	6.9%	13	13.3%

These data demonstrate that there are two factors that make first and second person pronouns more frequent in news articles on the mid- and downmarket sites. On the one hand, direct speech, in which most of these pronouns occur, is generally more frequent there than on the up-market sites (see chapter 7.3.2). On the other hand, first person singular and second person forms occur with a higher frequency in direct speech from mid- and down-market news articles. Similarly, comparing the data from the *Times* from 1985 with those from the *Times Online*, it becomes clear that the increase in first and second person pronouns is not only due to a more frequent use of direct speech. Within direct speech first and second person pronouns are also clearly more frequent in the more recent online data. Consequently, the frequency of direct speech alone is not sufficient as an explanation for differences in the frequency of first and second person pronouns in news articles.

One factor that can be expected to correlate with the likelihood of first and second person pronouns appearing within direct quotes is the source type. Private sources tend to be quoted with statements about their personal experiences, opinions and emotions whereas official sources are often quoted in relation to general assessments of a situation or to provide information from official institutions. It therefore seems plausible that quotes by private sources contain more first and second person pronouns than quotes by official sources. In order to test this hypothesis, I determined how frequently first and second person pronouns occur in the quotes that I had classified with respect to their source type in chapter 7.4. 69 percent of the quotes that were classified as having a private source contain at least one first or second person pronoun. For quotes by official sources, the frequency is considerably lower, at about 44 percent (see Table 8.8). This difference can be found for all the pronouns, even though it is only significant for first person singular pronouns ($p < 0.001$). An exploratory analysis indicates that this pattern also seems to hold across all the online sites (see Table 8.9). For a systematic comparison, however, more quotes would need to be taken into account, especially since there were only few quotes by private sources in the data from the up-market news sites.

Other differences in the use of quotes can also have an effect on the frequency of first and second person pronouns. In particular the use of mixed quotation, combining indirect

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Table 8.9: Percentage of quotes containing any first or second person pronoun across five news sites

	Official sources		**Private sources**	
	in % of all official quotes	(abs.)	in % of all private quotes	(abs.)
Times Online	51.1%	(45/88)	85.7%	(6/7)
Guardian	47.3%	(44/93)	100%	(7/7)
BBC News	34.1%	(30/88)	55.6%	(5/9)
Mail Online	44.4%	(24/54)	74.4%	(32/43)
Sun	40.7%	(22/54)	56.3%	(18/32)

Author submitted manuscript

and direct speech, is closely related to the use of first person pronouns. I already presented examples in chapter 7.2 in which mixed quotation served to present the subject of the quoted sentence in indirect speech, while the remaining part is quoted directly. I have argued that presenting the main part of the quote in direct speech can be a strategy to distance the news report from the content of the statement. At the same time, presenting the subject of the statement in indirect speech can increase the distancing effect, especially when the source is the subject of the sentence. In direct speech, the subject would be realised in the first person, which increases the immediacy of the statement. In the case of first person plural subjects, it is also possible to interpret *we* as a vague pronoun, which would allow the audience to include themselves in the referent of the pronoun. To present the subject in indirect speech and third person makes the reported statement appear less immediate and precludes the possibility for an inclusive reading. The effects of mixed quotation can, therefore, be illustrated with the following example:

(8.15) The Taliban have claimed responsibility for this attack so clearly aimed at the Afghan government and innocent civilians. Their disregard for Afghan lives is deplorable. *We* will continue to stand with the Afghan people and their government and with *our* allies and partners around the world to defeat *our* common enemy and build a more secure and prosperous future. (US embassy statement on attacks in Kabul, 18 January 2010[3])

(8.16) The US embassy said: "The disregard for Afghan lives is deplorable." *It* would continue to work with allies and partners "to defeat *our* common enemy and build a more secure and prosperous future". (guardian-wn-100118)

Example 8.15 contains part of a statement by the US embassy in Kabul on attacks that took place in the city, and example 8.16 contains the representation of this statement in a news report by the *Guardian*. What is interesting is the representation of the second part of the statement in the form of mixed quotation. The subject of the sentence is included in the indirect part of the quote, so that *We* is changed to *It*. The first possessive pronoun in *our allies* is dropped completely, but the second possessive pronoun is included in the direct part of the quote (*our common enemy*). The first two uses of the first person pronoun in the original statement refer to the United States. They can be interpreted to have referential

[3] http://kabul.usembassy.gov/pr_1801.html, accessed on 14/4/2011

Author submitted manuscript. Consult published version for correct page numbers!

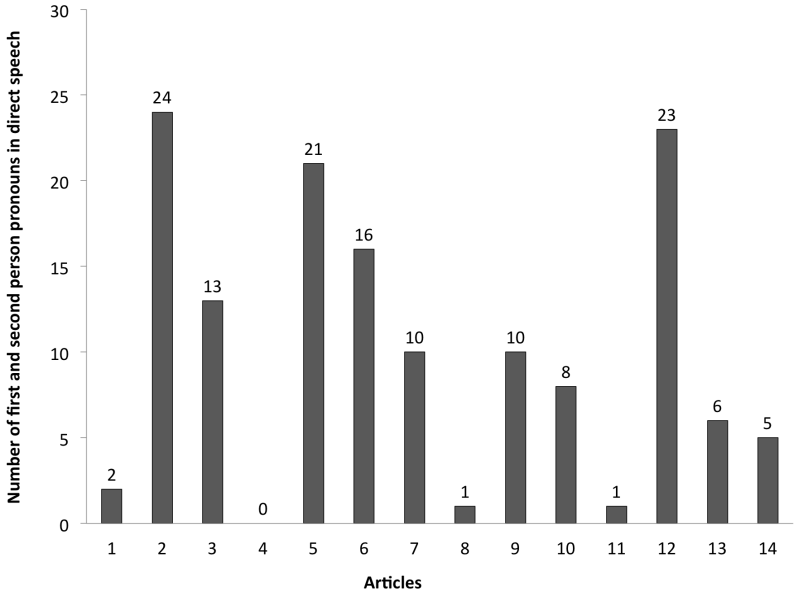


Figure 8.1: Distribution of first and second person pronouns in direct speech across 14 articles in the UK news section of the *Times Online*

meaning, referring to specific US government institutions and their representatives. They could, however, also be interpreted to refer to 'the US people and their government' in analogy to *the Afghan people and their government*. In this case, they have vague meaning and a US audience could possibly read these pronouns as including themselves. For the UK audience targeted by the *Guardian*, such an interpretation is less plausible, and indeed the *Guardian's* representation of the statement precludes such an interpretation by the use of the third person pronoun *It*. In contrast, the possessive pronoun in *our common enemy* still allows for a vague interpretation. Given that the UK is one of the US's allies in Afghanistan and that the Taliban are therefore also an enemy of the UK, it is possible for readers in the UK to interpret this pronoun inclusively.

Of course, direct speech sometimes contains first person plural pronouns which are clearly exclusive of the targeted audience of the news sites. However, examples like 8.16 illustrate that the choices involved in speech representation and, in particular, with respect to the use of pronouns can have an effect on the possibilities for the audience to identify with news actors. The use of first person pronouns lets the news actor appear more immediate and can even allow for inclusive readings, which is not the case for representations in the third person. The choice of pronouns in direct speech therefore depends not only on the type of source (e.g. official vs. private sources), but also on whether or not the source is someone with whom the news site invites their readers to identify. This also explains the tendency to present statements by extremist groups in the form of mixed quotation.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

The use of first and second person pronouns varies considerably across the individual articles. Figure 8.1 illustrates the absolute frequency of first and second person pronouns in direct speech in all 14 UK news articles from the *Times Online*. The number of first and second person pronouns varies between 0 and 24, with three articles containing almost half of all occurrences (68 of 140). These are also the three articles with the largest number of words in direct speech, whereas those with 0 to 1 occurrences contain almost no direct speech. However, articles with a similar number of words in direct speech also differ with respect to the number of first and second person pronouns. Article 7, for instance, contains 710 words in direct speech, which is slightly fewer than article 1 with 907 words. The number of first and second person pronouns is nevertheless five times higher in article 7 than in article 1. A closer look at the articles with many first and second person pronouns shows that the pronouns tend to occur in clusters. Thus, there are some quotes which are particularly rich with respect to first and second person pronouns. Five out of the ten occurrences of first and second person pronouns in article 7 occur in two quotes by the same source, immediately following each other (see example 8.17). Moreover, as examples 8.18 and 8.19 show, other articles contain direct speech with five or six pronouns in a single quote.

(8.17) "Of course *we* recognised that this would be a controversial appointment," an RMJM source said. "But *we* saw an advantage in recruiting someone with global business experience who had demonstrated a proven ability in business acquisitions to help *us* to develop *our* international markets. Do international people view Sir Fred in the same way as he is seen by the media? *I* don't think so." (times-uk-100116, article 7 in Figure 8.1)

(8.18) "*I* think *we* are now on a path to recovery. *I* am confident that the steps *we* have taken have put *us* on the right path." (times-uk-100126, article 12 in Figure 8.1)

(8.19) "While *I* have learned from Spotlight for the first time some alleged aspects of *my* wife's affair and her financial arrangements, *I* will be resolutely defending attacks on *my* character and contesting any allegations of wrongdoing. To that end, *I* will be addressing the media today after *I* have taken legal advice." (times-uk-100108, article 3 in Figure 8.1)

This makes it difficult to generalise claims about the personalising effects of the frequency of first and second person pronouns across the different news sites. A high frequency of first and second person pronouns in direct speech should be seen as an indication that this site has some news articles containing quotes with many pronouns. It does not necessarily mean that the articles in general contain more personalised quotes.

This difficulty is further increased by the fact that not all uses of these pronouns serve a personalising function to the same degree and with the same effects. How first and second person pronouns create personalisation depends crucially on their referents. In the following, I am therefore going to present and discuss selected examples of the different types of pronoun uses occurring in direct speech in news articles across the various news sites. As I will explain below, it will not be possible to establish frequencies for the different uses. Instead, my aim will be to illustrate for each pronoun its most typical uses, as well

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

as noteworthy exceptions, and to discuss their relevance for personalisation. This will then enable me to draw further conclusions from the quantitative data presented above.

As outlined in section 8.1, first person plural pronouns can generally be used with specific reference, vague reference or with generic reference. The generic use can hardly be found in the quotes from my news articles. The following two examples from *BBC News* could perhaps be considered for this category.

(8.20) "If *we* speak of the dead, the competition should have been cancelled. But the Confederation of African Football (Caf) has decided otherwise." (bbc-wn-100110)

(8.21) In summing up, the judge said the background of "tragedy and grief" will have struck a chord with all who had heard it. He said: "It would be extraordinary if *we* didn't feel empathy with the family and what Mrs Inglis had to face." (bbc-uk-100120)

In example 8.20, *we* does not refer to a specific or vague group of individuals, but rather to an ethical standpoint. The impersonal formulation *If one considers the dead* is more or less equivalent in meaning to *If we speak of the dead*. Therefore, this use of *we* can be classified as generic. Example 8.21 is more difficult to classify. One interpretation is that the pronoun *we* refers to those who were present at court. It might also be possible to interpret this pronoun generically, if one takes this statement to represent a generalised experience: 'anyone who hears about this case must feel empathy'. However, even with this interpretation, the example can at best be considered a borderline case.

Given that generic uses can hardly be found, practically all uses of *we* in news quotes are referential or vague. Since many cases are difficult to classify, it is not possible to present frequencies for the different categories of pronoun reference. The difficulty to classify first person plural pronouns according to their reference is also noted by Westin (2002, 44), who, consequently, decides to present only rough estimates about the different uses of *we* in her editorials. Likewise, I will only comment on the frequency of these different uses in very general terms. In this sense, it can be said that referential uses seem to be more frequent overall and that there are no clear trends with respect to the distribution of vague and referential meaning across the different news sites. The following examples illustrate the use of referential *we* in news quotes in my data.

(8.22) She [Kate McCann] said: "If I'm honest, *our* daughter's been taken and nothing's ever going to be as bad as that." (sun-100114a)

(8.23) "The PM and I met this afternoon and *we* discussed how *we* take forward economic policies to secure the recovery. [...]" (bbc-wn-100106)

(8.24) Lord Chief Justice Lord Judge, and two other judges, said the case was one of "true exceptionality" and overturned Hussain's 30-month jail sentence, replacing it with 12 months' imprisonment suspended for two years, with a supervision requirement for two years. [...] "*We* have come to the conclusion that *we* have ample justification for ordering that it [Hussain's sentence] should be suspended." (guardian-uk-100120)

(8.25) Adebayor told a French radio station: "*We* all decided to do something good for the country and play to honour those who died. Unfortunately, the head of state and the country's authorities have decided otherwise. *We* will pack up and go home." (times-wn-100110)

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

(8.26) Martin Narey, chief executive of children's charity Barnardo's, said he was encouraged by how the pilots went. He said: "I was very cautious about this experiment, and *we* at Barnardo's were worried about the possibility this scheme might drive sex offenders underground, away from police and probation supervision, and put children in danger." (bbc-uk-100124)

(8.27) Mr Michael Eaton, its [the National Coal Board's] chief spokesman, said: *We* are encouraged by the number of people returning to work. *We* have been disappointed by the attitude of some of the leaders of the NUM over the weekend, as they have restated their stance of 'no movement' from the March 6, 1984 position – and give *us* no opportunity to enter into a new round of negotiations." (times-uk-850108)

(8.28) A spokesman [for Northumbria Police] added: "*We* would also wish to remind people about the dangers of walking in the current weather conditions of deep snow and very cold temperatures, particularly at night." (mail-uk-100110)

As these examples show, referential uses of *we* occur in quotes from all news sites. The examples also illustrate the variety of referents the pronoun can have. The referent can consist of only two individuals, as in example 8.22, or of large groups, such as the Northumbria Police in example 8.28. Sometimes the referent of *we* is specified explicitly by the speaker, e.g. in example 8.26 *we at Barnardo's*. The examples can further be distinguished into cases in which the pronoun refers to a plural speaker and cases in which it refers to the speaker and a third party. In example 8.23, the referent includes the speaker and a third person, *The PM and I*. In examples 8.27 and 8.28, in which spokespersons are quoted, the pronoun refers to a collective, the National Coal Board and the Northumbria Police, respectively. For these last examples, in which *we* refers to a larger group, it is not entirely clear whether they should still be considered referential uses, or whether they are rather vague uses. Vague reference is given when the pronoun refers to a vaguely defined collective, e.g. a group sharing a common nationality or political orientation. Depending on the size of the group, its character and the way in which it is introduced in the text, it can be difficult to decide whether it functions as a specific referent or as a vague referent. In general, no clear dividing line between specific and vague reference can be drawn, which is one of the reasons why it is hardly possible to classify all occurrences of *we* systematically into referential, vague and generic uses.

Still, there are instances of *we* which can more or less clearly be considered vague, especially when the referent is defined through a common nationality. In example 8.29, nationality is explicitly mentioned as the defining characteristic of the referent of *we* in *we as a nation*. In 8.30, the defining characteristic also seems to be nationality, even though this is less explicit and can only be inferred from the context. In 8.31, *we* can be taken to refer to everyone located in the UK, since this is the group facing the cold temperatures that are announced. The referent in 8.32 is established through a common affiliation to a political party. In 8.33, the referent is less clearly defined. The quote occurs in an article announcing the newspaper's launch of a charity single to raise money for the victims of the earthquake. Therefore, it is possible that the implied referent corresponds to the target audience of the *Sun*. The pronoun could also be interpreted in a broader sense, referring to people in general. In this case, *we* would have generic reference.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

(8.29) "*We*, sadly, face a series of threats, the nature of which will require the projection of power beyond *our* borders to protect *our* national security," he [British Armed Forces Minister] said. "My great fear is that *we* as a nation will become so risk-averse, cynical and introverted that *we* will find *ourselves* in inglorious and impotent isolation by default." (times-uk-100114)

(8.30) "I am saying that *we* must come to a prosperous coal industry. There is coal for a prosperous industry in good pits, good seams, good machinery, safe conditions . . . the older uneconomic pits have to be close [sic] down." (times-uk-850125)

(8.31) "*We* are looking at temperatures of -2C in London with parts of England and Wales falling to -8C or -9C," he [spokesman for the Met Office] said. (mail-uk-100106)

(8.32) He [Geoff Hoon] said: "This is about ensuring that the Labour Party can go united into what is in any event going to be a difficult campaign supporting a strong leader that *we* can all back." (bbc-wn-100106)

(8.33) SuBo [Susan Boyle], who had the top-selling album in the UK and the US last year, said: "The Haiti earthquake has caused devastation and suffering that most of *us* will never experience nor could imagine." (sun-100124a)

In all these examples, the target audience is at least potentially included in the referent. Whether or not a reader interprets the *we* in example 8.32 inclusively depends on his or her political orientation. For 8.29 to 8.31, an inclusive interpretation is quite likely, since the implied reader of these news publications is British and located in the UK. However, anyone with an Internet connection obviously is able to read these articles and many actual readers might not interpret these uses inclusively. For a reader from the target audience of the news sites, an inclusive interpretation is, nevertheless, possible. In contrast, in the following two examples, the target audience is not included in the referent of *we*.

(8.34) In its report, the [French parliamentary] committee said requiring women to cover their faces was against the French republican principles of secularism and equality. "The wearing of the full veil is a challenge to *our* republic. This is unacceptable. *We* must condemn this excess," the report said. (bbc-wn-100126)

(8.35) Trecile, one of several inspectors general of Haiti's police force, which has 9,000 officers, said it was crucial for Haitians to remain calm and united. "No one is responsible for this, there is no one to blame, *we* are all in this together," he said. "Those of *us* who have survived need to be courageous to cross this difficult step, show solidarity and help one another." (guardian-wn-100116)

In these two examples, the referent of *we* is again vaguely defined through nationality. In 8.34, *we* refers to all French and, in 8.35, to all Haitians. In both cases, the implied reader of the British online news sites will interpret these pronouns exclusively.

First person plural pronouns can, thus, mostly be found with specific and vague reference in news quotes. Some of the vague uses allow the audience to include themselves in the referent, which is interesting from the point of view of personalisation. In this case the readers perceive themselves as being addressed by the news actor and align themselves with him or her. This increases the involvement on the part of the audience. In contrast, exclusive uses of vague *we* have a similar effect to referential uses. They refer to news actors in the first person and thus emphasise their status as persons who act, decide or

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

experience a situation personally. As the examples above have illustrated, all these uses can be found in all the news media and no clear tendencies with respect to the distribution of these different uses could be determined.

While first person plural pronouns in the news quotes of my data are hardly ever used with generic reference, this is quite common for the (overall less frequent) second person pronouns. The following examples 8.36 to 8.40 illustrate this use.

(8.36) Mr Holland said: "There is an apparent Catch-22 here. The prosecution has to be initiated within six months but *you* have to exhaust the university's complaints procedure before the commission will look at *your* complaint. That process can take longer than six months." (times-uk-100128)

(8.37) Pat Lang, a veteran former head of analysis and clandestine human intelligence for the Defence Intelligence Agency, echoed Johnson's criticism of the Khost operation. "A number of basic rules were violated. One that comes to mind is *you* never trust foreign agent assets," he said. (guardian-wn-100106)

(8.38) The Local Government Association warned last night that salt is actually ineffective against ice below -8c (18f). "If it gets colder than that at the weekend *you* could dump the entire world's supply of salt on the roads and it wouldn't help," a spokesman said. (mail-uk-100108)

(8.39) TV presenter Annabel Giles was among the hundreds of drivers stranded there. She was stuck in her car on the A3 for more than 22 hours en route to Brighton. She told the BBC News website: *You* can't sleep because the traffic keeps moving on a couple of inches and everyone starts beeping their horns and *you* have to wake up and get into gear and move on". (bbc-uk-100106)

(8.40) "I don't think it will ever be safe to release them. I am convinced they will do it again. They loved what they did. *You* could see it on their faces in court. They were smiling. They are bloodthirsty." (sun-100124b)

In examples 8.36 and 8.37, the second person pronoun is used to formulate rules that are generally valid. Similarly, in 8.38, *you* refers to no one in particular, but stands for a general truth. Here, as in the previous two examples, *you* could be replaced with *one* without changing the meaning of the sentences. Still, the effect would be different. Due to the presence of *you*, the reader is invited to imagine himself or herself in the position of the subject. This invitation for the identification with the subject becomes particularly salient in example 8.39, where generic *you* is used to generalise the personal experiences of the speaker. The audience is expected to sympathise with the news actor who describes how she was stuck in her car in heavy snow. Likewise, in example 8.40, the speaker generalises his personal experience by describing it in the second person. The source of this quote is the father of one of the victims in the Edlington case and he describes the offenders in court. By saying *You could see it on their faces* rather than *I could see it on their faces* he claims that the same observation could have been made by anyone, including the audience.

In contrast to generic uses, referential uses of *you* do not include the audience. In all cases, referential *you* in direct speech in news articles is used by one news actor to address another news actor. This can be a single addressee, as in 8.41 and 8.42, or a plural addressee, as in 8.43 and 8.44.

(8.41) Jean-Pierre turned to his brother by his bedside and told him: "When *you* are in a hole, I will try to reach out to *you*, too." (guardian-wn-100124)

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

(8.42) Day pleaded guilty to a charge read out in court as: "On 11 November 2009 *you* wore a decoration on a badge, on a wing stripe and or on an emblem so nearly resembling a military decoration, badge, stripe or emblem as recognised by defence council as to be calculated to deceive contrary to section 197 1B of the act 1955." (mail-uk-100112)

(8.43) "The bottom line for the two of *you* is that I'm sure *you* both pose a very high risk of serious harm to others." (bbc-uk-100122)

(8.44) "If *you* can persuade *yourselves* that activities like today's can enhance the miners' case," Mr Kinnock told the meeting, "if *you* are intent on going ahead, there is nothing I can do in the world to prevent fully grown men and women." (times-uk-850118)

This use can be found on all the online news sites, as well as in the printed *Times* from 1985. A frequent context in which it occurs are reports from court cases, as in 8.42 and 8.43. However, as example 8.41 illustrates, referential *you* is also used in the narration of personal experiences by private news actors. A special case of referential *you* is given in the following example from the *Sun*.

(8.45) Howard Stalker, 43, who ferried our two-man team in on his quad bike, said: "We can't thank *you* enough." (sun-100108b)

Here the addressee is the newspaper itself – or its representatives, – since the *Sun* appears as a news actor in its own story when bringing supplies to a group of people who were snowed in at a pub. Self-referentiality and the presentation of the newspaper as an actor in the news are typical features of tabloid journalism (Conboy 2006, 37-40) and it is probably no coincidence that this example was found in the down-market *Sun*. Such uses are therefore highly relevant for the personalisation of the news site, even if they occur only rarely.

Vague uses of *you* refer to a group of addressees that is not clearly defined. Again, it is often difficult to distinguish between referential and vague uses. In example 8.46, for instance, *you* could either refer to the US president Obama, to whom the tape with the message was sent, or to all Americans, to whom the message later refers. In the first case, *you* would have specific reference, in the second case it would have vague reference based on nationality. Examples 8.47 and 8.48 are clearer cases of vague *you*. In 8.47, the vague referent is defined on the basis of party affiliation, and, in 8.48, the basis is a combination of nationality and being a parent.

(8.46) He [a man claiming to be Osama bin Laden] said: "The message sent to *you* with the attempt by the hero Nigerian Umar Farouk Abdulmutallab is a confirmation of our previous message conveyed by the heroes of September 11." "If it was possible to carry our messages to *you* by words we wouldn't have carried them to *you* by planes." (mail-wn-100124)

(8.47) He also lashed out at Republicans even as he appealed yet again for a new era of bipartisanship. In a reference to the Massachusetts seat and the need for a "supermajority" to pass any legislation, he said: "If the Republican leadership is going to insist that 60 votes in the Senate are required to do any business at all in this town, then the responsibility to govern is now *yours* as well. Just saying no to everything may be good short-term politics, but it's not leadership." (times-wn-100128)

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

(8.48) "We are saying in law we will set out a guarantee to parents: if *your* child falls behind, our investment will ensure *your* child gets that extra help," Balls said. (guardian-uk-100104)

As with vague *we*, whether or not readers include themselves in the referent of vague *you* depends on who they are. For a reader of the British target audience, an inclusive interpretation is less likely in 8.46 and 8.47, since the potential referent in these examples is American. A case could possibly be made for an inclusive interpretation in 8.46 for British readers who might identify with the American position, given that the UK has also been a target of terror attacks from related groups. In 8.48, readers from the target audience will be quite likely to include themselves in the addressee if they have children or plan to have children in the future. A special case in which readers are explicitly addressed can be found in the *Sun*.

(8.49) Lescinel, 23, said: "To see them give so generously is very important. To all Sun readers, please make a prayer and send anything *you* can to help make a difference." (sun-100124a)

In this example, the referent is still vague, but, since the referent is defined as everyone reading the *Sun*, all readers are part of the addressee. Similar to example 8.45 above, the *Sun* appears as an actor in the news by launching a charity single for the victims of the earthquake in Haiti. Directly addressing the *Sun* readers also turns them into news actors by suggesting that they can *help make a difference*. Again, despite the low frequency of this use, it is nevertheless an important factor in the personalisation of the audience of the *Sun*.

In terms of personalisation, uses of *you* that include the audience can again increase the involvement of readers. This is particularly true for generic uses that invite readers to identify with the described positions. Referential uses and vague uses excluding the audience, in contrast, point to news actors as addressees in news events and emphasise the interaction between different news actors. Again, the distribution of these different forms cannot be clearly determined since there are too many ambiguous cases. In addition, the printed *Times* from 1985 contains only 5 occurrences of second person pronouns in news quotes, which does not allow for any generalisations – apart from the fact that *you* is generally quite infrequent in these data. What can be said is that all online news sites contain all groups of uses, with vague uses generally being least frequently observed. Moreover, there seems to be a tendency for the *Sun* to contain more referential uses and fewer generic uses than the other sites.

In contrast to *we* and *you*, first person singular pronouns can only be used with specific reference to a single speaker – with the exception of very rare generic uses, which were not found in my news data. Their interpretation, therefore, poses fewer problems than the interpretation of the other pronouns. They always refer to a single news actor and it can be expected that direct speech in the first person will result in a highly personalised presentation of this news actor.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

A closer look at the first person singular pronouns that occur in direct speech in my news articles shows that they are, indeed, highly relevant for personalisation. This is not only the case because they refer to an individual speaker, rather than a group, but also because of the context in which they are used. In most cases, first person singular pronouns are used to introduce a personal perspective on the part of the speaker. This can take place in the form of a personal judgement of a situation, expressing commitment and responsibility, presenting personal experiences or by expressing the speaker's emotions. In the following, I am going to present and discuss examples of each of these contexts.

Examples 8.50 to 8.55 illustrate the use of first person singular pronouns to express the opinion or judgement of the speaker.

(8.50) "Obama is a smart guy. *I* don't have a problem with his foreign policy, but it's his approach to the economy that worries *me*. *I* am a businessman and a free-market guy and *I* don't think Obama sees it like *me*," Rocco Gianni, a salesman, said. "*I* think we" need healthcare reform but *I*'m concerned about the form the current bill is in." (times-wn-100120)

(8.51) "In *my* opinion, that use of force had not been authorised by the security council, and had no other basis in international law." (guardian-uk-100126)

(8.52) Chancellor Alistair Darling said: "As far as *I*'m concerned we should be concentrating on the business of government and getting through the recession." (bbc-wn-100106)

(8.53) David Cameron said such a march would be unacceptable. "*I* think this group are just saying that because they want to get some cheap publicity. Their views are completely reprehensible to the overwhelming majority of not just the British public, but British Muslims as well." (mail-uk-100104)

(8.54) "*I* wouldn't say she's a mum. *I*'m embarrassed to say she's *my* mother. She doesn't clean or cook, she's not capable of doing anything." (sun-100122b)

(8.55) Mr Peter Rogers, head of Manufacturers' Hanover Trust's foreign exchange department in New York said: "The British Government has got a breathing space rather than a permanent turnaround for the pound. *I* can't believe, the Government would allow it, but the market here believes parity can happen." (times-uk-850116)

In this use, *I* most frequently occurs in the constructions *I think* and *I don't think* (e.g. in examples 8.50 and 8.53). This is not surprising, given that *I think* has been identified as one of the most frequent markers of (epistemic) stance in spoken interaction (Baumgarten and House 2010, 1186; Biber et al. 1999, 984-985; Kärkkäinen 2003, 5; see also Aijmer 1997; Fetzer and Johansson 2010). As the examples show, many other formulations can also be found, including *in my opinion* (8.51), *as far as I'm concerned* (8.52), as well as less routinely used expressions of speaker judgement, such as *I wouldn't say* (8.54). The speakers are often politicians or experts, but there are also private sources who express their opinion, for instance the speaker in example 8.54.

The second context in which *I* frequently occurs in news quotes is when speakers are taking over responsibility and/or expressing commitment. This occurs almost exclusively with politicians, which again is not surprising, since this is the main group of actors who are interviewed in news articles with respect to events in which they carry responsibility.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Of the following examples, only 8.61 and 8.60 are not statements by a politician. However, as the national co-ordinator of the WHO in Ethiopia, the source in 8.61 is also clearly a public news actor. Example 8.60, by contrast, is a quote from a private source, a woman accused of killing someone while she was drunk.

(8.56) Bill Clinton is due to arrive today. "As UN Special Envoy for Haiti, *I* feel a deep obligation to the Haitian people to visit the country. . . to ensure our response continues to be co-ordinated and effective," he said. (times-wn-100118, ellipsis in original)

(8.57) With his party still reeling from the loss of a crucial Massachusetts Senate seat Mr Obama was forced to restate the soaring themes of his election victory in the language of a fighter. "We have finished a difficult year," he said. "We have come through a difficult decade. But a new year has come. A new decade stretches before us. We don't quit. *I* don't quit. Let's seize this moment to start anew." (times-wn-100128)

(8.58) While he said he would defend the September dossier "till the end of *my* days", Campbell held his hands up with respect to the "dodgy dossier" published in February 2003, which included large chunks lifted wholesale from an academic paper, blaming it on a "mistake" by a member of the Coalition Information Committee. (guardian-uk-100112)

(8.59) The home secretary said: "Protecting children and families from sex offenders is one of *my* top priorities and the UK already has one of the most robust systems of managing sex offenders in the world." (bbc-uk-100124)

(8.60) "What I did terrifies me," Davies told the court. "*I* can guarantee that *I* will never again touch a drop of alcohol." (mail-wn-100112)

(8.61) Following yesterday's article in the Times on the outbreaks, which are reported in 12 different camps, the World Health Organization has become involved. Yesterday Dr Mesfin Dmise, the national co-ordinator, said: "*I* asked the Ministry of Health to let *me* know the results of their inspection of the camps last week." They promised to let him have them today, he said, and "if the tests are positive *I* will inform the WHO in Geneva. If the result is negative *I* might well carry out independent laboratory tests". (times-wn-850124)

Again, the examples vary with respect to the extent to which the commitment and responsibility is explicitly mentioned. This is, for instance, quite explicit in *I feel a deep obligation* (8.56), *I can guarantee that I will never again* (8.60) and *one of my top priorities* (8.59). In example (8.61), by contrast, the commitment of the source is only implied through the announcement of future actions. Particularly interesting from a rhetorical point of view is example 8.57, reporting on Obama's State of the Union address. After four uses of vague first person plural pronouns, he repeats the sentence *We don't quit* with the singular pronoun: *I don't quit*. Contrasting *we* and *I* in this way, the speech particularly emphasises Obama's personal commitment.

While uses of *I* to express commitment or take on responsibility occur mostly in quotes by official news actors, the use of *I* to recount personal experiences is often found with private sources (e.g. in examples 8.63 and 8.64). Again, this corresponds to the topics about which different types of sources are interviewed. Private sources often relate personal experiences and this is, therefore, also the context in which uses of *I* by private sources are found. However, *I* can be used to express personal experience with other types of sources

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

too. In example 8.65, for instance, a medical doctor speaks in the role of an expert about his personal motivation to investigate the spontaneous bleeding of an Indian girl.

(8.62) "*I*'ve been up and down this country and every door is closed," the priest said. (times-wn-100124)

(8.63) "Only a couple of weeks ago *I* had eggs thrown at *me* by white guys in a car as *I* was leaving," said Abdul Rahman, 52, the mosque's caretaker, who emigrated to the UK from the Caribbean. "They hit *me* in the chest and just drove off." (guardian-uk-100116)

(8.64) The 23-year-old said: "We went through hell. *I* am eight months pregnant, *I* couldn't go to the toilet all night, *I* couldn't warm the bottle up for *my* baby daughter. It was very frightening." (mail-uk-100106)

(8.65) He said: "*I*'ve never seen a case of someone who bleeds spontaneously from their scalp or their palms, or read about it in medical history."
"*I* was interested to see if *I* could help Twinkle." (sun-100110b)

(8.66) "[. . .] When they landed in Japan *I* would stand next to the immigration officer, helping with translations," Mr Alex Triguboff said. "Sometimes they'd just photographed one person's visa. Again and again *I* saw the same name and the same date on the visa." (times-wn-850119)

Unlike for some of the other uses of *I*, there are no specific formulations that are repeatedly used to express personal experience.

The last group of occurrences of *I* in direct speech in news articles is used to express emotions (8.67 to 8.72).

(8.67) "You came here to play football but unfortunately two of your brothers were killed," Mr Hayatou said. "*I* am feeling terribly bad and short of words." (times-wn-1001010)

(8.68) Mr McGuinness said: "*I* was shocked at the financial revelations made in last night's Spotlight programme." (times-uk-100108)

(8.69) The prime minister said it was "extraordinary" that so much money had been donated to the appeal in such a short time. "The generosity of people's spirit, our willingness to give even when some people have so little to give, humbles *me*," he told a meeting in London. (bbc-uk-100116)

(8.70) "The idea that anyone would stage this kind of demonstration in Wootton Bassett fills *me* with revulsion," said the Home Secretary. (mail-uk-100104)

(8.71) The decision was condemned by police, politicians and Sharon's widower Paul, who said: "*I* am appalled. *I* don't understand how these do-gooders who sit on the board can sleep at night." (sun-100116a)

(8.72) Mr Ford added: "*I* was delighted when Mr MacGregor said that he would back working miners to the hilt. [. . .]". (times-uk-850126)

What is particularly noteworthy about this use is that many of the sources are public news actors, in particular politicians. In the examples above, only the quote in 8.71 has a private source. This may be surprising insofar as politicians and other official news actors are usually not expected to comment on their emotions, but to give factual accounts and represent their party, institution or nation. However, public actors can express feelings in the first person, but still do this in a representative function. In example 8.67, Mr Hayatou, the president of the Confederation of African Football, uses a statement of personal emotion

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

to express concern over the death of two football players. Making this statement in his official role, this can be understood as being not only his personal concern, but as representing the concern of the association. Similarly, the emotion expressed in example 8.69 has a representative function, officially acknowledging the donations by people in the UK. Expressing personal emotions in news media can therefore be one way for public actors to fulfil representative functions related to their official roles.

Expressing judgement, commitment, emotions and personal experiences are, of course, not mutually exclusive and several of the examples above fall into more than one of these categories. For instance, making commitments is often done through expressing emotions, as in example 8.56 *I feel a deep obligation.* Similarly, expressing emotions often also serves to present judgements, for instance in *I am appalled* in example 8.71. For this reason, a classification of all occurrences of the first person singular pronoun into the four groups is difficult. However, as the examples above show, the four main contexts can generally be found across all the news media in my data. Moreover, these contexts seem to cover the large majority of occurrences of first person singular pronouns. In all these uses, *I* refers to the speaker and introduces a personal perspective. Thus, it can be concluded that first person singular pronouns occurring in direct speech in news articles serve a highly personalising function for the news actors who use them.

Summing up, one can say that the close investigation of individual uses of first and second person pronouns in quotes in news articles has resulted in some interesting observations. Second person pronouns, which are the least frequent forms in this context, can be found in all news media, even though they are quite rare in the *Times* from 1985. On all the online sites, they are regularly used with generic reference to formulate general rules or to generalise personal experience. I have argued that this use invites the audience to identify with the position that is described and, thus, can increase the involvement of the reader. It is noteworthy that this use is not restricted to particular news sites. Including quotes that describe rules or generalise personal experience in the second person, therefore, seems to be a personalisation strategy that is common to all the news sites. The only clear way in which the down-market *Sun* differs from the other news sites with respect to its use of second person pronouns in quotes is that it contains two special occurrences that cannot be found in this way elsewhere. In one case, *you* refers to the newspaper itself, which appears as an actor in its own news story; in the other case, it refers to the readers of the *Sun*, who are directly addressed in one of the quotes.

First person pronouns hardly ever occur with generic reference. Instead, they refer to news actors and an identification with the referent is only possible for the audience in the case of some of the vague uses of the plural pronoun. I have argued that both singular and plural forms of the first person pronoun tend to present news actors as actors in and experiencers of news events. Consequently, they have a personalising effect for news actors. This effect is stronger in the case of first person singular pronouns, which refer to single actors and are used to present their opinions, commitment, experiences and emotions in a very personal way. While it was not possible to identify clear differences in the uses of

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Table 8.10: Ratio of first person plural pronouns to first person singular pronouns inside direct speech in five online news sites and the *Times* from 1985

News site	1st pl.	1st sg.	Ratio
Times Online	178	72	2.5
Guardian	155	111	1.4
BBC News	105	86	1.2
Mail Online	164	163	1.0
Sun	84	138	0.6
Times 1985	70	20	3.5

Author submitted manuscript

first person pronouns across the different news publications, the ratio between plural and singular uses within each publication shows some interesting tendencies (see Table 8.10). There are clear differences across the sites with respect to the ratio of plural to singular uses. In all the up-market sites, first person plural pronouns are more frequent than first person singular pronouns. For the *Times Online*, the plural pronouns are 2.5 times more frequent than the singular forms, for the *Guardian*, the ratio is 1.4 and for *BBC News* it is 1.2. With 3.5, the ratio is even higher for the *Times* from 1985. The mid-market *Mail Online* contains about the same number of singular and plural pronouns. The down-market *Sun*, finally, has a ratio of 0.6, meaning that first person singular pronouns occur 1.6 times more often than first person plural pronouns. Statistical testing indicates that the differences between the *Guardian*, *BBC News*, and the *Mail Online* are not significant. Likewise the difference between the *Times Online* and the *Times* from 1985 is not significant. All the other differences are significant; in particular the differences between the *Sun* and the up-market sites are highly significant at $p < 0.001$ (see appendix B).

Thus, one can conclude from these figures that first person pronouns in quotes of news articles from the *Sun* show a tendency to refer to individual news actors, whereas first person pronouns in up-market sites tend to refer to groups. This means that overall the personalising effect of first person pronouns tends to be stronger in the *Sun* than on the up-market sites. It also reflects the *Sun's* focus on private news actors who speak for themselves, in contrast to official actors who represent a larger group.

8.5 First and second person pronouns outside of direct speech in news articles

Table 8.11 lists all the occurrences of first and second person pronouns that were found in online news articles outside of direct speech. There are only few occurrences of first person pronouns, and none at all of second person pronouns. First person singular pronouns occur once in the *Sun*, 25 times in the *Times Online* and three times in the *Times* from 1985. The occurrence in the *Sun* appears in a headline, reproduced in example 8.73. As becomes clear when reading the article (see example 8.74), the headline represents the statement of

Author submitted manuscript. Consult published version for correct page numbers!

Table 8.11: Frequency of pronouns in news articles occurring outside of direct speech in five online news sites and the *Times* from 1985, normalised per 10,000 words (and absolute)

		1st sg.		1st pl.		2nd		Total	
News site	Words	norm.	(abs.)	norm.	(abs.)	norm.	(abs.)	norm.	(abs.)
Times Online	(17,264)	14.5	(25)	1.7	(3)	0.0	(0)	16.2	(28)
Guardian	(16,065)	0.0	(0)	0.0	(0)	0.0	(0)	0.0	(0)
BBC News	(14,898)	0.0	(0)	5.4	(8)	0.0	(0)	5.4	(8)
Mail Online	(22,148)	0.0	(0)	1.8	(4)	0.0	(0)	1.8	(4)
Sun	(11,983)	0.8	(1)	5.8	(7)	0.0	(0)	6.7	(8)
Times 1985	(19,968)	1.5	(3)	0.5	(1)	0.0	(0)	2.0	(4)

a news actor, Wyclef Jean. Thus, this occurrence of *I* outside of quotation marks can be attributed to free direct speech with missing quotation marks in a headline.

(8.73) *I* saw apocalypse (sun-100130b)

(8.74) Wyclef, who flew to the country as soon as he heard news of the earthquake, said: "It looked like the apocalypse - there were bodies everywhere." (sun-100130b)

The 25 occurrences of first person singular forms in the *Times Online* all occur in a single article, which also contains the three occurrences of first person plural pronouns that were found in the data from this site. This is an article on the earthquake in Haiti which was presented in the world news section of the *Times Online*. Uncharacteristically for news articles on this site, it changes into first person narration in the sixth paragraph. The following example reproduces paragraphs five and six of the article, the passages immediately before and after the shift.

(8.75) A 24-year-old man was freed from beneath the collapsed Napoli Inn. The record for surviving an earthquake is 14 days set by a man trapped beneath a gym in the 1990 earthquake in the Philippines who had access to rain water. On Friday *I* had paid *my* second visit to the New Hope Ministry, hoping that the misery *I* had witnessed earlier in the week would have been brought under control: that the aid supplies spreading through the fast-rotting shantytowns at the centre of Port-au-Prince would have reached this forlorn outpost only a few miles out of town. (times-wn-100124)

The first two sentences in 8.75 are written in a style that is typically used in news reports on the site. A recent development is reported and it is contextualised with background information on a similar event in a previous catastrophe. After that the text suddenly shifts into a personal narration of the journalist's experiences when he did his background investigation in Haiti. This is the only case in which an article that I collected from the news section of an online site integrates personal experiences of the journalist in this way. Perhaps surprisingly, the only similar uses can be found in the printed *Times* from 1985, in which three occurrences of the first person singular pronoun refer to the journalist.

(8.76) Medical teams from the Ministry of Health last week inspected the infected areas, but so far they have issued no statement on their findings. Yesterday *I*

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

> visited Harbo, a new camp in a part of Wollo region which has only recently been affected by the more severe effects of a famine that is slowly spreading south. [. . .]
> The camp's senior doctor, Dr Zelaham Hawaz, told *me*: "It is just acute diarrhoea and acute vomiting. [. . .]". (times-wn-850123)

This example is in some ways similar to 8.75 above. It also reports on a dramatic situation in a non-European country, namely the outbreak of cholera in a refugee camp in Ethiopia. In this case, too, the experiences of the journalist are presented, even though in a much more marginal way.

The *Times Online* article that contains the first person singular occurrences also contains three occurrences of the first person plural pronoun. Apart from these, there are also 19 first person plural forms in news articles from *BBC News*, the *Mail Online* and the *Sun*, as well as one occurrence in the *Times* from 1985. This last occurrence in the printed *Times* does not occur in the main body of the article, but in a byline given as *By Our Foreign Staff*. Similarly, the eight instances on *BBC News* all occur in the formulation *our correspondent*, but in this case they appear within the main text.

> (8.77) "We can say that Britain has just crossed the line in coming out of recession," said BBC chief economics correspondent Hugh Pym. [. . .]
> *Our* correspondent said the move out of recession had been greatly boosted by the government car scrappage scheme. (bbc-uk-100126)

The formulation *our correspondent* points out that *BBC* conducted its own investigation and did not only publish a press release. In this particular instance, this effect is even increased through the fact that the correspondent is presented as *BBC chief economics correspondent Hugh Pym* earlier in the same article. These uses of *our correspondent*, which seem to be a regular feature in articles on *BBC News*, contribute to the personalisation of the news organisation by emphasising their active role and their networks with specific informants and experts.

An even stronger personalising effect is achieved through five of the seven uses of first person plural forms in the *Sun*. They occur in the same two articles that have been mentioned in the previous section and in which the *Sun* appears as a news actor in its own story. In the first case (8.78), a team from the *Sun* delivered food to a village cut off by the snow. In the second case (8.79), the *Sun* initiated fundraising for the victims of the earthquake.

> (8.78) Villagers in remote Cow Ark, Lancs – cut off for three weeks since heavy snow on December 17 – were overjoyed yesterday when The Sun came to their rescue.
> *Our* mercy crew delivered bread, milk, cheese, ham and other vital provisions - including Britain's favourite newspaper.
> Howard Stalker, 43, who ferried *our* two-man team in on his quad bike, said: "We can't thank you enough." (sun-100108b)
>
> (8.79) Football stars led by England captain John Terry have pledged thousands to *our* Helping Haiti appeal. Terry, 29, said: "These people desperately need our help. I urge every Sun reader to give their support if they possibly can." (sun-100124a)

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Again, such self-referential uses of *we* can only be found in the *Sun*. However, even though the up-market news sites do not use self-referential pronouns in their news articles, they still use self-referential expressions.

(8.80) An investigation by *The Times* has revealed that the trainee doctor was an open and public supporter of al-Qaeda, and that his months of secret work for Jordanian intelligence, feeding a stream of information about low-level al-Qaeda operatives to his handler, was a means to establish his credibility. (times-wn-100106)

(8.81) *The Guardian* has uncovered evidence that when the BNP wins elections, reports of racially and religiously aggravated crime in the local area often rise. It has happened to varying extents in BNP strongholds in the West Midlands, London and Essex despite falls in hate crime in the surrounding areas. (guardian-uk-100116)

In 8.80 and 8.81, the *Times* and the *Guardian* refer to themselves and present themselves as actors in the news event. Unlike the *Sun*, they exclusively use their names and no first person pronouns, which is one of the features that contribute to the less personal style of these examples, compared to 8.78 and 8.79.

Among the remaining six occurrences of first person plural instances, there are two instances of *we* in headlines of the *Mail Online* articles with missing quotation marks for direct speech. The other four pronouns occur in the *Sun* and the *Mail Online* and they all have vague reference, as in the following examples.

(8.82) But today, as hundreds of mourners lined the patriotic town of Wootton Bassett to pay a respectful tribute to *our* war dead, friends of a British soldier killed in Afghanistan spoke of their anger at Choudary's plan. (mail-uk-100104)

(8.83) Next week looks just as grim. And by then *we* could be out of grit - and GAS. (sun-100108b, capitalisation in the original)

These uses can be regarded as cases of patriotic *we* for which the referent is vague and implicitly defined through a common nationality.

In sum, first person pronouns occur very infrequently outside of direct speech in the news articles in my data. They do not occur at all in the *Guardian* and all occurrences in the *Times Online* can be attributed to a single article. In the *BBC News* articles, the formulation *our correspondent* is regularly used and serves a personalising function with respect to the news organisation. A few even more personalising uses of first person plural pronouns can be found in the *Sun*, which presents itself as a news actor dedicated to helping those in need. In addition, the *Sun* as well as the *Mail Online* contain some uses of patriotic *we*. Impersonal uses of first person pronouns were not found, and there were also no occurrences of second person pronouns outside of direct speech in the news articles.

8.6 First person singular pronouns across different types of articles

While first and second person pronouns are used rarely outside of direct speech in news articles, this is quite different for soft news articles and columns (see Table 8.12). First

Author submitted manuscript

person singular forms are particularly frequent, indicating that the authors of soft news articles and columns have a high degree of presence in these texts. A comparison with Table 8.7 even shows that the frequency of first person singular and second person forms outside of direct speech in soft news and columns is similar to their frequency within direct speech in news articles. First person singular forms are used 139-282 times per 10,000 words in direct speech in news articles. The frequency of first person singular forms outside of direct speech is within this range for soft news, whereas it is only slightly lower for columns. The frequency of the second person pronouns falls within the range of 48-82 instances per 10,000 words which was observed for second person pronouns in direct speech in news articles.

Table 8.12: Frequency of pronouns outside of direct speech in different types of articles in the *Times Online* and the *Mail Online*, normalised per 10,000 words (and absolute)

News site		**1st sg.**		**1st pl.**		**2nd**		**Total**	
Category	Words	norm.	(abs.)	norm.	(abs.)	norm.	(abs.)	norm.	(abs.)
Times Online									
News	(17,264)	14.5	(25)	1.7	(3)	0.0	(0)	16.2	(28)
Columns	(24,457)	125.9	(308)	72.0	(176)	54.8	(134)	252.7	(618)
Soft news	(23,624)	218.8	(517)	31.3	(74)	49.1	(116)	299.3	(707)
Mail Online									
News	(22,148)	0.0	(0)	1.8	(4)	0.0	(0)	1.8	(4)
Columns	(25,996)	114.2	(297)	65.4	(170)	51.8	(133)	230.8	(600)
Soft news	(9,091)	212.3	(193)	11.0	(10)	61.6	(56)	284.9	(259)

The quantitative distribution in Table 8.12 shows not only that the use of first and second person pronouns outside of direct speech differs markedly across the three types of articles.[4] It also indicates some striking parallels between the two news sites. On both news sites, first person singular forms are particularly frequent in soft news. Second person forms are similar in frequency in soft news and columns across both sites. Finally, first person plural forms are more frequent in columns than in soft news. Statistical tests for each type of article across the two news sites resulted in no significant differences for columns. For soft news, only the use of first person plural pronouns differed significantly at $p < 0.01$ (see appendix B).

In the following, I am going to focus on the various uses of first person singular pronouns outside of direct speech in columns and soft news. The aim is to investigate to what extent the personalising functions of *I* differ across the three types of articles, news, columns, and soft news. In addition, the comparison of the data from the *Times Online* with those from the *Mail Online* will show to what extent these functions differ across the two newspapers.

The distribution of first person singular pronouns across the various columns is not entirely even. On both news sites, columns on soft news topics tend to contain more first person singular pronouns than columns on hard news topics. On the *Times Online*, 9 out of 14 columns on hard news topics do not contain any first person singular pronouns outside of direct speech, whereas all the columns on soft news topics from the same site

[4] For first person pronouns, these differences are significant at $p < 0.001$. See appendix B.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

contain the pronoun. For the *Mail Online*, the difference between the two subcategories is less marked. While columns on hard news topics also tend to have fewer occurrences of *I* outside of direct speech, there is only one column in which the pronoun does not occur at all.

The use of first person singular pronouns in columns has some similarities with their use in direct speech in news articles. Given that columns are opinionated texts, most of the occurrences stand in a context in which they express the opinion of the author. Sometimes this is stated explicitly, such as in *My own guess is* (8.84) and *I (do not) think* (8.85). In addition, personal experiences and emotions of the text producer are also used for argumentative purposes (8.86 to 8.88).

(8.84) *My* own guess is that all of this means that the wind is blowing in the direction of a continued recovery, to borrow a phrase from my colleague David Smith, and one more rapid than most observers are expecting. (times-cn-100104)

(8.85) The point they miss is that the culprits in this case are almost certainly not freaks. *I* do not think they were born without consciences. *I* think they have had their consciences shrivelled by the lives they have led. (mail-cn-100124)

(8.86) Personally, *I* am very much in favour of marriage. As a widower, now remarried, *I* have had two very successful and happy marriages. (mail-cn-100104)

(8.87) *I* was delighted but not surprised that Prince William received a warm and enthusiastic welcome in Sydney this week. *I* would have been astonished and very disappointed if his welcome had been anything else. (times-cn-100122)

(8.88) Bottom line? *I* love *my* job. It took 20 years of hard work to get to the top of it, and *I* am good at it. (Heck, *I*'ve even won awards.) (mail-cs-100114)

In 8.86, the columnist refers to his personal experiences with marriage before beginning his criticism of political plans for changes in the British tax system that would benefit married couples. To start this criticism by stating that he is *very much in favour of marriage* serves to clarify that he is not attacking marriage or married couples, and his own positive experience with marriage is presented as evidence. In a similar way, the author of 8.87 introduces his analysis of why Prince William received a warm welcome in Australia by presenting his own emotional reaction to this welcome, and the hypothetical reaction he would have felt had the welcome been different. Likewise, the columnist in 8.88 refers to her positive work experiences and her emotions about work in order to explain why she decided to continue working full time after she had had children.

Personal experiences and emotions are, thus, rhetorical resources which are employed both by news actors who are interviewed in news articles and by columnists. While in news articles first person formulations expressing opinions, experiences and emotions are restricted to direct speech, such formulations occur in columns also outside of direct speech.

In the soft news articles from my data, first person singular pronouns are very unequally distributed. The large majority of the 517 occurrences in soft news articles from the *Times Online* occur in only three articles (times-sn-100110, times-sn-100118, and times-sn-100130) with 158, 109 and 229 instances of the pronoun, respectively. All the other

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

articles in the *Times Online* contain no more than 7 occurrences each, and several articles contain none at all. In the *Mail Online*, the distribution is even less balanced. 170 of the 193 instances of first person singular pronouns occur in a single article (mail-sn-100104), while the majority of soft news articles do not contain any first person singular pronouns outside of direct speech. Considering this, the close similarity in the normalised frequency of first person singular pronouns on the two sites is to some extent a matter of chance. Exchanging a single article would have been sufficient to lead to a very different picture.

Still, there are strong parallels in the use of first person pronouns in soft news articles across the two sites. For both sites, the articles with many first person singular pronouns are texts in which the author reports on his or her personal experiences in the first person. The following examples illustrates this use of *I* in each of the four articles.

(8.89) So what am *I* going to do, stand here and scoff? Hell, no. *I*'m getting on a plane and getting over there to give it a try (albeit noting the roughly 11 million-1 chance that the plane will crash and end *my* bid for immortality on the spot). *I* am going to go and stay with Paul McGlothin and Meredith Averill, creators of a calorie restriction lifestyle they call "the CR Way" (and authors of a book of the same name) which is followed, they claim, by 100,000 people worldwide. (times-sn-100130)

(8.90) It is 7.30am. *I*'m showered and dressed, as if for a day at the office. Morning habits die hard. Instead of dashing off in the car with the Today programme blasting on the radio, *I* am staring at *my* 20-month-old son, who is merrily spooning Weetabix and banana into his bib. If *I*'m wondering what *I*'m doing here, he clearly is too. (times-sn-100110)

(8.91) *I* will spare you tales of exploits in the gloaming world of fast gay encounters. You would simply not believe what *I* have seen and done. You would not want to know. (times-sn-100118)

(8.92) Like many women, at 38 *I* was a fat girl struggling to be in a slim body, engaged in a constant war with *my* weight which had been going on since *I* was a teenager. (mail-sn-100104)

In 8.89, the author reports in the first person on his self-experiment with a particular diet. It can be assumed that the self experiment took place for the purpose of this article. To report on the diet in the form of a first person narration of the author's personal experience leads to a much more personalised account than if he had given a more traditional review of the diet. In contrast to this article, the other three articles do not report on experiences that have been made for the purpose of writing an article. Instead the authors make their life experiences the topic of the text. In 8.90, the author reports on her experience of becoming a full-time mother after having lost her job. The author of 8.91 writes about his decision to start a heterosexual relationship after having been involved in homosexual ones for twenty years. Finally, in 8.92, the author talks about how she managed to reduce her weight – and, thus, advertises her newly published book on the same topic.

The soft news articles in which first person singular pronouns occur less frequently can be distinguished into three groups. First, there are articles which report in detail on an individual the journalist has interviewed. In these cases, first person singular pronouns emphasise the interaction between the journalist and the person who is interviewed. In 8.93,

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

for instance, *Ramsay tells me* refers to the presence of the journalist in the interview situation. First person references to the interviewer can often be found at the beginning of such articles, describing the first contact between journalist and interviewee (8.94) or the first impression of the journalist when they meet (8.95).

(8.93) Ramsay tells *me* that there was pressure on him to apologise publicly. (times-sn-100116)

(8.94) The first time *I* try to contact Rachel Johnson, new editor of The Lady magazine, she is on a shoot; not the type requiring cameras and sample-size frocks, but being entertained by Holland & Holland, makers of bespoke guns for the grouse-moors set. (times-sn-100124)

(8.95) [...] but while she looks the part, June, now 84, is a lot more fun than, to be honest, *I*'d expected. (mail-sn-100118)

On the one hand, this use personalises the journalist by referring to him or her as a person who is addressed, who acts, and who has opinions and expectations. At the same time, however, this use also personalises the interviewee. In most cases, this is a prominent person who is usually seen in an official capacity, but who in these texts is presented in a much more private fashion. Presenting the interaction between journalist and interviewee in this way, thus, serves to introduce these actors "as they are when you meet them in person". Due to the use of first person singular pronouns, readers can adopt the perspective of the journalist, who is meeting these public figures in person.

The second group of articles in which *I* is occasionally used are texts that give advice. In 8.96, the text producer introduces a podcast that can be downloaded from the news site and which is supposed to help one lose weight through hypnosis. Similarly, in 8.97, the reader is given advice on what pose to adopt in order to look slim on photographs.

(8.96) *I*'m simply going to teach you techniques that allow you to reach a state of deep, but lucid, relaxation. Ultimately, you are in charge of this process, not *me*. (times-sn-100104)

(8.97) If you want to see how it should be done, check out Keira Knightley and Paris Hilton – the queens of the red carpet pose. *I* heard that Paris had lessons in it and, when you consider how often she is photographed, it makes perfect sense. [...] Whenever *I*'ve worked as a stylist with people like Emma Bunton, Carmen Electra and Sienna Miller, *I* always take a digital camera to the fitting – you want to see how a dress photographs and what it looks like from all angles. (mail-sn-100122)

First person reference to the text producer is used here to present her as an expert. It serves to present her role in the process of hypnosis in 8.96 and it introduces her professional experience in 8.97. While the previously discussed uses of *I* invited the audience to identify with the text producers, the *I* in these texts represents an interlocutor who directly addresses the reader with *you*.

The last group of soft news articles use *I* to express the opinion of the text producer.

(8.98) *I* can't be the only one feeling queasy at our Prime Minister's wife holding hands with a former cocaine user with a string of convictions for assault. (mail-sn-100130)

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

(8.99) While it ill behoves this column to favour one child over another (what am *I* saying? it's called exercising critical judgment), the Fitflop boot also offers the pronating foot more support than the average Ugg, so that even walking to the bus offers the illusion that it's vaguely compensating for all those missed gym sessions. (time-sn-100106)

These articles in many ways resemble columns, and the author of example 8.99 in fact refers to the text as *this column*. This points to the fact that the articles which I refer to as soft news in this study and which were presented in the "Life & Style" section of the *Times Online* or the "Health" and "Femail" sections of the *Mail Online* far from constitute a coherent category. This lack of homogeneity in the soft news articles is not primarily an effect of data collection decisions. Rather, it is a consequence of the fact that soft news sections generally include a large variety of formats (see also Niblock 2008, 50).

8.7 Summary and conclusion

In this chapter, I investigated how first and second person pronouns on news sites contribute to personalisation. The analysis showed that these pronouns are used in a variety of ways, with different effects that are relevant for personalisation. Many occurrences of first person pronouns refer to specific news actors or text producers and emphasise their role as actors or experiencers of the described events. Vague and generic uses can invite the audience to identify with the position of the speaker. In some cases, direct address of the audience and reference to the newspaper as an actor serve to bring an interactional element into the (mass) communicative situation.

I argued for the need to investigate the use of first and second person pronouns within direct speech and outside of direct speech separately. Through a quantitative analysis of first and second person pronouns in direct speech from news articles, I was able to show that differences in pronoun frequency across time and across different publications are not – as was sometimes suggested – only an effect of differences in the frequency of direct speech. Instead, there are also clear differences across the news sites in how frequently first and second person pronouns occur within direct speech quotes.

A first difference was found when comparing the news articles from the *Times* from 1985 with the news articles from the *Times Online* from 2010. First and second person pronouns are clearly more frequent in the more recent online data than in the older print data. A second difference can be observed across the online sites. First person singular (and, to a lesser extent, second person) pronouns are more frequent on mid- and down-market sites than on up-market sites. If one looks at all first person pronouns, the tendency of the down-market *Sun* to focus on private individuals in contrast to official actors, groups or collectives is reflected in a higher proportion of singular forms than plural forms, whereas this is the opposite for the up-market sites.

Finally, the investigation of the use of pronouns, in particular *I*, outside of direct speech pointed to marked differences across the different types of articles. In news articles, *I* can almost exclusively be found within direct speech, where it tends to be used to express the

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

opinion, commitment, personal experience or emotions of the source of the quote. Very similar uses of *I* can be found in columns, but here it occurs outside of quotes and expresses the opinion, personal experience or emotion of the text producer. Investigating the use of *I* outside of quotes in soft news articles resulted in the observation that this group of articles does not constitute a coherent category. Instead, several subcategories can be identified, some of which correspond stylistically to columns.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Author submitted manuscript

9. Conclusion

In the previous five chapters, I studied individual features of articles on online news sites. On the one hand, I discussed how each feature contributes to personalisation. On the other hand, I investigated their use on five British online news sites from 2010 and the *Times* from 1985 in a contrastive way. The main results for each feature were already summarised at the end of each chapter. What still needs to be discussed is the relation between these features and, in particular, the relation between the three dimensions of communicative setting, content, and linguistic realisation. In the following, I will address this relation in two steps. First, I will summarise the main descriptive results for all investigated features by establishing personalisation profiles of news articles for each site included in this study. For this purpose, I will place a (hypothetical) prototypical article in the three-dimensional personalisation model developed in chapter 2. Second, I will discuss the relation between the three dimensions on a more theoretical level, by revisiting some of the interrelations between features that were encountered in the previous five chapters.

I suggested in chapter 2 that differences in personalisation strategies between individual articles can be visualised with the help of a three-dimensional model (Figure 2.3). The three axes stand for the three dimensions of communicative setting, content, and linguistic realisation, with articles being placed relative to other articles, rather than by calculating absolute positions. The model is only suitable for placing individual articles, not for placing entire online news sites or newspapers, since articles within each publication can differ considerably with respect to their characteristics for each dimension. In order to still apply this model to illustrate the main differences between news sites, I am using the construct of a hypothetical prototypical news article. In what follows, I will discuss for each site the characteristics of such a prototypical news article and place it in this model. This prototypical article is derived from the findings of chapters 4 to 8 and integrates the most frequent and the most salient characteristics of the site in question. It is important to note that these prototypical articles are not necessarily articles that have actually been published and included in my data. They are hypothetical in the sense that they are abstractions from my data and, thus, they can combine features from several articles. Figure 9.1 illustrates the placement of these prototypical articles.

The prototypical article from the printed *Times* from 1985 deals with a political topic, concentrating on negotiations among heads of states or between heads of states and other major official actors. It does not contain any news actors or sources who represent their own

Author submitted manuscript

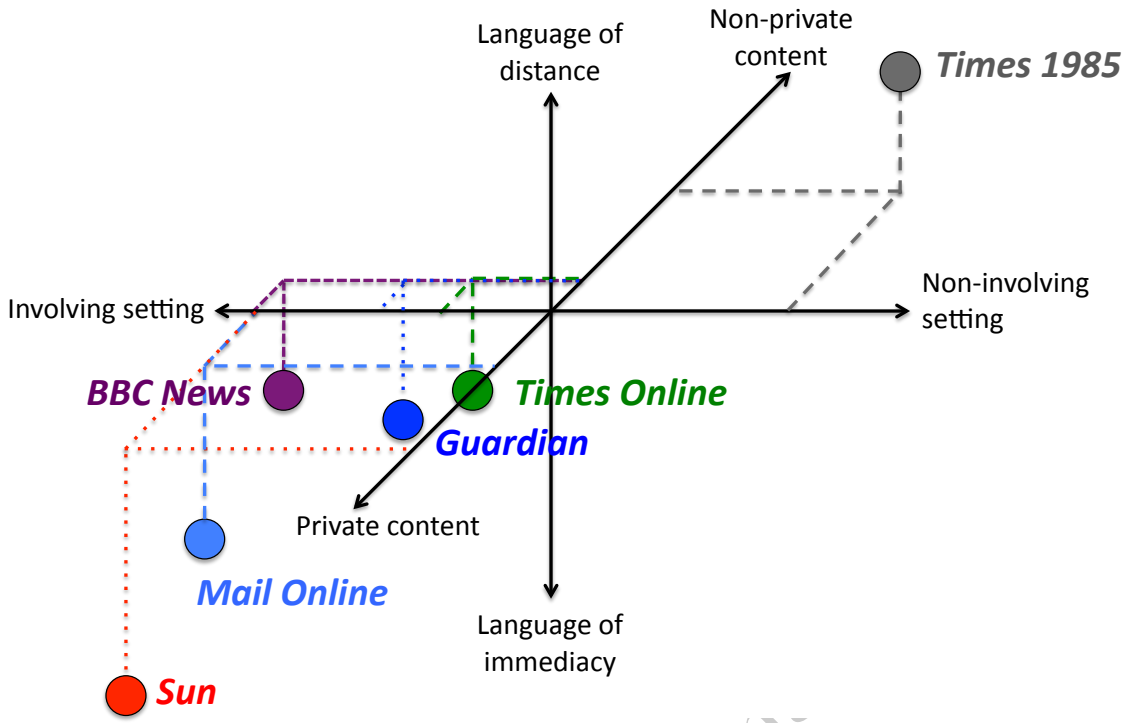


Figure 9.1: Placement of six (hypothetical) prototypical news articles in a three-dimensional model of personalisation strategies

Author submitted m

personal interests, nor does it contain information about the private lives of news actors. In terms of content, the article is, thus, close to the non-private end of the scale. Concerning personalising features on the level of linguistic realisation, only a few brief direct speech quotes are included, most quoting the main official actors in the events and some quoting other media reports. The article does not contain any second person pronouns and the first person pronouns are restricted to one first person singular pronoun and two first person plural pronouns, most likely occurring within a single quote. The first person singular pronoun is used to express the opinion or commitment of the speaker, whereas the plural pronouns have vague reference, referring to the members of the source's political party or nation. Thus, there are very few features that express linguistic immediacy. Concerning the features of the communicative setting, it is more likely for the article not to have any embedded pictures. If it has a picture, it is a single black and white photograph showing the main political actor or actors. The only way for the audience to interact with the text producers and with each other is to send in a letter to the editor, which could appear in a revised and shortened version several days later, but, due to the selection process, might never be published. Indeed, for the prototypical article from the *Times* from 1985, no letters to the editor are published at all. Thus, the setting also contributes very little to

Author submitted manuscript. Consult published version for correct page numbers!

the personalisation of this prototypical article and it would, therefore, be placed close to the non-involving end of the scale.

The prototypical articles for the three up-market news sites from 2010, the *Times Online* the *Guardian* and *BBC News*, also deal with a topic from the category of politics and finance. All three articles mostly focus on official actors, such as politicians and representatives of interest groups. The main reason for positioning these articles closer to the private end of the content axis than the article from the *Times* from 1985 is that they contain more emotions, experiences, opinions and commitment expressed by official actors. This is closely related to the fact that they contain much more direct speech than the article from the *Times* from 1985. The article from the *Guardian* contains more than twice as much direct speech than the earlier print article, while for the *Times Online* and *BBC News* this figure is a bit lower, at around 1.6 times the amount found in the *Times* from 1985. Similarly, they contain more than 6 times as many second person pronouns and more than 4 times as many first person singular pronouns as the article from the *Times* from 1985. In terms of linguistic features, the three articles are, thus, considerably closer to the pole of linguistic immediacy than the earlier print article.

Likewise, the communicative setting for all the prototypical articles from the three up-market online sites is more involving than for the article from the printed *Times* from 1985. Nevertheless, this is the dimension for which most variation across the up-market online sites can be observed. The article on the *Times Online* allows one to submit comments below the article, an option which is, however, less frequently used by readers than on the other up-market sites. It also contains one coloured photograph at the beginning of the article. Similarly, the article on the *Guardian* is illustrated with one photograph at the top and it may contain a user comment section at the bottom of the article, although it is actually more likely not to. Readers exploring the *Guardian* website would notice that there are other sections providing many options for interacting with a rather active community of readers. However, the prototypical article does not contain any links to these sections. What makes the setting of the *Guardian* article more involving compared to the article from the *Times Online* is the fact that its byline is formatted as a link that leads to the profile of the journalist, complete with his or her image, professional background and main interests, as well as information about how to contact him or her via e-mail and through Twitter.

The prototypical article on *BBC News* has by far the most involving communicative setting of the three up-market sites. It does not have a comment section below the article and it may not even have a link to a related discussion in the debate section. However, it contains generic links to debate sections and it ends with a request to submit personal experiences and visual material relating to the topic of the article. It might also integrate previously submitted user-generated content, for instance in the form of links leading to sites with selected user contributions. Moreover, while the other two up-market online articles contain only one image each, the article on *BBC News* contains two images and a video embedded in the article, as well as links to three additional videos in the sidebar.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Based on my data, it is not entirely clear what the most typical topic for the *Mail Online* article should be. Three topic categories are up for discussion: politics and finance, catastrophe, and crime. Each of these categories account for 21 percent of the news articles from the *Mail Online*. Given that politics and finance is by far the most frequent topic of the up-market news sites, and that crime is clearly the most frequent topic of the *Sun*, I chose the remaining category, catastrophe, in order to emphasise what is specific to this news site. The prototypical article from the *Mail Online* of January 2010, thus, deals with a major catastrophe, involving many official actors such as heads of states and rescue teams, but also many private actors who are directly involved in the role of victims and witnesses. Consequently, private actors are more frequent than on the up-market news sites. Due to this as well as the strong presence of personal experiences and emotions, the content of this article is much closer to the private end of the content scale compared to the up-market articles. The amount of direct speech in this article is slightly lower than for the *Guardian* article, but higher than for the *Times Online* and *BBC News*. Given that the two pronouns with the strongest personalising potential, namely first person singular and second person pronouns, are more frequent than on the up-market sites, the overall linguistic immediacy is slightly higher. In terms of features for interaction, there is a comment section below the article. The most notable feature of the communicative setting is, however, the high number of embedded visual elements, consisting of 8 photographs. This means that the article can be positioned very closely to the *BBC News* article on the communicative setting dimension.

The prototypical *Sun* article is clearly a crime article. It contains official actors in the role of judges, lawyers and representatives of the police and public administration, but it also presents private individuals in the role of victims, offenders, relatives and witnesses. Indeed, given that prominent elements like the headline and the images focus on the victims and the offenders, the private actors are more prominently represented than the official actors. Consequently, this is the article placed closest to the private end of the content axis. It contains almost as much direct speech as the *Guardian* article and both first person singular and second person pronouns are more frequent than in any of the other articles. It is also noteworthy in being the only article in which first person singular pronouns occur more frequently than first person plural pronouns, reflecting the strong focus on individuals rather than groups, also on the level of linguistic realisation. Thus, it is the article placed closest to the pole of linguistic immediacy. Like most prototypical articles, this one also has a section with user comments below the article text, which, however, attracts fewer comments than comment sections on the other sites. The main feature for interaction is given through requests to submit stories and images. These occur both at the top of the site as well as at the end of the article and provide readers with e-mail addresses and phone numbers specifically for this purpose. The article contains four embedded images and a link to an additional video, which is comparable to the article on *BBC News* but less than the article on the *Mail Online*. All in all, the position on the dimension of the communicative setting is very close to these two articles.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

The placement of these six prototypical articles in the three dimensional model in Figure 9.1 is, of course, based on several simplifications. First, as explained above, these prototypical articles are abstractions of my data and do not stand for any particular articles that were actually published. Second, the exact placement of each point is certainly open to debate. Obvious points of contention are the relative weight of features located on the same dimension (e.g. direct speech and first and second person pronouns) and the question of where the zero point of each axis should be located (e.g. should the up-market online news articles be located on the non-private or on the private half of the content axis?). Of course the exact placement of articles might also change when more features of a dimension are taken into account. Third, the news articles from each of the sites are quite varied and can only insufficiently be represented by a particular point in this three-dimensional model. Given the variation, the news articles from each site would actually occupy a zone that is roughly centred around the prototypical article. However, the variation is not constant over the sites; it is larger for the mid- and down-market sites than for the up-market sites. The data from the *Sun*, for instance, contains one article on a political topic which, in terms of content and linguistic realisation, is very close to the political articles on the up-market news sites and, thus, would be placed close to the prototypical article on *BBC News*. The opposite, i.e. an up-market news article with a focus on private content and linguistic immediacy typical of the *Sun*, is more difficult to find. Fourth, the visualisation also simplifies the actual content of the news sites, given that it is only based on news articles. When analysing individual personalisation features in previous chapters, in particular when discussing direct speech and first and second person pronouns, I sometimes included data from soft news and columns from two of the online sites. These very selective analyses have indicated that different types of articles show marked differences with respect to linguistic personalisation strategies. The fact that they also differ with respect to how much focus on private actors and topics they contain is quite obvious. Given that news articles are not necessarily the most central type of articles on online news sites – in particular on mid- and down-market sites – the exclusion of soft news, columns and other types of articles in Figure 9.1 is, of course, highly relevant.

Despite these simplifications and shortcomings, Figure 9.1 is helpful in visualising some important descriptive results of this study. First, there are clear differences between the news articles from the printed *Times* from 1985 and the news articles from all the online sites from 2010, and these differences could be observed with respect to all three dimensions. Second, two of the dimensions, namely content and linguistic realisation, repeatedly show clear differences across market-orientation. The up-market sites contain less private content and less linguistic immediacy than the down-market *Sun*, while the mid-market *Mail Online* takes a place in between. Even if the data for the mid- and down-market sites are only based on one news site each, the results are so consistent both across different features of my analysis and also in comparison with previous research that it seems plausible to conclude that market-orientation plays an important role regarding personalisation features on the levels of content and linguistic realisation. Third, it is noteworthy that market-

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

orientation does not seem to correlate with the degree to which the communicative setting is involving. For this dimension, the main variation occurs within the three up-market sites, with the *Times Online* showing the least involving setting and *BBC News* showing the most involving setting. The mid-market *Mail Online* and the down-market *Sun* show a frequency of visual elements and features for interacting with the audience which put them close to *BBC News*. Still, one factor which is not reflected in Figure 9.1 and which appears to be related to market-orientation is the way in which user-generated content is combined with editorial content. As I argued in chapter 4.3, the mid- and down-market sites are more likely to blur the boundaries between user-generated content and editorial content than the up-market news sites are.

So far, I have summarised some of the main results with respect to the level on which personalisation features were observed, i.e. the communicative setting, the content and the linguistic realisation. One aspect that has not been addressed yet is which entity is personalised, i.e. whether the presence of the above-mentioned features leads to the personalisation of the news event and its actors, the audience, or the text producers. In chapter 2.4 I argued that, at least in principle, all entities can be personalised through features on any of the three levels. However, the study of some of the most salient features on each level showed that most dimensions are much more likely to lead to the personalisation of one entity than the others. Table 9.1 illustrates the main relations between the different dimensions of personalisation features and their effects for the personalisation of different entities. The content of articles mostly leads to the personalisation of news events by focussing on actors and presenting them as private individuals. The audience can be personalised as an exception, for instance when user-generated content is presented that shows how readers of the news site are personally affected by an event. Likewise, text producers can be personalised when their own experiences are integrated into an article. For opinion columns and soft news articles this is quite common, whereas it is rare for news articles. Examples were found in one article each from the *Times* from 1985 and the *Times Online*, in which journalists write about their experiences in investigating a news event. In a slightly different manner, the *Sun* was found to personalise itself by reporting about events in which it appears as an actor, for instance its launch of a charity single for the victims of the Haiti earthquake.

Moreover, the linguistic realisation mostly contributes to the personalisation of news events and actors, at least for news articles. The use of direct speech gives a voice to news actors and represents their statements in their own words. With very few exceptions, the use of first person pronouns in news articles is restricted to direct speech, thus again contributing to the personalisation of news actors. The audience can be personalised through certain vague or generic uses of first person plural and second person pronouns. Second person pronouns with specific reference to readers, a feature known as synthetic personalisation, can mainly be found in elements surrounding articles, such as requests for submitting user-generated content, as well as in soft news articles and opinion columns. These non-news articles are also the main place in which linguistic features contribute to

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Table 9.1: Main relations between dimensions of features and personalised entities

	Communicative setting	Content	Linguistic realisation
News event and news actor	+++	+++	+++
Audience	+++	+	++
Text producer and media institution	++	+(1)	+(1)

key

+++	major effect
++	sometimes present
+	rarely observed

(1) quite common for soft news and opinion columns

Author submitted manuscript

the personalisation of text producers. In contrast, news articles only very rarely contain self-references to journalists. Somewhat more common are self-references to the media institution, such as *The Guardian has uncovered evidence* (guardian-uk-100116).

For features of the communicative setting, generalisations are more difficult. The visual material integrated into news articles often focuses on news actors and, thus, contributes to their personalisation. I have, however, also argued that images and videos increase the involvement of the audience and thus can also be seen to contribute to their personalisation. For interactive features, the main entity that is personalised is the audience, but text producers can also be personalised by presenting them as an interactive partner. In particular, this is true when the bylines of articles are complemented with the journalist's profile, providing readers with personal information and contact details.

In the discussion above, I treated the three dimensions of the communicative setting, the content and the linguistic realisation as if they were independent from each other. However, the analyses in chapters 4 to 8 repeatedly showed that there are many interdependencies between these dimensions. In the following, I would like to discuss some of these connections in some more detail.

The first of these connecting elements is user-generated content, a feature discussed in chapter 4. Requests for the audience to submit content address readers directly and, thus, contain features of linguistic immediacy. At the same time, they are an invitation for the audience to interact with text producers and they are closely tied to features of the communicative setting allowing such interaction. If readers follow this invitation and submit content which subsequently is combined with news articles, this leads to personalisation on the level of content, by presenting the personal experiences of private individuals. It may also increase linguistic immediacy, namely when text by readers is presented verbatim. If the submitted content consists of visual material, its presentation affects the communicative setting by leading to an increase in visualisations of a news event. Thus, all three dimensions of personalisation are closely interwoven in the case of user-generated content. Therefore, it is not surprising that articles containing many requests for user-generated

Author submitted manuscript. Consult published version for correct page numbers!

content, such as the articles dealing with the heavy snowfall in the UK, also tend to have a strong presence of private actors, many illustrations with images and videos, and – at least in some cases – direct quotes expressing personal experiences (see chapter 4.3 as well as the examples at the beginning of the introduction).

Likewise, visual elements relate to more than one dimension of personalisation. If seen in contrast to text, they are mainly a feature of the communicative setting, allowing the presentation of information in a different mode, which is not available on all mediums. However, the content of the images and the way in which the content is presented (i.e. the visual realisation) are also of interest for personalisation. In this study, I have mainly discussed the first of these two factors, the content, when analysing which actors were represented in the visual elements of reports on the Edlington hearing (see chapter 5.3). The results of this case study show very strong parallels to the results of the case study on the presence of actors and topics in the text from the same articles (see chapter 6.4). This finding is in itself not very surprising, given that text producers will try to find illustrations that fit the article content. It illustrates, however, that visual elements not only relate to the content dimension through their own content, but also through the fact that the content of visual elements is often indicative of the content of articles, in particular the emphasis on particular actors.

Different types of actors are the main indicator for the personalisation of content discussed in chapter 6. Analysing the role of different actors in the personalisation of news, it soon became clear that content and linguistic realisation closely interact. With respect to content, one can investigate whether an actor represents his or her personal interests or the interests of a larger group and one can also determine the amount of private information that is available about an actor. However, I have also shown that the linguistic realisation of referring expressions plays an important role in presenting an actor in a more or less private way. At the same time, the distinction between private, public and celebrity actors also plays an important role in the analysis of linguistic features. In chapter 7.4.3, I showed that different types of actors have a different likelihood of appearing as sources of direct quotes across the news sites. This may well be a direct consequence of the overall variation in the presence of different types of actors, which, in turn, is connected to the variation in the general topic categories presented as news (see chapter 6.2). However, in chapter 8.4, I also presented findings that indicate that the type of source of a quote correlates with features of its linguistic realisation. More specifically, my findings suggest that quotes by private sources have a higher likelihood of containing first and second person pronouns than quotes by official sources. These results are based on a very small sample and further research on this aspect is desirable. They are still interesting for pointing to another interaction between the two dimensions of content and linguistic realisation of news articles.

A final way in which content and linguistic realisation were shown to interact concerns the different types of articles that can be found on online news sites. I have already mentioned above that the analysis of linguistic features of personalisation showed very

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

marked differences for news articles, soft news and opinion columns. First and second person pronouns are almost exclusively restricted to direct speech in news articles, but they occur quite frequently outside of direct speech in soft news and columns (see chapter 8.6). In particular, the use of first person singular pronouns seems to be indicative of the type of article in which it appears. Given that the distinction between news and soft news is not very stable and that it is realised differently on different news sites, this finding has important implications for the study of linguistic features. It raises the question of whether the higher presence of linguistic immediacy in news articles from mid- and down-market sites is directly related to a stronger presence of content that would be considered soft news on up-market sites. This is a question that cannot be answered by the present study, but which would certainly deserve further investigation. It also points to the need for more research that takes into account both the market-orientation of news sites and the section (news, opinion, life & style, sport, etc.) in which articles are published.

This study also identified other areas that invite further research. My investigation of personalisation was based on data from online news sites, but it seems clear that most of the features can also be found in print newspapers. A systematic comparison of print newspapers and their online counterparts would, therefore, be an obvious next step. Similarly, the diachronic dimension would certainly deserve further attention, both for print newspapers and for online news sites. The selective reference to a small set of articles from the printed *Times* from 1985 indicated important changes in personalisation strategies within the last 25 years, but more detailed research would be needed to understand how these changes occurred in detail. The comparison with earlier print data from mid- and down-market newspapers, in particular, would be a promising research project, that could provide more information about the influence of these newspapers on the development of personalisation in up-market publications. The study of personalisation could also be extended to additional features. On the level of the communicative setting, the increasing integration of online news sites with social media and micro blogging platforms provides an interesting starting point for further research. On the level of content, the argumentative structure of texts, in particular opinionated articles like columns, could lead to further insights into how personal experience and emotions are used as a communicative resource in mass media texts. On the level of the linguistic realisation, stance expressions are one of the more obvious candidates for extending the study of personalisation.

Different theoretical and methodological approaches would certainly also lead to valuable complementary perspectives on personalisation. One aspect of personalisation which was not covered in this study, for example, concerns its wider political and social implications. Personalisation is often associated with sensationalism, tabloidisation, and trivialisation, which are taken to have negative effects on the social and political sphere. In this study, I avoided characterisations of personalisation as a detrimental phenomenon. Instead, I tried to provide a differentiated view on all facets of personalisation, including aspects that do not necessarily have negative effects. However, a detailed investigation into the social and political consequences of all aspects of personalisation would certainly pro-

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

vide a very welcome complementation to my research. Moreover, while this study adopted a purely product-based approach, ethnographic approaches to the newsroom could reveal more about the level of consciousness and the reasoning with which personalisation strategies are applied by text producers. Similarly, reception studies would complement the view by providing insight into how readers respond to personalisation.

The main aim of this study was to shed more light on the many facets of personalisation, to investigate its various dimensions and features and to relate them to each other in a systematic way. By integrating perspectives on the communicative setting, the content and the linguistic realisation of news texts, I was able to show that all these dimensions can contribute to presenting news as something personal. I also demonstrated that personalisation is not only restricted to one of the main entities in the mass communication process, but that it can affect the presentation of news events as well as the audience and text producers. Combining all these perspectives, I could show that personalisation deserves to be studied not only as a secondary effect of tabloidisation or shifting boundaries between the private and the public sphere, but as a characteristic of mass media communication in its own right.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

References

Abrams, M. H. and Stephen Greenblatt (eds.). 2000. *The Norton Anthology of English Literature*, vol. 2. New York / London: Norton, 7th ed.

Ágel, Vilmos and Mathilde Hennig. 2006a. "Praxis des Nähe- und Distanzsprechens." In: Vilmos Ágel and Mathilde Hennig (eds.), *Grammatik aus Nähe und Distanz. Theorie und Praxis am Beispiel von Nähetexten 1650–2000*. Tübingen: Max Niemeyer. 33–74.

Ágel, Vilmos and Mathilde Hennig. 2006b. "Theorie des Nähe- und Distanzsprechens." In: Vilmos Ágel and Mathilde Hennig (eds.), *Grammatik aus Nähe und Distanz. Theorie und Praxis am Beispiel von Nähetexten 1650–2000*. Tübingen: Max Niemeyer. 3–31.

Aijmer, Karin. 1997. "'I think' – an English modal particle." In: Toril Swan and Olaf Jansen Westvik (eds.), *Modality in Germanic Languages: Historical and Comparative Perspectives*. Berlin: Mouton de Gruyter. 1–47.

Allan, Stuart and Einar Thorsen. 2011. "Journalism, public service and BBC News Online." In: Graham Meikle and Guy Redden (eds.), *News Online: Transformations and Continuities*. Basingstoke: Palgrave Macmillan. 20–37.

Alonso Belmonte, Isabel. 2009. "Positioning the reader: A study on the use of interactive textual patterns in English written newspaper editorials and articles of opinion." *English Text Construction* 2(1), 48–69.

Altman, Douglas G. 1991. *Practical Statistics for Medical Research*. London: Chapman & Hall.

Armstrong, Cameron B. and Alan M. Rubin. 1989. "Talk radio as interpersonal communication." *Journal of Communication* 39(2), 84–94.

Baldry, Anthony and Paul J. Thibault. 2006. *Multimodal Transcription and Text Analysis. A multimedia toolkit and coursebook with associated on-line course*. London: Equinox.

Baldry, Anthony P. 2007. "The role of multimodal concordancers in multimodal corpus linguistics." In: Terry D. Royce and Wendy L. Bowcher (eds.), *New Directions in the Analysis of Multimodal Discourse*. Mahwah, NJ: Lawrence Erlbaum Associates. 173–193.

Bateman, John, Judy Delin, and Renate Henschel. 2007. "Mapping the multimodal genres of traditional and electronic newspapers." In: Terry D. Royce and Wendy L. Bowcher (eds.), *New Directions in the Analysis of Multimodal Discourse*. Mahwah, NJ: Lawrence Erlbaum Associates. 147–172.

Author-submitted manuscript

Bateman, John A. 2008. *Multimodality and Genre. A Foundation for the Systematic Analysis of Multimodal Documents*. Basingstoke: Palgrave Macmillan.

Baumgarten, Nicole and Juliane House. 2010. "'I think' and 'I don't know' in English as lingua franca and native English discourse." *Journal of Pragmatics* 42(5), 1184–1200.

Baynham, Mike. 1996. "Direct speech: What's it doing in non-narrative discourse?" *Journal of Pragmatics* 25(1), 61–81.

Bednarek, Monika and Helen Caple. 2012. *News Discourse*. Continuum Discourse, London/New York: Continuum.

Beißwenger, Michael. 2000. *Kommunikation in virtuellen Welten: Sprache, Text und Wirklichkeit. Eine Untersuchung zur Konzeptionalität von Kommunikationsvollzügen und zur textuellen Konstruktion von Welt in synchroner Internet-Kommunikation, exemplifiziert am Beispiel eines Webchats*. Stuttgart: ibidem-Verlag.

Beißwenger, Michael. 2010. "Chattern unter die Finger geschaut: Formulieren und Revidieren bei der schriftlichen Verbalisierung in synchroner internetbasierter Kommunikation." In: Vilmos Ágel and Mathilde Hennig (eds.), *Nähe und Distanz im Kontext variationslinguistischer Forschung*. Berlin/New York: De Gruyter. 247–294.

Bell, Allan. 1991. *The Language of News Media*. Oxford: Blackwell.

Bell, Philip. 2001. "Content analysis of visual images." In: Theo van Leeuwen and Carey Jewitt (eds.), *Handbook of Visual Analysis*. London: SAGE Publications. 10–34.

Benn, Stanley I. and Gerald F. Gaus. 1983. "The public and the private: concepts and action." In: Stanley I. Benn and Gerald F. Gaus (eds.), *Public and Private in Social Life*. London: Croom Helm. 3–27.

Biber, Douglas. 1988. *Variation across speech and writing*. Cambridge: Cambridge University Press.

Biber, Douglas. 1995. *Dimensions of Register Variation. A Cross-Linguistic Comparison*. Cambridge: Cambridge University Press.

Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad, and Edward Finegan. 1999. *Longman Grammar of Spoken and Written English*. London: Longman.

Bierig, Jeffrey and John Dimmick. 1979. "The late night radio talk show as interpersonal communication." *Journalism Quarterly* 56, 92–96.

Bingham, Adrian. 2009. *Family newspapers? Sex, Private Life and the British Popular Press 1918-1978*. Oxford: Oxford University Press.

Bird, Elizabeth S. 2009. "Tabloidization: What is it, and does it really matter?" In: Barbie Zelizer (ed.), *The Changing Faces of Journalism. Tabloidization, Technology and Truthiness*. London/New York: Routledge. 40–50.

Blum, Joachim and Hans-Jürgen Bucher. 1998. *Die Zeitung: Ein Multimedium. Textdesign – ein Gestaltungskonzept für Text, Bild und Grafik*. Konstanz: UVK Verlagsgesellschaft.

Boczkowski, Pablo J. 2004. *Digitizing the News. Innovation in Online Newspapers*. Cambridge, Mass: MIT Press.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Bös, Birte. 2010a. "Dialogic quotation patterns in historical news reports." In: Nicholas Brownlees, Gabriella Del Lungo, and John Denton (eds.), *The Language of Public and Private Communication in a Historical Perspective*. Newcastle upon Tyne: Cambridge Scholars Publishing. 228–244.

Bös, Birte. 2010b. "People's voices. Eine diachrone Betrachtung persönlicher Zitate in der britischen Presse." In: Martin Luginbühl and Stefan Hauser (eds.), *MedienTextKultur. Linguistische Beiträge zur kontrastiven Medienanalyse*. Beiträge zur Fremdsprachenvermittlung, Sonderheft 16, Landau: Verlag Empirische Pädagogik. 67–93.

Boyce, Tammy. 2006. "Journalism and expertise." *Journalism Studies* 7(6), 889–906.

Bredel, Ursula. 2002. "'You can say "you" to yourself'. Establishing perspectives with personal pronouns." In: Carl F. Graumann and Werner Kallmeyer (eds.), *Perspective and Perspectivation in Discourse*. Human Cognitive Processing 9, Amsterdam/Philadelphia: John Benjamins. 167–180.

Brendel, Elke, Jörg Meibauer, and Markus Steinbach. 2011a. "Exploring the meaning of quotation." In: Elke Brendel, Jörg Meibauer, and Markus Steinbach (eds.), *Understanding Quotation*. Mouton Series in Pragmatics 7, Berlin: De Gruyter Mouton. 1–33.

Brendel, Elke, Jörg Meibauer, and Markus Steinbach (eds.). 2011b. *Understanding Quotation*. Mouton Series in Pragmatics 7, Berlin: De Gruyter Mouton.

Brinker. 1988. "Thematische Muster und ihre Realisierung in Talkshowgesprächen." *Zeitschrift für germanistische Linguistik* 16, 26–45.

Brüngel-Dittrich, Melanie. 2006. *Speech Presentation in the British and German Press*. Frankfurt am Main: Peter Lang.

Bublitz, Wolfram and Christian Hoffmann. 2011. "'Three men using our toilet all day long without flushing - This may be one of the worst sentences I've ever read': Quoting in CMC." In: Joachim Frenk and Lena Steveker (eds.), *Anglistentag 2010 Saarbrücken. Proceedings*. Trier: Wissenschaftlicher Verlag Trier. 433–447.

Bucher, Hans-Jürgen. 2007. "Textdesign und Multimodalität. Zur Semantik und Pragmatik medialer Gestaltungsformen." In: Kersten Sven Roth and Jürgen Spitzmüller (eds.), *Textdesign und Textwirkung in der massenmedialen Kommunikation*. Konstanz: UVK Verlagsgesellschaft. 49–76.

Bucher, Hans-Jürgen. 2010. "Multimodalität – eine Universalie des Medienwandels: Problemstellungen und Theorien der Multimodalitätsforschung." In: Hans-Jürgen Bucher, Thomas Gloning, and Katrin Lehnen (eds.), *Neue Medien – neue Formate: Ausdifferenzierung und Konvergenz in der Medienkommunikation*. Frankfurt a. M.: Campus. 41–79.

Bucher, Hans-Jürgen. 2011. "Multimodales Verstehen oder Rezeption als Interaktion. Theoretische und empirische Grundlagen einer systematischen Analyse der Multimodalität." In: Hajo Diekmannshenke, Michael Klemm, and Hartmut Stöckl (eds.), *Bildlinguistik. Theorien – Methoden – Fallbeispiele*. Berlin: Erich Schmidt Verlag. 123–156.

Burger, Harald. 1989. "Diskussion ohne Ritual oder: der domestizierte Rezipient." In: Werner Holly, Peter Kühn, and Ulrich Püschel (eds.), *Redeshows. Fernsehdiskussionen in der Diskussion*. Tübingen: Max Niemeyer. 116–141.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Burger, Harald. 1991. *Das Gespräch in den Massenmedien*. Berlin / New York: Walter de Gruyter.

Burger, Harald. 2005. *Mediensprache eine Einführung in Sprache und Kommunikationsformen der Massenmedien*. Berlin: Walter de Gruyter, 3rd ed.

Caldas-Coulthard, Carmen Rosa. 1994. "On reporting reporting: the representation of speech in factual and factional narratives." In: Malcolm Coulthard (ed.), *Advances in Written Text Analysis*. London/New York: Routledge. 295–308.

Calsamiglia, Helena and Carmen López Ferrero. 2003. "Role and position of scientific voices: reported speech in the media." *Discourse Studies* 5(2), 147–173.

Cappelen, Herman and Ernie Lepore. 1997. "Varieties of quotation." *Mind* 106(423), 429–450.

Catenaccio, Paola, Colleen Cotter, Mark De Smedt, Guiliana Garzone, Geert Jacobs, Felicitas Macgilchrist, Lutgard Lams, Daniel Perrin, John E. Richardson, Tom Van Hout, Ellen Van Praet, and NewsTalk&Text Research Group. 2011. "Towards a linguistics of news production." *Journal of Pragmatics* 43(7), 1843–1852.

Chafe, Wallace L. 1982. "Integration and involvement in speaking, writing, and oral literature." In: Deborah Tannen (ed.), *Spoken and Written Language: Exploring Orality and Literacy*. Norwood, NJ: Ablex. 35–53.

Chicago Manual of Style. 2010. Chicago/London: The University of Chicago Press, 16th ed.

Chovanec, Jan. 2000. *Interpersonal aspects of printed news discourse*. Doctoral dissertation, Masaryk University, Faculty of Arts, Department of English and American Studies.

Chung, Deborah Soun. 2007. "Profits and Perils: Online News Producers' Perceptions of Interactivity and Uses of Interactive Features." *Convergence* 13(1), 43–61.

Claridge, Claudia. 2003. *'Life is ruled and governed by opinion': The Lampeter Corpus of Early Modern English Tracts. Manual of Information*. 2nd ed. URL `http://khnt.aksis.uib.no/icame/manuals/LAMPETER/LC-manual.pdf`.

Claridge, Claudia. 2007. "Constructing a corpus from the web: message boards." In: Marianne Hundt, Nadja Nesselhauf, and Carolin Biewer (eds.), *Corpus Linguistics and the Web*. Language and Computers: Studies in Practical Linguistics, 59, Amsterdam/New York: Rodopi. 87–108.

Claridge, Claudia. 2010. "News discourse." In: Andreas H. Jucker and Irma Taavitsainen (eds.), *Historical Pragmatics*. Handbooks of Pragmatics 8, Berlin/New York: De Gruyter Mouton. 587–620.

Clark, Herbert H. and Richard J. Gerrig. 1990. "Quotations as demonstrations." *Language* 66(4), 764–805.

Cohen, Stanley and Jock Young (eds.). 1973. *The Manufacture of News. Social Problems, Deviance and the Mass Media*. London: Constable.

Collins, Daniel E. 2001. *Reanimated Voices. Speech Reporting in a Historical-Pragmatic Perspective*. Pragmatics & Beyond New Series 85, Amsterdam/Philadelphia: John Benjamins.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Conboy, Martin. 2003. "Parochializing the global. Language and the British tabloid press." In: Jean Aitchison and Diana M. Lewis (eds.), *New Media Language*. London/New York: Routledge. 45–54.

Conboy, Martin. 2006. *Tabloid Britain: Constructing a Community Through Language*. London: Routledge.

Cotter, Colleen. 2010. *News Talk. Investigating the Language of Journalism*. Cambridge: Cambridge University Press.

Coulmas, Florian. 1986. "Reported speech: Some general issues." In: Florian Coulmas (ed.), *Direct and Indirect Speech*. Berlin: Mouton de Gruyter. 1–28.

Crisell, Andrew. 1986. *Understanding Radio*. London/New York: Methuen.

Crittenden, John. 1971. "Democratic functions of the open mike radio forum." *The Public Opinion Quarterly* 35(2), 200–210.

Crystal, David. 2006. *Language and the Internet*. Cambridge: Cambridge University Press, 2nd ed.

Czarniawska, Barbara. 2011. *Cyberfactories. How News Agencies Produce News*. Cheltenham / Northampton MA: Edward Elgar.

Derrida, Jacques. 1982. "Sending: On representation." *Social Research* 49(2), 294–326.

Diekmannshenke, Hajo, Michael Klemm, and Hartmut Stöckl (eds.). 2011. *Bildlinguistik. Theorien – Methoden – Fallbeispiele*. Berlin: Erich Schmidt Verlag.

Donsbach, Wolfgang. 1991. *Medienwirkung trotz Selektion. Einflussfaktoren auf die Zuwendung zu Zeitungsinhalten*. Köln: Böhlau.

Dürscheid, Christa. 2003. "Medienkommunikation im Kontinuum von Mündlichkeit und Schriftlichkeit. Theoretische und empirische Probleme." *Zeitschrift für angewandte Linguistik* 38, 37–56.

Dürscheid, Christa. 2004. "Netzsprache – ein neuer Mythos." *Osnabrücker Beiträge zur Sprachtheorie* 68, 141–157.

Dürscheid, Christa. 2005. "Medien, Kommunikationsformen, kommunikative Gattungen." *Linguistik online* 22(1). URL `http://www.linguistik-online.de/22_05/duerscheid.html`.

Dürscheid, Christa. 2007a. "Private, nicht-öffentliche und öffentliche Kommunikation im Internet." *Neue Beiträge zur Germanistik* 6(4), 22–41.

Dürscheid, Christa. 2007b. "Schrift – Text – Bild: Ein Brückenschlag." *Zeitschrift für germanistische Linguistik* 35(3), 269–282.

Enli, Gunn Sara. 2008. "Redefining Public Service Broadcasting: Multi-Platform Participation." *Convergence* 14(1), 105–120.

Esser, Frank. 1999. "'Tabloidization' of news. A comparative analysis of Anglo-American and German press journalism." *European Journal of Communication* 14(3), 291–324.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Evert, Stefan. 2006. "How random is a corpus? The library metaphor." *Zeitschrift für Anglistik und Amerikanistik* 54(2), 177–190.

Fairclough, Norman. 1988. "Discourse representation in media discourse." *Sociolinguistics* 17, 125–139.

Fairclough, Norman. 1992. *Discourse and Social Change*. Cambridge: Polity Press.

Fairclough, Norman. 1995. *Media Discourse*. London: Arnold.

Fairclough, Norman. 2001. *Language and Power*. Language in Social Life, Harlow: Longman, 2nd ed.

Ferrarotti, Laura. 2009. "The news presenter and the television audience: a comparative perspective of the use of 'we' and 'you'." In: Louann Haarman and Linda Lombardo (eds.), *Evaluation and Stance in War News: A Linguistic Analysis of American, British and Italian Television News Reporting of the 2003 Iraqi War*. London/New York: Continuum. 72–96.

Fetzer, Anita and Marjut Johansson. 2010. "Cognitive verbs in context. A contrastive analysis of English and French argumentative discourse." *International Journal of Corpus Linguistics* 15(2), 240–266.

Floyd, Alan. 2000. "The reporting verbs and bias in the press." *Revista Alicantina de Estudios Ingleses* 13, 43–52.

Forceville, Charles. 1999. "Educating the eye? Kress and Van Leeuwen's Reading Images: The Grammar of Visual Design (1996)." *Language and Literature* 8(2), 163–178.

Fowler, Roger. 1991. *Language in the News. Discourse and Ideology in the Press*. London/New York: Routledge.

Franklin, Bob, Martin Hamer, Mark Hanna, Marie Kinsey, and John E. Richardson. 2005. *Key Concepts in Jouralism Studies*. London: SAGE Publications.

Frehner, Carmen. 2008. *Email – SMS – MMS. The Linguistic Creativity of Asynchronous Discourse in the New Media Age*. Linguistic Insights. Studies in Language and Communication 58, Bern: Peter Lang.

Fries, Udo. 2001. "Text classes in early English newspapers." *European Journal of English Studies* 5(2), 167–180.

Galtung, Johan and Mari Ruge. 1973. "Structuring and selecting news." In: Stanley Cohen and Jock Young (eds.), *The Manufacture of News. Social Problems, Deviance and the Mass Media*. London: Constable. 62–72.

Gans, Herbert J. 1979. *Deciding What's News. A Study of CBS Evening News, NBC Nightly News, Newsweek, and Time*. New York: Pantheon Books.

Garaventa, Andreas. 1993. *Showmaster, Gäste und Publikum. Über das Dialogische in Unterhaltungsshows*. Bern: Peter Lang.

Garcia, Mario R. and Pegie Stark. 1991. *Eyes on the News*. St. Petersburg, FL: The Poynter Institute for Media Studies.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Garretson, Gregory and Annelie Ädel. 2008. "Who's speaking? Evidentiality in US newspapers during the 2004 presidential campaign." In: Annelie Ädel and Randi Reppen (eds.), *Corpora and Discourse. The Challenges of Different Settings*. Studies in Corpus Linguistics 31, Amsterdam/Philadelphia: John Benjamins. 157–187.

Greger, Volker. 1998. "Privatisierung politischer Berichterstattung im Fernsehen? Zur Veränderung der Akteursstruktur in politischen Informationssendungen von 1986 bis 1994." In: Kurt Imhof and Peter Schulz (eds.), *Die Veröffentlichung des Privaten – Die Privatisierung des Öffentlichen*. Opladen/Wiesbaden: Westdeutscher Verlag. 251–282.

Gripsrud, Jostein. 1992. "The aesthetics and politics of melodrama." In: Peter Dahlgren and Colin Sparks (eds.), *Journalism and Popular Culture*. London: SAGE Publications. 84–95.

Gripsrud, Jostein. 2000. "Tabloidization, popular journalism, and democracy." In: Colin Sparks and John Tulloch (eds.), *Tabloid Tales. Global Debates over Media Standards*. Lanham: Rowman & Littlefield. 285–300.

Grösslinger, Christian, Gudrun Held, and Hartmut Stöckl (eds.). 2012. *Pressetextsorten jenseits der 'News'. Medienlinguistische Perspektiven auf journalistische Kreativität*. Sprache im Kontext 38, Frankfurt am Main: Peter Lang.

Gruber, Helmut. 2008. "Specific genre features of new mass media." In: Ruth Wodak and Veronika Koller (eds.), *Handbook of Communication in the Public Sphere*. New York: Mouton de Gruyter. 363–382.

Gutzmann, Daniel and Erik Stei. 2011. "How quotation marks what people do with words." *Journal of Pragmatics* 43(10), 1650–2663.

Habscheid, Stephan. 2005. "Das Internet - ein Massenmedium?" In: Torsten Siever, Peter Schlobinski, and Jens Runkehl (eds.), *Websprache.net. Sprache und Kommunikation im Internet*. Linguistik – Impulse & Tendenzen 10, Berlin/New York: Walter de Gruyter. 46–66.

Hall, Stuart. 1973. "The determinations of news photographs." In: Stanley Cohen and Jock Young (eds.), *The Manufacture of News. Social Problems, Deviance and the Mass Media*. London: Constable. 176–190.

Hampton, Mark. 2008. "The 'objectivity' ideal and its limitations in 20th-century British journalism." *Journalism Studies* 9(4), 477–493.

Harcup, Tony and Deirdre O'Neill. 2001. "What is news? Galtung and Ruge revisited." *Journalism Studies* 2(2), 261–280.

Heller, Mária. 2006. "New ICTs and the Problem of 'Publicness'." *European Journal of Communication* 21(3), 311–329.

Hennig, Mathilde. 2010. "Mündliche Fachkommunikation zwischen Nähe und Distanz." In: Vilmos Ágel and Mathilde Hennig (eds.), *Nähe und Distanz im Kontext variationslinguistischer Forschung*. Berlin/New York: De Gruyter. 295–324.

Hermida, Alfred and Neil Thurman. 2008. "A clash of cultures. The integration of user-generated content within professional journalistic frameworks at British newspaper websites." *Journalism Practice* 2(3), 343–356.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Herring, Susan C. 2003. "Media and language change: Introduction." *Journal of Historical Pragmatics* 4(1), 1–17.

Herring, Susan C. 2007. "A Faceted Classification Scheme for Computer-Mediated Discourse." *Language@Internet* 4. URL `http://www.languageatinternet.de/articles/2007/761`.

Hodge, Bob. 1979. "Newspapers and communities." In: Roger Fowler, Bob Hodge, Gunther Kress, and Tony Trew (eds.), *Language and Control*. London / Boston: Routledge & Kegan Paul. 157–174.

Holly, Werner. 1997. "Zur Rolle von Sprache in Medien. Semiotische und kommunikationsstrukturelle Grundlagen." *Muttersprache* 97(1), 64–75.

Holly, Werner. 2008. "Tabloidisation of political communication in the public sphere." In: Ruth Wodak and Veronika Koller (eds.), *Handbook of Communication in the Public Sphere*. New York: Mouton de Gruyter. 317–344.

Hundt, Marianne and Christian Mair. 1999. "'Agile' and 'uptight' genres: The corpus-based approach to language change in progress." *International Journal of Corpus Linguistics* 4(2), 221–242.

Ikeo, Reiko. 2009. "An elaboration of the faithfulness claims in direct writing." *Journal of Pragmatics* 41(5), 999–1016.

Imhof, Kurt and Peter Schulz (eds.). 1998a. *Die Veröffentlichung des Privaten – die Privatisierung des Öffentlichen*. Opladen/Wiesbaden: Westdeutscher Verlag.

Imhof, Kurt and Peter Schulz. 1998b. "Einleitung: Die Veröffentlichung des Privaten – die Privatisierung des Öffentlichen." In: Kurt Imhof and Peter Schulz (eds.), *Die Veröffentlichung des Privaten – die Privatisierung des Öffentlichen*. Opladen/Wiesbaden: Westdeutscher Verlag. 9–14.

Janoschka, Anja. 2004. *Web Advertising. New Forms of Communication on the Internet*. Pragmatics & Beyond New Series 131, Amsterdam/Philadelphia: John Benjamins.

Jewitt, Carey (ed.). 2009. *The Routledge Handbook of Multimodal Analysis*. London/New York: Routledge.

Jewitt, Carey and Rumiko Oyama. 2001. "Visual meaning: a social semiotic approach." In: Theo van Leeuwen and Carey Jewitt (eds.), *Handbook of Visual Analysis*. London: SAGE Publications. 134–156.

Jucker, Andreas H. 1992. *Social Stylistics. Syntactic Variation in British Newspapers*. Berlin / New York: Mouton de Gruyter.

Jucker, Andreas H. 1996. "News actor labelling in British newspapers." *Text – Interdisciplinary Journal for the Study of Discourse* 16(3), 373–390.

Jucker, Andreas H. 2000. "Adressatenbezug und Formen der interaktiven Kommunikation in den Massenmedien." In: Gerd Richter, Jörg Riecke, and Britt-Marie Schuster (eds.), *Raum, Zeit, Medium – Sprache und ihre Determinanten. Festschrift für Hans Ramge zum 60. Geburtstag*. Darmstadt: Hessische Historische Kommission. 637–660.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Jucker, Andreas H. 2003. "Mass media communication at the beginning of the twenty-first century. Dimensions of change." *Journal of Historical Pragmatics* 4(1), 129–148.

Jucker, Andreas H. 2006. "'but 'tis believed that...': Speech and thought presentation in Early English newspapers." In: Nicholas Brownlees (ed.), *News Discourse in Early Modern Britain. Selected Papers of CHINED 2004*. Bern: Peter Lang. 105–125.

Jucker, Andreas H. 2010. "'Audacious, brilliant!! What a strike!': Live text commentaries on the Internet as real-time narratives." In: Christian R. Hoffmann (ed.), *Narrative Revisited. Telling a Story in the Age of New Media*. Pragmatics & Beyond New Series 199, Amsterdam/Philadelphia: John Benjamins. 57–77.

Jucker, Andreas H. 2011. "Mass media." In: Jan-Ola Östman (ed.), *Pragmatics in Practice*. Handbooks of Pragmatics Highlights 9, Amsterdam/Philadelphia: John Benjamins. 248–263.

Jucker, Andreas H. and Christa Dürscheid. 2012. "The linguistics of keyboard-to-screen communication. A new terminological framework." *Linguistik online* 56, 35–60. URL `http://www.linguistik-online.org/56_12/juckerDuerscheid.html`.

Kappas, Arvid and Marion G. Müller. 2006. "Bild und Emotion – ein neues Forschungsfeld." *Publizistik* 51(1), 3–23.

Kärkkäinen, Elise. 2003. *Epistemic Stance in English Conversation. A Description of its Interactional Functions, with a Focus on 'I think'*. Pragmatics & Beyond New Series 115, Amsterdam/Philadelphia: John Benjamins.

Karlsson, Michael and Jesper Strömbäck. 2010. "Freezing the flow of online news. Exploring approaches to the study of the liquidity of online news." *Journalism Studies* 11(1), 2–19.

Katriel, Tamar and Marcelo Dascal. 1989. "Speaker's commitment and involvement in discourse." In: Yishai Tobin (ed.), *From Sign to Text. A Semiotic View of Communication*. Amsterdam/Philadelphia: John Benjamins. 275–295.

Keppler, Angela. 1998. "Das Private ist politisch. Die Veröffentlichung des Privaten – eine ambivalente Medienstrategie." In: Kurt Imhof and Peter Schulz (eds.), *Die Veröffentlichung des Privaten – die Privatisierung des Öffentlichen*. Opladen/Wiesbaden: Westdeutscher Verlag. 157–164.

Kitagawa, Chisato and Adrienne Lehrer. 1990. "Impersonal uses of personal pronouns." *Journal of Pragmatics* 14(5), 739–759.

Knieper, Thomas. 2006. "Die Flut im Wohnzimmer. Die Tsunami-Berichterstattung als traumatischer Stressor für die bundesdeutsche Bevölkerung." *Publizistik* 51(1), 52–66.

Koch, Peter. 1999. "Court Records and Cartoons. Reflections of Spontaneous Dialogue in Early Romance Texts." In: Andreas H. Jucker, Gerd Fritz, and Franz Lebsanft (eds.), *Historical Dialogue Analysis*. Pragmatics & Beyond New Series 66, Amsterdam/Philadelphia: John Benjamins. 399–429.

Koch, Peter and Wulf Oesterreicher. 1985. "Sprache der Nähe – Sprache der Distanz: Mündlichkeit und Schriftlichkeit im Spannungsfeld von Sprachtheorie und Sprachgeschichte." *Romanistisches Jahrbuch* 36, 15–43.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Koch, Peter and Wulf Oesterreicher. 2011. *Gesprochene Sprache in der Romania: Französisch, Italienisch, Spanisch*. Romanistische Arbeitshefte 31, Berlin: De Gruyter, 2nd ed.

Kress, Gunther. 1986. "Language in the media: the construction of the domains of public and private." *Media, Culture & Society* 8, 395–419.

Kress, Gunther. 2010. *Multimodality. A social semiotic approach to contemporary communication*. London/New York: Routledge.

Kress, Gunther R. and Theo van Leeuwen. 1996. *Reading Images. The Grammar of Visual Design*. London: Routledge.

Kress, Gunther R. and Theo van Leeuwen. 2001. *Multimodal Discourse. The Modes and Media of Contemporary Communication*. London: Arnold.

Laberge, Suzanne and Gillian Sankoff. 1979. "Anything 'you' can do." *Syntax and Semantics* 12, 419–440.

Lakoff, Robin Tolmach. 1982. "Some of my favorite writers are literate: The mingling of oral and literate strategies in written communication." In: Deborah Tannen (ed.), *Spoken and Written Language: Exploring Orality and Literacy*. Norwood, NJ: Ablex. 239–260.

Landert, Daniela and Andreas H. Jucker. 2011. "Private and public in mass media communication: From letters to the editor to online commentaries." *Journal of Pragmatics* 43(5), 1422–1434.

Landis, Richard J. and Gary G. Koch. 1977. "The measurement of observer agreement for categorical data." *Biometrics* 33(1), 159–174.

Langer, John. 1998. *Tabloid Television: Popular Journalism and the "Other News"*. London: Routledge.

Langlotz, Andreas and Miriam A. Locher. 2012. "Ways of communicating emotional stance in online disagreements." *Journal of Pragmatics* 44(12), 1591–1606.

Langton, Loup. 2009. *Photojournalism and today's news. Creating visual reality*. Chichester: Wiley-Blackwell.

Lauerbach, Gerda. 2010. "Manoeuvring between the politica, the personal and the private: Talk, image and rythm in TV dialogue." *Discourse & Communication* 4(2), 125–159.

Leech, Geoffrey N. and Michael H. Short. 1981. *Style in Fiction. A Linguistic Introduction to English Fictional Prose*. English Language Series 13, London: Longman.

Lehmann, Hans Martin, Caren auf dem Keller, and Beni Ruef. 2006. "ZEN Corpus 1.0." In: Roberta Facchinetti and Matti Rissanen (eds.), *Corpus-based Studies of Diachronic English*. Linguistic Insights 31, Bern: Peter Lang. 135–155.

Leitner, Gerhard. 1983. *Gesprächsanalyse und Rundfunkkommunikation. Die Struktur englischer phone-ins*. Hildesheim: Georg Olms.

Linke, Angelika. 1985. *Gespräche im Fernsehen. Eine diskursanalytische Untersuchung*. Bern: Peter Lang.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Linke, Angelika. 2000. "Informalisierung? Ent-Distanzierung? Familiarisierung? Sprach(gebrauchs)wandel als Indikator soziokultureller Entwicklungen." *Der Deutschunterricht* 3/00, 66–77.

Ljung, Magnus. 2000. "Newspaper genres and newspaper English." In: Friedrich Ungerer (ed.), *English Media Texts Past and Present*. Pragmatics & Beyond New Series 80, Amsterdam/Philadelphia: John Benjamins. 131–149.

Locher, Michael A. and Stanton E.F. Wortham. 1994. "The cast of the news." *Pragmatics* 4(4), 517–533.

Lorenzo-Dus, Nuria. 2009. *Television Discourse. Analysing Language in the Media*. Basingstoke: Palgrave Macmillan.

Lüders, Marika. 2008. "Conceptualizing personal media." *New Media & Society* 10(5), 683–702.

Luginbühl, Martin. 2012. "'I felt the glass windows of my parents' apartment shaking'. Ein intermedialer und interkultureller Vergleich zur Inszenierung von Nähe in Zeitungsberichten und Fernseh-Korrespondentenberichten." In: Christian Grösslinger, Gudrun Held, and Hartmut Stöckl (eds.), *Pressetextsorten jenseits der 'News'*. Sprache im Kontext 38, Frankfurt am Main: Peter Lang. 249–264.

Luginbühl, Martin, Thomas Baumberger, Kathrine Schwab, and Harald Burger. 2002. *Medientexte zwischen Autor und Publikum. Intertextualität in Presse, Radio und Fernsehen*. Zürich: Seismo.

Luginbühl, Martin and Daniel Perrin. 2011. "'das, was wir in der Tagesschau den Rausschmeißer nennen': Altro- und Ethno-Kategorisierungen von Textsorten im Handlungsfeld journalistischer Fernsehnachrichten." In: Stephan Habscheid (ed.), *Textsorten, Handlungsmuster, Oberflächen. Linguistische Typologien der Kommunikation*. Berlin: De Gruyter. 577–596.

Macdonald, Myra. 2000. "Rethinking personalization in current affairs journalism." In: Colin Sparks and John Tulloch (eds.), *Tabloid Tales. Global Debates over Media Standards*. Lanham: Rowman & Littlefield. 251–266.

MacGregor, Phil. 2007. "Tracking the audience. Metric data start a subtle revolution." *Journalism Studies* 8(2), 280–298.

Machin, David. 2009. "Multimodality and theories of the visual." In: Carey Jewitt (ed.), *The Routledge Handbook of Multimodal Analysis*. London/New York: Routledge. 181–190.

Mair, Christian and Marianne Hundt. 1997. "The corpus-based approach to language change in progress." In: Hans Sauer and Uwe Böker (eds.), *Anglistentag 1996. Proceedings*. Tübingen: Niemeyer. 71–82.

Manovich, Lev. 2001. *The Language of New Media*. Cambridge, Mass: MIT Press.

Mayes, Patricia. 1990. "Quotation in Spoken English." *Studies in Language* 14(2), 325–363.

McCullough, David. 1992. *Truman*. New York: Simon & Schuster.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

McQuail, Denis. 1987. *Mass Communication Theory. An Introduction.* London: Sage Publications, 2nd ed.

Mininni, Giuseppe. 1990. "Le 'Talk-show' comme forme de la communication de masse." *Rassegna italiana di linguistica applicata* 22-23, 29–46.

Montgomery, Martin. 1986. "DJ Talk." *Media, Culture & Society* 8(4), 421–440.

Montgomery, Martin. 2007. *The Discourse of Broadcast News. A Linguistic Approach.* London/New York: Routledge.

Montgomery, Martin. 2010. "Rituals of personal experience in television news interviews." *Discourse & Communication* 4(2), 185–211.

Müller, Marion G. and Arvid Kappas. 2011. "Visual emotions – emotional visuals. Emotions, pathos formulae, and their relevance for communication research." In: Katrin Döveling, Christian von Scheve, and Elly A. Konijn (eds.), *The Routledge Handbook of Emotions and Mass Media.* Routledge International Handbooks, London/New York: Routledge. 310–331.

Nacey, Susan. 2012. "Scare quotes in Norwegian L2 English and British English." In: Sebastian Hoffmann, Paul Rayson, and Geoffrey Leech (eds.), *English Corpus Linguistics: Looking back Moving forward. Papers from the 30th International Conference on English Language Research on Computerized Corpora (ICAME 30).* Amsterdam/New York: Rodopi. 117–130.

Neurauter-Kessels, Manuela. 2011. "Im/polite reader responses on British online news sites." *Journal of Politeness Research* 7(2), 187–214.

Neurauter-Kessels, Manuela. 2013. *Impoliteness in Cyberspace: Personally Abusive Reader Responses in Online News Media.* Doctoral dissertation, University of Zurich, Zurich.

Niblock, Sarah. 2008. "Features." In: Bob Franklin (ed.), *Pulling Newspapers Apart: Analysing Print Journalism.* London: Routledge. 48–57.

Ochs, Elinor. 1986. "From feelings to grammar: A Samoan case study." In: Bambi B. Schieffelin and Elinor Ochs (eds.), *Language Socialization across Cultures.* Cambridge: Cambridge University Press. 251–272.

O'Keeffe, Anne. 2006. *Investigating Media Discourse.* London/New York: Routledge.

Page, Norman. 1988. *Speech in the English Novel.* Basingstoke/London: Macmillan, 2nd ed.

Paterson, Chris and David Domingo (eds.). 2008. *Making Online News: The Ethnography of New Media Production.* New York: Peter Lang.

Petersen, Thomas. 2006. "Lasswells Frage und Hovlands Problem. Feldexperimente zur Wirkung potenziell emotionalisierender Bildelemente in der Medienberichterstattung." *Publizistik* 51(1), 39–51.

Posner, Roland. 2003. "Kultursemiotik." In: Ansgar Nünning and Vera Nünning (eds.), *Konzepte der Kulturwissenschaften.* Stuttgart: J. B. Metzler. 39–72.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Prentice, Sheryl and Andrew Hardie. 2009. "Empowerment and disempowerment in the Glencairn Uprising. A corpus-based critical analysis of Early Modern English news discourse." *Journal of Historical Pragmatics* 10(1), 23–55.

Puschmann, Cornelius. 2010. "'Thank you for thinking we could'. Use and function of interpersonal pronouns in corporate web logs." In: Heidrun Dorgeloh and Anja Wanner (eds.), *Approaches to Syntactic Variation and Genre*. Berlin/New York: De Gruyter Mouton. 167–194.

Quandt, Thorsten. 2008. "(No) news on the world wide web? A comparative content analysis of online news in Europe and the United States." *Journalism Studies* 9(5), 717–738.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik. 1985. *A Comprehensive Grammar of the English Language*. London: Longman.

Rademann, Tobias. 1999. "Newspapers on the Internet." In: Hans-Jürgen Diller, Erwin Otto, and Gert Stratmann (eds.), *English Via Various Media*. anglistik & englischunterricht 62, Heidelberg: Universitätsverlag C. Winter. 251–269.

Ross, Dieter. 1998. "Die Regression des Politischen: Die Massenmedien privatisieren die Öffentlichkeit." In: Kurt Imhof and Peter Schulz (eds.), *Die Veröffentlichung des Privaten – Die Privatisierung des Öffentlichen*. Opladen/Wiesbaden: Westdeutscher Verlag. 149–156.

Royce, Terry D. and Wendy L. Bowcher (eds.). 2009. *New Directions in the Analysis of Multimodal Discourse*. Mahwah, NJ: Lawrence Erlbaum Associates.

Sanders, José. 2010. "Intertwined voices: Journalists' modes of representing source information in journalistic subgenres." *English Text Construction* 3(2), 226–249.

Santini, Marina. 2010. "Review of John Bateman. 2008. Multimodality and Genre. A Foundation for the Systematic Analysis of Multimodal Documents. Basingstoke: Palgrave Macmillan." *Linguist List* 21(1606). URL `http://linguistlist.org/pubs/reviews/get-review.cfm?SubID=2623234`.

Scannell, Paddy. 2005. "The meaning of broadcasting in the digital era." In: Gregory Ferrell Lowe and Per Jauert (eds.), *Cultural Dilemmas in Public Service Broadcasting*. Goteborg: Nordicom. 129–142.

Scheufele, Bertram. 2001. "Visuelles Medien-Framing und Framing-Effekte. Zur Analyse visueller Kommunikation aus der Framing Perspektive." In: Thomas Knieper and Marion G. Müller (eds.), *Kommunkation visuell. Das Bild als Forschungsgegenstand – Grundlagen und Perspektiven*. Köln: Herbert von Halem Verlag. 144–158.

Schlobinski, Peter. 2006. "Die Bedeutung digitalisierter Kommunikation für Sprach- und Kommunikationsgemeinschaften." In: Peter Schlobinski (ed.), *Von *hdl* bis *cul8r*. Sprache und Kommunikation in den Neuen Medien*. Thema Deutsch 7, Mannheim: Dudenverlag. 26–37.

Schmitt, Christine. 1998. "Nachrichtenproduktion im Internet. Eine Untersuchung von MSNBC Interactive." In: Jürgen Wilke (ed.), *Nachrichtenproduktion im Mediensystem. Von den Sport- und Bilderdiensten bis zum Internet*. Köln: Böhlau. 293–330.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Schmitz, Ulrich. 2007. "Sehlesen. Text-Bild-Gestalten in massenmedialer Kommunikation." In: Kersten Sven Roth and Jürgen Spitzmüller (eds.), *Textdesign und Textwirkung in der massenmedialen Kommunikation*. Konstanz: UVK Verlagsgesellschaft. 93–108.

Schneider, Kristina. 1999. "Exploring the roots of popular English news writing: A preliminary report on a corpus-based project." In: Hans-Jürgen Diller, Erwin Otto, and Gert Stratmann (eds.), *English Via Various Media*. anglistik & englischunterricht 62, Heidelberg: Universitätsverlag C. Winter. 201–222.

Schneider, Kristina. 2000. "Popular and quality papers in the Rostock Historical Newspaper Corpus." In: Christian Mair and Marianne Hundt (eds.), *Corpus Linguistics and Linguistic Theory. Papers from the Twentieth International Conference on English Language Research on Computerized Corpora (ICAME 20)*. Amsterdam: Rodopi. 321–337.

Schneider, Kristina. 2002. *The Development of Popular Journalism in England from 1700 to the Present. Corpus compilation and selective stylistic analysis*. Doctoral dissertation, Unversität Rostock, Rostock.

Schröder, Thomas. 2010. "Die Bilder-Zeitung: Wie ein Textmedium zu einem Medium der visuellen Kommunikation wird." In: Hans-Jürgen Bucher, Thomas Gloning, and Katrin Lehnen (eds.), *Neue Medien – neue Formate: Ausdifferenzierung und Konvergenz in der Medienkommunikation*. Frankfurt a. M.: Campus. 169–188.

Schudson, Michael. 2003. *The Sociology of News*. Contemporary Societies, New York / London: W. W. Norton & Company.

Schuelke, Gertrude L. 1958. "'Slipping' in indirect discourse." *American Speech* 33(2), 90–98.

Schultz, Tanjev. 1999. "Interactive options in online journalism: A content analysis of 100 U.S. Newspapers." *Journal of Computer-Mediated Communication* 5(1).

Schultz, Tanjev. 2000. "Mass media and the concept of interactivity: an exploratory study of online forums and reader email." *Media, Culture & Society* 22(2), 205–221.

Selting, Margret. 1985. "Levels of style-shifting. Exemplified in the interaction strategies of a moderator in a listener participation programme." *Journal of Pragmatics* 9(2-3), 179–197.

Short, Michael. 1988. "Speech presentation, the novel and the press." In: Willie Van Peer (ed.), *The Taming of the Text. Explorations in Language, Literature and Culture*. London/New York: Routledge. 61–81.

Short, Mick, Elena Semino, and Martin Wynne. 2002. "Revisiting the notion of faithfulness in discourse presentation using a corpus approach." *Language and Literature* 11(4), 325–355.

Short, Mick, Martin Wynne, and Elena Semino. 1999. "Reading reports: Discourse presentation in a corpus of narratives, with special reference to news reports." In: Hans-Jürgen Diller, Erwin Otto, and Gert Stratmann (eds.), *English Via Various Media*. anglistik & englischunterricht 62, Heidelberg: Universitätsverlag C. Winter. 39–65.

Sieber, Peter. 1998. *Parlando in Texten. Zur Veränderung kommunikativer Grundmuster in der Schriftlichkeit*. Germanistische Linguistik 191, Tübingen: Max Niemeyer.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Simon-Vandenbergen, Anne-Marie. 1986. "Investigating readability in qualities and tabloids." In: Anne-Marie Simon-Vandenbergen (ed.), *Aspects of Style in British Newspapers*. Gent: Studia Germanica Gandensia. 10–43.

Simonelli, Giorgio and Paolo Taggi. 1985. *I fantasmi del dialogo. Il telefono nella radio e nella televisione*. Roma: Bulzoni.

Slembrouck, Stefaan. 1986. "Towards a description of speech presentation and speech reportage in newspaper language." In: Anne-Marie Simon-Vandenbergen (ed.), *Aspects of Style in British Newspapers*. Gent: Studia Germanica Gandensia. 44–115.

Söll, Ludwig. 1974. *Gesprochenes und geschriebenes Französisch*. Grundlagen der Romanistik 6, Berlin: Schmidt.

Stark Adam, Pegie, Sara Quinn, and Rick Edmonds. 2007. *Eyetracking the News. A Study of Print and Online Reading*. St. Petersburg, FL: The Poynter Institute for Media Studies.

Stenvall, Maija. 2008. "On emotions and the journalistic ideals of factuality and objectivity – Tools for analysis." *Journal of Pragmatics* 40(9), 1569–1586.

Stewart, Graham. 2005. *The History of The Times. Volume VII 1981-2002. The Murdoch Years*. London: HarperCollins.

Stöckl, Hartmut. 2004a. *Die Sprache im Bild – Das Bild in der Sprache. Zur Verknüpfung von Sprache und Bild im massenmedialen Text. Konzepte – Theorien – Analysemethoden*. Linguistik – Impulse & Tendenzen 3, Berlin/New York: Walter de Gruyter.

Stöckl, Hartmut. 2004b. "In between modes. Language and image in printed media." In: Eija Ventola, Cassily Charles, and Martin Kaltenbacher (eds.), *Perspectives on Multimodality*. Document Design Companion Series 6, Amsterdam/Philadelphia: John Benjamins. 9–30.

Stöckl, Hartmut and Christian Grösslinger (eds.). 2010. *Mediale Transkodierungen – Metamorphosen zwischen Sprache, Bild und Ton*. Heidelberg: Universitätsverlag Winter.

Storrer, Angelika. 2009. "Rhetorisch-stilistische Eigenschaften der Sprache des Internets." In: Ulla Fix, Andreas Gardt, and Joachim Knape (eds.), *Rhetoric and Stylistics / Rhetorik und Stilistik. An International Handbook of Historical and Systematic Research / Ein internationales Handbuch historischer und systematischer Forschung. Part 2*. Handbooks of Linguistics and Communication Science 31, Berlin/New York: Mouton de Gruyter. 2211–2226.

Tannen, Deborah. 1982. "The oral/literate continuum in discourse." In: Deborah Tannen (ed.), *Spoken and Written Language: Exploring Orality and Literacy*. Norwood, NJ: Ablex. 1–16.

Tannen, Deborah. 1986. "Introducing constructed dialogue in Greek and American conversational and literary narrative." In: Florian Coulmas (ed.), *Direct and Indirect Speech*. Berlin: Mouton de Gruyter. 311–332.

Thomassen, Lasse. 2007. "Beyond Representation?" *Parliamentary Affairs* 60(1), 111–126.

Thurman, Neil. 2008. "Forums for citizen journalists? Adoption of user generated content initiatives by online news media." *New Media & Society* 10(1), 139–157.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Troesser, Michael. 1983. "Rundfunkmoderation bei Beteiligungssendungen – im Spannungsfeld zwischen Offenheit und Kontrolle." *International Journal of the Sociology of Language* 40, 70–91.

Tuchman, Gaye. 1978. *Making News. A Study in the Construction of Reality.* New York: The Free Press.

Tuomarla, Ulla. 2000. *La citation mode d'emploi : Sur le fonctionnement discursif du discours rapporte direct*. Helsinki: Academia Scientiarum Fennica.

Tuor, Nadine. 2009. *Online-Netzwerke. Eine kommunikationstheoretische, sozialpsychologische und soziolinguistische Analyse.* Networx 55, Hannover / Zürich: Mediensprache. URL `http://www.mediensprache.net/de/networx/docs/networx-55.asp`.

Turner, Graeme. 2010. *Ordinary People and the Media. The Demotic Turn.* Theory, Culture & Society, London: SAGE.

Turow, Joseph. 1974. "Talk show radio as interpersonal communication." *Journal of Broadcasting* 18(2), 171–180.

Valdeón, Roberto A. 2006. "The discursive construction of anti-European sentiment in the times.co.uk texts." *Revista Alicantina de Estudios Ingleses* 19, 405–431.

van der Wurff, Richard. 2005. "Print and online newspapers in Europe. Introduction." In: Richard van der Wurff and Edmund Lauf (eds.), *Print and Online Newspapers in Europe. A Comparative Analysis in 16 Countries.* Amsterdam: Het Spinhuis Publishers. 1–12.

van Dijk, Teun A. 1988a. *News Analysis. Case Studies of International and National News in the Press.* Hillsdale, New Jersey: Lawrence Erlbaum Associates.

van Dijk, Teun A. 1988b. *News as Discourse.* Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Van Hout, Tom and Geert Jacobs. 2008. "News production theory and practice: Fieldwork notes on power, interaction and agency." *Pragmatics* 18(1), 59–85.

van Leeuwen, Theo. 1996. "The representation of social actors." In: Carmen Rosa Rosa Caldas-Coulthard and Malcolm Coulthard (eds.), *Texts and Practices. Readings in Critical Discourse Analysis.* London/New York: Routledge. 32–70.

Vandelanotte, Lieven. 2009. *Speech and Thought Representation in English. A Cognitive-Functional Approach.* Topics in English Linguistics 65, Berlin/New York: Mouton de Gruyter.

Ventola, Eija. 2005. "Multimodality, multimediality and multiliteracies: A new era for the press and using the press." In: Hartmut E. H. Lenk and Andrew Chesterman (eds.), *Pressetextsorten im Vergleich – Contrasting Text Types in the Press.* Germanistische Linguistik Monographien 17, Hildesheim: Georg Olms. 351–387.

Ventola, Eija, Cassily Charles, and Martin Kaltenbacher (eds.). 2004. *Perspectives on Multimodality.* Document Design Companion Series 6, Amsterdam/Philadelphia: John Benjamins.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Wagner, Jörg. 2002. *Mensch-Computer-Interaktion. Sprachwissenschaftliche Aspekte.* Frankfurt am Main: Peter Lang.

Wales, Katie. 1996. *Personal pronouns in present-day English.* Studies in English language, Cambridge: Cambridge University Press.

Wardle, Claire. 2006. ""It could happen to you". The move towards "personal" and "societal" narratives in newspaper coverage of child murder, 1930-2000." *Journalism Studies* 7(4), 515–533.

Wardle, Claire. 2007. "Monsters and angels. Visual press coverage of child murders in the USA and UK, 1993–2000." *Journalism* 8(3), 263–284.

Wardle, Claire. 2008. "Crime reporting." In: Bob Franklin (ed.), *Pulling Newspapers Apart: Analysing Print Journalism.* London: Routledge. 144–152.

Wardle, Claire and Andrew Williams. 2008. *User Generated Content. Understanding its impact upon contributors, non-contributors and BBC News.* Cardiff: Cardiff School of Journalism, Media and Cultural Studies. URL `http://www.bbc.co.uk/blogs/knowledgeexchange/cardiffone.pdf`.

Waugh, Linda R. 1995. "Reported speech in journalistic discourse." *Text* 15(1), 129–173.

Wehner, Josef. 1998. "Öffentliche Meinung und Person – Zur Darstellung der Politik in den Medien." In: Kurt Imhof and Peter Schulz (eds.), *Die Veröffentlichung des Privaten – Die Privatisierung des Öffentlichen.* Opladen/Wiesbaden: Westdeutscher Verlag. 318–331.

Weintraub, Jeff. 1997. "The Theory and Politics of the Public/Private Distinction." In: Jeff Weintraub and Krishan Kumar (eds.), *Public and Private in Thought and Practice. Perspectives on a Grand Dichotomy.* Chicago/London: University of Chicago Press. 1–42.

Weiss, Ralph. 2002. "Vom gewandelten Sinn für das Private." In: Ralph Weiss and Jo Groebel (eds.), *Privatheit im öffentlichen Raum. Medienhandeln zwischen Individualisierung und Entgrenzung.* Schriftenreihe Medienforschung der Landesanstalt für Rundfunk Nordrhein-Westfalen (43), Opladen: Leske + Budrich. 27–87.

Westin, Ingrid. 2002. *Language Change in English Newspaper Editorials.* Language and Computers: Studies in Practical Linguistics 44, Amsterdam/New York: Rodopi.

Wiener, Morton and Albert Mehrabian. 1968. *Language within Language: Immediacy, a Channel in Verbal Communication.* New York: Appleton, Century & Crofts.

Włodarczyk, Matylda. 2007. *Pragmatic Aspects of Reported Speech. The Case of Early Modern English Courtroom Discourse.* Frankfurt am Main: Peter Lang.

Wolf, Claudia Maria. 2006. *Bildsprache und Medienbilder. Die visuelle Darstellungslogik von Nachrichtenmagazinen.* Wiesbaden: VS Verlag für Sozialwissenschaften.

Xekalakis, Elefteria. 1999. *Newspapers Through the Times. Foreign Reports from the 18th to the 20th Centuries.* Doctoral dissertation, University of Zurich, Zurich.

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Author submitted manuscript

Appendix A: List of articles

Key

ID: Identifier used throughout the study to refer to this article. The first part identifies the news publication from which the article was collected. The second part identifies the article category: world news (wn), UK news (uk), soft news (sn), news column (cn), soft news column (cs). The third part identifies the date on which the article was selected (date A), in the format yymmdd.

Headline: Headline as it appeared in the article.

Date A: Date on which article appeared as a top-listed item

Date B: Date on which collected article version was published or last edited. If date B is earlier than date A, then the article was published a day (or more) before it appeared as a top-listed item. If date B is later than date A, then the article had to be collected a second time for technical reasons, and had been updated in the meantime. For the articles from the printed *Times* from 1985, the two dates are necessarily identical and, thus, only one date is given.

Author's submitted manuscript

ID	Headline	Date A	Date B
Times Online, world news			
times-wn-100104	Western embassies in Yemen shut down after al-Qaeda threatens attack	04/01/2010	04/01/2010
times-wn-100106	Triple agent was CIA's best hope for years	06/01/2010	06/01/2010
times-wn-100108	Obama: Our security failed and the buck stops with me	08/01/2010	08/01/2010
times-wn-100110	Togo team arrive home as cup kicks off in chaos	10/01/2010	11/01/2010
times-wn-100112	Nuclear scientist killed by bomb in Iran	12/01/2010	12/01/2010
times-wn-100114	Race against time for Haiti earthquake aid	14/01/2010	14/01/2010
times-wn-100116	Security concerns hampers Haiti aid effort	16/01/2010	16/01/2010
times-wn-100118	Lynch mobs turn on looters amid Haiti aid crisis	18/01/2010	18/01/2010
times-wn-100120	Republicans take Ted Kennedy's seat in Massachusetts in historic upset	20/01/2010	20/01/2010
times-wn-100122	Bank shares tumble on Obama crackdown	22/01/2010	22/01/2010
times-wn-100124	Desperate Haitians learn to tackle earthquake aftermath alone	24/01/2010	24/01/2010
times-wn-100126	Desperate Haitians over-run UN food hand-out	26/01/2010	26/01/2010
times-wn-100128	Barack Obama: I don't quit - let's start again	28/01/2010	28/01/2010
times-wn-100130	Hamas promises vengeance as Mahmoud al-Mabhouh is electrocuted in Dubai	30/01/2010	30/01/2010
Times Online, UK news			
times-uk-100104	Chancellor sparks dodgy dossier spending row and admits 50% tax rate is here to stay	04/01/2010	04/01/2010

ID	Headline	Date A	Date B
times-uk-100106	Geoff Hoon and Patricia Hewitt launch new challenge to Brown leadership	06/01/2010	06/01/2010
times-uk-100108	Peter Robinson defiant amid rule breach fallout from wife's affair	08/01/2010	08/01/2010
times-uk-100110	Met Office warns drivers: expect icy roads for days	10/01/2010	10/01/2010
times-uk-100112	Alastair Campbell: Blair pledged UK to war in notes to Bush	12/01/2010	12/01/2010
times-uk-100114	Risk-averse Britain may lose stomach for war, warns minister	14/01/2010	14/01/2010
times-uk-100116	World's worst banker Sir Fred Goodwin quietly returns to business	16/01/2010	16/01/2010
times-uk-100118	Islamic sect's plan to build mega-mosque next to Olympics site collapses	18/01/2010	18/01/2010
times-uk-100120	Have-a-go hero businessman Munir Hussain freed by Court of Appeal	20/01/2010	20/01/2010
times-uk-100122	Edlington child torturers jailed for at least five years	22/01/2010	22/01/2010
times-uk-100124	Indian hijack plot caused new UK terror alert	24/01/2010	24/01/2010
times-uk-100126	Britain exits longest recession on record – just	26/01/2010	26/01/2010
times-uk-100128	Scientists in stolen e-mail scandal hid climate data	28/01/2010	28/01/2010
times-uk-100130	Car boot mother Fiona Donnison charged with children's murders	30/01/2010	30/01/2010
Times Online, soft news			
times-sn-100104	Think yourself slim with our hypnodiet podcast	04/01/2010	04/01/2010
times-sn-100106	Ugg or Fitflop? The battle of the boots	06/01/2010	06/01/2010
times-sn-100108	Milk: do we need a dairy godmother?	08/01/2010	07/01/2010
times-sn-100110	Maternal Affairs: Back to square mum	10/01/2010	10/01/2010
times-sn-100112	Iris Robinson: will her public forgive her?	12/01/2010	12/01/2010
times-sn-100114	Operation Mincemeat: full story of how corpse tricked the Nazis	14/01/2010	14/01/2010
times-sn-100116	Gordon Ramsay's escape to India – with a TV crew	16/01/2010	16/01/2010
times-sn-100118	The day I decided to stop being gay	18/01/2010	18/01/2010
times-sn-100120	Why we are still binge drinking	20/01/2010	20/01/2010
times-sn-100122	Celebrity Watch: American Idol doesn't need Simon Cowell	22/01/2010	22/01/2010
times-sn-100124	The Lady and the Rachel Johnson revamp	24/01/2010	24/01/2010
times-sn-100126	Our mother's drinking was wrecking us	26/01/2010	26/01/2010
times-sn-100128	How one man gave Congo's women hope	28/01/2010	28/01/2010
times-sn-100130	Does eating less make you healthier?	30/01/2010	30/01/2010
Times Online, news column			
times-cn-100104	Consumers and businesses lead US into recovery	04/01/2010	03/01/2010
times-cn-100106	The art of controlling contrition	06/01/2010	06/01/2010
times-cn-100108	Housebuilders aren't home and dry	08/01/2010	08/01/2010
times-cn-100110	Put paid to pay consultants	10/01/2010	09/01/2010
times-cn-100112	Chinese car companies are catching up fast	12/01/2010	12/01/2010
times-cn-100114	Low exports aren't grist to the mill	14/01/2010	14/01/2010

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

ID	Headline	Date A	Date B
times-cn-100116	Google beards Chinese tiger	16/01/2010	16/01/2010
times-cn-100118	A Greek crisis may well become Germany's problem	18/01/2010	18/01/2010
times-cn-100120	Glass and a half full for some, not for others	20/01/2010	20/01/2010
times-cn-100122	The Prince charms us, but he hasn't moved us	22/01/2010	22/01/2010
times-cn-100124	Humbled Barack Obama focuses on banks and the deficit	24/01/2010	24/01/2010
times-cn-100126	The Edlington boys are not beyond redemption	26/01/2010	26/01/2010
times-cn-100128	We will stay until Afghanistan is secure	28/01/2010	28/01/2010
times-cn-100130	Blair's world view: simply goodies v baddies	30/01/2010	30/01/2010
Times Online, soft news column			
times-cs-100104	The Queen's life is bo-ring	04/01/2010	02/01/2010
times-cs-100106	Why we must treasure our regional accents	06/01/2010	06/01/2010
times-cs-100108	We can't just vote Big Brother out	08/01/2010	07/01/2010
times-cs-100110	Cassandra: Frankly, I'm getting fed up with jargon	10/01/2010	10/01/2010
times-cs-100112	Since when do we judge Cameron or Brown by how they dress?	12/01/2010	12/01/2010
times-cs-100114	Misogyny that lurks behind the chuckles over Sarah Palin	14/01/2010	13/01/2010
times-cs-100116	Haiti and its 200 years of tragedy	16/01/2010	14/01/2010
times-cs-100118	Using ROFL for lulz? Hey, man! It's just the way I roll	18/01/2010	18/01/2010
times-cs-100120	A natural catastrophe that shames us all	20/01/2010	20/01/2010
times-cs-100122	Debenhams' divorce gift list gets my vote	22/01/2010	21/01/2010
times-cs-100124	Cassandra: It has finally come to an end	24/01/2010	24/01/2010
times-cs-100126	The moral is: question your motives, parents	26/01/2010	25/01/2010
times-cs-100128	In a burka you're cutting me off as well as you	28/01/2010	27/01/2010
times-cs-100130	Giles Coren reviews Dean Street Townhouse	30/01/2010	30/01/2010
Guardian, world news			
guardian-wn-100104	US maintains extra airport checks for countries on watch list	04/01/2010	04/01/2010
guardian-wn-100106	Barack Obama criticises CIA failures over Detroit bomb plot	06/01/2010	06/01/2010
guardian-wn-100108	British troops tried to rescue hostages at Iran border	08/01/2010	07/01/2010
guardian-wn-100110	Sunday Mirror journalist Rupert Hamer killed in Afghanistan	10/01/2010	10/01/2010
guardian-wn-100112	Nigeria's president breaks silence to quash death rumours	12/01/2010	12/01/2010
guardian-wn-100114	Haiti earthquake: International teams begin rescue operation	14/01/2010	14/01/2010
guardian-wn-100116	Haiti earthquake: Aid effort ramps up as 200,000 dead predicted	16/01/2010	16/01/2010
guardian-wn-100118	Taliban militants attack Afghan capital	18/01/2010	18/01/2010
guardian-wn-100120	Haiti hit by second strong earthquake	20/01/2010	20/01/2010
guardian-wn-100122	China hits back at US over internet censorship	22/01/2010	22/01/2010
guardian-wn-100124	Miracle survivor found as Haiti rescue teams ordered to stand down	24/01/2010	23/01/2010
guardian-wn-100126	Suicide car bomber strikes Baghdad police forensics office	26/01/2010	26/01/2010

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

ID	Headline	Date A	Date B
guardian-wn-100128	French court acquits Dominique de Villepin of slander and forgery	28/01/2010	28/01/2010
guardian-wn-100130	Haiti revival after quake could take generations says UN chief	30/01/2010	29/01/2010
Guardian, UK news			
guardian-uk-100104	Tories reject Alistair Darling's dodgy claims about their spending plans	04/01/2010	04/01/2010
guardian-uk-100106	Geoff Hoon and Patricia Hewitt call for secret ballot to settle leadership debate	06/01/2010	06/01/2010
guardian-uk-100108	Worry of exam delay for pupils	08/01/2010	08/01/2010
guardian-uk-100110	Threat of more heavy snow easing, say forecasters	10/01/2010	10/01/2010
guardian-uk-100112	Campbell: Blair wrote letters to Bush in 2002 saying we are with you on Iraq	12/01/2010	12/01/2010
guardian-uk-100114	Flood alerts issued as snow begins to melt	14/01/2010	14/01/2010
guardian-uk-100116	Fears of unrest as BNP in drive for east London votes	16/01/2010	16/01/2010
guardian-uk-100118	British Airways braced for March strike by cabin crew	18/01/2010	18/01/2010
guardian-uk-100120	Appeal court frees man jailed for attacking burglars	20/01/2010	20/01/2010
guardian-uk-100122	Edlington brothers jailed for torture of two boys	22/01/2010	22/01/2010
guardian-uk-100124	Miliband warns of terror threat after Bin Laden tape aired	24/01/2010	24/01/2010
guardian-uk-100126	Jack Straw's evidence to Iraq inquiry challenged by former legal adviser	26/01/2010	26/01/2010
guardian-uk-100128	40 days that made illegal attack into legal war on Iraq	28/01/2010	28/01/2010
guardian-uk-100130	Tony Blair accused of putting war with Iran on the electoral agenda	30/01/2010	30/01/2010
BBC News, world news			
bbc-wn-100104	Dubai to open world's tallest building	04/01/2010	04/01/2010
bbc-wn-100106	Cabinet ministers rebuff Hewitt and Hoon ballot call	06/01/2010	06/01/2010
bbc-wn-100108	Obama announces security overhaul over Detroit plot	08/01/2010	08/01/2010
bbc-wn-100110	Africa Cup of Nations begins under cloud after attack	10/01/2010	10/01/2010
bbc-wn-100112	Explosives alert closes US port	12/01/2010	12/01/2010
bbc-wn-100114	US rushes troops to Haiti earthquake zone	14/01/2010	14/01/2010
bbc-wn-100116	Clinton visits quake-hit Haitians	16/01/2010	16/01/2010
bbc-wn-100118	Afghan capital Kabul hit by Taliban attack	18/01/2010	18/01/2010
bbc-wn-100120	Obama to stick to agenda despite Massachusetts defeat	20/01/2010	20/01/2010
bbc-wn-100122	China condemns groundless US criticism of web control	22/01/2010	22/01/2010
bbc-wn-100124	Afghanistan parliamentary election postponed	24/01/2010	24/01/2010
bbc-wn-100126	France MPs' report backs Muslim face veil ban	26/01/2010	26/01/2010
bbc-wn-100128	Afghanistan summit: Gordon Brown says tide must turn	28/01/2010	28/01/2010

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

ID	Headline	Date A	Date B
bbc-wn-100130	China hits back at US over Taiwan weapons sale	30/01/2010	30/01/2010
BBC News, UK news			
bbc-uk-100104	British soldier killed in Helmand explosion	04/01/2010	04/01/2010
bbc-uk-100106	Schools closed and travellers hit as snow continues	06/01/2010	06/01/2010
bbc-uk-100108	No respite as UK hit by coldest night of winter	08/01/2010	08/01/2010
bbc-uk-100110	Sunday Mirror's Rupert Hamer killed in Afghan blast	10/01/2010	10/01/2010
bbc-uk-100112	Alastair Campbell defends every word of Iraq dossier	12/01/2010	12/01/2010
bbc-uk-100114	Met Office warns of icy roads amid road salt rationing	14/01/2010	14/01/2010
bbc-uk-100116	DEC's Haiti earthquake appeal nets £12m	16/01/2010	16/01/2010
bbc-uk-100118	Blair to face Iraq inquiry on 29 January	18/01/2010	18/01/2010
bbc-uk-100120	Mother gets life for heroin death	20/01/2010	20/01/2010
bbc-uk-100122	Edlington torture attack brothers detained	22/01/2010	22/01/2010
bbc-uk-100124	Sarah's Law sex offender alert scheme may be expanded	24/01/2010	24/01/2010
bbc-uk-100126	UK economy emerges from recession	26/01/2010	26/01/2010
bbc-uk-100128	Mother held over Sussex car boot children's murders	28/01/2010	28/01/2010
bbc-uk-100130	WWI war dead reburied in special service	30/01/2010	30/01/2010
Mail Online, world news			
mail-wn-100104	The world's tallest skyscraper opens in a blaze of glory... and it's been renamed after Arab ruler who bailed out Dubai with $25bn	04/01/2010	05/01/2010
mail-wn-100106	Clueless CIA were conned by suicide bomber's Bin Laden lies	06/01/2010	06/01/2010
mail-wn-100108	Obama: We're at war with Al Qaeda and we'll do whatever it takes to defeat them	08/01/2010	08/01/2010
mail-wn-100110	British Sunday newspaper journalist killed by roadside bomb in Afghanistan	10/01/2010	10/01/2010
mail-wn-100112	Minister's niece jailed for 15 years for knifing lover to death after drinking spree	12/01/2010	12/01/2010
mail-wn-100114	Haiti earthquake in pictures: The unimaginable horror which has torn a country apart	14/01/2010	14/01/2010
mail-wn-100116	Haiti earthquake in pictures: Shell-shocked survivors roam streets scavenging for scraps	16/01/2010	16/01/2010
mail-wn-100118	Haiti's Ground Zero: 30,000 dead and almost every building flattened in town at epicentre of earthquake	18/01/2010	18/01/2010
mail-wn-100120	Shattered Haiti hit by second earthquake measuring 6.1	20/01/2010	20/01/2010
mail-wn-100122	Haiti earthquake miracle boy Kiki: I smiled because I was alive... but now I'm sad for my dead brothers and sister	22/01/2010	22/01/2010
mail-wn-100124	From Osama to Obama: Chilling message from Al Qaeda leader warns U.S. of more terror strikes	24/01/2010	24/01/2010

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

ID	Headline	Date A	Date B
mail-wn-100126	Barack Obama: I'd rather be a good one-term president than a mediocre two-term leader	26/01/2010	26/01/2010
mail-wn-100128	We deserve these setbacks: Apologetic Obama tells America he's listening... but vows he won't quit in first State of the Union address	28/01/2010	28/01/2010
mail-wn-100130	Greece denies bailout deal as Darling promises to provide whatever assistance is appropriate	30/01/2010	30/01/2010
Mail Online, UK news			
mail-uk-100104	Top officer says Muslim hate preacher DOES have right to march as 400,000 join Facebook group against Wootton Bassett protest	04/01/2010	05/01/2010
mail-uk-100106	Army rescues 1,000 drivers stranded in cars for 12 HOURS as UK is paralysed by heavy snow (with more on its way)	06/01/2010	06/01/2010
mail-uk-100108	Now we're running out of grit: Councils forced to ration dwindling supplies as Britain faces weekend of chaos	08/01/2010	08/01/2010
mail-uk-100110	Panic buying hits supermarkets as Met Office warns of eight inches MORE snow today and tomorrow	10/01/2010	10/01/2010
mail-uk-100112	Fantasist wore haul of fake medals on Remembrance Day march to impress his young wife	12/01/2010	12/01/2010
mail-uk-100114	Haiti earthquake: 16 Brits missing as horrifying images reveal extent of destruction	14/01/2010	14/01/2010
mail-uk-100116	UKIP vows to ban the burka to win over working class voters	16/01/2010	16/01/2010
mail-uk-100118	Frustrated air passenger arrested under Terrorism Act after Twitter joke about bombing airport	18/01/2010	18/01/2010
mail-uk-100120	Life in jail for mother who injected disabled son with heroin to end his living hell	20/01/2010	20/01/2010
mail-uk-100122	Torture victim's mother vents her rage at evil brothers as parents face child neglect charges	22/01/2010	22/01/2010
mail-uk-100124	Name the Devil Boys: We must not let them hide	24/01/2010	24/01/2010
mail-uk-100126	Britain crawls out of recession as longest downturn in history finally comes to an end (but only just)	26/01/2010	26/01/2010
mail-uk-100128	Mother who killed her two young children and left their bodies in car boot had lost baby to cot death	28/01/2010	28/01/2010
mail-uk-100130	Why High Court judge says England captain John Terry can't keep his affair with team mate's girlfriend secret	30/01/2010	30/01/2010
Mail Online, soft news			
mail-sn-100104	I'll never be that chubby little girl again! How Anna Richardson overcame the childhood fear that drove her to comfort eat	04/01/2010	04/01/2010
mail-sn-100106	Curves ahead! The plus-size models that prove fashion is finally ready to embrace larger women	06/01/2010	05/01/2010

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

ID	Headline	Date A	Date B
mail-sn-100108	Young widow gives birth to miracle twins 20 months after husband died from cancer	08/01/2010	08/01/2010
mail-sn-100110	Matthew Taylor famously gave up his glittering career as an MP to be a better father. Little did he know his son's devastating illness was about to test his love and endurance to the very limit	10/01/2010	10/01/2010
mail-sn-100112	Does my bum look big in this? 'It better be... because it's good for your health	12/01/2010	12/01/2010
mail-sn-100114	How to look stylish inside and stay warm enough when you're outside	14/01/2010	14/01/2010
mail-sn-100116	EastEnders star Jo Joyner gives birth to miracle twins	16/01/2010	15/01/2010
mail-sn-100118	Kenneth Williams? A terrible flasher, dear. Frankie Howerd? VERY naughty: Jane Fryer meets the deliciously unshockable June Whitfield	18/01/2010	18/01/2010
mail-sn-100120	Breast screening benefits a myth as 7,000 women a year are wrongly told they have cancer	20/01/2010	19/01/2010
mail-sn-100122	Strike a pose - and drop a dress size!	22/01/2010	22/01/2010
mail-sn-100124	I want to become a doctor... like the one who saved me as a baby	24/01/2010	24/01/2010
mail-sn-100126	Paris Haute Couture: Christian Dior, Armani Privé and Chanel kick off week of high-octane fashion	26/01/2010	26/01/2010
mail-sn-100128	Woman saves herself from life-threatening anorexia with photo diary which charted battle from 4st 11lb to health	28/01/2010	28/01/2010
mail-sn-100130	I have always admired Sarah Brown, but is she making herself look a Twit with such dubious friends?	30/01/2010	30/01/2010
Mail Online, news column			
mail-cn-100104	The surprising lessons I learned on a journey to the heart of feckless Britain	04/01/2010	04/01/2010
mail-cn-100106	Guess who's in a lather about Dame Suzi Leather	06/01/2010	06/01/2010
mail-cn-100108	Sir Ian Kennedy claims he will listen to our views on MPs' expenses... it's now time to put fountain pen to paper!	08/01/2010	08/01/2010
mail-cn-100110	Here comes the bribe... but it'll all end in tears David Cameron	10/01/2010	10/01/2010
mail-cn-100112	A Whip hovered wearing snow shoes. His name? Bob Blizzard	12/01/2010	12/01/2010
mail-cn-100114	Red's definitely my colour - or is it pale blue? A commentary, as told by Peter Mandelson to Craig Brown	14/01/2010	14/01/2010
mail-cn-100116	Here come the cuts... and the new class war	16/01/2010	16/01/2010
mail-cn-100118	There's plenty of fuel, but it's going to cost us	18/01/2010	18/01/2010
mail-cn-100120	One by one, the men who bankrupted Britain have returned to lucrative jobs. Don't they have a shred of shame?	20/01/2010	20/01/2010

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

ID	Headline	Date A	Date B
mail-cn-100122	Have the fat cats finally gone too far?	22/01/2010	22/01/2010
mail-cn-100124	What if the horrors of Edlington were visited on cosy Notting Hill?	24/01/2010	24/01/2010
mail-cn-100126	Muzzle him, catapult him to the South Pole - whatever it takes to keep Ed Balls away!	26/01/2010	26/01/2010
mail-cn-100128	Harmanism: It's not the middle classes but social engineering zealots like Ms Harman who are to blame for Britain's inequality gap	28/01/2010	28/01/2010
mail-cn-100130	Messianic yet mellifluous... Baron Bellicose, the war guy in a beautiful suit	30/01/2010	30/01/2010
Mail Online, soft news column			
mail-cs-100104	Down into the Celebrity Big Brother quagmire we go (again)	04/01/2010	04/01/2010
mail-cs-100106	The age of luxury is upon us again	06/01/2010	06/01/2010
mail-cs-100108	Car for Carey Mulligan! New British It-girl and girlfriend of Shia LaBeouf is destined for big things in Hollywood	08/01/2010	08/01/2010
mail-cs-100110	When does a trendy taste for toyboys become something nastier?	10/01/2010	10/01/2010
mail-cs-100112	The verdict on Chris Evans: A rude awakening for some Togs, but at least the bad boy of broadcasting has grown up	12/01/2010	12/01/2010
mail-cs-100114	As my youngest goes to nursery... was I right to be a working mother?	14/01/2010	14/01/2010
mail-cs-100116	Health notes: The source of a better life	16/01/2010	16/01/2010
mail-cs-100118	Would I lie like Mo Mowlam to get the job of my dreams? You bet I would	18/01/2010	18/01/2010
mail-cs-100120	Veg out in style: Go meat free like Sir Paul McCartney	20/01/2010	20/01/2010
mail-cs-100122	Sofas: armed and dangerous (especially if you're a breakfast TV presenter)	22/01/2010	22/01/2010
mail-cs-100124	Bryan McFadden to make Delta Goodrem his Westwife... after the final whistle	24/01/2010	24/01/2010
mail-cs-100126	Screeching, Spandex and the very Super Troupers: Jan Moir joins the world's biggest Abba singalong	26/01/2010	26/01/2010
mail-cs-100128	The week I tried to be Wonder Woman – and ended up utterly pooped	28/01/2010	28/01/2010
mail-cs-100130	How was Loose Women voted best factual TV show when it is fronted by a gaggle of sexual incontinents?	30/01/2010	30/01/2010
Sun, news			
sun-100104a	Choudary demands PM debate	04/01/2010	05/01/2010
sun-100104b	Happy New Beer	04/01/2010	04/01/2010
sun-100106a	Hoon and Hewitt bid to oust PM	06/01/2010	06/01/2010
sun-100106b	Snow go zone	06/01/2010	06/01/2010
sun-100108a	Dubai waiter rapes Brit girl – then SHE and fiance are held for illegal sex	08/01/2010	08/01/2010
sun-100108b	Ice Age Britain	08/01/2010	08/01/2010
sun-100110a	Journalist killed in Afghanistan	10/01/2010	10/01/2010

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

ID	Headline	Date A	Date B
sun-100110b	Indian girl cries tears of blood	10/01/2010	10/01/2010
sun-100112a	Set free. . . to rant against freedom	12/01/2010	12/01/2010
sun-100112b	Scared gritless	12/01/2010	12/01/2010
sun-100114a	Vile cop: F*** the McCanns	14/01/2010	14/01/2010
sun-100114b	Aid floods into wrecked Haiti	14/01/2010	14/01/2010
sun-100116a	Somali who hid WPC Sharon's killer is allowed to stay in the UK	16/01/2010	16/01/2010
sun-100116b	He DID go to Spec-engravers	16/01/2010	16/01/2010
sun-100118a	UK trebles aid sent to Haiti	18/01/2010	18/01/2010
sun-100118b	Snooker Ronnie's dad out	18/01/2010	18/01/2010
sun-100120a	Dad who fought burglar is freed	20/01/2010	20/01/2010
sun-100120b	Eight killed in US shooting	20/01/2010	20/01/2010
sun-100122a	Devil brothers are locked away	22/01/2010	22/01/2010
sun-100122b	Kids see mum sniff 12 gas cans a day	22/01/2010	22/01/2010
sun-100124a	SuBo and Take That to sing on Sun charity single	24/01/2010	23/01/2010
sun-100124b	Devil boys will kill next time	24/01/2010	23/01/2010
sun-100126a	13-year-old rapist is jailed	26/01/2010	26/01/2010
sun-100126b	Elintrapment	26/01/2010	26/01/2010
sun-100128a	Mum confesses: I've killed my children	28/01/2010	28/01/2010
sun-100128b	Doherty's ex dies of drug overdose	28/01/2010	28/01/2010
sun-100130a	Dad of Year John Terry did dirty on best pal Wayne	30/01/2010	30/01/2010
sun-100130b	I saw apocalypse	30/01/2010	30/01/2010
The Times from 1985, world news			
times-wn-850107	Airlines volunteer to rescue Falasha operation		07/01/1985
times-wn-850108	Hint of US compromise on anti-satellite weapons		08/01/1985
times-wn-850109	Superpowers agree on arms talks after Geneva accord		09/01/1985
times-wn-850110	Reagan gets what he wanted in Geneva		10/01/1985
times-wn-850111	US space plans still sour accord		11/01/1985
times-wn-850112	Rift deepens between Russia and US over Star Wars		12/01/1985
times-wn-850114	US dismayed by Soviet warning over Star Wars		14/01/1985
times-wn-850115	Lebanon pullout agreed by Israel		15/01/1985
times-wn-850116	Khomeini approves suicide hit-squad		16/01/1985
times-wn-850117	Israel washes hands of Lebanon		17/01/1985
times-wn-850118	US reaffirms principle of joint action to curb dollar rates		18/01/1985
times-wn-850119	Israel honours Japanese Holocaust hero		19/01/1985
times-wn-850121	Reagan parade wiped out by freeze		21/01/1985
times-wn-850122	Reagan says US poised for greatness		22/01/1985
times-wn-850123	Ethiopia hushes up outbreak of colera in refugee camps		23/01/1985
times-wn-850124	Refugees flee cholera camp		24/01/1985
times-wn-850125	Sharon loses libel suit		25/01/1985
times-wn-850126	Botha promises new voting and property rights to blacks		26/01/1985
The Times from 1985, UK news			
times-uk-850107	Fowler to act over surrogate births		07/01/1985
times-uk-850108	NCB attitude to talks hardens as miners return		08/01/1985
times-uk-850109	Rail strike threat on January 17		09/01/1985
times-uk-850110	Return to chemical weapons denied		10/01/1985
times-uk-850111	NUM chiefs vote to expel defiant Notts area		11/01/1985
times-uk-850112	Pound still sliding despite base rate rise		12/01/1985

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

ID	Headline	Date A	Date B
times-uk-850114	Thatcher prepared for another interest rate rise		14/01/1985
times-uk-850115	Higher base rates and sliding pound may halt tax cuts		15/01/1985
times-uk-850116	Crisis moves by breathing space for weak sterling		16/01/1985
times-uk-850117	Legal threat to unions as rail strike goes ahead		17/01/1985
times-uk-850118	Left angers Kinnock by pit debate demand		18/01/1985
times-uk-850119	Notts miners in secret coal board meetings		19/01/1985
times-uk-850121	Pit strike talks about talks likely this week		21/01/1985
times-uk-850122	Ministers leave pit strike to run on until it collapses		22/01/1985
times-uk-850123	Thatcher orders three-year spending squeeze		23/01/1985
times-uk-850124	NUM to discuss draft deal to end strike		24/01/1985
times-uk-850125	Thatcher demands straight answer on pit closures		25/01/1985
times-uk-850126	NUM and coal board agree on talks about talks		26/01/1985

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Appendix B: Statistical tests

The results of the statistical tests are presented in the form of p-values for chi-square tests. The tests were calculated with the statistical software R. Yates' continuity correction was used where appropriate. Note that the lowest possible value returned for p is 2.2e-16. The values are in most cases calculated for pairs of data sets in order to allow comparisons of differences between various combinations of data sets. The inverted p-value can be taken to indicate the relative difference between data sets: the lower the p-value, the higher the difference. All p-values lower than 0.001 are marked with **bold font**, p-values between 0.01 and 0.001 are marked with *italics*. However, all these results should be interpreted with care (see section 3.5).

For convenience, the quantitative data that are tested are included before the results. In most cases, this is identical to the data presented in the respective chapter. In some cases, adaptions such as the exclusion of infrequent categories were necessary. These are specified for each case below.

Author submitted manuscript

Results of chi-square test for data of Table 5.1

Data for Table 5.1: Visual elements in news articles from five online news sites and the printed *Times* from 1985

		embedded			linked	
news site	**articles**	**images**	**videos**	**per article**	**galleries**	**videos**
Times Online	28	17	12	1.0	23	6
Guardian	28	24	6	1.1	7	3
BBC News	28	51	28	2.8	11	71
Mail Online	28	221	7	8.1	3	0
Sun	28	104	13	4.2	2	15
Times 1985	36	15	–	0.4	–	–

p-values of chi-square test with Yates' continuity correction for embedded visual elements (images plus videos), calculated for all pairs of sites

	Guardian	*BBC News*	*Mail Online*	*Sun*	*Times 1985*
Times Online	0.9237	*0.005451*	**3.807e-11**	**4.398e-05**	0.03842
Guardian		*0.007119*	**7.154e-11**	**6.401e-05**	0.03032
BBC News			**0.0004591**	0.2538	**2.683e-07**
Mail Online				0.02970	**2.2e-16**
Sun					**6.028e-11**

p-values of chi-square test with Yates' continuity correction for linked galleries, calculated for all pairs of sites

	Guardian	*BBC News*	*Mail Online*	*Sun*
Times Online	0.03009	0.156	*0.001951*	**0.0007605**
Guardian		0.5823	––[1]	––[1]
BBC News			0.1043	0.05028
Mail Online				––[1]

[1] Because the value for one of the expected frequencies is too low, the chi-square test cannot be applied

p-values of chi-square test with Yates' continuity correction for linked videos, calculated for all pairs of sites

	Guardian	*BBC News*	*Mail Online*	*Sun*
Times Online	––[1]	**1.108e-07**	––[1]	0.1531
Guardian		**4.113e-09**	––[1]	0.02649
BBC News			**6.455e-11**	**8.156e-05**
Mail Online				*0.001275*

[1] Because the value for one of the expected frequencies is too low, the chi-square test cannot be applied

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Results of chi-square test for data of Table 7.2

Data for Table 7.2: Number of words within direct speech in news articles from five online news sites and the *Times* from 1985

News site	Words in ds	Words outside ds	Total words
Times Online	5,179	17,264	22,443
Guardian	6,686	16,065	22,751
BBC News	4,425	14,898	19,323
Mail Online	9,648	24,755	34,403
Sun	4,886	11,983	16,869
Times 1985	3,220	19,968	23,188

p-values of chi-square test with Yates' continuity correction for number of words within and outside of direct speech, calculated for all pairs of sites

	Guardian	*BBC News*	*Mail Online*	*Sun*	*Times 1985*
Times Online	**2.2e-16**	0.6784	**2.2e-16**	**2.2e-16**	**2.2e-16**
Guardian		**2.2e-16**	**0.0005187**	0.3654	**2.2e-16**
BBC News			**2.2e-16**	**2.2e-16**	**2.2e-16**
Mail Online				0.03061	**2.2e-16**
Sun					**2.2e-16**

Results of chi-square test for data of Table 7.3

Data for Table 7.3: Number of words within direct speech in world news and UK news articles from four online news sites and the *Times* from 1985

	World news			UK news			p-value[1]
News site	In ds	Out ds	Total	In ds	Out ds	Total	
Times Online	2,283	9,589	11,872	2,896	7,675	10,571	**2.2e-16**
Guardian	2,549	8,183	10,732	4,137	7,882	12,019	**2.2e-16**
BBC News	2,140	7,352	9,492	2,285	7,546	9,831	0.2558
Mail Online	2,445	8,153	10,598	7,203	16,602	23,805	**2.2e-16**
Times 1985	1,368	8,804	10,172	1,852	11,164	13,016	0.09195

[1] p-values of chi-square test with Yates' continuity correction for number of words within and outside of direct speech

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Results of chi-square test for data of Table 7.4

Data for Table 7.4: Number of words within direct speech in news articles, soft news articles and columns from two online news sites

	News			Soft news			Columns		
News site	**In ds**	**Out**	**Total**	**In ds**	**Out**	**Total**	**In ds**	**Out**	**Total**
Times Online	5,179	17,264	22,443	5,421	23,624	29,045	1,614	24,457	26,071
Mail Online	9,648	24,755	34,403	3,652	9,091	12,743	1,592	25,996	27,588

p-values of chi-square test with Yates' continuity correction for pairs of values:

- *Times Online*, news vs. soft news: **2.2e-16**
- *Times Online*, soft news vs. columns: **2.2e-16**
- *Times Online*, news vs. columns: **2.2e-16**
- *Mail Online*, news vs. soft news: 0.1916
- *Mail Online*, soft news vs. columns: **2.2e-16**
- *Mail Online*, news vs. columns: **2.2e-16**
- *Times Online*, news vs. *Mail Online*, news: **2.2e-16**
- *Times Online*, soft news vs. *Mail Online*, soft news: **2.2e-16**
- *Times Online*, columns vs. *Mail Online*, columns: 0.04194

Results of chi-square test for data of Table 7.6

Data for Table 7.6, modified for testing: Named and anonymous sources in randomly selected direct speech quotes from the news sections of five online news sites and the *Times* from 1985

	Times Online	***Guardian***	***BBC News***	***Mail Online***	***Sun***	***Times 1985***
named	74	86	87	65	72	64
anonymous	26	14	12	13	20	35

Note: Identified sources are not included, because of zero values.
p-value for named versus anonymous sources across all news sites: **0.0004655**

p-values of chi-square test with Yates' continuity correction for named and anonymous sources, calculated for all pairs of sites

	Guardian	*BBC News*	*Mail Online*	*Sun*	*Times 1985*
Times Online	0.05183	0.02088	0.1898	0.6018	0.2015
Guardian		0.8549	0.7783	0.2247	**0.000864**
BBC News			0.5191	0.1131	**0.0002382**
Mail Online				0.523	*0.009159*
Sun					0.05532

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Results of chi-square test for data of Table 7.7

Data for Table 7.7, modified for testing: Official and private sources in randomly selected direct speech quotes from the news sections of five online news sites and the *Times* from 1985

	Times Online	***Guardian***	***BBC News***	***Mail Online***	***Sun***	***Times 1985***
official	88	93	88	54	54	90
private	7	7	9	27	32	1

Note: The categories celebrity, impersonal and undecided are not included, because of zero values. The data from the *Mail Online* article mail-uk-100124, which lead to distorted values, are also excluded.

p-values of chi-square test with Yates' continuity correction for official and private sources, calculated for all pairs of sites

	Guardian	*BBC News*	*Mail Online*	*Sun*	*Times 1985*
Times Online	0.8588	0.8277	**3.222e-05**	**2.663e-06**	−−[1]
Guardian		0.7456	**1.569e-05**	**1.142e-06**	−−[1]
BBC News			**0.0001498**	**1.391e-05**	−−[1]
Mail Online				0.7175	**3.608e-08**
Sun					**2.341e-09**

[1] Because the value for one of the expected frequencies is too low, the chi-square test cannot be applied

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Results for chi-square test for data of Table 8.5

Data for Table 8.5: Frequency of pronouns in news articles from five online news sites and the *Times* from 1985

News site	Words	**1st sg.** (abs.)	**1st pl.** (abs.)	**2nd** (abs.)
Times Online	22,443	97	181	34
Guardian	22,751	111	155	32
BBC News	19,323	86	113	28
Mail Online	29,607	163	168	51
Sun	16,869	139	91	40
Times 1985	23,188	23	71	5

First person singular pronouns

p-values of chi-square test with Yates' continuity correction, comparing first person singular pronouns and total number of words, calculated for all pairs of sites

	Guardian	*BBC News*	*Mail Online*	*Sun*	*Times 1985*
Times Online	0.4232	0.902	0.06808	**1.060e-06**	**8.24e-12**
Guardian		0.571	0.3582	**4.386e-05**	**2.644e-14**
BBC News			0.1260	**7.492e-06**	**4.978e-12**
Mail Online				**0.0005734**	**2.2e-16**
Sun					**2.2e-16**

First person plural pronouns

p-values of chi-square test with Yates' continuity correction, comparing first person plural pronouns and total number of words, calculated for all pairs of sites

	Guardian	*BBC News*	*Mail Online*	*Sun*	*Times 1985*
Times Online	0.1381	*0.008675*	*0.001226*	*0.002079*	**1.192e-12**
Guardian		0.2418	0.1135	0.08882	**1.604e-08**
BBC News			0.8526	0.6162	**2.013e-05**
Mail Online				0.7469	**1.343e-05**
Sun					**0.0004049**

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Second person pronouns

p-values of chi-square test with Yates' continuity correction, comparing second person pronouns and total number of words, calculated for all pairs of sites

	Guardian	*BBC News*	*Mail Online*	*Sun*	*Times 1985*
Times Online	0.8586	0.9627	0.638	0.06917	**4.561e-06**
Guardian		0.9883	0.4302	0.03522	**1.487e-05**
BBC News			0.5351	0.05804	**1.254e-05**
Mail Online				0.1589	**2.758e-07**
Sun					**5.66e-10**

Results for chi-square test for data of Table 8.6

Data for Table 8.6: Frequency of pronouns in different types of articles in the *Times Online* and the *Mail Online*

News site		**1st sg.**	**1st pl.**	**2nd**	**Total**
Category	Words				
Times Online					
News	22,443	97	181	34	312
Columns	26,071	340	186	174	700
Soft news	29,045	792	134	204	1,130
Mail Online					
News	29,607	163	168	51	382
Columns	27,588	343	188	149	680
Soft news	12,743	388	79	86	553

p-values of chi-square test with Yates' continuity correction for pairs of values

	1st st.	**1st p.**	**2nd**	**all pronouns**
Times Online, news vs. soft news	**2.2e-16**	**9.996e-07**	**2.2e-16**	**2.2e-16**
Times Online, soft news vs. columns	**2.2e-16**	**0.0001389**	0.659	**2.786e-14**
Times Online, news vs. columns	**2.2e-16**	0.2636	**2.2e-16**	**2.2e-16**
Mail Online, news vs. soft news	**2.2e-16**	0.5636	**2.2e-16**	**2.2e-16**
Mail Online, soft news vs. columns	**2.2e-16**	0.5238	0.1157	**2.2e-16**
Mail Online, news vs. columns	**2.2e-16**	0.0952	**1.994e-13**	**2.2e-16**
news, *Times Online* vs. *Mail Online*	0.06808	*0.001226*	0.638	0.3509
soft news, *Times Online* vs. *Mail Online*	0.08484	0.04442	0.8062	0.04172
columns, *Times Online* vs. *Mail Online*	0.5607	0.6965	0.06601	0.1231

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Results for chi-square test for data of Table 8.7

Data for Table 8.7: Frequency of pronouns in news articles occurring inside of direct speech in five online news sites and the *Times* from 1985

News site	Words	**1st sg.** abs.	**1st pl.** abs.	**2nd** abs.
Times Online	5,179	72	178	34
Guardian	6,686	111	155	32
BBC News	4,425	86	105	28
Mail Online	7,459	163	164	51
Sun	4,886	138	84	40
Times 1985	3,220	20	70	5

First person singular pronouns

p-values of chi-square test with Yates' continuity correction, comparing first person singular pronouns in quotes and total number of words in quotes, calculated for all pairs of sites

	Guardian	*BBC News*	*Mail Online*	*Sun*	*Times 1985*
Times Online	0.2755	0.04452	*0.001753*	**1.194e-06**	*0.001623*
Guardian		0.3103	0.03093	**4.064e-05**	**4.163e-05**
BBC News			0.4207	*0.008357*	**2.347e-06**
Mail Online				0.03266	**2.783e-08**
Sun					**9.152e-12**

First person plural pronouns

p-values of chi-square test with Yates' continuity correction, comparing first person plural pronouns in quotes and total number of words in quotes, calculated for all pairs of sites

	Guardian	*BBC News*	*Mail Online*	*Sun*	*Times 1985*
Times Online	**0.0004663**	*0.003444*	**5.201e-05**	**1.898e-07**	*0.001544*
Guardian		0.906	0.6811	0.03341	0.7116
BBC News			0.5895	0.03444	0.6284
Mail Online				0.07922	0.9948
Sun					0.1751

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

Second person pronouns

p-values of chi-square test with Yates' continuity correction, comparing second person pronouns in quotes and total number of words in quotes, calculated for all pairs of sites

	Guardian	*BBC News*	*Mail Online*	*Sun*	*Times 1985*
Times Online	0.2459	0.9872	0.9423	0.4076	*0.001892*
Guardian		0.3435	0.1402	0.03052	0.02218
BBC News			0.8324	0.3562	*0.003124*
Mail Online				0.4574	**0.0009352**
Sun					**0.0001689**

Results of chi-square test for data of Table 8.8

Data for Table 8.8: First and second person pronouns in quotes by private and public sources

	Official sources abs.	**Private sources** abs.	**p-value**[1]
Number of quotes	377	98	
Containing any pronoun	165	68	
Containing no pronoun	212	30	**1.050e-05**
Containing 1st sg.	57	39	
Containing no 1st sg. pronoun	320	59	**1.304e-07**
Containing 1st pl.	108	30	
Containing no 1st pl. pronoun	269	68	0.7973
Containing 2nd	26	13	
Containing no 2nd	351	85	0.06584

[1] p-values of chi-square test with Yates' continuity correction for official and private sources

Results for chi-square test for data of Table 8.10

Data for Table 8.10: Ratio of first person plural pronouns to first person singular pronouns inside direct speech in five online news sites and the *Times* from 1985

News site	**1st pl.**	**1st sg.**	**Ratio**
Times Online	178	72	2.5
Guardian	155	111	1.4
BBC News	105	86	1.2
Mail Online	164	163	1.0
Sun	84	138	0.6
Times 1985	70	20	3.5

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!

p-values of chi-square test with Yates' continuity correction, comparing first person plural and first person singular pronouns occurring in direct speech for all pairs of sites

	Guardian	*BBC News*	*Mail Online*	*Sun*	*Times 1985*
Times Online	*0.002920*	**0.0006238**	**5.347e-07**	**6.629e-13**	0.2864
Guardian		0.5444	0.05887	**1.056e-05**	*0.001420*
BBC News			0.3329	**0.0007089**	**0.0003882**
Mail Online				*0.005813*	**5.197e-06**
Sun					**3.660e-10**

Results for chi-square test for data of Table 8.12

Data for Table 8.12: Frequency of pronouns outside of direct speech in different types of articles in the *Times Online* and the *Mail Online*

News site Category	Words	**1st sg.**	**1st pl.**	**2nd**	**Total**
Times Online					
News	17,264	25	3	0	28
Columns	24,457	308	176	134	618
Soft news	23,624	517	74	116	707
Mail Online					
News	22,148	0	4	0	4
Columns	25,996	297	170	133	600
Soft news	9,091	193	10	56	259

p-values of chi-square test with Yates' continuity correction for pairs of values

	1st st.	**1st p.**	**2nd**	**all pronouns**
Times Online, news vs. soft news	**2.2e-16**	**2.250e-11**	$--^{1}$	**2.2e-16**
Times Online, soft news vs. columns	**1.671e-14**	**1.068e-09**	0.4243	*0.002642*
Times Online, news vs. columns	**2.2e-16**	**2.2e-16**	$--^{1}$	**2.2e-16**
Mail Online, news vs. soft news	$--^{1}$	$--^{2}$	$--^{1}$	**2.2e-16**
Mail Online, soft news vs. columns	**2.158e-11**	**8.274e-10**	0.2799	*0.005776*
Mail Online, news vs. columns	$--^{1}$	**2.2e-16**	$--^{1}$	**2.2e-16**
news, *Times Online* vs. *Mail Online*	$--^{1}$	$--^{2}$	$--^{1}$	**1.576e-06**
soft news, *Times Online* vs. *Mail Online*	0.7536	*0.001777*	0.1912	0.5275
columns, *Times Online* vs. *Mail Online*	0.2501	0.4046	0.6192	0.1252

[1] Because one of the values is zero, the chi-square test cannot be applied

[2] Because the value for one of the expected frequencies is too low, the chi-square test cannot be applied

Author submitted manuscript

Author submitted manuscript. Consult published version for correct page numbers!