



Universiteit
Leiden
The Netherlands

Multimodal Literacies and Digital Longform Journalism: Affordances and Meaning-Making

Heddam, C.S.

Citation

Heddam, C. S. (2023). Multimodal Literacies and Digital Longform Journalism: Affordances and Meaning-Making. *Txt, 10*, 49-62. Retrieved from <https://hdl.handle.net/1887/4290873>

Version: Publisher's Version

License: [Creative Commons CC BY 4.0 license](https://creativecommons.org/licenses/by/4.0/)

Downloaded from: <https://hdl.handle.net/1887/4290873>

Note: To cite this publication please use the final published version (if applicable).

Christine Stein Hededam

Multimodal Literacies and Digital Longform Journalism

Affordances and Meaning-Making



Christine Stein Hededam is a student and soon-to-be graduate of Book and Digital Media Studies at Leiden University. Originally from Denmark, she has spent the past four years pursuing her academic goals and putting down roots in the Netherlands. In the summer of 2023, she accepted a position as Associate Editor for Brill's Education list.

DIGITAL MULTIMODAL LONGFORM journalism builds on a long analogue tradition of immersive deep-dive stories. Various scholars have noted the genre's similarity to the New Journalism movement of the 1960s and 70s, which broke with the journalistic conventions of the time.¹ Characteristics of New Journalism included literary devices previously associated with fiction writing and the expression of subjective or even auto-ethnographic perspectives. The writers of New Journalism thus combined genres to create an immersive experience for their readers. Today, journalists continue to blend genres, not just of text, but media forms, as they aim to create the 'completeness of experience' that Dowling claims is characteristic of immersive longform.² But although digital multimodal longform journalism builds on an analogue tradition, the text's placement in the digital space changes its affordances. This paper thus asks: What affordances contribute to effective meaning-making in digital multimodal longform articles?

CHARACTERISTICS OF DIGITAL MULTIMODAL LONGFORM

Digital multimodal longform journalism is often called an 'emerging genre' which 'seeks to capture its audience by combining text, photographs, looping videos, dynamic maps and data visualizations into a unified whole'.³ The crucial thing here is that the many different modalities complement rather than distract from one another. Hiippala further observes that 'simplified navigation and user interfaces, together with smooth transitions between multimedia content, slow down the readers' interaction with the longform'.⁴ While easily overlooked, transitions are important because they organize an article's content and can be used 'to mark a switch between different semiotic modes'.⁵ Thus, they

guide the reader's attention from one modality to another without losing the flow of the narrative. This is often achieved by imbuing transitions with cinematic qualities to keep the reader immersed in the story.⁶ Immersion, as used in this paper, is taken to signify a completeness of experience, during which readers find themselves in a 'cognitive container', in which different modalities 'work to hold reader attention rather than scatter it'.⁷ Moreover, immersion means that the reader is emotionally, intellectually, and/or critically engaged in the narrative aesthetics of the piece.⁸

While writers have experimented with uploading longform texts to digital environments since the early years of the internet, the genre has only recently found its place. This is largely due to the discrepancy between the linear narrative style of the genre and the hypertext nature of the internet. For while the digital environment frees journalists from the space limits of analogue publishing, and hyperlinks promise an infinitely flexible and interconnected space for stories to unfold, hyperlinks can be distracting and do not lend themselves to a cohesive narrative. Moreover, ads and various notifications from the computer's other programs vie for readers' attention. In fact, in the heydays of the internet, 'the best web writing was thought to be short and direct'.⁹ More than two decades later, we are still faced with the paradox of the rise of longform in a world of ever-shortening texts. However, a surge in recent research on the topic is gradually providing a deeper understanding of the genre. Especially the last ten years have seen an explosion of digital multimodal longform journalism (duly followed by academic analyses). In academia and journalism alike, the genre's breakthrough is largely attributed to *The New York Times* article 'Snow Fall: The Avalanche at Tunnel Creek', writ-

ten by John Branch and published online on December 20, 2012.¹⁰ The article, which consists of six chapters of captivating writing, video interviews, interactive graphics, and animated simulations, ‘helped open the door for more compelling combinations of multimedia and text in *The New York Times* newsroom—and at publications across the country’.¹¹

THE PARADOX OF THE RISE OF LONGFORM IN A WORLD OF EVER-SHORTENING TEXTS

Nonetheless, the challenges posed by the digital environment persist. The early idea that web writing should be short and direct was effectively summarized in the title of Steve Krug’s 2000 bestselling guide to web design and usability, *Don’t Make Me Think*. Later, such sentiment has been supported by eye-tracking studies, such as a 2006 study which discovered that readers often scan web content in an F-shaped pattern, revealing that ‘exhaustive reading is rare’.¹² More fundamentally, popular discourse on how new media are worsening attention spans continues to exist. Tracking the American discourses on attention, Newman writes that ‘the idea of a connection between a culture’s media and its collective habits and patterns of paying attention has been appealing to a number of influential thinkers, including Walter Benjamin, Theodore Adorno, Max Horkheimer, and Marshall McLuhan’.¹³ However, Newman argues that ‘while it might bear the influence...of some scholarly voices such as McLuhan’s, the circulation of the notion of media-shortened attention has proceeded in popular discourse in the absence of compelling, expert-produced data, as a lay theory of media effects’.¹⁴

So, if our attention spans are not decreasing, then why do we skim-read web content? Researchers have suggested that ‘screen-

based reading on the Web is more likely to involve skim reading' because it is characterized by 'more time spent browsing and scanning, keyword spotting, one-time reading, non-linear reading, and reading more selectively'.¹⁵ However, a 2016 eye-tracking study focused on digital multimodal longform journalism found that 'users do not look at longform digital journalism as a sum of identifiable parts. They look at them all when they look at, read, watch, scroll, and share that story'.¹⁶ This suggests that the articles which participants read for the study succeeded in combining different modalities into a unified whole. Following Hiippala's earlier stated observation, this unification effectuates a slower interaction between the reader and the longform and, hence, encourages deep reading. In other words, the way that the different modalities work together to construct a seamless narrative allows the reader to immerse themselves in the story rather than skim it. This stands in contrast to other web content where the modalities, say a main text and images or videos in the form of ads, scatter the reader's attention. As noted, the uniformity of digital multimodal longform journalism is important because it 'creates a cognitive container characterized by an internally coherent news package'.¹⁷ In this way, 'digital longform maintains the feel of a container associated with print newspapers', so that the different modalities complement one another rather than distract the reader from the narrative.¹⁸

AFFORDANCES AND MULTIMODAL LITERACIES

In 1966, the ecological psychologist James J. Gibson coined the term 'affordances', explaining that they are 'attributes of an object enabling the perceiver to take action'.¹⁹ Importantly, his theory showed that 'the qualities of an object cannot be character-

ized without considering the abilities, context and needs of the perceiver'.²⁰ Gibson's ideas have later been applied to a number of fields, but have been especially prominent in media studies. There, the term 'affordances' was popularized by Don Norman who applied it to the (human-computer) interaction and the relationship between design and user experience. Appropriating Gibson's theory to digital multimodal journalism, we can interpret the ability, context, and needs of the perceiver as the reader's literacy levels. Thus, a digital multimodal longform article has certain affordances that encourage specific interactions between the reader and the article. For example, navigational cues determine how the reader scrolls or clicks through the article and how and when they interact with the different modalities. Furthermore, for a digital multimodal longform article to retain the reader's attention and avoid skim reading or distraction from the narrative, the many modes must encourage the reader to carry meanings across modalities. To this end, it has been argued that the different modalities in digital multimodal longform share similarities with 'the visual storytelling techniques of cinema'.²¹ Thus, they afford a more cohesive narrative as they 'deepen the way we engage with narrative, transforming news consumption from article reading to an immersive multimedia experience'.²²

To achieve an immersive multimedia experience, readers must be able to 'combine modes' through meaning-making processes. These processes are at the heart of literacies. Duffy explains that 'to be literate in a domain is to be able to evaluate, interpret and critique what one is witnessing, reading or experiencing in order to make meaning from it and ultimately to guide decision making'.²³ To make meaning from a digital multimodal longform article

thus requires multiple literacies as one is witnessing, reading, and experiencing multiple modalities at the same time. Therefore, literacy cannot only be understood as the ability to read the text in a digital multimodal longform article. The reader must also be able to connect the different modalities into a coherent narrative, thereby making meaning. This is an individual process and hence ‘the outcome may vary between individuals’.²⁴ Furthermore, Potter emphasizes that media literacy ‘is not a natural state’ but rather a multidimensional perspective which is developed cognitively, emotionally, aesthetically, and morally.²⁵ This perspective, he argues, is built on knowledge structures about ‘media effects, media content, media industries, real world, and the self’.²⁶ These structures aid the meaning-making process as they allow people to reflect critically on the media they consume. Thus, affordances can contribute to effective meaning-making if they encourage sustained interaction with the longform (creating a so-called cognitive container) and guide the reader towards immersion in the narrative to create emotional, intellectual, and/or critical engagement with the article.

CASE STUDIES

To investigate which affordances contribute to effective meaning-making in digital multimodal longform articles, this paper presents two case studies. Focusing on traditional news outlets and honoring the legacy of ‘Snow Fall’, both articles are from *The New York Times*’ digital platform. The articles were selected from the site’s yearly round-up of ‘selected Times graphics, visualizations and multimedia stories,’ called ‘2022: The Year in Visual Stories and Graphics’.²⁷ The two articles were chosen as they display different modalities and transitions and, therefore, are

somewhat representative of the genre when considered through the limited scope of this paper.

‘What It’s Like to Ski Nearly Blind’ was published in February 2022 as part of *The New York Times*’ project ‘Athletes and their Olympic-level fears’.²⁸ Asides from the byline credited cinematographer, editor, and producer, Emily Rhyne, the multimedia project involved a total of seven producers, two cinematographers, two designers and developers, two people assisting with additional production, three project editors, and three companies or persons who provided addition video. While the project description does not mention how many people worked on each article, this summary still shows the scale on which digital multimodal longform journalism can unfold, and the amount of expertise required to produce stories told through multiple modalities.

When the reader lands on the opening screen of the article, its title and lead are superimposed on a muted video. Below the lead, a prompt encourages the reader to ‘scroll to continue’. Doing so takes the reader to the same full-screen video, now fully visible with the option to ‘replay with sound’. After watching the one-minute and two-second video, the reader is again prompted to scroll further. Doing so lands them on six short paragraphs of text. The text is displayed on a simple background without any distracting elements other than a small minimalistic illustration that breaks up the text block. This structure of interspersed, screen-by-screen, text and video persists as the reader continues vertically down the article. One of these videos merits a closer look. One of these videos in the shape of ski goggles: First, the caption ‘100% vision’ demonstrates the athlete’s view of had she been fully sighted. As the reader con-

tinues to scroll, explanatory text boxes flow in from the bottom of the screen before the caption changes to ‘Millie’s 5% vision’ and the ski goggles blacken to represent the athlete’s limited peripheral sight. Another transparent text box moves over the screen before the video is overlapped with audio as well as animation that signals the incoming sound waves to the goggles. Eventually, another three text boxes appear from the bottom of the screen to explain how important audio cues are for the athlete’s completion of the course. Ultimately, at the end of the article, the reader has been guided through an Olympic-level slalom ski course by means of text, video, audio, animation, and illustration.

As demonstrated above, ‘What It’s Like to Ski Nearly Blind’ operates in a vertical space as the reader transitions between modalities by scrolling, so that each screen enters and exits at once. This corroborates Hiippala’s 2017 finding that ‘the longform genre prefers to organize the content into a linear structure’ in such a way that the entire screen is dedicated to ‘a single semiotic mode at a time.’²⁹ Importantly, this sense of linearity makes it easy to become immersed in the narrative, as the reader does not need to resolve ‘discourse relations across the layout space’ which ‘allows the reader to remain focused on the unfolding narrative’.³⁰ Furthermore, the different modalities employed in ‘What It’s Like to Ski Nearly Blind’ work together to create a ‘cognitive container’ as the text flow guides the reader’s attention directly from one modality to the next. Modes such as text, video, and audio also complement each other as they allow the reader to both experience the narrative from a third-person perspective in the text and, thereafter, from a first-person perspective in the video interview. This likely sparks an emotional involvement in the narrative as

the reader reads and hears about the athlete's relationship fear. Meanwhile, the superimposed textual, visual, and auditory modes employed in the screen with the ski goggles likely engage the reader critically or intellectually as they work to understand the technique used to ski nearly blind.

Looking to the affordances in this first case study, the text boxes saying 'scroll to continue' encourage the reader to embark on the journey into the athlete's experience of skiing nearly blind. The scroll transitions themselves afford effective meaning-making as they structure the story by slowly guiding the reader along a linear narrative which makes it easier for the reader to become immersed in it. As previously suggested, the different modalities complement each other by letting the reader experience the story from multiple perspectives and thus afford meaning-making through an emotional, intellectual, and/or critical engagement with the story. This is especially true for the animation with the ski goggles, which gives the reader a chance to, literally, see the world through the athlete's eyes. Ultimately, it is clear that 'What It's Like to Ski Nearly Blind' requires multiple literacies. This, however, does not imply that meaning-making is more laborious than it is when reading a text-only article. Because, even without highly developed knowledge structures, the transition explanations and the linearity of the narrative makes it easy for the reader to immerse themselves in, and reflect on, the story.

'Inside the Apocalyptic Worldview of "Tucker Carlson Tonight"' was published in April 2022 and is part of *The New York Times*' series 'American Nationalist', which uncovers 'the rise of Tucker Carlson.'³¹ On the '2022: The Year in Visual Stories and Graphics' landing page, the article was attributed to the investi-

gative reporter, Karen Yourish, who works on the Graphics desk of *The New York Times*. However, on the article's third screen, she shares the byline with eight other writers. Later, additional credits are also given to four other reporters, while video clips and transcripts are credited to external websites. Again, this shows the scope of digital multimodal longform journalism and confirms Dowling's assertion that the emergent genre marks a 'departure from the print tradition of the single-byline story'.³²

The article's opening screen shows a collage of muted looping videos, all taken from various episodes of *Tucker Carlson Tonight*. The full-screen collage is superimposed with a lead, a small box informing that 'this story contains audio,' and, after a few seconds, a text in the bottom right corner appears, saying 'click to continue or use your arrow keys'. The next screen brings the reader 16 video clips with audio in a row. The clips gradually expand, giving the impression that Carlson is moving intimidatingly closer while the background collage fades. The article continues, alternating between black screens with white text, videos with audio, video stills with superimposed text or audio, and a variety of multimedia graphics which sometimes include audio and/or text. The entire article spans 64 screens, divided into six chapters. First appearing at the start of the second chapter, a progress bar remains visible throughout the rest of the article to indicate to the reader how far they are in the story. Every time the screen features a video with audio, a progress circle also appears in the bottom right corner to indicate the length of the video. Once a video has finished playing or an estimated reading time has passed, the 'click to continue or use your arrow keys' text reappears, nudging the reader to the next

screen. The last screen in chapter five is a video of one of Carlson's outros, signaling the end of the article.

'Inside the Apocalyptic Worldview of "Tucker Carlson Tonight"' contains combinations of text, video, audio, animation, and graphs. Unlike the first case study, this article operates in a horizontal space and the reader must resolve 'discourse relations across the layout space'. In the academic literature on multimodal literacies, this is said to make effective meaning-making more difficult as readers do not have the same sense of linearity as in a vertically oriented story. But while a scroll transition is the 'appropriate choice' for 'moving into textual content,' a horizontal transition can lend the story cinematic qualities. Hiippala argues that this 'shows that media convergence is not restricted to content, but also extends to multimodal structures'.³³ With this in mind, the horizontal layout and click/arrow key transitions thus seem fitting for the story as video is a much more predominant mode in 'Inside the Apocalyptic Worldview of "Tucker Carlson Tonight"' than in 'What It's Like to Ski Nearly Blind'. Moreover, the lack of linearity is mitigated by the presence of the progress bar at the top of the screen in chapters two through six and the progress circle in the bottom-right corner during video content. Firstly, these elements give the reader a better understanding of where they are in the narrative. Secondly, the progress bar helps the reader make meaning by organizing each screen into greater narratives which fall under headings such as 'The Ruling Class,' 'Replacement,' 'Show's Format,' and 'Destruction of Society'.

In terms of affordances, the 'click to continue or use your arrow keys' text that appears on each screen enables the reader to take action. It reminds them to continue reading the article and makes

it more interactive, as using arrow keys is reminiscent of playing a video game. The few seconds which the reader must wait before the text appears also afford a short break in between screens. This gives the reader time to reflect on the content they have seen and is likely to aid in meaning-making processes. Furthermore, the different modalities transport the reader into the story and cast them in the role of one of Carlson's viewers. In the textual content, the words "you" and "they" are always written in separate colors to stand out from the otherwise white text. The video and audio content complements the text and substantiates the writer's arguments using scenes from Carlson's show. The interaction between these modalities is quite repetitive but serves to create a very clear and coherent narrative that is easy to follow. In fact, it is at times so easy to follow that it seems almost hypnotic. However, prompted by superimposed text, graphs with statistics and illustrative graphics encourage the reader to critically reflect on the clips and audio from the show. Overall, the different modalities all represent different ways of telling the same story, making it easy to carry meaning from one mode to the other. Thus, while the longform operates in an unconventional space for textual content, it exploits the cinematic possibilities afforded by the horizontal layout. This demonstrates perhaps one of the most important characteristics of digital multimodal longform journalism: that text is one of the most flexible modalities available to storytellers.

CONCLUSION

To conclude, we have seen that affordances can contribute to effective meaning-making if they encourage sustained interaction with the longform. In the case studies, this is aided by transitions which provide linearity and flow to the story, or which invite the reader

to interact with the article and provide them with a short break during which meaning-making processes can occur. Moreover, the full-screen layout of the articles, whether vertical or horizontal, captures the reader's attention in a distraction-free 'cognitive container'. Meanwhile, the complementary relationship between different modalities guides the reader towards greater immersion in the narrative by engaging several of their senses. Furthermore, the case studies demonstrated that different modalities are often used to express different perspectives, thus allowing the reader to engage emotionally, intellectually, and/or critically with the article as they carry meaning from one modality to the other. Thus, affordances that contribute to effective meaning-making in digital multimodal longform articles include transitions, layout and structure, and intentionally designed modalities, including but not limited to text, video, audio, graphics, illustrations, and animation. Finally, it is important to remember that 'the qualities of an object cannot be characterized without considering the abilities, context and needs of the perceiver'.³⁴ In other words, making meaning is an individual process informed by a person's literacy level.