2.   News waves in a changing media landscape 1950-2014

Wouter van Atteveldt, Nel Ruigrok, Kasper Welbers, and Carina Jacobi

Abstract

Over time the number of news waves seem to be increasing. Due to intensified competition, journalists are increasingly responding to stories in other media and audience preferences, at the expense of stories on political issues. By using the digitalized archive of the Dutch newspaper *De Telegraaf*, news waves were automatically identified in the four million articles from 1950-2014. Contrary to expectations, the amount of news waves has not increased linearly since the 1950s, but shows a U-curve, with more waves in the 1950s than in the 1970s. Political content in news waves peaked in the 1990s. The general shape of news waves consists of a small lead, a sharp and narrow peak, and a slow decline.

Keywords: measuring news waves, automatic text analysis, political communication, Dutch journalism

Introduction

In recent years, numerous studies have investigated the dynamics of media hypes, media storms, and news waves (Boydstun et al., 2014; Vasterman, 2005; Wien & Elmelund-Præstekær, 2009; Djerf-Pierre, 2012). Although most authors agree on a loose definition of a wave, storm, or hype as a sudden burst of attention for a topic, there is no consensus on one of the most important criteria of 'hypes', namely that the amount of attention is excessive or unwarranted by the real events. This lack of an objective norm about how much media attention should be devoted to an event, makes it difficult to define when media attention is excessive.

A key assumption in the study of news waves is the self-referential nature of media attention. If media coverage were a pure mirror of reality, the amount of attention to an issue would naturally follow events and real-world developments around that issue. Media attention, however, is also guided by the media's need to 'tell a good story' (Wolfsfeld, 2014), leading to issue attention cycles where an issue bursts into attention after a key event, later fading again as all relevant arguments are made and new issues demand attention (Downs, 1972; Zhu, 1992). This intra-media cycle is reinforced by intermedia effects as journalists look to each other to reinforce their sense of news (Cook, 2006). Finally, there are positive feedback loops with audience and (political) sources, as journalists want to write about issues high on the public and political agenda, but this agenda is shaped (and gauged) by the media attention itself. These three positive feedback processes of intra-media, intermedia, and extra-media dynamics reinforce each other, leading to the self-referential spikes of attention defined as news waves.

In keeping with mediatization theory (Strömbäck and Esser, 2014), the self-referential nature of journalism is expected to have increased in the last decade. The storytelling, intermedia, and audience feedback loops identified above are all essential parts of media logic (Altheide, 2004), which is assumed to have increased over time with increasing media independence and commercialization (Asp, 2014). The fierce competition among (new) media outlets means that resources and time investments shrink and journalists increasingly tend to focus on each other, afraid to miss a scoop. This reinforces the imitation behaviour among journalists who are put under pressure to publish news quickly (Anderson, 2011). As a result, journalism is said to have transformed from a trustee model, with journalists providing news they thought the public needed, into a market model, where journalists cater to the audiences preferences to decide what is newsworthy (Schudson, 2003). This leads to media covering news stories that are likely to please large audiences and advertisers (McManus, 2009), meaning those stories highest in news value, such as a focus on elite persons and negativity (Takens et al., 2013). Looking for confirmation (Donsbach, 2004) media become more self-referential, leading to more events that trigger huge and fast developing news waves.

Due to the difficulty of defining and identifying news waves, however, many existing studies focus either on a single topic for which it is easier to compare attention (e.g. Boydstun et al., 2014; Vasterman, 2005; Ruigrok et al., 2009) or on a single (series of) waves, arguing the excessive nature of attention on qualitative grounds (e.g. Vasterman, 2005). As a result, similar to many predictions from mediatization theory, there is no sound

empirical evidence that news waves have indeed increased in frequency and intensity.

Research into the dynamics of news waves is important because people depend heavily on the media for information on societal and political developments. News waves concentrate media attention on a single event or issue. In certain cases, this can be beneficial, such as the 'burglar alarm' function that Zaller ascribes to the media (2003): if something really bad is happening, the media need to sound the alarm, using the sudden increase of attention of the news wave to 'wake everyone up' and force attention on the issue. On the other hand, media logic can sometimes be shaped more by what the public finds interesting than by the public interest (Takens, 2013). If news waves take away attention from real issues by focusing on scandals or other 'good stories' (Wolfsfeld, 2014), it is possible that real problems do not get addressed because they do not fit the right patterns or news values required for a wave to occur. Thus, it is important to understand what kind of events can trigger news waves and under which circumstances waves form and end. By providing an automatic way of identifying and measuring news waves and presenting longitudinal results on the frequency, shape, and political content of waves, this study provides a starting point for a more quantitatively informed discussion of how we can identify waves and understand their dynamics.

In this chapter, we propose a method for automatically identifying news waves based on the attention to clusters of co-occurring words using topic modelling. By applying this method to about four million newspaper articles in seventy years of coverage in the Dutch newspaper *De Telegraaf*, we can show whether the amount, duration, and size of news waves has changed during this period. This will show whether news production routines that lead to news waves have changed under the influence of professionalization and commercialization of the press, as predicted by mediatization theory (Strömbäck, 2008). Moreover, we will measure whether a news wave is political in nature, giving insight on whether news waves have shifted from hard news to soft news, as would be expected from a market model of journalism (Schudson, 2003).

Theoretical framework

The term 'media hype' is often used in popular writing to indicate a sudden and presumably unwarranted amount of attention for an issue. In the scientific literature, this term is defined as 'a media-generated, wall-to-wall news

wave, triggered by one specific event and enlarged by the self-reinforcing processes within the news production of the media' (Vasterman, 2005: 515; Wien & Elmelund-Praestekaer, 2009). Other authors prefer other terms than hype. Wolfsfeld and Sheafer (2006) use the term 'media wave' to denote periods in which an issue suddenly gets a lot of attention, using the metaphor of politicians 'riding the wave' to use the media coverage to their advantage. In a study comparing such news waves with non-wave periods, Boydstun, Hardy, and Walgrave (2014: 511) talk about 'media storms', defined as 'an explosive increase in news coverage of a specific item (event or issue) constituting a substantial share of the total news agenda during a certain time'. Since we will not make any comparisons with real-world indicators of newsworthiness, we will use the term 'news wave' in this study, loosely defined as a sudden, short-lived, and explosive increase of attention for an issue.

News waves as a result of positive feedback loops

People use the media to provide 'information that people can trust and act upon' (Strömbäck, 2005: 339). News, however, is not a direct reflection of reality. Rather, journalists construct 'good stories', and this construction inevitably includes choices on what to cover and how to cover it (Wolfsfeld, 2014; Tuchman, 1978; Gans, 1979). Events that are negative in nature, unexpected, and happen in elite countries, for example, are more likely to be covered than more complex, far away events (Galtung & Ruge, 1965; Harcup & O'Neill, 2001, 2016). Although such news values can explain which events are turned into news stories, they cannot explain why some events get the internal momentum required to turn into news waves. This 'sudden change' in the newsworthiness of events can only be explained by looking at the internal dynamics of media production, in which the influence from other journalists is often at least as important as the intrinsic characteristics of the events. News waves, then, are triggered by an external newsworthy event, but they grow into waves by self-reinforcing positive feedback processes inherent in the logic of the media (cf. Vasterman, 2005). In particular, we can identify three levels at which these processes occur: within a single news organization (intra-media); between organizations (intermedia), and between journalists and their sources and audiences (extra-media).

To start with the intra-media level, in 1978 Fishman found, when looking at the news coverage of both newspapers and local television stations and their reports on crime, that news coverage is presented within a 'theme',

such as violence against the elderly, providing 'presentational order' to the news: 'Items are presented in groups organized around a theme, and items that fit the current theme are covered more extensively with media applying a "consistency rule"' (Fishman, 1978: 534-535). In other words, news determines news. Vasterman (2005) found this same pattern in his study into coverage of random violent acts in the Netherlands. This all-inclusive attention to the topic reduces the news thresholds to other events and information related to the central theme (Vasterman, 2005: 513). As Paimre and Harro-Loit (2011: 435) state: 'This self-referential nature of journalistic performance is an inductive factor for any type of news wave'.

A second dimension of the self-referential character of news selection deals with the intermedia influence. As stated by Cook (2006: 78), competition 'does not push reporters toward the exclusive "scoop" but instead toward risk-averse consensus, on the presumption that the glory they get from the former is less than the trouble they might face if the scoop came into question or if they missed out on the big news story everyone else covered'. The nature of such pack journalism is also an economic choice: it is easier and cheaper to follow-up on an existing story than to start from scratch and attach value to a new event (Lund, 2002). The increasing number of media outlets and commercialization creates further incentives for journalists to compete by copying (Boczkowski, 2010; Redden & Witschge, 2010). Journalists also depend more strongly than ever on press agencies' news feeds, especially online news outlets that need a steady supply of fresh and reliable content to compete in the high-paced 24/7 news cycle (Welbers et al., 2016). Ultimately, the tendency of journalists to monitor their colleagues and institutional sources to avoid missing big stories or making errors can give rise to hypes (Ruigrok et al., 2016: 12).

The third dimension of the self-referential character of news selections deals with the relationship between the media and their audiences and sources. The changing media landscape led to a shift from a 'trustee model' (Schudson, 2003), where professional judgements by journalists determine news selection choices, towards a 'market model', implying that journalists cater to the preferences of the audience (Hamilton, 2004; McManus, 2009; Strömbäck et al., 2012). As Brants and Van Praag (2006: 30) put it: 'the assumed wishes and desires of the public have become more decisive for what the media select and provide'. This is especially true for online news sites. Several studies found that online audience metrics affect the news selection choices of journalists (Anderson, 2011; Welbers et al., 2015). Rather than the inherent importance, the perceived appeal to target audiences determines the newsworthiness of events (Niblock & Machin, 2007: 191; Strömbäck et

al., 2012: 726). The interest of the audience, however, is determined partially by the media themselves (McCombs & Shaw, 1993; McCombs, 2005). By reporting about an issue, public salience of the issue increases, making it more attractive for journalists to follow up on the story, until the curiosity of the public is sated and interest diminishes (Downs, 1972; Djerf-Pierre, 2012).

Similar to their audience, journalists and (political) sources can reinforce each other in what Wolfsfeld dubbed the Politics-Media-Politics cycle (Wolfsfeld, 2014). Van Aelst and Vliegenthart (2014) show how parliamentary questions are often based on media coverage, but these questions, in turn, generate more coverage of the event. Moreover, Ruigrok et al. (2009) show how Geert Wilders, a Dutch populist anti-Islamist politician, plays a central role in two news waves, initiating a first wave by publishing an anti-Islam film and using his political standing to generate press attention for it; and accelerating a second wave by asking inciting parliamentary questions about incidents with Moroccan youths in the town of Culemborg, turning a local incident into a national news wave (Korthagen, 2011).

Evolution of Dutch journalism: *De Telegraaf,* 1950-2014

In this study, we focus on the Dutch newspaper *De Telegraaf* in the period of 1950-2014. After the Second World War, Dutch society was still 'pillarized' along religious and political lines. The media were largely incorporated into the socio-political system and would aim to influence their (particular) readership with the same opinions as reigned in the pillar (Broersma, 1999, 2007; Wijfjes, 2004). With the process of depillarization starting in the 1960s newspapers were no longer affiliated with a particular party, but showed an internal pluriformity in terms of content (Van der Eijck, 2000: 329; Hallin & Mancini, 2004). Journalism became guided by a strong sense of professional norms, especially independence and objectivity, which manifested itself in a critical watchdog role, scrutinizing the political elite and informing citizens (Schudson, 2003; Strömbäck & Esser, 2014). After 'depillarization' in the 1960s and 1970s and the weakening of the ties between newspaper companies and readers, a strong increase in mergers and takeovers was observed within the newspaper market. Media outlets have become profit oriented rather than politically oriented (Harcourt, 2005). Together with an increasing journalistic independence, politics journalism showed 'a move away from reliance on craft norms defining what is newsworthy and how to report, toward a journalism based on serving the marketplace' (McManus, 1995: 301). Although *De Telegraaf* was not attached to a specific pillar or political

party, the newspaper went through the same phases of growing independence followed by stronger commercialization and increased competition (Wijfjes, 2004) fitting the pattern described as mediatization (Strömbäck & Esser, 2014; Brants & Van Praag, 2006).

Media content became more strongly guided by 'media logic', which can be defined as 'the assumptions and processes for constructing messages within a particular medium. This includes rhythm, grammar, and format' (Altheide, 2004: 294). Moreover, commercial pressure and competition for an ever more fickle readership caused commercial interests to become a stronger factor within the media logic, especially with the onset of commercial television (1989 in the Netherlands) and the spread of online news sources. As defined by Schudson (2003), the trustee logic of independent journalists writing in the public interest was replaced by a market logic of journalists writing what sells best (Brants & Van Praag, 2015).

Empirical support for these shifts in media logic has been more difficult to find, however, partly because of the long time scales involved (Strömbäck & Esser, 2014). A study into news coverage of political campaigns revealed a clear shift towards more independent and critical reporting between 1956 and 1986 (Brants & Van Praag, 2006). Although they do report a shift towards more contest coverage between 1986 and 2003, news was mostly still substantive and free of cynicism in 2003. Takens et al. (2013) investigated negativity, contest news, and personalization in Dutch election coverage from 1998 to 2006 and found a significant increase only in the level of presidentialization, and a small peak in 2002 followed by decline in the other indicators.

Although the empirical evidence on the general mediatization thesis is mixed, and there are no general, longitudinal studies of news waves, based on the trends reviewed above we expect the amount of newspaper coverage determined by news waves to have increased from 1949-2014.

H1. From 1950-2014, news coverage in De Telegraaf *is increasingly determined by news waves.*

Moreover, due to increased focus on the interest of the audience, especially after the 1980s, we also expect the amount of 'hard news' in news waves, here exclusively operationalized as political news, to have decreased:

H2. From 1949-2014, the political content of news waves in De Telegraaf *decreases*

A final interesting question pertains to the shape of news waves. News waves are assumed to show a 'sharp and continuous increase of reporting on a specific issue for a limited period of time' (Geiß, 2011: 272). Moreover, it is

often found that waves can go through different phases, where a first wave of attention about the facts of the news wave is followed by smaller wavelets as sources discuss the (political) implications of these facts (cf. Ruigrok et al., 2009). Vasterman's (2005) seminal study contains a description of the shape of the news waves studied, showing an increase during the first week and a total length of around three weeks, but these are mostly schematic and based on single events. To our knowledge, there are no studies that systematically analysed the shape of news waves and whether this shape changed over time, so we formulated the following exploratory research question:

RQ1: *What is the general shape of a news wave, and does this change from 1950-2014?*

A method for automatically identifying news waves

A news wave is defined as an abnormally large amount of attention for an issue in a relatively short span of time. In order to identify news waves automatically, we used automatic topic modelling and operationalized a news wave as a time span in which the issue is covered much more intensely than normal.

Data and preprocessing

All digitally available news articles published from 1950 until 2014 in *De Telegraaf* were downloaded and analysed. The articles from 1950 until 1994 were retrieved from the Dutch Royal Library (http://www.delpher.nl).The articles from 1999 to 2008 were retrieved from LexisNexis, and the final years were scraped directly from the e-edition, which is an online edition that is identical to the print edition. There is no digital archive available for the years 1995-1999, so these articles are missing. In total, this yields a corpus of almost four million news articles.

These articles were uploaded to the Amsterdam Content Analysis Toolkit (AmCAT; Van Atteveldt, 2008) and lemmatized and POS-tagged using the Frog parser (Van den Bosch et al., 2007), meaning that all words are reduced to their dictionary form and the part of speech (i.e. noun, verb) of each word is determined. Finally, all words except for nouns and proper names

were removed, as these are the parts of speech that are most indicative of the topic of a text.

Defining a news wave

For each year in the corpus, news waves were identified in four steps:

1) A vocabulary of potentially interesting words was identified by computing how often each word occurs in a six-day, rolling interval, and using these frequencies to compute the time entropy of each word using the RNewsFlow R package (Welbers & Van Atteveldt, 2016). From this, we selected the 30,000 words with the lowest entropy, that is, the words that are most typical of specific periods in the year.

2) An LDA topic model was made with the selected words, using 250 topics. This clusters words that frequently occur together, essentially allowing a wave to consist of multiple related terms (Jacobi et al., 2016). The high number of topics was chosen to make sure that news waves were about a specific event, not a more general issue like the economy (cf. Ibid.).

3) A six-day rolling average of the number of words in each topic per day was computed. Potential news waves were defined as consecutive periods in which the attention for a topic was more than three standard deviations higher than the mean attention for that topic.

4) Based on the qualitative validation of the 2014 data described below, a threshold was chosen for both long and intense news waves: A long news wave was defined as one in which news on an issue was published on at least seven consecutive days, of which on at least one day attention was four standard deviations higher than the mean. An intense news wave was defined as a wave containing at least one day during which attention was six standard deviations higher than the mean. Note that these criteria are not mutually exclusive, and some waves are both long and intense.

This yields a list of waves per year. For each wave, the length, amount of articles per day, and the total amount of articles was computed. Finally, the political content of news was determined by looking for mentions of Dutch political parties: an article that contained at least one mention of a party was considered political.

Validation and model estimation

Since there is no agreed definition of exactly what a news wave is, there is also no gold standard against which to validate the automatically identified waves. To provide a validation for the method, we compared the automatically identified waves to our earlier analysis of news waves in 2014 (Ruigrok & Van Atteveldt, 2015). In that study, news waves were identified using a different method and checked manually, so the comparison provides a triangulation that gives an indication of the validity of the automatic method. Moreover, this comparison was used to determine the threshold criteria used in step four, outlined above.

Overall, our qualitative conclusion is that the automatically identified waves are plausible and correspond, more or less, to the waves identified in our earlier study. Examples of intense waves found in this sample are the political affair around the Dutch Intelligence Service (AIVD) and the minister responsible (Plasterk), and the incident around the footballer Suarez, who bit an opponent at the World Championships. Examples of long waves are the Russian invasion of the Crimean region and the Dutch municipal elections. The most important wave that was not found automatically was about the right-wing politician Wilders, who made a statement at a rally that he wanted 'fewer, fewer, fewer Moroccans'. This was not found because the involved words are relatively common and there were no specific nouns or names that typified that wave. Automatically identified waves that were not seen as waves in our earlier study included the coverage of annual events such as Easter and Christmas, due to the very typical vocabulary and the specific time period. These differences notwithstanding, most of the waves did correspond between both methods, lending credence to the automatic method.

Results

Are news waves increasing over time?
Our first hypothesis stated that, between 1950 and 2014, the amount of news in waves increased linearly as media logic and especially commercialization increasingly dominated the journalistic profession (cf. Wijfjes, 2004; Strömbäck & Esser, 2014). Figure 2.2 shows the percentage of news that was within a news wave over time. To facilitate the testing of our hypotheses, we divided the data into three periods, based on Wijfjes (2004) and Brants and Van Praag (2006). The latter defines a period of partisan logic until 1970, a
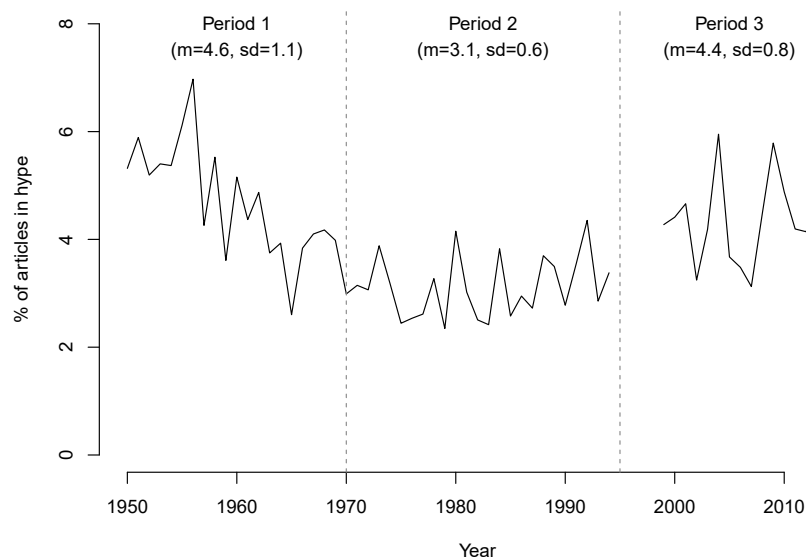
Figure 2.1.   Percentage of articles found in news waves over time

period of public logic from 1970-1990, and a period of media logic since 1990. To deal with the missing data from 1994-1999, and because technical and commercial changes to the journalistic profession occurred most strongly in the 2000s, we take the third period to start after 1999.

As can be seen from the picture, the importance of news waves did not increase linearly. Rather, it shows a U-curve, dropping from around 4.5% of news in the 1950s to around 3% in the 1980s, before climbing to almost 4.3% again in the 2000s. Note that the years 1994-1999 are missing since there are no digital archives for this period.

This development is also summarized in Table 2.1, which lists a number of characteristics of news waves per period. The period right after the Second World War is characterized by the longest news waves (8.3 days) and also contains the highest number of articles (86 per wave, on average). As a result, the percentage of all news that was contained in a news wave was highest in this period at 4.5%.

The second period, although it had most news waves per year, is characterized by smaller waves, containing only 63 articles each. As a result, only 3.1% of all news in this period was part of a news wave.

In the third period there are less news waves per year (33). The waves are shorter (7.6 days, on average) and contain fewer articles (62). However,

Period	Number of news waves per year	Average size of news waves (#articles)	Average length of news wave (days)	% of all articles in news waves
1950-1970	36	86	8.3	4.5%
1971-1994	42	63	8.1	3.1%
1999-2014	32	62	7.6	4.3%

Table 2.1.   Overview of news waves per period

because the total amount of published news articles also decreased, the proportion of news contained in news waves actually increased to 4.3%, almost as high as in the first period.

Chi-squared statistics were calculated to see if the percentage of articles in a wave is significantly different per period. Results show that all differences are significant: The percentage of news in waves in the first period (4.5%) is significantly higher than in the second (3.1%, $X_{(1)}$=4758.1, p<0.001) and third period (4.3%, $X_{(1)}$=697.6, p<0.001). Moreover, the third period is significantly higher than the second ($X_{(1)}$=1294, p<0.001), confirming the U-curve apparent from Figure 2.1. Since H1 stated that the proportion of news in waves increased linearly, the hypothesis is rejected: even though this proportion had increased since the 1970s, it climbed back to its level of the 1950s, rather than a linear increase.

Hard or soft news: Political articles in news waves

Our second hypothesis dealt with the content of waves rather than their frequency. In particular, we expected the nature of waves to shift away from political (hard) news over time under influence of commercialization (cf. Brants, 2015). Figure 2.3 shows two proportions: the solid line shows the percentage of political articles that is contained in a wave, while the dashed line shows what percentage of articles in a wave is political. For example, in 2000, around 5% of all political news was contained in a news wave. In all news waves combined, around 12% of news was political.

The dashed line in the figure shows that there is certainly not a linear decrease in the political content of news waves. On the contrary, news waves in the third period are most political, containing 9% political articles on average, significantly more than in the second (average 4.9%, $X_{(1)}$=543.2, p<.001) and first period (average 5.4%, $X_{(1)}$=415.1, p<.001). Although the
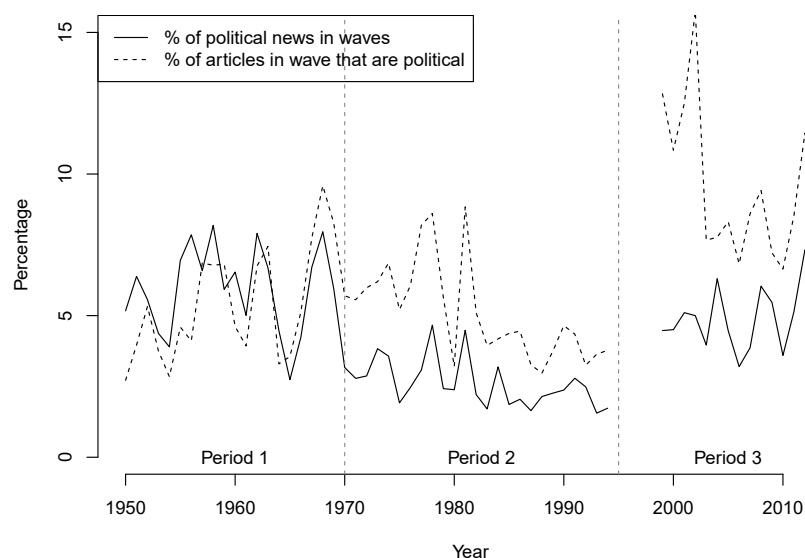
Figure 2.2.   Political content of news waves

difference between the first and second period is smaller, waves in the second period are significantly less political than in the first ($X(1)$=12.1, p<.001).

The solid line shows the reverse proportion, namely what percentage of all political news is contained in a news wave. Thus, rather than showing whether news waves are political, it shows to what extent political news is characterized by waves. This line is similar to the U-curve in Figure 2.1, with waves in the first period containing the largest proportion of political news (5.8%), significantly more than in the second (2.6%, $X(1)$=1155.8, p<.001) and in the third period (5.1%, $X(1)$=30.7, p<.001).

These results mean the second hypothesis must also be rejected: news waves in the last period are actually the most political, and the percentage of political news in waves decreases until the 1980s, but then actually increases again.

Dynamics of news waves

The research question of this study (RQ1) concerns the general shape of news waves. Earlier research suggests that the intensity of waves should increase explosively, followed by a longer decline. Moreover, it suggests that

waves are often characterized by multiple local maxima as the waves goes through different stages (cf. Vasterman, 2005; Ruigrok et al., 2009).

Figure 2.4 shows the shape of news waves per period by calculating what percentage of articles in each wave occurs on which day of the wave, centred around the day with the most news. As discussed in the method section, we included two types of waves: intensive waves and long waves. Since it can be expected that these have different shapes, Figure 2.4 shows both types of waves separately (waves that are both long and have a very high peak are counted in both lines). Moreover, the figure shows separate lines for the three periods discussed above.

From the figure, it is apparent that the general shape of waves is similar for all three periods, and is also mostly similar for intense and long waves: There is a very clear and narrow peak (day 0), which accounts for 20% to over 35% of news. Before the peak there is a small lead, but most days have less than 5% of total news. The peak is followed by an immediate and steep decline, with the day after the peak receiving less than 10% of news. After this, there is a slower decline for about a week, after which there is almost no news. Not surprisingly, intense waves have the highest peak, containing over 30% of articles on the day with the most coverage and declining steeply to less than 5% on day seven. Long waves have a lower peak with just above 20% of news on one day and decline a bit more gradually and
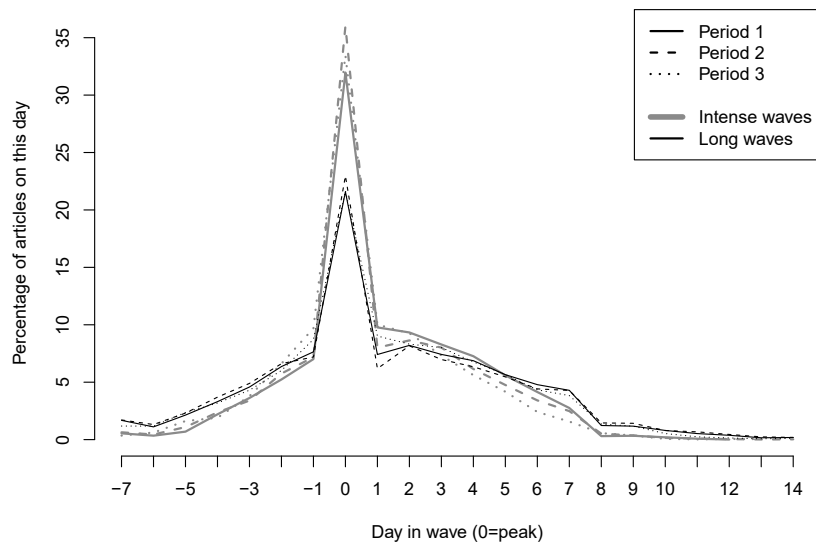

Figure 2.3.   The shape of news waves: Percentage of news articles per day in a steep news wave and a long news wave

also have some more news before the peak. Overall, however, the shape for long and intense waves are very similar and vary little between the three periods.

Conclusion/discussion

Contrary to our expectations, we did not find a linear increase of news waves over time. Rather, we found a U-curve, with the amount of news in waves decreasing from the 1950s until the 1970s and then increasing again in recent decades to almost the same level as the 1950s. It is possible that the relatively large role of news waves in the immediate post-war period is due to the specific situation of *De Telegraaf*. Whereas most newspapers in the Netherlands were closely linked to a specific political party, *De Telegraaf* was neutral and had always aimed for a large market share, rather than a specific ideological niche, focusing on topics such as crime and fraud and stories high on human interest, and introducing new genres such as the interview and the special report (Wolf, 2009).

In line with the general movement towards a more critical and socially responsible press (Wijfjes & Bardoel, 2015), we see that in the 1970s and 1980s the share of news in news waves decreased, reflecting a more professional perspective on journalism and possibly a move towards a trustee model, in comparison to the immediate post-war period. In recent decades, this trend has reversed, in line with the general movement of Dutch media towards a more audience-centred perspective (Wijfjes, 2004; Brants, 2015). This reflects a change from a trustee model to a market model (Schudson, 2003; Welbers et al., 2015), focusing on what the public is interested in under competitive pressure and the need to perform well commercially (cf. Johnston, 2009).

Our second hypothesis stated that, over time, news waves would contain less political news, as journalists focused more on scandals and soft news (cf. Bardoel & Wijfjes, 2015; Brants, 2015). This hypothesis was also rejected, with news waves in the most recent period, in fact, being most political. News waves actually became more political over time, but, overall, politics did not become more dominated by news waves. Similarly, political news was least dominated by news waves in the 1970s and 1980s, consistent with the notion of this period being the high watermark of responsible and professional journalism.

Our final results dealt with the overall shape of news waves. We were unable to find evidence for the idea of a news wave as 'a sharp and continuous

increase of reporting on a specific issue for a limited period of time' (cf. Geiß, 2011: 272). Rather, the average news wave turned out to be strongly centred on a single peak day, containing almost a third of total news on that day, and the decline after the peak was approximately as sharp as the increase before the peak and attention dropped almost to zero after a week. We also found no evidence of waves consisting of multiple peaks as the story evolved from facts to commentary (cf. Ruigrok et al., 2009; Vasterman, 2005), but it is possible that such waves would be split into multiple 'wavelets' by the topic modelling algorithm.

A strong limitation to the automatic method presented in this study is the lack of a 'gold standard' definition of what news waves are, making it impossible to formally validate our method for identifying them. Presumably, the main reason for this lack of a gold standard is that the theoretical definitions of news waves are not sufficiently concrete to be used to distinguish waves from non-waves in an objective manner. The lack of a formal validation also makes it difficult to judge how sensitive the method is to noise in the data, such as text digitization errors. Furthermore, it makes it difficult to compare and adapt the methods, e.g. comparing different settings for the number of topics or by filtering out sections such as sports, lifestyle, or financial reports, which could be said to be of a different kind than regular news. These problems are compounded for studying 'hypes', which have an added normative (and subjective) criterion of there being 'too much' news on a topic.

If we are to move to a more robust and quantitative programme of analysing news waves (or hypes), the first order of business should be to construct a gold standard consensus on which topics and periods can and cannot be characterized as waves. Creating this consensus would force us to adopt a sufficiently concrete conceptualization of news waves, and allow us to formally validate methods for identifying and analysing them. This study is an important step towards creating this gold standard by providing a list of potential 'real' waves, based on a publicly available newspaper archive.

Since pack journalism is an important aspect of news waves, a second limitation is that this study is based on only a single newspaper. Out of the three positive feedback loops that lead to news waves, we limited ourselves to the intra-media loop. Moreover, a more comparative analysis would show whether the patterns discussed above are specific to *De Telegraaf* or indicative of the evolution of Dutch journalism in general. The reason for concentrating on a single newspaper was simply that there were no other digital newspaper archives available for the period before the 1990s.

Fortunately, the Royal Library recently published the digitized archives for *Algemeen Handelsblad/NRC Handelsblad*, a daily newspaper generally seen as a quality or elite newspaper compared to the more market-oriented newspaper *De Telegraaf*. This provides an opportunity to expand this study in future research and show whether patterns and news waves are shared between media outlets, as assumed in Vasterman's (2005) definition of of news hypes as 'wall-to-wall' media coverage of an issue, and whether the convergence of outlets has changed over time. Finally, taking the intermedia and audience/source loops into account as well can give a better understanding of the internal dynamics of news waves, although these dynamics might be difficult to trace in daily publication data given that news waves are strongly concentrated on the peak day of the hype.

A final limitation of this study is that we only looked at the frequency and length of news waves, and did not consider the content of the wave or key event, apart from whether it contained politics. In future research, it would be interesting to analyse the content of the identified waves, for example by looking at the general topic and whether a wave was about important events, such as the Prague Spring or fall of the Berlin Wall, or rather about private scandals of public figures.

In sum, this chapter paves the way for a programme of quantitative analysis of the changing role and nature of news waves in Dutch journalism. We provided an automatic method for identifying news waves and applied this to a digital archive of newspaper material spanning almost the entire post-war period. This showed that the idea of a linear increase of news waves is overly simplistic, and provides a starting point for more research into the nature and dynamics of news waves and their role in journalism.

References

Aelst, Peter van, and Rens Vliegenthart (2014). Studying the tango: An analysis of parliamentary questions and press coverage in the Netherlands. *Journalism Studies*, 15(4): 392-410.

Altheide, David L. (2004). 'Media logic, social control, and fear'. *Communication Theory*, 23 (3): 223-238.

Anderson, C.W. (2011). 'Between creative and quantified audiences: Web metrics and changing patterns of newswork in local US newsrooms'. *Journalism* 12 (5): 550-566.

Asp, Kent (2014). 'News media logic in a new institutional perspective'. *Journalism Studies*, 15 (3): 256-270.

Atteveldt, van, Wouter (2008). *Semantic network analysis: Techniques for extracting, representing, and querying media content.* Charleston SC: BookSurge Publishers.

Bardoel, Jo & Huub Wijfjes (2015). 'Journalistieke cultuur in Nederland: Een professie tussen traditie en toekomst'. In: Jo Bardoel & Huub Wijfjes (eds), *Journalistieke Cultuur in Nederland.* Amsterdam: Amsterdam University Press.

Boczkowski, Pablo (2010). *News at work: Imitation in an age of information abundance.* Chicago, IL: University of Chicago Press.

Boydstun, Amber E., Hardy, Anne & Stefaan Walgrave (2014). 'Two faces of media attention: Media storm versus non-storm coverage'. *Political Communication*, 31 (4): 509-531.

Brants, Kees (2015). 'Van medialogica naar publiekslogica?' In: Jo Bardoel & Huub Wijfjes (eds), *Journalistieke Cultuur in Nederland*, 237-255. Amsterdam: Amsterdam University Press.

Brants, Kees & Philip van Praag (2006). 'Signs of media logic half a century of political communication in the Netherlands'. *Javnost-the public*, 13 (1): 25-40.

Brants, Kees & Philip van Praag (2015). 'Beyond media logic'. *Journalism Studies 18(4): 395-400.*

Broersma, Marcel (1999). 'Botsende stijlen. De Eerste Wereldoorlog en de Nederlandse journalistieke cultuur'. *Tijdschrift voor Mediageschiedenis* 2 (2): 40-68.

Broersma, Marcel (2007). 'Form, style and journalistic strategies. An introduction'. In: Idem (ed.). *Form and style in journalism. European newspapers and the representation of news, 1880-2005*, ix-xxix. Leuven: Peeters.

Cook, Timothy E. (2006). 'The news media as a political institution: Looking backward and looking forward'. *Political Communication*, 23 (2): 159-171.

Djerf-Pierre, Monika (2012). 'When attention drives attention: Issue dynamics in environmental news reporting over five decades'. *European Journal of Communication*, 27 (3): 291-304.

Donsbach, Wolfgang (2004). 'Psychology of news decisions factors behind journalists' professional behavior'. *Journalism* 5 (2): 131-157.

Downs, Anthony (1972). 'Up and down with ecology-the issue-attention cycle'. *The public interest*, (28): 38-50.

Eijk van der, Cees (2000). The Netherlands: Media and politics between segmented pluralism and market forces. *Democracy and the media: a comparative perspective*, 303-342.

Fishman, Mark (1978). 'Crime waves as ideology'. *Social Problems* 25 (5): 531-543.

Galtung, Johan & Mari Ruge (1965). 'The structure of foreign news: The presentation of the Congo, Cuba and Cyprus crises in four Norwegian newspapers'. *Journal of International Peace Research* 2 (1): 64-90.

Gans, Herbert J. (1979). *Deciding what's news: A study of CBS evening news, NBC nightly news, Newsweek, and Time.* Evanston, IL: Northwestern University Press.

Geiß, Stefan (2011). 'Patterns of relationships between issues: An analysis of German prestige newspapers'. *International Journal of Public Opinion Research* 23 (3): 265-286.

Hallin, Daniel C. & Paolo Mancini (2004). *Comparing media systems: Three models of media and politics.* Cambridge: Cambridge University Press.

Hamilton, James T. (2004). *All the news that's fit to sell: How the market transforms information into news.* Princeton, NJ: Princeton University Press.

Harcourt, Alison (2005). *The European Union and the regulation of media markets, European Policy Research Series.* Manchester: Manchester University Press.

Harcup, Tony & Deirdre O'Neill (2001). 'What is news? Galtung and Ruge revisited'. *Journalism Studies* 2 (2): 261-280.

Harcup, Tony & Deirdre O'Neill (2016). 'What is news? News values revisited (again)'. *Journalism Studies 18(12): 1470-1488*

Jacobi, Carina, Van Atteveld, Wouter & Kasper Welbers (2016). 'Quantitative analysis of large amounts of journalistic texts using topic modelling'. *Digital Journalism*, 4 (1): 89-106.

Johnston, Jane & Susan Forde (2009). "Not wrong for long': The role and penetration of news wire agencies in the 24/7 landscape'. *Global Media Journal-Australian Edition*, 3 (2): 1-16.

Korthagen, Iris (2011). 'Media aandacht maakt het verschil: Invloeden van mediatisering op het beleid in Gouda'. *Bestuurswetenschappen*, 65 (5): 14-31.

Lund, Anker Brink (2002). 'News ecology and media society: Formative views on the production of news'. In: Madeleine Hurd, Tom Olsson, Patrik Åker (eds), *Storylines: Media power and identity in modern Europe*, 83-87. Stockholm: Hjalmarson and Högberg Bokförlag.

McCombs, Maxwell E. (2005). 'A look at agenda-setting: Past, present and future'. *Journalism studies*, 6 (4): 543-557.

McCombs, Maxwell E. & Donald L. Shaw (1993). 'The evolution of agenda-setting research: Twenty-five years in the marketplace of ideas'. *Journal of communication*, 43 (2): 58-67.

McManus, John H. (1995). 'A market-based model of news production'. *Communication Theory*, 5 (4): 301-338.

McManus, John H. (2009). 'The commercialization of news'. In: Karin Wahl-Jorgensen & Thomas Hanitzsch (eds), *The handbook of journalism studies*, 218-235. New York: Routledge.

Niblock, Sarah & David Machin (2007). 'News values for consumer groups: The case of Independent Radio News, London, UK'. *Journalism*, 8 (2): 184-204.

Paimre, Marianne & Harro-Loit Halliki (2011). 'Media generated news waves-catalysts for discursive change: The case study on drug issues in Estonian print media'. *Journalism* 12 (4): 433-448.

Redden, Joanna & Tamara Witschge (2010). 'A new news order? Online news content examined'. In: Natalie Fenton, *New media, old news*, 171-186. London: Sage.

Ruigrok, Nel & Wouter van Atteveldt (2015). '2014 in nieuwsgolven (2014 in news waves)'. Research commissioned by The Dutch Journalism Fund. Available at: https://www.svdj.nl/de-stand-van-de-nieuwsmedia/2014-in-nieuwsgolve/.

Ruigrok, Nel, Scholten, Otto, Krijt, Martijn & Joep Schaper (2009). 'Fitna in de media: Een brongerichte media hype'. *Tijdschrift voor Communicatiewetenschap*, 37 (3): 238-253.

Ruigrok, Nel, Van Atteveldt, Wouter, Gagestein, Sarah & Carina Jacobi (2016). 'Media and juvenile delinquency: A study into the relationship between journalists, politics, and public'. *Journalism* 18(7): 907-925.

Schudson, Michael (2003). *The sociology of news*. New York: W. W. Norton & Company.

Strömbäck, Jesper (2005). 'In search of a standard: Four models of democracy and their normative implications for journalism'. *Journalism Studies*, 6 (3): 331-345.

Strömbäck, Jesper (2008). 'Four phases of mediatization: An analysis of the mediatization of politics'. *The International Journal of Press/Politics*, 13 (3): 228-246.

Strömbäck, Jesper & Frank Esser (2014). *Mediatization of politics*. London: Palgrave.

Strömbäck, Jesper, Karlsson, Michael & David Nicholas Hopmann (2012). 'Determinants of news content'. *Journalism Studies* 13 (5-6): 718-728.

Takens, Janet (2013). *Media logic and electoral democracy*. Amsterdam: VU Uitgeverij.

Takens, Janet, Van Atteveldt, Wouter, Van Hoof, Anita & Jan Kleinnijenhuis (2013). 'Media logic in election campaign coverage'. *European Journal of Communication*, 28(3): 277-293.

Tuchman, Gaye (1978). *Making news: A study in the construction of reality*. New York: Free Press.

Van Aelst, Peter & Rens Vliegenthart (2014). 'Studying the tango: An analysis of parliamentary questions and press coverage in the Netherlands'. *Journalism Studies*, 15(4): 392-410.

Van Atteveldt, Wouter (2008). *Semantic network analysis: Techniques for extracting, representing, and querying media content*. Charleston, SC: BookSurge Publishers.

Van den Bosch, Antal, Busser, Bertjan, Canisius, Sander & Walter Daelemans (2007). 'An efficient memory-based morphosyntactic tagger and parser for Dutch'. In: Frank Van Eynde, Liesbeth Augustinus, Ineke Schuurman, Vincent Vandeghinste (eds), *Selected papers of the 17th computational linguistics in the Netherlands meeting, Leuven, Belgium,* 99-114. LOT Occasional Series 7: 191-206.

Van der Eijk, Cees (2000). 'The Netherlands: Media and politics between segmented pluralism and market forces'. In: Gunter, R., Mughan, A. (Eds.), Democracy and the Media: A Comparative Perspective. Cambridge University Press, Cambridge.

Vasterman, Peter (2005). 'Media-hype. Self-reinforcing news waves, journalistic standards and the construction of social problems'. *European Journal of Communication*, 20 (4): 508-530.

Welbers, Kasper & Wouter van Atteveldt (2016). 'Package 'RNewsflow' tools for analyzing content homogeneity and news diffusion using computational text analysis'. Available at: https://cran.r-project.org/web/packages/RNewsflow/ RNewsflow.pdf.

Welbers, Kasper, Van Atteveldt, Wouter, van Atteveldt, Jan Kleinnijenhuis, Jan, Ruigrok, Nel & Joep Schaper (2015). 'News selection criteria in the digital age: Professional norms versus online audience metrics'. *Journalism*, 17 (8): 1037-1053.

Wien, Charlotte & Christian Elmelund-Præstekær (2009). 'An anatomy of media hypes: Developing a model for the dynamics and structure of intense media coverage of single issues'. *European Journal of Communication*, 24 (2): 183-201.

Wijfjes, Huub (2004). *Journalistiek in Nederland. Beroep, cultuur en organisatie 1850-2000*. Amsterdam: Boom.

Wolf, Mariëtte H. B. B. (2009). *Het geheim van De Telegraaf: Geschiedenis van een krant.* Amsterdam: Boom.

Wolfsfeld, Gadi (2014). *Making sense of media and politics five principles in political communication.* New York: Taylor and Francis.

Wolfsfeld, Gadi & Tamir Sheafer (2006). 'Competing actors and the construction of political news: The contest over waves in Israel'. *Political Communication,* 23 (3): 333-354.

Zaller, John (2003). 'A new standard of news quality: Burglar alarms for the monitorial citizen'. *Political Communication*, 20 (2): 109-130.

Zhu, Jian-Hua (1992). 'Issue competition and attention distraction: A zero-sum theory of agenda-setting'. *Journalism & Mass Communication Quarterly*, 69 (4): 825-836.

About the authors

Wouter van Atteveldt is Associate Professor at the Department of Communication Science, Vrije Universiteit Amsterdam (http://vanatteveldt. com). His research focuses on political communication, especially the antecedents and consequences of mass media coverage of political discourse, and has a strong methodological focus on developing AI/Computational NLP techniques to improve automatic text analysis. He has developed the Amsterdam Content Analysis Toolkit (AmCAT), an open source infrastructure that makes it easy to do large scale automatic and manual text analysis for the social sciences. Van Atteveldt has given numerous workshops in R

and his research has been widely published in communication sciences journals.

Carina Jacobi holds an MSc in Sociology and a PhD in Communication from the University of Amsterdam. Her dissertation explored the questions of how to define and measure news quality in the current digitalized media environment. Her research interests include popular news, longitudinal changes in news content and journalism, and (digital) research methods. She currently works as an information analyst for the Dutch Inspection for Environment and Transport.

Nel Ruigrok (PhD) is owner of LJS Media Research, a company that focuses on scientific research for organizations such as government Ministries and NGOs in the Netherlands. She also provides workshops and training session for communication professionals. Academically she is affiliated to the Erasmus Research Centre for Media, Communication and Culture, Erasmus University, Rotterdam where she teaches a number of courses at the Master Media and Journalism.

Kasper Welbers holds a PhD in Communication Science from the Vrije Universiteit Amsterdam, where he defended his dissertation on Gatekeeping in the Digital Age in 2016. His research focuses on changes in the news diffusion process due to the proliferation of new media technologies. In particular, he uses computational text analysis to trace how news messages diffuse trough networks of old and new media channels. He is currently employed as a postdoctoral researcher at the University of Leuven, where he studies how professional news organizations use social media channels.