

**IFLA International News Media Conference:
“Collecting, Preserving, and Transforming the News – for Research and
the Public”**

27-28 April 2017

Reykjavik

Investigating Journalistic Time over Time

Henrik Bødker

Department of Media and Journalism Studies,

Aarhus University, Aarhus, Denmark.

E-mail address: hb@cc.au.dk



Copyright © 2017 by Henrik Bødker. This work is made available under the terms of the Creative Commons Attribution 4.0 International License: <http://creativecommons.org/licenses/by/4.0>

Abstract:

Journalistic products are central in the constitution of both individual and collective time; and this happens within complex relations between media technologies and broader cultural patterns. Based on two earlier studies investigating the textual constitution of time in digital media (both live and archived) this paper discusses the parameters of a longitudinal study of journalism’s textual constitution of time over time.

The empirical base of this is the Danish newspaper Berlingske Tidende, which has been partly digitised from its inception in 1749 by the Royal Danish Library. The immediate challenge is develop and expand the analytical approach employed in the earlier study of the development of online news — based on websites held at the Internet Archive (<https://archive.org/index.php>) — from a hand-held qualitative analysis to a semi-automatic analysis of a larger body of digital data. The starting point is an deductive, semi-automated study that draws out explicit markers of temporality (dates, times, weekdays and modes of verbs in headlines) from the texts on the front page (of a digitised newspaper or an archived website). Patterns revealed at the morphological level need, however, to be complemented by an analysis of the syntactical level, e.g. of the actual and relative positioning and size of temporal markers on the page and, perhaps, within the discrete articles.

The study discussed in this paper is a pilot study to be conducted together with my colleague Professor Niels Brügger in close collaboration with the Royal Danish Library in Aarhus; this study is simultaneously embedded within the newly formed special interest group focused on digital methodologies with regard to print and journalism (headed by the author) under Digital Humanities Lab, Denmark. In addition to the specificities of the pilot study, this paper thus also raises issues linked to inter-institutional collaboration and trans-disciplinarity.

Keywords: journalism, media, temporality, methodology and digital archives.

1 Introduction

Journalism and its products are deeply intertwined with the constitution of the rhythms of the community in which the journalistic institution is embedded; and this happens through an intricate interplay between media technologies and broader patterns linked to economic transactions, political processes, patterns of work, family life and leisure. A significant part of the discussion about how such processes have evolved over time has been centred upon notions of speed, i.e. how life in western societies in general have been sped up through the interaction between digitalisation and market economies and within that how journalism has become (even more) obsessed with breaking and constantly updating news.

While there at some levels there is plenty of evidence for such an analysis it must be emphasised that digital technologies also turn news websites into evolving repositories of yesterday's news. This is a view that addresses the temporality of journalism from its textual manifestations rather than the related practices of the newsroom, the field or how time is constituted through processes of consumption. From the perspective of journalism as text (in a broad use of that term) one can thus argue that as much as the digital allows speed of transmission and updating it is also about the possibilities of accumulation and retrieval. Put differently, the news website allows for a complex and interactive construction of historical contexts for the latest news. In some ways, the news of the present thus drags along its own genealogy. The different ways in which this can be conceptualised were the focus of my article "The Time(s) of News Websites" (Bødker, 2016). An important aspect of this analysis was to stress that the temporalities constituted textually through the news website must be analysed as relations between different temporal markers on the site, e.g. relations between the masthead "The Guardian" (in a blue colour signalling calm trustworthiness earned over time) and a flashing red "breaking news" banner.

The specific frame that guided the analysis of this article was derived from Brügger's (2010) approach to the textual study of websites; within this frame websites (on different levels) acquire meaning through the morphological and syntactical characteristics of textual elements (e.g. the relations between breaking news signifiers and makers of institutional longevity). In the article "The Shifting Temporalities of Online News: The Guardian's news site from 1996-2015" (Bødker and Brügger, 2017) the framework developed by Brügger (2010) as well as the insights from its application to online news in Bødker (2016) were applied to a selected range of news websites. The broader aim of this was, firstly, to conduct an analysis that might nuance more general discussions of speed as the main concept through which to understand online news and their development; and, secondly, to test the analytical approach on historical, digital data. With regard to the first aim, the article tells a complex story of how different temporalities intersected on the website and that speed was not the main characteristic of either the early or late versions of the site. While, in terms of the second aim, the applicability of the analytical framework was largely confirmed the research also revealed some complexities of working with archived web data, complexities that in some ways were accentuated through a study aimed at revealing relations between website elements in terms of temporality.

Yet, in addition to such methodological difficulties (which are discussed at more length in the article) the results also raised interesting questions about how journalism has constructed time textually over longer time periods and, in relation to that, whether it would be possible to transform aspects of the analytical framework into semi-automatic

procedures that could inform an analysis of broader developments of how the textual form of journalism has constituted time over time. Given the exploratory character of such a study (the contours of which will be outlined below) it should be stressed that such an endeavour would be just as much focused on experimenting with approaches to digitally archived journalism as it would be aimed at revealing historical patterns.

As a parallel to the study on the archived websites discussed above such a broader study would (initially) focus on how various textual elements on the front page of newspapers collectively constituted time. Time is, of course, also constructed within discrete articles (through temporal markers, events, verb tenses etc.) and this has been the focus of some research not least in relation to questions of the past and/or memory, e.g. Neiger, Zandbjerg and Meyers (2014) and Tenenboim-Weinblatt (2013). Individual journalistic texts are, however, situated within a broader framework that constitute time in terms of seriality, temporal makers, positioning on the page etc. In terms of the construction of meaning these two levels — the individual texts and their broader contextual forms — cannot be fully separated. But given that much journalism research has been focused on analysing meaning within articles Bødker (2016) and Bødker and Brügger (2017) were precisely focused on calling more attention to how the constitution of meaning (here temporality) also is a product of the textual form of the newspaper and/or webpage. It is this perspective and related insights that the study proposed here will attempt to apply to a longer historical development.

A very obvious question here relates to how the addition of the digital in the mid-1990s affected how print constituted time. At the level of articles Tenenboim-Weinblatt and Neiger (2014) recently did a study of the main temporal directions of print and digital journalism in the Israeli media and concluded (broadly) that the former was mainly future oriented and the latter (immediate) past oriented. While this helps nuance the broader discussed focused on speed it is also important to stress that such differences are tied to the specific cultural and social contexts of the news outlets studied. This is obviously also the case if one wants to apply the analytical framework outlined above to a longer historical development.

A semi-automatic study of how journalism textually constitutes time over time consequently necessitates a number of important considerations: 1) what specific journalistic institution(s) is/are to be studied? 2) what are the broader cultural, social and economic contexts of these institutions? 3) What texts are available and in what form? and, 4) how can the answers to the first three questions inform the construction of a research design that will allow an analysis of how the textual constitution of time has evolved over time? Below these considerations will be developed and nuanced through a progression towards a preliminary research design.

2 The Journalistic Institution *Berlingske Tidende* and its wider contexts

Berlingske Tidende has been chosen for a number of reasons for this preliminary study. First, this is Denmark's longest running newspaper (1749 to the present), which (at least in theory) allows the researcher an almost unbroken trajectory through which developments can be assessed. It is, however, important to acknowledge that the paper in its different incarnations constitute somewhat different institutions, which need to be understood in their respective contexts. When, why and in what contexts did the paper for instance

change name (e.g. to *Den til Forsendelse med de Kongelige Brevposter privilegerede Berlingske politiske og Avertissements-Tidende* in 1831); ownership (when the Berling family bought it in 1831); printing technology (e.g. the first rotation press in 1881); issues (e.g. twice daily six days a week in 1847), modes of distribution etc.? It must also be taken into account the extent to which the newspaper is conceived as local, regional and/or national since this is related to the various rhythms underlying the selection and presentation of news.

While these contexts will not be explored in more detail here they are indeed important not only for designing the study but also for interpreting its results. And, so are aspects from the broader political, social and cultural contexts within which markers of temporality must be understood. While this may seem obvious such considerations are important to forefront in order to legitimize the use of aggregated digitised/digital data within broader discussions of their de-contextualisation (see for instance Jensen, 2016). In addition to such questions about broader contextualisations there are related questions about which journalistic texts have been preserved in a digital form, how this transformation has taken place and through which channels and structures they are available.

3 What texts are available and in what form?

The State and University Library in Aarhus — recently merged with the Royal Danish Library in Copenhagen — holds newspapers dating back to 1666 and is in the process of digitising their print newspaper holdings; and *Berlingske* is one of the newspapers that the library decided to digitise from its inception to the present. From 2005 the royal library began to harvest the Danish web domain and with that a lot of the websites of the Danish newspapers. Stored versions of the website of Berlingse from the mid-1990s to 2005 will thus have to be found in the Internet Archive (<https://archive.org/index.php>). Taken together, the newspaper collection and the (national) web archive at the Royal Danish Library and the Internet Archive thus contain texts from *Berlingske*.

The digitised newspaper pages are available as images with a corresponding page “behind” the image in which the articles and other items have been segmented, the text converted to a searchable text file, and each word and item have been given coordinates in relation to their position on the scanned page. A search for a specific word will find the work in this file and mark the word on the image of the newspaper page so you can read the context. In addition to the possibility of keyword search the coordinates of items can be useful for understanding how specific markers of time are positioned and grouped on the page (something which I will return to below). This is also possible for the archived webpages by drawing on data from the stored HTML-code.

4 A Possible Research Design

Given the considerations briefly touched upon above the next step is to outline and discuss a possible research design that can help grasp some of the shifting ways in which the front page of *Berlingske* has constituted time over time. At an overall level the time period, media and “depth” of the study needs to be determined. With regard to period one could

look at the publication from its inception to today and thus, from this is possible, include the website. In terms of depth one could restrict the analysis to the front page of both the newspaper and the website. In relation to this, it also needs to be considered whether this would include all editions, i.e. every page available, or whether a specific weekdays/editions should be singled out. Another possibility would be to focus on selected time periods and/or events (e.g. specific national crises, wars etc.) in relation to which broader questions of (national) temporality might be emphasised. Given some of the issues raised below it might, however, be advisable to select a shorter time period in which to test the design.

In continuation of the analytical frame outlined above the study will basically attempt to draw out two types of data, which very broadly can be termed morphological and syntactical. While the former is focused on the qualities of the isolated textual elements, the latter is concerned with their relations. At the morphological level one could, in a relatively simple fashion, search for temporal markers such as day, time, weekday, month and year, yesterday, tomorrow, now, present, past, future and phrases such as “published since” (either in the masthead or as ways to begin articles) and/or the tense of verbs in headlines. Data from a number of searches could be accumulated in a semi-automatic fashion. From a list of pre-selected temporal markers one could design a script that could search for these in various combinations, i.e. day, day and month, day or month etc. The accumulated data would then have to be stored in way that shows and allows analysis of the appearance of terms in time. At a basic level this would, however, be akin to deductive “dictionary-based techniques” (Bouman and Trilling 2015: 5) applied within other fields;¹ and, while the dictionary manually put together in this instance would not be directly transferable to other studies it would, on the other hand, have a fairly high validity.

Applying such a dictionary of temporal terms could reveal when certain terms appear or disappear and which terms they appear with or without and also how events are framed temporally on an overall level. This could give some indication of how this particular publications aims to situate itself and the covered events within broader temporal patterns. Based on a coding and accumulation of results specific periods of change could be singled out for more scrutiny. Such changes could, as was the case in the study of *The Guardian* Website, correspond to major changes in layout and format. Changes could, however, also be linked to specific and major news events or the advent of new competitors. Singling out periods of transformation could help reveal some of the ways in which media technologies and journalistic practices interact with broader contexts.

An deductive, dictionary-based approach on the morphological could be complemented by an analysis at the syntactical level. Initially, this could look how the terms applied in the first study appear relative to each other on the page. Are, for instance, the majority of the temporal markers clustered in the upper left corner of the page? And, do such patterns change over time? In relation to temporal markers on the semantic level one could here also look at the relative size of headlines (with specific temporal markers) as well as illustrations and photographs, which introduce a different kind of temporality into journalistic coverage. In relation to this, one could for instance look at which and how close specific temporal markers appear in relation to photographs.

¹) Bouman and Trilling (2015) distinguish between deductive and inductive digital approaches where the former apply “predefined categories” and the latter where “the computer rather than the researcher ... makes the decisions about what is meaningful in the dataset” (3).

While a dictionary-based study would rely on a simple (yet repeated) word searches an analysis at the syntactical level would need to draw out and work with data that are not directly accessible to the ordinary user of the archive. Such an analysis would thus require collaboration with the people at the Royal Library, which could help design a series of scripts that could draw out the data of relative positioning. Such data would, however, also have to be analysed in terms co-occurrence and of relative proximity and distance with regard to the different markers and/or textual elements investigated. This could, as Boumans and Trilling (2015) mention in relation to framing analysis and “co-occurrence of words” be done by “applying statistical techniques like *principal component analysis* or *cluster analysis*” (8; emphasis in the original), which then could be “graphically visualised” (8) as (shifting) relations on a newspaper page or website.

A inductive analysis of how terms and other textual elements linked to temporality have been related on the page over time may reveal something about the importance given to temporal markers in terms of positioning as well as how new media technologies (e.g. photography) played into that. Yet, as pointed out above with regard to the appearance of selected terms, one needs to revert to knowledge about technologies of newspaper production as well as other contexts (economy, important events etc.) in order to understand overall patterns and shifts revealed by the inductive analysis. Most likely, then, such an analysis would also be helpful for formulating new questions.

5 Conclusions and Considerations

The considerations outlined above relate to a possible study. The inspiration for this was, as was pointed out, an earlier study done on archived websites, which revealed results that might be worth examining over longer time periods given the increasing amount of digital data available at Royal Danish Library of Denmark. In broader terms such growing bodies of data may help unearth new evidence and new questions for a broader longitudinal understanding of *the form of news*, which was the title of Nerone and Barnhurts’ seminal book-length study of the various elements that played into making the news look as they did and how this changed over time (Nerone and Barhurst, 2001). Utilising the available digital data can reveal historical patterns that may help nuance contemporary discussions about speed and through that also put into perspective some of the broader shifts in cultural and/or societal temporalities.

Such studies do, however, not only on the availability of data but also on intermediaries that can assist scholars of journalism and journalism history in conducting analysis that rely on methods that scholars within this field not normally have in their toolbox. In this particular case what is needed is permission as well as resurces that can collaborate on the designing of scripts that can draw out data as well as making these available in a format suitable for later anlysis; and such collaborations may be futhered by appropriate fora, which is why the author of this article has taken the initiative to form a special interest group under the Digital Humanities Lab, Denmark. This group, which consists of journalism scholars from Danish universities as well as people form the Royal Danish Libraru is aimed at developing knowledge about some of the possibilities of research based on in the growing body of digitally stored journalism. The group will host seminars where foreign scholars, who have conducted larger studies of journalism based on digital archives, present and discuss issues related to research and metodologies. Another activity

in the group will be to initiate smaller pilot studies; the study outlined above is meant to be one of these.

Acknowledgements

I wish to thank my colleague Professor Niels Brügger for discussing some of the issues raised in this article with me.

References

Boumans JW & Trilling D (2014) Taking Stock of the Toolkit. *Digital Journalism*. 4(1), 8-23.

Brügger N (2010) *Website analysis: Elements of a conceptual architecture*. Aarhus: Papers from The Centre for Internet Studies.

Bødker H (2016) The Time(s) of News Websites. In: Franklin B and Eldridge II S (eds) *The Routledge Companion to Digital Journalism Studies*, Abingdon: Routledge, pp. 55-63.

Bødker, H & Brügger (2017) The shifting temporalities of online news: The Guardian's website from 1996-2015. *Journalism*, e-pub ahead of print, Feb. 7 2017.

Jensen, Helle Strandgaard (2016) Doing media history in a digital age: change and continuity in historiographical practices. *Media, Culture and Society*. 38(1), 119-128.

Neiger, M Zandbjerg, E and Meyers, O (2014) Reversed Memory: Commemorating the Past through Coverage of the Present. In: Zelizer, B and Tenenboim-Weinblatt, K. (eds.) *journalism and memory*. Basingstoke: Palgrave Macmillan, pp113-127.

Nerone, J & Barnhurst, KG (2001) *The Form of News — A History*. New York: The Guilford Press.

Tenenboim-Weinblatt, K (2013) Bridging Collective Memories and Public Agendas: Toward a Theory of Mediated Prospective Memory. *Communication Theory*. 23(2), 91–111.

Tenenboim-Weinblatt, K. and Neiger, M. (2014) Print Is Future, Online Is Past: Cross-Media Analysis of Temporal Orientations in the News. *Communication Research*. 42 (8), 1047-1067.