



Live Text Coverage of Political Events: Combining Content and Corpus-based Discourse Analysis

Hendrik Michael, Valentin Werner

University of Bamberg, Germany

E-mail: hendrik.michael@uni-bamberg.de, valentin.werner@uni-bamberg.de

Abstract

Live text (LT) has emerged as a web-native CMC alternative to traditional forms of live broadcasting. Through a combined content and corpus-based discourse analysis of the LT coverage of a major political event (the 2020 US presidential debates), the present study tests (i) how current LTs emphasize transparency and accountability, and (ii) how they are a form of journalistic communication that normalizes professional norms of objectivity in hybrid media settings. Political LT emerges as multi-layered and multi-authored discourse that places strong emphasis on accountability and disclosure transparency by updating and linking information, while maintaining the journalistic gatekeeping/gatewatching function. Linguistically, it is characterized by an informal tone but also by a continuation of traditional news media practices as regards objectivity, as instantiated by the salience of debate topics and political terms and – unlike the more widely studied sports LT – by a clear delineation of information from opinion and contextualization.

Keywords: online news, live text, live blogging, hybridity, political journalism, digital journalism, media studies

1. Introduction: (Digital) Journalism and Politics

Journalism is the main source of political information in deliberative democracies (McNair, 2009; Schudson, 2008) and it functions as gatekeeper and agenda-setter (Shoemaker et al., 2009), specifically during election periods (Schulz, 2011). In this context, television became the main channel through which the public receives political information (e.g. Farnsworth & Lichter, 2010; Gottfried et al., 2016). Previous analysis has shown that broadcast journalism is particularly important for voters as it puts campaigns into the public spotlight. A prime example for this are the presidential debates, a cornerstone of every campaign season (Kraus, 2020). However, the advent of hybrid media systems (Chadwick, 2017) during the last decade(s) has markedly altered how journalists cover politics and how audiences consume relevant content. An increasing number of recipients, and particularly younger segments of the audience, use online news and social media as their primary sources of political information now (Mitchell et al., 2020).

Political journalism today has defining characteristics such as multimodality, connectivity and hypertextuality (Nuernbergk & Neuberger, 2019), which foster media convergence and hybrid journalistic practices. Multimodality and connectivity allow cross-media production by which news content is effectively played out on different platforms and establishes connections with different target audiences. Connectivity means that online journalism follows and implements what other media are publishing, making news media producers gatewatchers that scan the flow of information in digital publics and provide context by linking and fact-checking (Bruns, 2018). By linking and updating (inaccurate) information online news media can create transparency and accountability in ways that TV and print media lack. Despite the evolution of these hybrid media practices, online journalism is also subject to a professional normalization. Studies by Singer (2005) and Lasorsa et al. (2012) have shown that long-established roles and practices of professional news media are still in operation. For instance, they shape to a significant degree how journalists share information by relying on media of record

or their host media when providing links. In addition, journalists have been found to mostly adhere to standards of objective reporting by using impersonal and informative language.

A paradigmatic example of hybrid media practices, illustrating how political information is mediated and consumed, is live text (LT), alternatively labeled “live blogging”. LT has emerged as a popular web-native alternative to TV and radio broadcasting for live reporting on pre-scheduled events with a limited duration (Thurman & Walters, 2013). This form of CMC is characterized by modularity (various content and navigational zones), multimodality (embedded multimedia content) and its dynamic nature (see Figure 1 for an example).

Usually, LT is multi-authored and reports on events in reverse chronological order. As the discourse emerges as the events reported upon unfold, it has also been termed a “text-in-motion” (Hauser, 2008) “text-in-process” (Chovanec, 2018) or “open news discourse” (Thorsen & Jackson, 2018), delineating it from traditional news practices.

To explore whether and in which forms the above-mentioned trends in hybrid media systems are empirically traceable in current journalism, the present study focuses on LT in political reporting (cf. Tereszkievicz, 2014; Van Driel, 2020) and aims to illustrate current journalism practice in a CMC format.

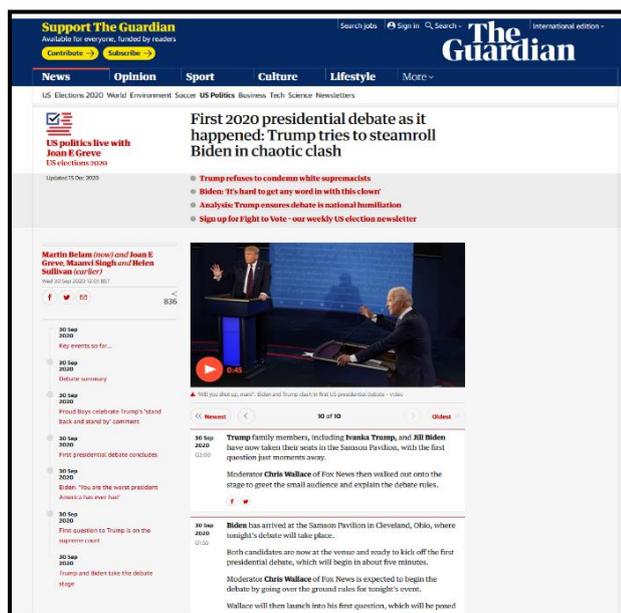


Figure 1: LT from *The Guardian*¹

2. Data, Method, and Research Hypotheses

This study focuses on a specific form of CMC, namely political LT produced by professional journalists. Thus, the analysis relies on a purpose-built corpus of LT coverage of two televised US presidential debates (Donald Trump vs. Joe Biden), which both lasted for approximately 90 minutes and were screened on 29 September (73 million viewers) and 22 October 2020 (63 million viewers), respectively. Data were collected from four popular media outlets (*The Guardian*, *Daily Mirror*, *New York Times*, *Wall Street Journal*) to represent both British and US-American perspectives, to include news organizations associated with the tabloid and the quality press, and to represent media that have liberal and conservative editorial stances. The overall corpus size amounts to 61,490 tokens (excluding audience contributions through Tweets, etc.). By comparison, the LT corpus compiled is small, but was designed for a close analysis of current media practice, taking into account the multimodal nature of LT as an understudied CMC manifestation. The LT contained in the corpus was created either by a single reporter (*Daily Mirror*) or by teams of up to 16 different journalists (*New York Times*). The difference in contributors is also reflected in the number of individual posts per reporting (ranging from $n = 35$ to $n = 319$), while the length of individual posts varied between 33 and 199 words. Generally speaking, a higher post frequency and a higher number of commentators was observed to correlate with shorter individual posts. The aforementioned findings are clearly indicative of varying practices in different media outlets. To facilitate a discourse-oriented mixed-methods approach (O’Keeffe, 2006; Bednarek & Carr, 2020), data was stored both in its original HTML format to allow the consideration of multimodal aspects and manual annotation in *MAXQDA* (MAXQDA, 2019), as well as in

¹<https://www.theguardian.com/us-news/live/2020/sep/29/presidential-debate-latest-news-tonight->

TXT format to allow quantitative analysis (wordlists, n -grams) in the corpus software *AntConc* (Anthony, 2020). Manual data annotation was conducted by a research assistant and double-checked by the co-authors. With the help of the aforementioned data and given the general characteristics provided in Section 1, the present study tackles the following broader hypotheses:

- H₁: Presidential debate LTs emphasize transparency and accountability by updating and linking information, while maintaining the journalistic gatekeeping/gatewatching function.
- H₂: Presidential debate LTs are a form of journalistic communication that normalizes professional norms of objectivity in hybrid media settings.

For the operationalization of H₁, we conduct a content analysis and, adapting categories established in media studies (e.g. Singer, 2005; Bruns, 2018), we annotate (i) sources of information used, such as political agents, media of record, ordinary people, etc., as well as (ii) indications of transparency and accountability in terms of updates, links, and quotes. For H₂, we test how criteria for journalistic objectivity, such as the lack of personal stance, fair representation, fair skepticism, identification of better arguments, and sticking to hard facts, etc., originally postulated in a classic study by Donsbach & Klett (1993), are reflected linguistically. This may materialize, for instance, in terms of salient content words, (non-)usage of personal pronouns, mental verbs, and reporting verbs.

3. Results

3.1. Content Analysis

The data from the content analysis suggest a clear pattern, namely that most information was sourced from media of record, host media, and political agents (Figure 2). Note that numbers for the last category would be much higher (>10) if statements by the presidential candidates, which naturally are very salient in the discourse (see also Section 3.2.1), were included as sources of information as well. The difference in sources from media of record between the first and the second debate can be explained by the fact that shortly before the first debate, the *New York Times* published an investigative exposé on Trump’s tax record that was referred to repeatedly. Notably, only few scientific or public service organizations were referenced as sources of information, and neither are ordinary people often given a voice.

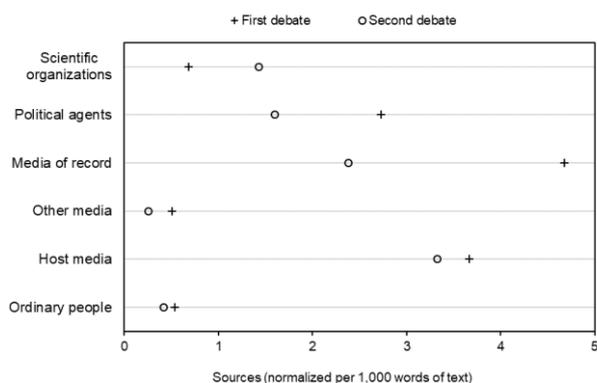


Figure 2: Source of information in political LT

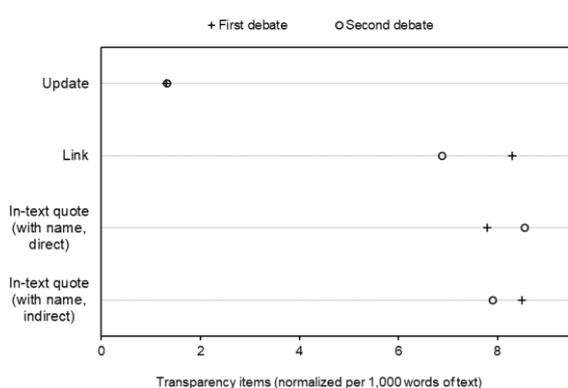


Figure 3: Transparency items in political LT

With respect to linking and updating, the data yield a mixed picture (Figure 3). Updates are relatively scarce, given that for example *The Wall Street Journal* and the *Daily Mirror* do not indicate updates at all, while this practice of creating transparency is only institutionalized by *The Guardian*. By contrast, linking emerges as a routine that is broadly used as an efficient strategy to create transparency and accountability. Similar rates of usage are observable for using named direct and indirect quotes to foster accountability. Unnamed quotes (not shown in Figure 3) are rare and are exclusively used in the *New York Times* LT.

3.2. Corpus Analysis

3.2.1. Content Words

Among the highly frequent (top 100) content words, the following categories emerge:

1. Names of participants: *Trump, Biden, Joe, Wallace, moderator*
2. Political discourse: *(First/presidential) debate, president, election, people, campaign, Americans, voters, candidates*
3. Debate topics: *Coronavirus, pandemic, (Supreme) court, health, states, US, China, climate, the Affordable Care Act*
4. Liveness: *Now, here, before, tonight*

5. Accountability: *Fact (check)*
6. Proclamatory character: *Say*
7. Personal stance: *(I) think*

Categories 1–3 demonstrate the fact-based/event-based nature of the LT coverage, while 4 is illustrative of the dynamic nature of the medium (see Section 1). 5 directly relates to the issue of accountability (see Section 2), while 6 and 7 are the most frequent items to report facts and personal stance, respectively, which merit a closer look (see further Section 3.2.2). In the corpus, *I think* is typically used to delineate opinions and value judgements by the reporters from information, as in *Adam, I think you've got it right here (New York Times)*, simultaneously supporting the construction of an expert identity as a typical journalistic practice (Ekström & Lundell, 2011). In addition, *I think* serves the purpose to represent the vox pop, a strategy conducive to the objectivity principle of fair representation.

3.2.2. N-grams

Among the most frequent 3/4-grams, we find *Mr Trump/Mr Biden said*, often preceded (and/or followed) by a direct quote as in *Mr Trump said he wants "crystal clean water and air"*, which could be interpreted as the mere reporting of hard facts through a proclamatory verb.

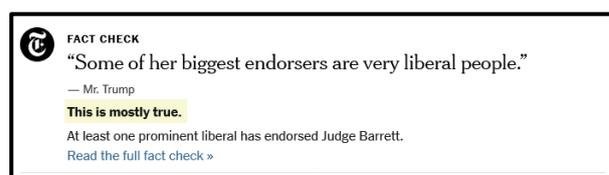


Figure 4: "Fact check" post with link from the *New York Times*²

In addition, *Mr Trump this is* (preceded by quote, followed by evaluative adjective) is highly frequent, as in *"I'm going to release them [his tax return forms] as soon as we can. I wan't to do it" - Mr Trump. This is misleading....* This recurrent pattern is indicative of the fact checking/contextualization function that fosters objectivity (see Figure 4 for another example as appearing in the LT).

3.2.3. Reporting Verbs

A dedicated look at reporting verbs (following the list from Garretson & Ädel, 2008) reveals that the most common neutral reporting verb lemma *say* is used with reference to either candidate in the same manner (see Table 1). However, among other patterns, it also emerges (i) that *accuse*, an item usually employed by challengers rather than incumbents, surprisingly is used in relation to Trump, while *criticize* is associated with Biden and (ii) that the disfavoring item *claim* mostly collocates with statements by Trump.

²<https://www.nytimes.com/interactive/2020/09/29/us/politics/debate-live-stream-analysis.html>

LEMMA	Trump	Biden	LEMMA	Trump	Biden
SAY	164	166	CLAIM	59	2
TELL	16	17	PROMISE	5	1
ASK	5	3	ADD	11	14
CALL	8	17	CITE	4	1
TALK	20	12	ANSWER	3	3
ARGUE	9	8	RESPOND	12	25
ACCUSE	15	7	MENTION	4	2
SUGGEST	10	2	DEFEND	10	3
CRITICIZE	5	12	INSIST	8	3
SPEAK	4	7	ACKNOWLEDGE	5	1
NOTE	2	7	REPEAT	10	2

Table 1: Reporting verbs (lemmas)

4. Discussion and Conclusion

Overall, LT can be viewed as a clear instantiation of CMC that exemplifies media convergence (or hybridity), in which formerly separate communicative practices are consciously combined to make their points (Bateman et al., 2017). Its multi-layered and multi-authored discourse is illustrative of dynamic news creation through discourse (Karlsson, 2011), which is facilitated in a CMC environment. Arguably, this process orientation, which involves constant updating, creates accountability. Further support for H₁ was evident. Through gatowatching, LT relies on a mix of sources, including opinion statements and evaluation (Bruns, 2018), and thus creates “disclosure transparency” (Pantic et al., 2017). At the same time, it emerges that elite sources dominate and that journalists apparently continue to fulfill their gatekeeping or “filtering” function, which reduces the participatory transparency, very much despite the affordances of LT as an interactive medium.

While previous studies (Tereskiewicz, 2014) have diagnosed an overall informal tone in political LT, as regards objectivity and its linguistic representation, political LT appears to continue traditional news media practices concerning objectivity principles (cf. Donsbach & Klett, 1993), as stated in H₂. This is traceable linguistically, for instance pertaining to the salience of debate topics and political terms, as well as in terms of the delineation of information (*X said...*) vs. opinion (*I think...*) vs. contextualization (*This is exaggerated...*). However, also some bias in reporting verbs, occasionally boosted by evaluative items (*weirdly claims...*), could be observed in the data.

The latter finding notwithstanding, the present data are supportive of a general trend toward normalization of online news discourse (Singer, 2005; Lasorsa et al., 2012) and imply that matter-of-fact-style political LT should be clearly differentiated from other LT types, such as sports LT, where the merging of reporting, commenting and glossing and a dedicated infotainment function is much more explicit (cf. Werner, 2016, 2019).

Overall, by way of a case study, the present analysis was intended (i) to provide an insight into current mediated journalistic language (Van Hout & Burger, 2017), and (ii) to highlight the potential of a combined media-linguistic approach toward political reporting through a from of CMC. Therefore, it may also inform the wider discussion in neighboring fields, such as communication and (digital) journalism studies.

5. References

- Anthony, L. (2020). *AntConc* (3.5.9). Waseda: Waseda University.
- Bateman, J., Wildfeuer, J. and Hiippala, T. (2017). *Multimodality: Foundations, research and analysis*. Berlin: Mouton de Gruyter.
- Bednarek, M., Carr, G. (2020). Computer-assisted digital text analysis for journalism and communications research: Introducing corpus linguistic techniques that do not require programming. *Media International Australia*. doi.org/10.1177/1329878X20947124
- Bruns, A. (2018). *Gatowatching and news curation: Journalism, social media, and the public sphere*. Frankfurt: Lang.
- Chadwick, A. (2017). *The hybrid media system: Politics and power*. Oxford: Oxford University Press.
- Chovanec, J. (2018). Participating with media: Exploring online media activity. In C. Cotter & D. Perrin (Eds.), *The Routledge handbook of language and media*. London: Routledge, pp. 505-522.
- Donsbach, W., Klett, B. (1993). Subjective objectivity: How journalists in four countries define a key term of their profession. *Gazette*, 51, pp. 53-83.
- Ekström, M., Lundell, Å.K. (2011). The joint construction of a journalistic expert identity in studio interactions between journalists on TV news. *Text & Talk*, 31(6), pp. 661-681.
- Farnsworth, S.J., Lichter, R.S. (2010). *The nightly news nightmare: Media coverage of U.S. presidential elections, 1988–2008*. Lanham: Rowman & Littlefield.
- Garretson, G., Ädel, A. (2008). Who’s speaking? Evidentiality in US newspapers during the 2004 presidential campaign. In A. Ädel & R. Reppen (Eds.), *Corpora and discourse: The challenges of different settings*. Amsterdam: Benjamins, pp. 157-187.
- Gottfried, J., Barthel, M., Shearer, E. and Mitchell, A. (2016). The 2016 presidential campaign - a news event that’s hard to miss. *Pew Research Center – Journalism & Media*. www.journalism.org/2016/02/04/the-2016-presidential-campaign-a-news-event-thats-hard-to-miss/
- Hauser, S. (2008). Live-Ticker: Ein Neues Medienangebot zwischen medienspezifischen Innovationen und stilistischem Trägheitsprinzip [Live ticker: A new media type between media-specific innovations and stylistic lag]. *kommunikation@gesellschaft*, 9(1), pp. 1-10.
- Karlsson, M. (2011). The immediacy of online news, the visibility of journalistic processes and a restructuring of journalistic authority. *Journalism*, 12(3), pp. 279-295.
- Kraus, S. (2020). *Televised presidential debates and public policy*. London: Routledge.
- Lasorsa, D.L., Lewis, S.C. and Holton, A.E. (2012). Normalizing Twitter: Journalism practice in an emerging communication space. *Journalism Studies*, 13(1), pp. 19-36.
- MAXQDA. (2019). *MAXQDA*. Berlin: VERBI Software.
- McNair, B. (2009). Journalism and democracy. In K. Wahl-Jorgenson & T. Hanitzsch (Eds.), *The handbook of journalism studies*. London: Routledge, pp. 237-249.
- Mitchell, A., Jurkowitz, M., Baxter Oliphant, J. and Shearer, E. (2020). Americans who mainly get their news on social media are less engaged, less knowledgeable. *Pew Research Center – Journalism & Media*. www.journalism.org/2020/07/30/americans-

- who-mainly-get-their-news-on-social-media-are-less-engaged-less-knowledgeable/
- Nuernbergk, C., Neuberger, C. (2018). *Journalismus online: Profession – Partizipation – Technisierung* [Journalism online: Profession – participation – technologization]. Heidelberg: Springer VS.
- O’Keeffe, A. (2006). *Investigating media discourse*. London: Routledge.
- Pantic, M., Whiteside, E. and Cvetkovic, I. (2017): Politics, conflict generate more live-blog comments. *Newspaper Research Journal*, 38(3), pp. 354-365.
- Shoemaker, P., Vos, T. and Reese, S. (2009). Journalists as gatekeepers. In K. Wahl-Jorgenson & T. Hanitzsch (Eds.), *The handbook of journalism studies*. London: Routledge, pp. 73-87.
- Schudson, M. (2008). *Why democracies need an unlovable press*. Malden: Wiley.
- Schulz, W. (2011). *Politische Kommunikation: Theoretische Ansätze und Ergebnisse empirischer Forschung* [Political communication: Theoretical approaches and results of empirical studies]. Wiesbaden: Springer VS.
- Singer, J.B. (2005). The political j-blogger: ‘Normalizing’ a new media form to fit old norms and practices. *Journalism*, 6(2), pp. 173-198.
- Tereskiewicz, A. (2014). “I’m not sure what that means yet, but we’ll soon find out”: The discourse of newspaper live blogs. *Studia Linguistica Universitatis Jagellonicae Cracoviensis*, 131, pp. 299-319.
- Thorsen, E., Jackson, D. (2018). Seven characteristics defining online news formats: Towards a typology of online news and live blogs. *Digital Journalism*, 6(7), pp. 847-868.
- Thurman, N., Walters, A. (2013). Live blogging: Digital journalism’s pivotal platform. *Digital Journalism*, 1(1), pp. 82-101.
- Van Driel, M. (2020). “It makes it more real”: A comparative analysis of Twitter use in live blogs and quotations in older news media from a reader response perspective. In C. Tagg & M. Evans (Eds.), *Message and medium*. Berlin: Mouton de Gruyter, pp. 306-328.
- Van Hout, T., Burger, P. (2017). Mediatization and language of journalism. In O. Gar N. Flores & M. Spotti (Eds.), *The Oxford handbook of language and society*. Oxford: Oxford University Press. doi.org/10.1093/oxfordhb/9780190212896.013.9
- Werner, V. (2016). Real-time online text commentaries: A cross-cultural perspective. In C. Schubert & C. Sanchez-Stockhammer (Eds.), *Variational text linguistics: Revisiting register in English*. Berlin: Mouton de Gruyter, pp. 271-305.
- Werner, V. (2019). A multimodal analysis of football live text commentaries. In M. Callies & M. Levin (Eds.), *Corpus approaches to the language of sports: Texts, media, modalities*. London: Bloomsbury, pp. 201-240.