
Online Methodology: Analysing News Flows of Online Journalism

Robert Kautsky and Andreas Widholm

Department of Journalism, Media and Communication
Stockholm University

Keywords: Regular Interval Content Capture (RICC), methodology, online news, web journalism, versionality

Abstract

In the present global media climate, speed and immediacy are increasingly prioritised characteristics of news production. As online news has developed, the idea of a single news item has been replaced by fast-changing content and new repertoires of constructing 'Breaking News'. Whereas most research of online news has used synchronic rather than diachronic methods, this article introduces a new approach, which we choose to call Regular Interval Content Capture (RICC). The data produced by RICC enables dynamic online media texts to be studied as they are produced, edited, and changed using both quantitative and qualitative methods. In our study, the US 'Crucial Tuesday' primary elections serve as the empirical example. From a discourse analytical perspective, we analyse a total of 64 hours of online news flows collected from the US and International editions of CNN.com. The RICC approach allows us to find major representational differences between the two editions. Three different modes of writing, characterising different stages of CNN's reporting, were identified.

Introduction

During the past few years, there has been a growing interest in media globalisation and digital communication, and as the media have become increasingly global, speed and immediacy have become main characteristics of news production. The emergence of online news as a third live news medium besides television and radio, has arguably played an important role in these changes. Stuart Allan (2006) has described the rise of online news as a global journalistic revolution in its own right, with blogging, citizen journalism and participatory journalism as bearing examples. Others such as Terry Flew (2007) have concluded that the digitisation of the media industry has led to new kinds of globally distributed symbolic forms. In this article, we aim to present a novel approach to the study of online news journalism. In the past, the temporal dimension of a studied newspaper text was quite easily ignored, as any potential problem of different versions of the same news item was quickly solved by deciding on using one of them and disposing of the others. However, as online news has developed over the past decade, it has

become increasingly apparent that the way an online news article develops and changes over time cannot be as easily disregarded. In fact, it can be argued that it is its development over time that is one of its foremost *raisons d'être* as data to a researcher of online news, that which distinguishes it as news instead of mere content in the online media flow.

Studying online news through the same methods used for printed newspapers is unsatisfactory and has likely led many researchers to stick with the paper editions. The most markedly different aspects of online versus print journalism are the new possibilities for updating, changing and reshaping news that online publications have brought. Whereas print journalism is mono-linear, from writing, via editing to printing of a final version, news online can be published, edited and republished again. Current research suggests that this may pose a problem to the researcher as the same news article may exist in 'up to 20 different versions' (Karlsson 2006, 194). The existence of multiple versions of a text might seem simple enough to solve by choosing one of them, usually the last updated version available at the time of data collection, and then submitting it to the scrutiny of our analysis. But what about the previous 19 versions which are dismissed commonly on the basis of being 'too much data'? As part of the history of text version 20, do they not deserve at least some attention by the researcher? Could, in fact, the study of versions, histories and developments of 'unfinished' texts tell us more about the texts produced by contemporary journalism, than the to-date default study of the finished products? If we decide to try to study them, how should they be conceptualised? Are earlier versions different texts or the same text, caught in a different stage of development? Online news clearly opens up new areas of scientific inquiry, as well as posing new problems of methodology.

The overall purpose of this article is to develop and test a method for the study of online news that enables us to analyse different versions of websites as they appear in sequence, arguing that the versionality of online news opens a window into the text which allows us to peek into the discourse practice of online news. In essence, Regular Interval Content Capture (RICC) is merely a method for collecting time-series of data from online websites and as such is usable for any online research, not just specifically online news. However, as we will show in this article, employing this approach in a critical analysis of online news texts certainly opens up new possibilities for media research.

The empirical material for the article consists of a time-series of data comprising online top news stories as they are shown on the main pages of the US and International editions respectively of the online news site CNN.com. The data, covering the U.S. 'Crucial Tuesday' primary elections of March 2008 and described in greater detail below, was selected as an example for several reasons. Firstly, news organisations such as CNN, and increasingly online journalism in general, often

operate on both national and transnational markets which makes it an interesting object of study regarding how news is discursively shaped and modified to fit national and transnational publics. Secondly, since we knew in advance when 'Crucial Tuesday' would begin, the collection of data was easily demarcated in time. Thirdly, the online coverage was expected to be updated regularly, displaying some of the main characteristics of online breaking news. In addition, the U.S. presidential elections cannot be isolated to a strictly national context, as the articulation and impact of U.S. politics clearly illustrate how the national is intertwined with the global. Thus, the case was strategically selected with the expectation to provide us with a complex mix of national and global modes of journalistic writing, while at the same time keeping some factors constant since the news originates from the same sources at CNN, and differences between the two editions can largely be attributed to editorial choices.

While the term 'Super Tuesday' commonly refers to the day when the greatest number of states pick their Democratic or Republican presidential candidates, several other 'Tuesday-names' have been created in order to maintain the democratic drama. The elections on the 4th of March 2008, mostly occurred under the name 'Crucial Tuesday' or as CNN anchor John Roberts put it in the American Morning show: 'Call it what you want. Super Tuesday, crucial Tuesday, critical Tuesday, the last stand of the Democratic candidates. The first polls opened in Vermont just a couple of seconds ago' (CNN Transcripts).

Initial Remarks on Journalism and Online News Texts

Theories of globalisation have formed the core of many studies of Internet and online news in recent years. As journalism has been increasingly globalised, crossing both cultural and geographical borders, the collapse of time and space seems to be an academic argument that is here to stay. Neil Thurman (2007) has illustrated the rather paradoxical development of online journalism by comparing national and international readers of British and US newspapers online. As he argues, many newspapers and media organisations in Britain have gone from a rather modest interest in international readers, to an active adaptation of news materials and advertisements to fit transatlantic audiences. The UK's *Guardian* newspaper for example, receives 78 percent of its online readers from countries outside Britain, and similarly the BBC news attracts large audiences all over the globe. On the other hand, US newspapers are more national with regards to content as well as audience. *The New York Times*, for example, receives less than 30 percent of its readers 'overseas' (Thurman 2007, 287). The development of several editions of the same media platform can also be seen in the television industry where channels such as CNN and BBC produce news for national as well as international audiences. Regarding the former, our coming analysis will particularly focus on the consequences it has brought for online media texts, or as Thurman

puts it, how nationally produced news ‘could be re-presented to the international audience in a different way’ (Ibid, 301).

In order to conceptualise online news, we might first inquire in what ways the Internet differs from newspapers, radio or television as a mass medium, if it is merely a new channel for dissemination of news in already established formats, and to what extent online news has acquired new characteristics that somehow make it fundamentally different from other products of journalism. There are several reasons why we might consider the Internet as a hybrid medium. News websites are not simply digital versions of newspapers but a fusion of radio, television and traditional print media. Some authors have argued that the Internet is more than a hybrid medium, considering online news as a fourth kind of journalism (Pleijter et al. 2002, 28 in Deuze 2003, 206). As we will argue in this article, the Internet is clearly offering new possibilities for live reporting with the potential of making – to paraphrase Ingrid Volkmer (2006) – the ‘breaking news ideology’ of global news journalism even more powerful.

To the extent generalisations are possible, contemporary mainstream web journalism is a largely text-based medium; more so than traditional television, but less so than print media. As news sites take advantage of the Web’s specificity to distinguish themselves from traditional news formats, they are often doing so by providing new dimensions of journalistic immediacy through the production of news texts. Already in 2000, Gerd Kopper et al. made an ambitious review of ongoing research about Journalism and the Internet, which still in 2008 has some bearing on the topic. As online journalism in 2000 was a ‘very new and very fast developing branch of journalism’ (Kopper et al. 2000, 449) most research was lagging behind, something which is probably still very much the case. They discussed this problem in particular, as once the time-consuming process of setting up a research project and obtaining funding is done, the object of study might have changed or disappeared. Even today, these problems are related to the three commonly described characteristics of the World Wide Web – its *multimediality*, *hyperrelativity*, and *interactivity* (Stoval 2004, 5ff). During the last decade, online as well as offline media have increased their dependency on visual content tremendously. Where researchers in the past often reduced papers to their textual part only, the reduction of any audiovisual content from the unit of analysis is becoming harder and harder to make a case for. Increasingly, still and moving images, sounds, and the contexts in which the news is produced and read, must be acknowledged, just as in many cases readers are encouraged to participate as co-producers in reporting the news, using their cell-phones as audiovisual recorders and transmitters to the editing room.

John Knox (2007) argues that online news represents new communicative and social practices but its form and content must also be seen in the light of the

historical development of journalism. Although he declares that online news has not yet led to any considerable changes in the practices of newsgathering and writing, 'the impact of the internet on the packaging, distribution, delivery and reception of news in newspapers has been profound' (Knox 2007, 19). Online news provides a challenge for media studies, particularly since there are as many versions of news texts as there are consumers of them (depending on time of publication and reading, type of computer and web browser, Internet connection capacity, country or region from which the content is read etc.). The Internet's capacity of linking textual and visual materials to each other has also increased the internal and external intertextuality of news journalism, partly as a consequence of the new possibilities for 're-producing' news that the escalating digitisation of the media industry has brought. Moreover, online news represents profound changes in the media output that has been more or less neglected in recent years, namely visual and textual transformations over time. As Knox (2007, 48) concludes, 'there is a pressing need to describe the evolving multimodal macrogenre of the online newspaper', something that has not been easily done as most research has been using *synchronic* rather than *diachronic* approaches to online media texts. As stated earlier, this article can be seen as a response to the lack of methodological solutions to the versionality of online news, and in the following we will suggest a research technique that provides us with empirical materials suitable for diachronic analyses of single news texts.

Towards an Online Methodology

In order to perform a critical analysis of online news texts and by that attempt to grasp the changeable dimension of online media content, some methodological considerations are necessary. As journalism on the Internet is subject to rapid and constant changes, we first need a method for sorting and describing these changes. For this, we use systematic content analysis. In this study, we consider content analysis as a quantitative and systematic analysis of mass media content, from which one can draw general and representative conclusions that summarise the content rather than focusing on every single unit (Neuendorf 2002, 15). In order to further expand the analysis, we employ method triangulation by combining quantitative content analysis meant to give an overview of the changes that occur on the studied webpages, with qualitative discourse analysis enabling us to study the texts in greater detail but also in the light of its processual development seen through the quantitative analysis. Regular-Interval Content Capturing (RICC), which is our proposed technique for gathering the empirical material, and from which we will develop our analysis, will be explained in detail later on.

Discourse Analysis epitomises one of the most popular configurations of theory and method in contemporary social research. In media studies, Discourse Analysis has been a Grand Method for a long time, and as a counterpart to the empiricism

of more quantitatively oriented approaches, the researchers using it have delivered a large body of literature, explaining the relationship between power and 'the social constructions of reality' we daily encounter through the language in the media. Critical Discourse Analysis (CDA) has arguably become the standard work for studying media texts from an approach like the one above. The three-dimensional model of a) *Texts*, b) *Discourse practices*, and c) *Sociocultural practices* suggested by Norman Fairclough (1995; 2003) is perhaps the most established for the study of media discourse. From Fairclough we can also understand discourse analysis as 'seeing texts in terms of the different discourses, genres and styles they draw upon and articulate together' and a key feature is that the link between texts and society/culture is seen as mediated by discourse practices (2003, 58). However, there are a number of problems and challenges, not just to CDA scholars, but to the media research community in general, which must be dealt with if we want to adapt this approach to perform a critical analysis of online news texts. Tae Kim & Weaver (2002) note that despite the Internet revolution, media researchers still predominantly focus on one specific communication medium, the print media. The empirical data might be collected from electronic databases, but as such it still represents the print versions of newspaper texts, seen as 'the only thing available'. Gerlinde Mautner (2005) presents a similar critique and points to the significance of the Internet in contemporary social life which reasonably should result in an increased interest for web-based material for projects related to linguistics generally, and CDA in particular:

...if you are the kind of linguist who is interested in how discourse reflects and constitutes social life, and then ignore one of the key sites at which social life is being played out in contemporary society, you risk 'losing the plot' (Mautner 2005, 812).

However, Mautner also concludes that *the size* of the web creates an *embarras de riches* for the researcher. The data richness of the Internet is also connected to its, in Mautner's words, 'dynamic and ephemeral quality'; the collection of data definitely involves certain problems of making and keeping research replicable as the text corpus is 'changing constantly under your eyes' (2005, 818). Another problem is the unstructured and essentially anarchic character of the Net. Its diversity of voices and genres can possibly give a clue as to why many media scholars have been hesitant to the idea of ploughing the Internet for research data, when on the contrary, the digital form of the Internet facilitates easy searching and retrieving of data. As we see it, using online content is in the end about broadening the researcher's outlook, from a static field of vision, to a more dynamic and processual approach.

However, even though there is much talk about the transformative properties of digital texts, Mautner's explanation is still short of a more profound solution to the

analysis of these qualities. Returning to Fairclough, there is basically a theoretical and methodological black spot on the subject of (online) textual transformations. This is not a coincidence as it reflects a lack of research on how web-texts are the outcome of the processes by which they were made and how they can be seen as 'reflecting the technology available for producing them' (Bell 2003, 2). Whereas traditional CDA seem to consider both texts and genres as static in a ready-to-analyse kind of way, Mautner raises interesting questions of their dynamic character, but without providing us with a comprehensive method to analyse them. A frozen version of the corpus, she argues, is inevitably a distortion of the dynamic original (Mautner 2005, 821). In general these effects have been seen as *archaeological* problems, rather than strictly *textual* ones. The dynamics in online media research, are most often described in terms of intertextuality/hypertextuality rather than temporality, with a strong interest in the phenomenon of hyperlinks in a majority of recent web studies (see for example Karlsson 2006; Stovall 2004; Schneider & Foot 2004; Mitra 1999), but often reproducing the idea that there are old texts, new texts and nothing in-between. Interactivity and internal/external connectivity of texts are important aspects to consider, but as the only analytical parameter, it does not necessarily convey any substantial insights into the practices of online media discourses. It would seem apparent that media researchers in general need to adapt their methodological outlooks to the present digital media climate.

Regular-Interval Content Capturing (RICC)

One of the unique characteristics of online news is its relation to time and especially to its regime of immediacy. News journalism in general has always been dependent on scheduled routines. Television, radio and increasingly online journalism are all in one way or another live broadcast mediums, and as such they need to create a sense of liveness – being 'on location' at the right time in order to get proper images, backgrounds and earn public reliability. Through the study of online news, we can follow the development of stories in new ways, as online journalists are publishing (unfinished) texts implying that there is more to come. This new form of elastic breaking news agenda is a more or less unexamined area of contemporary news media, and can be studied through an examination of textual transformations over time. By adopting a dynamic, temporal and processual view of online news, we will emphasise the results of the process in which online journalists transform communicative events from 'nows' into 'news' in the online news flow. Whereas the most common way of studying the Net, is through an examination of separate 'finished' news texts, we will concentrate on the birth-and-death logic of texts published online. In the end, we propose both new ways to collect web-based data and strategies for analysing news items as dynamic and changing, rather than fixed and static, as displayed in table 1.

Table 1. Media characteristics and methodological outlooks

	Print media	Online News
Distribution	Periodic	Parallel flow(s)
Presentation form	Yesterday's news	Extended 'now'
'Traditional' text analysis	Static	'Frozen moments'
RICC	Static	Dynamic

The approach developed for this study is a technique for saving and analysing a time-series of consecutive versions of a news story. Rather than only looking at a single version or the end product, Regular-Interval Content Capturing (RICC) was designed specifically to fit the continuous flowing character of online news. The theoretical definition of news as changing instead of static is implemented methodologically by making cross-sections in the online news flow of a given website at set intervals, and saving the captured content as a time-series of data in an archive for later analysis. In this first trial of the method, we used an off-the-shelf screen-capture software together with a command script which executed the screen-capture software at set intervals, to save the main page of the studied websites as offline PDF files, time-stamped with the date and time of creation. Although more efficient custom software solutions can certainly be developed, this somewhat rough set-up still provided us with a reliable series of saved versions (as time-stamped PDF files) of the news flow on two studied editions of the CNN website we chose as our example. Likening the online news flow to an endless meat-loaf of content slowly coming out of the ovens of journalist production on the conveyor-belt of the Internet, the saved PDF files which make up the unit of analysis can be likened to slices of meat-loaf put on a plate in front of the researcher. These slices, making up the resulting data from the RICC method, as we discuss further below in relation to the data we collected, lend themselves well both to quantitative content analysis and to discourse analysis.

The RICC approach, just as it clearly opens up new possibilities for quantitative as well as qualitative analyses, also poses new challenges to the media researcher. The analytical potential of saved versions of a webpage at various points in time depends largely at what temporal intervals the material was collected. If the versions are saved with an interval of days they will resemble data collected with a 'static' approach and the full potential of the Regular-Interval Content Capturing method would not be used. Similarly, at intervals of less than a minute, the RICC method would suffer from a number of technical problems; bandwidth congestion, slow response from hosting servers, and limited processing capacity of the computer hosting the capturing software would likely cause a large number of the saved files to be incomplete or corrupt. The main advantage of the RICC method lies in between these two extremes, where it makes it possible for a researcher to study changes in webpage content at intervals from minutes to hours.

Choosing a suitable interval for capturing must be based on knowledge of the website that is the object of study, as well as on the subsequent method for analysis, and the intended purpose of the study. We will demonstrate this by discussing the findings of a case study of two versions of CNN.com in which we employed the RICC approach to collect the data for analysis.

Empirical Material and Methodological Application

For the empirical study of the ‘Crucial Tuesday’ coverage on CNN.com, we were interested in seeing if by using the RICC method, it would be possible to discern differences between the national and international editions of CNN. Due to the short-lived and changing character of online news, it was necessary to try to capture the changes over time in both the visual and textual components of the news flow on the respective sites, something that previous methods have not facilitated very well. Therefore, we studied how images, headlines, preambles and texts of the CNN top news stories change over time as a way of revealing the traces of discourse practice as well as discursive turns in the news flow. To put it simply, how does online journalism produce political breaking-news stories over time; which candidate is placed at the centre of attention during different stages in the flow; and how is the spectacle of politics framed for the national and international audience respectively?

Starting with a description of the collected data, the ‘Crucial Tuesday’ primary elections held on the 4th of March 2008 was the lead news story for 30 hours and 58 minutes in the US edition, and 33 hours and 42 minutes in the International edition¹. During this time, in total 1119 PDF versions were saved from the two editions through the RICC method described above, at 2, 5 or ten-minute intervals depending on the time of day. On a few occasions, a file was delayed, incomplete or missing due to bandwidth congestions or software error, meaning that the interval varies between 1-10 minutes, at one occasion being twenty minutes². When observing the number of changes found between adjacent files in the time-stamped material, we found that the chosen interval was adequate for the studied editions, although a shorter interval means an increased sensitivity for recording small and frequent changes in the site when such changes occur.

The 1119 time-stamped files were first coded using standard content analysis procedures according to 44 variables, intended to capture any changes in the top news story, and further give detailed information about which of the candidates were referenced in image, headline, preamble and text. In the analysis below, the most interesting results from the quantitative analysis are presented together with the findings of the discourse analysis. For the overall analysis, we have adopted a separation between micro and macro levels of analysis respectively. A macro oriented analysis of web news develops a deeper understanding of such simple

things as updating frequencies, but it can also tell us much about the rhetoric dimensions of image editing, text corrections, and the temporal relation between communicative events and text publishing. From large-scale analyses down to intervals of minutes, we can compare linguistic choices over time and in relation to what happens live, in the real world.

A micro-oriented analysis can reveal smaller transformations of news text. The selection process by editors and other journalists becomes highly visible, as there is only one top story at a time, and which one of a number of developing stories is placed on top is a multi-faceted question. Linguistics choices, metaphors and variation in visual techniques can tell us much more about how news evolves over time than traditional research methods previously have offered. Ultimately, we can dissect the process when breaking news is made and observe how different discourses, genres and modes of public address are constructed and mixed. Since we consider the text as temporally dynamic, textual themes and styles included, the method also offers the potential of unveiling journalistic priorities of online text production.

The CNN Online News Flow during the ‘Crucial Tuesday’ Primary Elections

The US and International editions of the CNN websites have the same general layout, with a top news story placed in the upper left corner below the logo and navigation menu. The top story consists of a square image positioned to the left next to a headline, a short text and one or more links to articles, video clips and other content. In the quantitative analysis presented here we have focused attention on changes made in the top news item as it is shown on the first page.

As stated earlier, we will implement an analysis alternating between a micro and a macro level respectively. The two versions of CNN we have studied show both similarities and differences regarding textual transformations over time. The reporting can be coarsely divided into three journalistic modes of writing which, separately and together, illustrate how online news are the outcome of journalistic considerations regarding the raw materials present at the time of writing, as well as the consumption potential at the time of publication. The three modes are characterised by specific writing styles, methods for selecting proper images as well as rhetorical dimensions closely connected to the relation between the past, present and future of the news. Before describing the modes we have chosen to call *establishing mode*, *intensifying mode*, and *closing mode*, in greater detail, we will first give a brief account of the general characteristics of CNN’s reporting.

Online news pages are always constructed within a framework of style and layout that is essentially fixed – both by the conventions of representations of news online, and by the built-in structures of the technological platforms used for

publication. Within this framework, the layout editor is highly active in adapting and changing the form as content is updated and new texts and images added. The top story in the US edition underwent 35 image changes, 26 preamble changes and 24 text changes during the nearly 34 hours which the elections were positioned as main news on the site. The corresponding numbers for the International edition are 30, 24 and 20, indicating that the US edition was updated slightly more frequently. Looking at the variation in images used to illustrate the top story in quantitative terms, CNN clearly chose different angles in its representation of the election on the US and the International edition of the website. In the US version, Republican candidates were shown in picture 31% of the time (almost nine hours), Democrat candidates were shown 44% of the time, and 25% of the time was shared (with composite pictures showing candidates from both parties). In the International edition, Republicans were shown 16% of the time, Democrats 79% of the time, and 6% of the time was shared. Thus, by merely looking at the choice of picture to illustrate the news, we found that Democrats were given a much larger exposure on the International than in the National edition of the CNN site.

Broken down into time for the different candidates, the attention was clearly on Hillary Clinton and John McCain. Clinton was shown 65% or 1091 minutes (42% or 712 minutes by herself) in the US edition, compared to 76% of the time (54% or 1016 minutes by herself) in the International edition. Obama was visually present 18% of the time (2% by himself) in the US edition and 30% of the time (7% by himself) in the International edition. McCain was in the image 55% of the time (30% by himself) in the US edition but only 21% (16% by himself) in the International. Finally, Huckabee was represented 9% (1% on his own) of the time in the US edition, and 4% of the time in the International. Here we can see large differences in which a candidate is focused on the respective websites, which serves to demonstrate a simple way of using the RICC method; these differences would likely have been missed using a methodological approach that did not account for the flowing character of online news.

Among the variables coded in images that also showed interesting results for the two editions, were whether or not the images showed ordinary people, and whether or not they displayed an American flag, in whole or in part. Ordinary people were shown either on their own, as voters or election workers mainly, or in conjunction with a candidate, during approximately one third of the time in both the US and the International editions (28% and 32% of the time, respectively). Use of the US flag as part of the image, in most cases hanging on a pole or the wall in the background, was slightly more common in the US edition, 36% of the time, compared to the International (31%), although it was a commonly shown symbol in both editions. However, broken down according to the four candidates, there were large differences regarding who was shown with the flag in the two editions. When Clinton was shown alone in the image, she appeared together with the

American flag only 3% of the time in the US edition, compared with 30% of the time in the International edition. Obama was never shown in front of a flag in the US edition, but 27% of the time in the International. For McCain there was no difference between the editions, 73% of the time he was by himself he was shown together with a flag in both editions. With the RICC approach it was possible to quantify the time each candidate was placed at the centre of attention, whereas with a static approach we would not have been able to see neither differences in time given to each candidate, nor any of the other changes over time which are central to the development of online news texts discussed below.

As we have argued earlier, one of the most important characteristics of online news is its relation to time or what we have called a regime of immediacy. Digital breaking news that is discussed here is – similar to television news – formed and rhetorically situated within a framework exhorting the public to ‘stay tuned’ as there are always more to come. The primary elections during ‘Crucial Tuesday’ were, of course, not an exception, but on the contrary a clear example of the kind of horse race agenda characterising political journalism in general.

Relating to the three modes of reporting we have identified from the material, the first mode can be described as the *establishing mode*, by which the primary elections are launched as main news on both of the sites. The primary elections occurred as main news on CNN International at 06.12 EST and about five hours later, 12.20 EST, on the US edition. This means that both editions want to establish the election stories around lunchtime in Europe as well as in the US, also illustrating that media consumption is central for understanding global online news discourse as a process oscillating between the global and the local level. ‘Candidates face crucial Tuesday battle’ is the opening preamble used on the International edition, and the site proclaims ‘a good chance at making the difference in deciding the Democratic and Republican presidential nominees’. In a similar fashion, the US edition describes the political battle as ‘Another make-or-break, do-or-die-primary day’ but instead of using the more or less well established ‘crucial’-prefix, CNN asserts that ‘voters on four states head to the polls today in what some analysts are calling critical Tuesday for the Democrats’. The initial formation of the elections as the main news of the day involved both historicising perspectives and a language explicitly directed towards the future. This was primarily realised by focusing on Hillary Clinton and her chance of becoming the second ‘Come Back Kid’ in the history of the Democratic Party. ‘Clinton hopes for wins, McCain looks to clinch’ (US) and ‘Clinton looks for comeback victories’ (International edition) are examples of how the narrative drama soon will meet a discursive turn as proclamations and projections – the very foundation of political breaking news – are starting to dominate CNN’s attention. Yet, textual transformations are held at a minimum and the construction of ‘liveness’ is limited to constant image changes,

alternating between individual and group pictures of candidates, voters and election officials.

An *intensification* of the update frequency as well as the rhetorical dimension of the news characterises the second journalistic mode. Although polling stations are still open, the news is increasingly narrated and dramatised, as the first comments from the candidates are published. Here, CNN start the more apparent live coverage of the elections and we can see how both texts and images are part of a mediated discursive turn. The intensifying character of the news flow is first and foremost symbolised by projections of coming results but also by an increasing use of sport metaphors as the mediatised political battle has begun. 19.03 EST, the International edition publishes the first forecast of the results: 'Republican John McCain and Democrat Barack Obama will win party primaries in Vermont state, CNN projects... The democratic campaign remains tight but Hillary Clinton is looking for a good day to boost the run' the site confirms. Similarly, the US edition follows up the latest counts by concluding that 'McCain wins Ohio, Vermont, CNN Projects; Dems close in Ohio'. During the following hours, we can follow the created drama, as new results are presented on both of the sites, enhanced by projections of what to come. 'Breaking News' and 'Developing story' are visual attributes used to call for public attention to stay tuned, also illustrating that the intensification also means a temporal shift from the past and the future to the present and constantly extended 'now'.

This can also be seen as an example of how the recent changes of the global media climate in favour of digital news production have extended the discursive repertoires of online news. The semiotic potential of this fourth form of journalism has also increased the available modes of representation of traditional news professionalism. Pleijter (2002) for instance notes quite ironically that the introduction of web journalism has sped up the whole news process, 'sometimes even causing journalists to spend more time at their desks instead of going out on the street' (Pleijter et al. 2002, 28 in Deuze 2003, 206). Interestingly, we cannot observe a clear demarcation between *projections* on the one hand, and final results on the other. The way results are presented is instead a matter of journalistic consideration regarding layout and the relation between preambles and the article texts. As time goes by, the 'p-word' is moved from preambles to article texts and from main clauses to subordinate clauses, seemingly by using the potential of revealing the final results without taking the risk of designating losers as winners without a 'plan B'. 01.04 EST, the International edition moves the attention from McCain to Clinton, as her wins in Ohio and Texas becomes more or less evident. Portrayed in a rain of confetti, Hillary Clinton 'bounces back' into the campaign whereas John McCain 'seals victory' and 'clinches nod' (US edition). Consequently, the reporting has gone from an open and close battle signified by the present, to finished results, taking the form of 'Hard news'.

Finally, the *closing mode* can be described as the representational method for closing the story as Breaking News. As results are clear, the style and form of the sites are part of a third discursive turn as journalists begin to summarise and most importantly analyse the results. In the morning, the day after the elections, the primaries have been the main news for nearly 24 hours of CNN International and Clinton's wins are occupying most of the attention of the site. Served as lunchtime news for the European audience, the words of Hillary Clinton, 'Campaign has turned corner' are replaced an hour later by the president giving his support to the new Republican candidate: 'U.S. President George W. Bush formally endorses his 2000 foe John McCain for president, an event that has both Republicans and their Democratic opponents excited'. The US edition develops this take on the elections further, by announcing that 'Bush's blessing for McCain pleases Dems'. Most importantly, however, the US edition of CNN reminds their readers of the fact that 'McCain can move on, Dems can't'. In a final and entirely new article, CNN goes back to their former starting point – new elections are to be held, and by that new strategies for establishing the democratic campaign as a political thriller in its own right. We can see a temporal shift from 'nows' to 'news' back to the 'open battle' in which Hillary Clinton - seen in the light of yesterday's news as history – needs further solid victories to justify 'a potential triumph at the convention, with a super delegate strategy'. To put it simply: a news rhetoric that goes back to the future. At 8.55 PM GMT, the top story on the International edition changes to 'Iraq fights for hearts of oil smugglers', and at 16.13 EST, the news flow meets its end in the US edition as the top story changes to 'Abbas reverses course, will renew peace talks', thus marking the end of the thirty-odd hours of online reporting from the Crucial Tuesday primaries on CNN.

To sum up the analysis, we have found that the representational strategies of CNN could be divided into three modes of reporting – establishment, intensification, and closure. By moving back in the cycle, usually by shifting attention to another area, the shelf-life of the news event is prolonged. Only when the story cannot be resuscitated again is the attention shifted elsewhere. During the life cycle of the event, competing stories from other events are given room in the breaking news banner, but are not moved into the top news area, meaning that even if the top news is constantly evolving, it does not jump back and forth, and what is the top news is thus fairly constant throughout the day.

Conclusions

In this study, we have searched for new ways around the methodological obstacles of studying online journalism. Many of these problems were formulated ten years ago or more, and have certainly meant that media research of online news has been held back, since many media researchers chose to stick with the 'old' media, instead of fuelling an academic surge in methodological innovation. With the

current acceptance among scholars that the Internet has come to stay there is now a slowly growing body of literature on the theory of the Web. However there is a large gap in methodology, where we hope that the approach presented in this article can contribute to filling a small part. There are several advantages by using an approach similar to ours. Perhaps the most revealing is the potential to be able to follow how news stories develop over time; since online journalists attempt to produce news as quickly as possible, and online news is increasingly published first and edited later, each step in the correction process tends to be published rather than making them all, and publishing a completely revised article half an hour later. These are the same kinds of editing 'offline journalists' have always been making in their daily work – only now we have front-row seats to watch the process.

In our case study of the two versions of the CNN.com website, we were able to show how the top news stories change dramatically over time, finding major differences in the framing of the primary elections for the national and international audiences. The large changes that the news stories underwent during the time observed displayed several discursive turns – from establishing the news, via intensified reporting, to closure. Through these modes, the relation to time turns from an extended now, into the reframing of current events as history and simultaneously looking towards coming 'stay-tuned' moments. We interpret the differences in framing as traces of discourse practices; online journalists' using the same sources, images and texts but adapting the form and content to fit different audiences and markets. These differences would not have been made visible to a researcher using a method for data collection that did not account for the versionality of online news.

Online mainstream journalism started as what Matheson (2004) among others called 'shovelware' – a digital space that was filled with unedited stuff produced for the printed papers. But somewhere along the road this has started to change, and in more and more cases, news are now produced on the web and for the web. This has as consequence that the printed 'offline versions' of newspapers are increasingly made from a compilation of content which was first written and published online, just as video clips may be edited for the online audience first, and only later made into a segment of TV news when there is a slot available in the chart. When this shift in the site of output of journalism began is hard to determine and is little discussed by researchers, partly because of the relatively few studies of online news that are made. This in turn is part of the lack of a simple methodology which enables a researcher to reduce a part of the Net which is considered an interesting object of study, into something which is possible to grasp and analyse, and furthermore, possible to store and retrieve in archives that remain available and unchanging. By reducing the processual flow of online journalism into images and texts, our method enables diachronic studies of online news texts. Moreover, by storing all collected data, the method lives up to the demands of

scientific replicability. However, the approach we have outlined here also poses new methodological and technological challenges to the researcher. In this study, we have focused on the main pages of the studied sites. RICC can, of course, also be used to collect and organise thicker empirical materials by following subpages and externally linked pages, although its technological implications must be examined further. Another problem is how to acknowledge the multimodality of online news with an approach like RICC. Although we have developed a powerful tool for the study of discursive transformations in online news, our suggested method is better at tracking changes in texts and still images, than in moving images and sounds. As we have shown in our analysis, online news is constantly changing in front of our eyes. In order to better understand the evolving versionality of online news texts, we should find ways of adapting RICC to incorporate the audiovisual media content, which increasingly fill our Internet browsers. Still, we believe that RICC may provide a viable way forward in methodology – in a research area where media scholars in general have lagged far behind their objects of study.

Notes

¹ The Crucial Tuesday primaries were the top story on CNN US edition from 12.20 PM EST on March 4, to 1.17 AM EST on March 6. On CNN International edition, the primaries were top news from 11:12 GMT on March 4 to 20:55 GMT on March 5.

² The top story was found to have been updated or edited in some way on average 1.8 times per hour in the US material (1.4 times/hour in the International material), and 5.0 times/hour during the most intense period of the material, meaning that even with an interval of ten minutes, we find it unlikely that we have missed any substantial number of changes. On average, we sampled the respective editions every 3.5 minutes. At the occasion when a file was lost and the interval between saved files increased to twenty minutes, there were no changes between the two saved versions, meaning that the missing file did not cause any loss of changes (as the webpage had not been updated during the extended interval).

References

Allan, S. (2005) *Journalism: critical issues*, Maidenhead: Open University Press.

————— (2006) *Online news: journalism and the Internet*, Maidenhead: Open University Press.

Bell, A. (2003) 'The Discourse Structure of News Stories', in Bell, A. & Garret P. (eds.) *Approaches to Media Discourse*, 2nd edition, Oxford: Blackwell.

Bell, A. & Garret, P. (1998) *Approaches to media discourse*, Oxford: Blackwell.

CNN Transcripts of 'American Morning', March 4 2008, available at <http://transcripts.cnn.com/TRANSCRIPTS/0803/04/ltn.01.html> (accessed 14.05.2008).

- Deuze, M. (2003) 'The Web and its Journalisms: Considering the Consequences of Different Types of News media Online', *New Media and Society* 5: 203-229.
- Fairclough, N. (1995) *Media discourse*, London: Edward Arnold.
- (2003) *Analysing discourse: Textual analysis for social research*, New York: Routledge.
- Flew, T. (2007) *Understanding Global Media*, New York: Palgrave.
- Karlsson, M. (2006) 'Nätjournalistik – En explorativ fallstudie av digitala mediers karaktärsdrag på fyra nyhetssajter', Ph.D. thesis, Lund University.
- Knox, J. (2007) 'Visual-verbal communication on online newspaper home pages', *Visual Communication* 6: 19-53.
- Kopper, G., Kolthoff, A. and Csepek, A. (2000) 'Research Review: Online Journalism – a report on current and continuing research and major questions in the international discussion', *Journalism Studies*, 1(3): 499-512.
- Matheson D. (2004) 'Weblogs and the Epistemology of the News: Some Trends in Online Journalism', *New Media Society* 2004, 6: 443-468.
- Mautner, G. (2005) 'Time to get wired: Using web-based corpora in critical discourse analysis', *Discourse & Society*, No. 16(6): 809-825.
- McNair, B. (1999) *Journalism and Democracy: An Evaluation of a Political Public Sphere*, London and New York, Routledge.
- Mitra, A. (1999) 'Characteristics of the WWW Text: Tracing Discursive Strategies', *Journal of Computer-Mediated Communication [Online]*, vol 5(1), available at <http://jcmc.indiana.edu/vol5/issue1/mitra.html> (accessed 16 June 2008).
- Neuendorf, K. (2002) *The content analysis guidebook*, London: Sage.
- Pleijter, A., Tebbe, F. and Hermans, L. (2002) 'Nieuwe journalisten door nieuwe bronnen?' in Deuze, M. 2003 'The Web and its Journalisms: Considering the Consequences of Different Types of Newsmedia Online', *New Media and Society* 5: 203-229.
- Schneider, S. & Foot, K. (2004) 'The Web as an Object of Study', *New Media and Society* 6: 114-122.
- Stovall, J. (2004) *Online Journalism: Practice and Promise of a New Medium*, Boston: Allyn and Bacon.
- Tae Kim, S. and Weaver, D. (2002) 'Communication Research about Thematic Meta-analysis', *New Media & Society* 14(4): 518-538.
- Thurman, N. (2007) 'The globalisation of journalism online: A transatlantic study of news websites and their international readers', *Journalism* 8(3): 285-307.
- Volkmer, I. (2005) 'News in the global public space', in S. Allan (ed.) *Journalism: Critical Issues*, Maidenhead: Open University Press.