Making sense of *danmu*: Coherence in massive anonymous chats on Bilibili.com

Discourse Studies

1–20

© The Author(s) 2020

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/1461445620940051

journals.sagepub.com/home/dis **Leticia-Tian Zhang and Daniel Cassany**

Universitat Pompeu Fabra, Spain

Abstract

Although coherence has been widely studied in computer-mediated communication (CMC), insufficient attention has been paid to emergent multimodal forms. This study analyzes a popular commentary system on Chinese and Japanese video-sharing sites – known as *danmu* or *danmaku* – where anonymous comments are superimposed on and scroll across the video frame. Through content and multimodal discourse analysis, we unpack *danmu*-mediated communication analyzing the newest interface (on Bilibili.com), the comments, the interpersonal interactions and the unusual use of the second-person pronoun. Results show that despite the technological constraints (hidden authorship, unmarked sending date and lack of options to structure comments), users construct order in interactions through repetition, *danmu*-specific expressions and multimodal references, while using playful language to make fun. This study provides an up-to-date analysis on an increasingly popular CMC medium beyond well-studied social networking sites, and broadens the understanding of coherence in contemporary CMC.

Keywords

Barrage subtitles, bullet comments, Chinese social media, cohesion, collaborative video annotation, computer-mediated communication, conversation, *Danmaku*, digital discourse, fan studies, multimodal discourse analysis

Introduction

Massive chats on Twitch.tv, YouTube Live and Facebook Live have become increasingly popular as a research subject (e.g. Ford et al., 2017; Hamilton et al., 2014). On these platforms, videos are live-streamed and viewers participate in the event by

Corresponding author:

Leticia-Tian Zhang, Universitat Pompeu Fabra, Roc Boronat, 138, Office 53.606, Barcelona, 08018, Spain.

Email: leticiatian.zhang@upf.edu

Figure 1. Screenshot of a *danmu*-commented episode of the Japanese anime series *Detective Conan* on Bilibili.com. The central comments read: 'The murderer is him' (L Zhang and Cassany, 2019c: 9).

posting comments and reading real-time responses. Different from traditional chat spaces, live-streaming videos have an embedded or adjacent chat channel, where the interaction develops as the streamer's activity proceeds. In this way, new topics tend to be prompted by the video content (Herring, 2013) and conversation may seem chaotic and illegible when too many participants join in (Ford et al., 2017).

On Chinese and Japanese video-sharing sites, a new genre of chat, namely *danmu*, is emerging. *Danmu* is a system of superimposed comments running across the screen from right to left as the video plays (Figure 1). Compared to comments on popular Western live-streaming sites, both genres allow context-specific and multimodally-composed texts. However, *danmu* comments are completely anonymous and are temporally fixed to appear at the point of insertion. Future audience view the video with all the accumulated comments – without authorship or date of insertion – and experience a temporal feeling of live interaction (Johnson, 2013; Li, 2017).

The commentary system was created by Nico Nico (2006–), a Japanese video-sharing site catering to *otakus* (fans of anime, comics and video games) (Johnson, 2013; Nakajima, 2019; Nozawa, 2012; Steinburg, 2017). Nico Nico now ranks 10th among the most visited sites in Japan (Alexa, 2019), with over seventy-five million registered users (Kadokawa, 2019). The embedded comments – referred to as *komento* (an English loanword) – became so popular that sometimes videos are covered with comments, similar to barrage (a concentrated artillery bombardment), or *danmaku* in Japanese. As the technology spread in China, the translated term *danmu*, has become a synonym of the system. We prefer the Chinese term, in accordance with the context of the study.

Indeed, China is probably where the foreign interface has gained the widest acceptance. During the last ten years, *danmu* has transitioned from a niche entertainment to a common feature in most video-sharing platforms. The most popular *danmu*-themed site, Bilibili (2009–), claims around ninety million visitors per month and more than four million monthly paying users (Bilibili, 2019). It is also the only website requiring a membership test (on knowledge of *danmu* netiquette and fandom culture) for anyone wishing to

send *danmu*. Since 2004, some movie theaters have experimented with *danmu*, allowing the audience to text live comments onto the big screen (cf. MuVChat screenings in the U.S. in Dwyer, 2017). With the rapid growth of the streaming industry, the *danmu* interface has been adopted by Chinese live-streaming platforms (Cao, 2019; Chen and Chen, 2019; G Zhang, 2019). However, this paper is concerned with the most common use of *danmu*, that is, on video-sharing platforms.

A rapidly increasing number of scholars have explored the *danmu* phenomenon from different perspectives. Computer scientists are interested in user motivations and its promising applications in the online education industry (Chen et al., 2017, 2019; Lin et al., 2018). Communication scholars have argued that through *danmu*, platforms like Bilibili can become an alternative space for democratic discussions (Yin and Fung, 2017) and a virtual heterotopia to resist social pressure, repression and control (Chen, 2018; Gu, 2017; Zheng, 2016). Finally, linguists have illustrated users' appropriations of *danmu*, such as collaborative subtitling (Díaz-Cintas, 2018; Yang, 2019, 2020), humorous acts (Hsiao, 2015; L Zhang and Cassany, 2019c), heteroglossia and multilingualism (Y Zhang, 2017, 2020) and language and intercultural learning (L Zhang and Cassany, 2019a, 2019b). However, no study to date has looked into the interactional organization of *danmu* from a discourse analytical perspective.

For an ordinary viewer, it can be difficult to make sense of a fast-moving and ever-changing body of texts, especially if they are also frequently colorful, multidirectional and written in diverse scripts (e.g. Chinese characters, Latin alphabet) with multiple semiotic resources (icons, smileys, etc.). However, the texts must somehow appear meaningful and entertaining to *danmu* enthusiasts who not only enjoy the 'bullet curtain', but also made the effort to pass the membership test to join the community. Hence, drawing on data collected from previous studies on *danmu* (L Zhang and Cassany, 2019a, 2019b, 2019c, in press), we raise the following research questions:

RQ1: What procedures do users apply to create coherence in *danmu*?

RQ2: To what extent are *danmu* interactions coherent and meaningful?

Coherence in multimodal computer-mediated communication

Coherence in computer-mediated communication (CMC) has merited considerable investigation over the years. The classic work of Herring (1999) showed that text-based CMC can be both incoherent and enjoyable. Defining *interactional coherence* as sustained, topic-focused, person-to-person exchanges, she found that many CMC media incoherent features, for example, disrupted adjacency, overlapping exchanges, and topic decay. However, users adopt a series of compensatory strategies to minimize possible confusion, for example, backchannels and addressivity in IRC channels, and linking and quoting expressions in asynchronous group discourse. Moreover, incoherence also motivates language play and humor, contributing to one of the key features of online communication (Danet, 2001).

Herring's work inspired much subsequent research on interactional coherence in textual CMC genres, most notably in instant messaging (Berglund, 2009; Lam and Mackiewicz, 2007; Mackiewicz and Lam, 2009; Woerner et al., 2007) and multiparty chat rooms (Herring, 2013; Markman, 2013; Simpson, 2005). Revisiting Herring's claims on coherence, some researchers (Berglund, 2009; Lam and Mackiewicz, 2007) believed that conversations should not be considered incoherent if they do not generally lead to miscommunication. Simpson (2005: 342) also noted that speakers manage to 'accord meaning and unity to the text in the discourse process'. In other words, the texts are coherent to the participants, otherwise the medium would not be popular.

Since the last decade, text-based CMC has been increasingly incorporated into Web 2.0 platforms, for example, social networking, video-sharing and online gaming sites. In this vein, a small number of publications have investigated coherence in textual CMC in multimodal contexts. Zelenkauskaitė and Herring (2008) contributed the first study to analyze interactional coherence in television-mediated text messaging (iTV SMS); Herring et al. (2009) investigated coherence in text chat in the interface of a fast-paced multiplayer online game; Honeycutt and Herring (2009) looked into coherence of exchanges on Twitter; Bou-Franch et al. (2012) provided a meticulous linguistic analysis on coherence in YouTube comments; and Ford et al. (2017) analyzed 'practices of coherence' in massive chats on Twitch, a live-streaming platform dedicated to video games.

The above-mentioned research was conducted by scholars from different fields (informatics, linguistics, anthropology) using different approaches (Dynamic Topic Analysis and VisualDTA [a tool to visualize topic development], content and discourse analysis, ethnographic observation). Interestingly, their findings all revealed conversations that were coherent to a greater or lesser extent. Among others, iTV SMS, YouTube comments and Twitter messages seemed to display the most surprising degree of coherent (and dyadic) interactions. In addition to classic cohesive devices (reference, substitution, ellipsis, conjunction and lexical cohesion; Halliday and Hasan, 1976), the researchers identified medium-specific resources for coherence building, for example, the @ sign as a marker of addressivity (Nilsen and Mäkitalo, 2010; Werry, 1996) in tweets (Honeycutt and Herring, 2009).

In contrast, chat messages in online gaming tend to be less coherent, especially in games which demand more immediate attention than others. In games of such nature, Herring et al. (2009) found most chats were short and abbreviated reactions, concerning the game itself. However, the authors also observed 'periodically interspersed' actual conversations, 'which, although they tend to be limited in scope, can be engaging and unexpectedly coherent' (Herring et al., 2009: 8).

Finally, chat messages in live streams are the most problematic, especially when the audience scales. Ford et al. (2017) referred to Twitch videos attracting more than 10,000 concurrent viewers as massive chats or *crowdspeak*. *Crowdspeak* contains seemingly 'chaotic, meaningless or cryptic' messages (e.g. emotes or digital icons provided by Twitch). However, through practices of coherence (shorthand, bricolage, and voice-taking), users achieve and enjoy 'a different kind of coherence that prioritizes crowd-based reaction and interaction over interpersonal conversation' (Ford et al., 2017: 859).

Danmu comments could be comparable to Twitch chats: both are triggered by a video and are shown during the viewing session. However, on video-sharing sites, *danmu*

comments are not live, in other words, *danmu* is ‘a type of asynchronous CMC that is experienced synchronously, where users see messages (annotations) appear and disappear from the viewable screen at set times unless playback is interrupted’ (Howard, 2012: 5). This *pseudo-synchronicity* has been considered a key feature accounting for the popularity of *danmu* (Chen et al., 2017; Johnson 2013; Li, 2017; Steinberg, 2017; Yang, 2020). However, how the unique interface affects users’ interaction still needs empirical examination.

Following pragmatic and discourse analytical perspectives (Bou-Franch et al., 2012; Herring, 2013; Simpson, 2005), we understand coherence as ‘a general process of sense-making in which individuals engage whenever they communicate’ (Bou-Franch et al., 2012: 502). Regarding massive chats, Ford et al. (2017: 859) provided another useful definition: ‘by coherence, we simply mean that the chat makes sense to participants and is not experienced as a breakdown, overload, or other difficulty’. To sum up, there has been an unceasing interest in exploring coherence norms and practices in emerging online environments. We believe *danmu*-mediated communication constitutes a necessary case study that offers original data to tackle the problem of coherence.

Methodology

The data analyzed in this study forms part of a 4-year project on the *danmu* language (L Zhang, 2020). Our earlier investigation focused on creative usages through the distinctive affordances of *danmu* (multimodally designed and positioned comments to create humor, *in situ* discussions that develop metalinguistic and intercultural awareness). Based on our understanding of the potential of the medium, this study goes back to the most fundamental question: how do users make sense of *danmu* and subsequently achieve various purposes (making fun or parody, learning)? To answer this question, we drew on a primary dataset (a.) and a complementary dataset (b.):

- a. A *danmu*-commented episode of the historical science fiction series *El Ministerio del Tiempo* (MdT) (Spanish Television, 2015–). Critically acclaimed as the best Spanish TV series in history (*El País*, 2017), the series narrates time travels of three newly recruited agents (Julian, Amelia and Alonso) to preserve Spanish history. As a representative sample, we chose the 70-minute pilot episode, translated by Chinese fans (or *fansub*; L Zhang and Cassany, 2019d) and uploaded to Bilibili in September 2015. The chosen episode was the first and most popular publication of the series on Bilibili, attracting 1590 *danmu*.
- b. The discussion thread ‘What are the funny *danmu*?’ on Zhihu (a Chinese social networking site similar to Quora). The initial post was created in January 2016 and, with more than 5000 answers, it is by far the most popular thread about *danmu* on the platform. The dataset was compiled from the most liked answers (>1000 likes) and includes 327 screenshots of ‘funny *danmu*’, selected by users from different video-sharing/streaming platforms. For the present study, we chose 106 screenshots that contain some interaction triggered by the ‘funny *danmu*’.

The analysis was initially designed to identify cohesive devices in the datasets following Halliday and Hasan's model (1976). However, after being confronted with contradictory results, we realized that the categories do not seem to adapt to the Chinese language. Many researchers have remarked on the *zero pronoun* or *zero anaphora* of Chinese, that is, a situation in which an anaphor is lexically absent (Huang, 1994; Jiang, 2016; Okurowski, 1989). Instead, context or pragmatics plays a central role in communication, as opposed to grammar (or inflection) in most European languages. Consequently, necessary adjustments need to be made when studying cohesion in Chinese, as concluded by Yeh (2004: 258):

'Devices in Halliday and Hasan's model (1976), such as reference, lexical cohesion, and conjunction, may be present in most languages. However, the importance attached to various types of cohesive devices might be different. Some of them might be avoided in a particular language, while the others are preferred. In our comparison of Chinese and English, for example, the third person impersonal pronoun is generally avoided and another cohesive device, lexical repetition, is, in compensation, adopted'.

Given the particularities of Chinese, we adopted an interpretive or data-driven approach to make sense of *danmu*. First, we conducted a multimodal analysis of the latest *danmu* interface on Bilibili – not yet described in the existing literature – to illustrate the platform affordances. This knowledge is crucial as contextualization of the subject. Moreover, from a semiotic technology perspective, the multiple semiotic affordances can motivate or constrain users' ways of interaction (Zhao et al., 2014).

Second, we conducted a content analysis of the 1590 comments (dataset a.). We coded the comments inductively based on the context the authors referred to: 1) the film, 2) personal viewing situation, and 3) previous comments. Focusing on the last category, we identified the exchange sequences and examined the length and general structure of comment threads. Following Bou-Franch et al. (2012), we marked separately adjacent turns and non-adjacent turns. The coding was conducted by the first author in Chinese and reviewed by the second author through multiple discussions.

The coding was followed by a fine-grained multimodal discourse analysis of the interactions observed in both datasets. While the focus was how users create order or mark their responses in conversations, a distinctive usage pattern emerged and is discussed separately in the section 'second-person references'. The illustrations of *danmu* data (Figures 3–8) were inspired by Bai et al. (2019: 535510). By juxtaposing two consecutive screenshots and reproducing the translated *danmu* correspondingly, we managed to consider not only the verbal content, but also time, position or direction that influence the meaning making of each comment.

Making sense of *danmu*

The interface

The first step to make sense of *danmu* interactions is to understand the system and platform affordances. Figure 2 shows the latest *danmu* interface on Bilibili.com (accessed 31 October 2019). Although the data for this study came from an earlier

Figure 2. Screenshot of the main part of a video page on Bilibili.com. The video is called [Trump] Dear CHINA and shows Trump saying the word ‘China’ non-stop.

version (see L Zhang and Cassany, 2019a, 2019b), the main constituents, spatial arrangement and functions remain unchanged.

This version was released in July, 2018 and features four main components: a video player with superimposed *danmu*, a rectangular area below for editing *danmu*, a drop-down list of all comments or ‘*danmu* pool’ (with the date of insertion and time of appearance in the video) on the top right, and a section for recommended videos. In addition, there are a video description, ‘uploader’ information, and also a traditional comments section below the video (see Wu et al., 2018 for a comparison between *danmu* and common comments on Bilibili).

The *danmu* editor has multiple functions. On the bottom left, it shows the number of real-time connected viewers and *danmu*. However, there is a maximum number of viewable *danmu* for each video depending on the duration, for example, 1000 for a three minute video. In this example, out of almost 7000 *danmu* sent since the video was uploaded, only the latest 1000 are shown in the video and accessible through the drop-down *danmu* list.

Next to the numerical data, there are two icons for (de)activating the *danmu* function and customizing the visual effect of *danmu*, respectively. Users can adjust the font, amount, transparency and speed of viewable texts, activate the anti-block function (with 15% of the screen uncovered), and filter comments based on specific qualities (movement, color, type).

The comment box occupies the center of the *danmu* editor. A default text in grey invites viewers to write and send *danmu* without interrupting the viewing session. On its left, when moving the pointer onto the icon A, a text editor is activated (marked in red in Figure 2). *Danmu* authors can select the font size (small or standard), color (from the palette or in

HTML color codes) and movement (scrolling, popping up at the top or bottom center). Upon clicking the ‘send’ button, *danmu* will appear in one of the following ways:

- *Scrolling danmu*. Comments enter the screen at the top right corner, scroll left in a straight line across the screen, and exit on the top left corner in a few seconds. Scrolling *danmu* is the default setting and the most frequent type in many videos, for example, the white *danmu* on the right side of the Trump video.
- *Top danmu*. Comments appear at the top center of the screen and freeze for a few seconds before disappearing. When coinciding with scrolling *danmu*, top *danmu* overlay the scrolling text, for example, yellow and blue comments at the top center of the Trump video (counting how many times Trump has said ‘China’).
- *Bottom danmu*. Comments (dis)appear in the same way as top *danmu* but at the bottom center of the screen. Bottom *danmu* can be used to add subtitles for ‘raw videos’ (Yang, 2019) or to invent funny conversations for the purpose of play (L Zhang and Cassany, 2019c).
- *Advanced danmu*. While scrolling, top and bottom *danmu* are known as common *danmu* and available for all Bilibili members, advanced *danmu* is a reserved function for those having *tipped* to and acquired permission from the video’s uploader. Advanced *danmu* allows complex text configurations (duration, specific location, visual effect, transparency, typeface), which are also exploited by expert users to create *danmu* art (Johnson, 2013; L Zhang and Cassany, 2019c).

The comments

Figure 3 presents a visualization of eight scrolling *danmu* passing on the screen in 25 seconds. The screenshots are taken from the series MdT, and the excerpt reproduces the translated and numbered comments according to their *entrance time*. In the first screenshot, turn 453 is sent and enters the screen at 18m 16s; in the second screenshot, turn 457 has scrolled to the left corner at 18m 47s after entering at 18m 39s. As the number of comments increases, the text layer ‘thickens’ gradually. Turn 458 and turn 459 both are sent around 1s later than turn 457 and run beneath it so as not to overlap. However, because turn 459 contains more characters, the text travels slightly faster than turn 458.

The excerpt in Figure 3 contains three *film turns* (FT) where viewers react to the character (turn 453) or to the plot (turn 457 and 458); one *viewing turn* (VT) explaining the personal reason for viewing (turn 456), and four *responding turns* (RT) to previous comments, marked by arrows in the excerpt. The exact reference is unclear in turn 460, however, it did not affect the coding.

Table 1 shows three layers of contexts in *danmu*. First, a predominant number of comments refer to the film, sharing impressions, doubts or complaints about the plot, characters, soundtrack, etc. Many viewers are also interested in the sociocultural background depicted in the film, for example, historic events, figures and locations, or the Spanish people (799. ‘They are really Spanish, speaking dirty language all the time’).

Second, less than a third of the comments refer to the *danmu* system. The most common referent is previous *danmu*, through which users can construct (extended) interactions. Meanwhile, depending on real-time connected users shown in the interface, some

Table 1. Classification and frequency of *danmu* references.

Contexts	Topics	Examples	Frequency
Film	Plot	231. (FT) ‘The Ministry of Time rocks, the mobile phone has signal even after time travels’.	1123 (71%)
	Characters	24. (FT) ‘The soldier who just died looks quite handsome!’	
	Soundtrack	387. (FT) ‘Why does the background music sounds like farting. . .’	
	Subtitles	48. (FT) ‘Thank you fansub group’.	
	Sociocultural background	63. (FT) ‘Spain in 1850 was extremely powerful, it called itself <i>the empire on which the sun never sets</i> ’.	
Danmu system	Previous <i>danmu</i>	64. (RT to 63) ‘Isn’t <i>the sun never sets</i> the British Empire????’	412 (26%)
	Real-time viewers	14. (RT) ‘To the other buddy, hello! 15/2’.	
Viewing session	Situation	192. (VT) ‘Watching alone I feel very lonely’.	55 (3%)
	Reasons	26. (VT) ‘I need to learn Spanish’.	

users send greetings (detailing the date). However, due to the anonymous nature of the interface, the referred-to user would be unlikely to notice the greeting or join the chat synchronously.

Finally, a small number of comments refer to the viewing session, sharing when, where and why they are viewing the series. Many *danmu* of this type appear at the beginning of the video, while others are triggered by unexpected plot developments, for example, an erotic scene (1302. ‘Shit, I do not wear headphones in the subway, and there are people behind me. . .’; and 1288. ‘The day we can see porn on Bilibili is just around the corner!’).

The interaction

Danmu interaction, defined here as person-to-person communication through the *danmu* system, is not marked by date of insertion nor any form of authorship (similar to Wikipedia articles). Furthermore, the *danmu* system does not provide any option for users to structure their comments, such as the option ‘reply to comments’ in many social networking sites. Instead, the order of *danmu* messages exclusively follows the moment of insertion in the video’s timeline. This configuration is similar to chat messages in IRC, which Herring (1999) noted as a cause of ‘disrupted adjacency’.

Such distinctive features can have a paradoxical effect. On the one hand, the anonymous, spontaneous and democratic environment prompts participation from anyone at any moment. In dataset a., we identified 152 exchange sequences or comment threads, with an of 3.7 turns per thread (Table 2).

On the other hand, the participant is unaware of and likely unconcerned with possible responses from future viewers. If there are any, the participant will not be notified and

Table 2. Length of *danmu* interactions.

No. of turns	No. of threads
2~5	138
6~10	11
11~33	3
Total	152

Figure 3. *Danmu* synchronicity in 25 seconds of MdT.

can only review the video intentionally afterwards. This lack of knowledge causes delayed, redundant or even contradictory responses. Of 412 RTs in dataset a., only 53 are adjacent turns entering the screen immediately after the triggering comment. Instead, interactions through non-adjacent turns are much more frequent. To make their contribution meaningful, users adopt a variety of strategies to address each other and achieve coherence.

Figure 4 shows intertwined *danmu* interactions containing three exchange sequences, triggered by film turns 262, 264 and 265, respectively. The first sequence is the adjacent pair of turns 262 and 263, sharing the confusion caused by the rapidly-spoken Spanish (*vamos* misheard as *zou* or ‘let’s go’ in Chinese). The second sequence is prompted by turn 264 drawing on contextual knowledge about the actor, which is confirmed by turn 267. The third sequence starts with turn 265, whose comparison to another time-travel series triggers turns 269 and 270.

Figure 4 also illustrates if and how participants mark their response to a specific addressee. Turn 263, despite being an adjacent turn, repeats the predicate in turn 262. Turn 267 also relies on lexical cohesion, in particular, an intertextual reference that would be clear to people with the same interest. Turns 269 and 270 repeat the keyword

Figure 4. Intertwined *danmu* interactions.

(‘Warehouse’) in the triggering turn and further strengthen the reference with two *danmu*-specific expressions: ‘wait for me’ and ‘don’t go’. Here authors play with the visual effect of scrolling texts, where later comments seem to be ‘chasing’ earlier ones. Similarly, twenty users adopt the expression ‘+1’, an Internet slang for ‘I agree’ or ‘me too’, for example, 130. ‘Forgive me for maybe hearing FBI...’ and 131. ‘FBI +1’.

In contrast, turn 266, despite being clearly a response (written in Spanish), does not provide more cues to locate the referent. Two hypotheses can be made to explain the situation: 1) when turn 266 was sent, there were so few existing comments in the time frame that the reference was clear for the author (the same explanation could apply to ‘the one before’ in turn 460, Figure 3); and 2) the language choice of turn 266 echoes the discussion on Spanish pronunciation (*vamos* vs. *zou*), and thus turn 266 potentially pertains to the first sequence. In any case, the task of sense-making falls on the audience. Depending on the version of *danmu* accessible at the moment of viewing, each viewer could relate some *danmu* differently than others.

Figure 5 shows another common way of marking relation in *danmu*. As scrolling *danmu* travel across the screen from right to left, *danmu* inserted at an earlier point will appear left of those inserted later (except for lengthy comments that fly faster). As a result, viewers like turn 203 refer to previous comments as ‘the left’.

In the excerpt in Figure 5, turn 200 reflects Amelia’s speech on women’s ‘worst enemies’ from a male perspective. Turn 201 concerns another interaction initiated by turn 190 (‘Female protagonist from *Ángel o Demonio*’). Both triggered by turn 200, turn 202 counters the accusation by pointing out the male role in liberation movements, while turn 203 further emphasizes the vicious patriarchal plotting. The discourse analysis clarifies that ‘the left’ in turn 203 refers to turn 200 and not the adjacent turn 202, otherwise the argument would not make sense. Moreover, the author usually needs to read the whole

Figure 5. Position-based references.

danmu before drafting a response. However, turn 203 enters before the previous turn has finished, and is therefore indented in relation to turn 202.

Feminist topics seem to stir interest among many viewers. In dataset a., the most heated discussion (32 turns) developed around feminine hygiene products, triggered by a present-day tampon shown to the nineteenth-century Amelia. In response to some complaints about the discussion being 'unbearable' and 'without limit', several users refer to the commenters as 'male chauvinists' and 'living antiquity'.

The majority of the data analyzed so far appear in the default setting, that is, scrolling fashion, white font and written in sinograms or the Latin alphabet. However, in addition to the default option, the *danmu* interface makes available to users a variety of options for typographic manipulation, for example, color, movement, size (as detailed earlier).

Figure 6 shows multimodal forms to mark the position of the triggering turn. In the top screenshots, the mention of the name 'Blanca' triggers turn 134 (many Chinese language learners use another name in the target language). After three film turns describing the job as 'a great responsibility' (turn 135), 'too hasty' (turn 136) or dubious (137. 'Don't tell me the Ministry of Time already exist in that era'.), turn 138 responds to turn 134 and shares the naming coincidence. To make the reference explicit, turn 138 also employs Spanish, and adds a Japanese-style emoticon (*kaomoji*), where the eyes (arrows) look suspiciously towards the left.

Similarly, in the bottom screenshots, turn 1320 raises a question regarding the logic of the plot, and turn 1323 starts the explanation with a leftwards arrow. Two turns precede turn 1323 in the bottom-right screenshot: film turn 1321 reacting to the 'unexpectedly new pillow', and turn 1322 succinctly responding to turn 1320 ('Multiracial of time').

In dataset b., most 'funny *danmu*' are also created by users exploiting the multimodal system affordances. Given the creativity and originality, these comments invite and initiate an exchange rather than simply respond to the film. Correspondingly, the responses

Figure 6. Multimodal references.

to those funny *danmu* make reference to their multimodal qualities. Figure 7 shows two instances of *danmu*-specific references to funny comments. On the left, the funny *danmu* – the yellow text – calls for attention using a striking color and the central position. The humorous intervention is welcomed by other viewers and referred to as ‘the yellow font’. On the right, the ‘funny *danmu*’ – or rather an amazing piece of art – receives wide admiration from the audience. Fans acknowledge the use of color (blue, as the hair of the character) and type (advanced *danmu*). Fans also use Chinese subcultural expressions (合影 ‘take a group photo’, 大佬 ‘a big shot’) to praise and pay respect to the *danmu* artist (see also Johnson, 2013: 308).

Second-person references

Since Halliday and Hasan (1976), deictics have been considered a potential indication of discourse cohesion. However, some researchers also argued that a text with cohesive devices is not necessarily coherent, and a coherent text sometimes does not contain apparent cohesive devices, especially in computer-mediated communication (Herring, 2013; Simpson, 2005: 342). Furthermore, recent studies show that deictics can acquire new meanings and functions adapted to the medium (Collister, 2012; Cuenca, 2014).

The excerpt in Figure 8 shows multiple uses of the deictic ‘you’ in *danmu*. Triggered by a ‘door of time’ that leads to the past, turn 533 imagines a Chinese scenario with the Yongzheng Emperor (or the ‘Fourth Prince’ in a popular time travel series). The plural ‘you’ (marked by a suffix in Chinese) in turn 533 thus refers to the minister and Julian in the first screenshot. Turn 534 (四爷那个你留下!) is an adjacent turn consisting of repetition and the *danmu* specific expression ‘you stay’ (similar to ‘wait for me’ and

Figure 7. Color-based (left) and type-based (right) references.

Figure 8. Second-person references in *danmu*.

'don't go' in Figure 4), where 'you' means turn 533. Turn 535 shifts to the perspective of Spaniards during the Inquisition, which makes 'you' a cataphora for 'heretics scum' and 'we', 'the Spanish'. In turn 536, the viewer would be inside the film frame, speaking from behind a door of time. Therefore, the plural 'you' again refers to the minister and Julian.

Figure 8 shows original and playful means of participation, facilitated by the use of the second-person pronouns. As in turn 533, viewers sometimes *speak to* the characters, especially when surprised by the plot development. For example, viewers object to the erotic scene between Alonso and the waitress (1302. 'What about your wife, your wife. How long has it been? No wonder you told her to forget you'. and 1305. 'Wait you just accept it like this?'). When a man from 1808 steals a modern pistol from the police,

viewers question him (1805. ‘Even if you take it, you can’t make bullets’. and 1093. ‘Let’s see if anyone cares about you’.).

As in turn 535, viewers can also *speak for* the characters. In Chinese Internet slang, such voiceover is called an OS (overlapping sound) and reveals characters’ thoughts and feelings. For instance, when Julian is stunned by the offer to work for the ministry, viewers actively picture his painful realization of the job duties (603. ‘Main actor OS: Why does it have to be me?’ and 617. ‘OS: I choose to go die’.).

In intense scenes, viewers use the imperative voice to *command* the characters. When Julian secretly takes a photo of the suspect, the camera shutter sound irritates the audience (1259. ‘Mute your phone, idiot’.). When the suspects sneak a book out of the store, viewers shout: 782. ‘Hey!’ and 785. ‘Pay!’. Other times, viewers intervene just for fun. When an actress unexpectedly kisses another actress, a commenter pleads: 222. ‘Fuck! Let that girl go, let me do it!’, along with declarations of love (809. ‘Girl you are so beautiful! Marry me!’).

Discussion and conclusion

Regarding RQ1, we identified both verbal and nonverbal resources activated by users to create coherence in *danmu*. Nonverbal resources are the most characteristic of the *danmu* system. Most *danmu* can be interpreted based on: 1) the insertion time in the film; 2) the surrounding elements on the screen (actor, subtitles, other *danmu*); 3) multimodal properties of the text (the directionality [right-left or top-bottom], different colors and types); and 4) the additional deictic symbols (arrows).

As to verbal resources, the most common device seems to be lexical cohesion, realized through: 1) quotation or repetition (e.g. ‘that person who said...’); 2) indirect addressivity (e.g. the semantic anaphora ‘the male chauvinist’ and ‘living antiquity’ referring to the patriarchal comment); and 3) cultural references (e.g. the pragmatic anaphora ‘King Fernando’ to substitute ‘male protagonist from *Isabel*’). Second-person references also contribute to cross-turn cohesion (e.g. ‘you stay’). Compared to Romance and Germanic languages, the Chinese used in *danmu* contains fewer grammar-based cohesive resources (e.g. substitution, ellipsis) and mainly relies on vocabulary and context to achieve coherence (Huang, 1994; Yeh, 2004).

Regarding RQ2, we found that *danmu* exhibits both coherent and incoherent features. Like a bar where people chat loudly, multiple and unpredictable topics of conversation emerge, develop and die away at the same time. The disrupted adjacency, redundant participation, and frequent topic digression are all signs of incoherence in *danmu*. However, unlike traditional chat channels, *danmu* constitutes an example of event-based communication, like Twitch chats (Ford et al., 2017; Musabirov et al., 2018), online gaming chats (Herring et al., 2009) and television-mediated chats (Zelenkauskaitė and Herring, 2008). Most *danmu* comments center on the ongoing film (*prompt-focused participation*, Herring, 2013) and only make sense if situated in the specific film scene, viewing context or the *danmu* co-text.

In short, we could conclude that despite obvious technological factors that discourage coherence (anonymity, character limit, short visibility), many users still achieve meaningful interpersonal interactions. The fact that certain comments lack clear references

does not seem to affect the general popularity and utility of the medium. Furthermore, some users exploit the potential incoherence of *danmu* for playful purposes (Herring, 1999, 2013), for example, to make fun of the scrolling effect ('wait for me', 'don't go'), or to take part in a specific frame (dialogue, command or voice-over of the character).

Another important contribution of this study relates to the construction of the potential coherence of *danmu*. In dataset a., 26% of the comments are RTs, while the average length of the threads is 3.7 turns, and the longest thread extends to 32 turns (on female hygiene). Similarly, Honeycutt and Herring (2009) found coherent Twitter threads of up to 30 comments in a massively noisy textual environment, while Zelenkauskaitė and Herring (2008) found an eight-comment thread of text comments posted to an iTV program (superimposed over music videos). In this vein, we encourage future studies to continue exploring the depth and patterns of topic development in *danmu*-mediated interactions.

Finally, our data and discussions extend the understanding of digital culture. So far, many studies have stressed the ephemerality and nonsensicalness of *danmu* comments. Cao (2019: 10) was concerned at having 'gone too far ahead in assuming that these fleeting texts may indeed require interpretive efforts to produce an effect'. Zheng (2016) also pointed out that many *danmu* videos are replete with nonsensical memes, which help build up an alternative community for otakus in the virtual space. However, our data illustrated that with more people engaging in interaction, *danmu* becomes more relevant and significant. These comments not only represent a variety of reactions to the film, but also facilitate understanding of adjacent comments and guide future interpretation of the film.

Acknowledgements

The authors are grateful to the anonymous reviewer for the detailed and insightful suggestions.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was supported by a predoctoral grant from the Chinese Scholarship Council for the first author (CSC No. 201608390036) and the Spanish competitive research project 'ForVid: Video as a language learning format in and outside schools' (RT2018-100790-B-100; 2019–2021; Ministry of Science, Innovation and Universities).

Note

1. The repetition can also be argued as a quotation, since the Chinese expression 仓库的 is an abbreviated form of '(that person who say/said) Warehouse'.

References

- Alexa (2019) Nicovideo.jp traffic statistics. Available at: <https://www.alexa.com/siteinfo/nicovideo.jp?ver=classic> (accessed 15 May 2019).
- Bai Q, Hu Q, Ge L, et al. (2019) Stories that big danmaku data can tell as a new media. *IEEE Access* 7: 53509–53519.
- Berglund TO (2009) Disrupted turn adjacency and coherence maintenance in Instant Messaging conversations. *Language@Internet* 6. Available at: <http://www.languageatinternet.org/articles/2009/2106/Berglund.pdf> (accessed 11 December 2019).
- Bilibili (2019) Bilibili Inc. Announces fourth quarter 2018 and fiscal year 2018 financial results. Available at: <http://ir.bilibili.com/news-releases/news-release-details/bilibili-inc-announces-fourth-quarter-2018-and-fiscal-year-2018> (accessed 11 December 2019).
- Bou-Franch P, Lorenzo-Dus N and Blitvich PGC (2012) Social interaction in YouTube text-based polylogues: a study of coherence. *Journal of Computer-Mediated Communication* 17(4): 501–521.
- Cao X (2019) Bullet screens (*danmu*): texting, online streaming, and the spectacle of social inequality on Chinese social networks. *Theory, Culture & Society*. Epub ahead of print 29 October. DOI: 10.1177/0263276419877675.
- Chen TZ (2018) Poetic presumption of animation, comic, game and novel in a post-socialist China: a case of a popular video-sharing social media Bilibili as heterotopia. *Journal of Consumer Culture*. Epub ahead of print 12 July. DOI: 10.1177/1469540518787574.
- Chen X and Chen Z (2019) Between the marked and the unmarked: twin semiotic paradoxes of the barrage in China's livestreaming fandom. *Media, Culture & Society*. Epub ahead of print 8 November. DOI: 10.1177/0163443719876617.
- Chen Y, Gao Q and Rau PLP (2017) Watching a movie alone yet together: understanding reasons for watching danmaku videos. *International Journal of Human-Computer Interaction* 54(4): 731–743.
- Chen Y, Gao Q, Yuan Q, et al. (2019) Facilitating students' interaction in MOOCs through timeline-anchored discussion. *International Journal of Human-Computer Interaction* 35(19): 1781–1799.
- Collister LB (2012) The discourse deictics $\wedge$ and $\leq$ —in a World of Warcraft community. *Discourse, Context & Media* 1(1): 9–19.
- Cuenca M (2014) The use of demonstratives and context activation in Catalan parliamentary debate. *Discourse Studies* 16(6): 729–752.
- Danet B (2001) *Cyberpl@y*. London: Berg.
- Díaz-Cintas J (2018) 'Subtitling's a carnival': new practices in cyberspace. *The Journal of Specialised Translation* 30: 127–149.
- Dwyer T (2017) Hecklevision, barrage cinema and bullet screens: an intercultural analysis. *Participations: Journal of Audience & Receptions Studies* 14(2): 571–589.
- El País (2017) 'El Ministerio del Tiempo', mejor serie española de la historia según la prensa especializada. Available at: https://elpais.com/cultura/2017/12/15/television/1513342137_758333.html (accessed 11 December 2019).
- Ford C, Gardner D, Horgan L, et al. (2017) Chat speed op pogchamp: practices of coherence in massive twitch chat. In: *Proceedings of the 2017 CHI conference extended abstracts on human factors in computing systems*, Denver, CO, USA, 6–11 May, pp. 858–871. New York: Association for Computing Machinery.
- Gu J (2017) *Celebrating and discussing the queerly masculine: Hollywood superheroes reimaged in fan videos on Chinese barrage video websites*. Master dissertation, Georgetown University, USA.
- Halliday MAK and Hasan R (1976) *Cohesion in English*. London: Longman.

- Hamilton WA, Garretson O and Kerne A (2014) Streaming on Twitch: fostering participatory communities of play within live mixed media. In: *Proceedings of the SIGCHI conference on human factors in computing systems*, Toronto, Canada, April 26–May 1, pp. 1315–1324. New York: Association for Computing Machinery.
- Herring SC (1999) Interactional coherence in CMC. *Journal of Computer-Mediated Communication* 4(4): 1–13.
- Herring SC (2013) Relevance in computer-mediated conversation. In: Herring SC, Stein D and Virtanen T (eds) *Handbook of Pragmatics of Computer-Mediated Communication*. Berlin: Mouton de Gruyter, pp. 245–268.
- Herring SC, Kutz DO, Paolillo JC, et al. (2009) Fast talking, fast shooting: text chat in an online first-person game. In: *Proceedings of 42nd hawaii international conference on system sciences* (ed RH Sprague Jr), Big Island, USA, 5–8 January, pp. 1–10. Los Alamitos: IEEE Press.
- Howard CD (2012) *Higher order thinking in collaborative video annotations: Investigating discourse modeling and the staggering of participation*. PhD Thesis, Indiana University, USA.
- Honeycutt C and Herring SC (2009) Beyond microblogging: conversation and collaboration via twitter. In: *Proceedings of 42nd Hawaii international conference on system sciences* (ed RH Sprague Jr), Big Island, USA, 5–8 January, pp. 1–10. Los Alamitos: IEEE Press.
- Hsiao C (2015) The verbal art of tucao and face-threatening acts in *danmu* screening. *Chinese Language and Discourse* 6(2): 109–132.
- Huang Y (1994) *The Syntax and Pragmatics of Anaphora: A Study with Special Reference to Chinese*. New York: Cambridge University Press.
- Jiang Y (2016) Deixis and anaphora. In: Huang C and Shi D (eds) *A Reference Grammar of Chinese*. Cambridge: Cambridge University Press, pp. 484–517.
- Johnson D (2013) Polyphonic/pseudo-synchronic: animated writing in the comment feed of Nicovideo. *Japanese Studies* 33(3): 297–313.
- Kadokawa (2019) Financial results for the second quarter of fiscal year 2019. Available at: https://ssl4.eir-parts.net/doc/9468/ir_material_for_fiscal_ym1/55491/00.pdf (accessed 11 December 2019).
- Lam C and Mackiewicz J (2007) A case study of coherence in workplace instant messaging. In: *Proceedings of the 2007 International Professional Communication Conference*, Seattle, USA, 1–3 October, pp. 1–6. Los Alamitos: IEEE Press.
- Li J (2017) The Interface affect of a contact zone: danmaku on video-streaming platforms. *Asiascape: Digital Asia* 4(3): 233–256.
- Lin X, Huang M and Cordie L (2018) An exploratory study: using danmaku in online video-based lectures. *Educational Media International* 55(3): 273–286.
- Mackiewicz J and Lam C (2009) Coherence in workplace instant messages. *Journal of Technical Writing and Communication* 39(4): 417–431.
- Markman KM (2013) Conversational coherence in small group chat. In: Herring SC, Stein D and Virtanen T (eds) *Pragmatics of Computer-Mediated Communication*. Berlin: Mouton de Gruyter, pp. 539–564.
- Musabirov L, Bulygin D, Okopny P, et al. (2018) Event-driven spectators' communication in massive esports online chats. In: *Extended abstracts of the 2018 CHI conference on human factors in computing systems* (eds R Mandryk and M Hancock), Montréal, Canada, 21–26 April, pp. 1–6. New York: ACM.
- Nakajima S (2019) The sociability of millennials in cyberspace: a comparative analysis of barrage subtitling in Nico Nico Douga and Bilibili. In: Frangville V and Gaffric G (eds) *China's Youth Cultures and Collective Spaces: Creativity, Sociality, Identity and Resistance*, Abingdon: Routledge, pp. 98–115.

- Nilsen M and Mäkitalo Å (2010) Towards a conversational culture? How participants establish strategies for co-ordinating chat postings in the context of in-service training. *Discourse Studies* 12(1): 90–105.
- Nozawa S (2012) The gross face and virtual fame: semiotic mediation in Japanese virtual communication. *First Monday* 17(3): 1–23.
- Okurowski ME (1989) Textual cohesion in modern standard Chinese. *Language Sciences* 11(1): 89–104.
- Simpson J (2005) Conversational floors in synchronous text-based CMC discourse. *Discourse studies* 7(3): 337–361.
- Steinberg M (2017) Converging contents and platforms: Niconico video and Japan's media mix ecology. In: Neves J and Sarkar B (eds) *Asian Video Cultures*. Durham, London: Duke University Press, pp. 91–113.
- Werry CC (1996) Linguistic and interactional features of Internet Relay Chat. In Herring SC (ed) *Computer-mediated Communication: Linguistic, Social and Cross- Cultural Perspectives*. Amsterdam: John Benjamins, pp. 47–63.
- Woerner SL, Yates J and Orlikowski WJ (2007) Conversational coherence in Instant Messaging and getting work done. In: *Proceedings of the 40th annual Hawaii international conference on system sciences*, Big Island, USA, 3–6 January, pp. 77–78. Los Alamitos: IEEE Press.
- Wu Q, Sang Y, Zhang S, et al. (2018) Danmaku vs. forum comments: understanding user participation and knowledge sharing in online videos. In: *Proceedings of the 2018 ACM conference on supporting groupwork* (eds A Forte, M Prilla and A Vivacqua), Sanibel Island, USA, 7–10 January, pp. 209–218. New York: AMC.
- Yang Y (2019) Danmaku subtitling: An exploratory study of a new grassroots translation practice on Chinese video-sharing websites. *Translation Studies*. Epub ahead of print 25 September. DOI: 10.1080/14781700.2019.1664318.
- Yang Y (2020) The danmaku interface on Bilibili and the recontextualised translation practice: a semiotic technology perspective. *Social Semiotics* 30(2): 254–273.
- Yeh C (2004) The relationship of cohesion and coherence: a contrastive study of English and Chinese. *Journal of language and linguistics* 3(2): 243–260.
- Yin Y and Fung A (2017) Youth online cultural participation and Bilibili: an alternative form of democracy in China? In: Luppiciini R and Baarda R (eds) *Digital Media Integration for Participatory Democracy*, Hershey, PA: IGI Global, pp. 130–154.
- Zhang G (2019) *Zhibo: an ethnography of ordinary, boring, and vulgar livestreams*. PhD Thesis, RMIT University, Australia.
- Zelenkauskaitė A and Herring SC (2008) Television-mediated conversation: coherence in Italian iTV SMS chat. In: *Proceedings of the 41st annual Hawaii international conference on system sciences*, Big Island, USA, 3–6 January, pp. 1–11. Los Alamitos, CA: IEEE Press.
- Zhang L (2020) *Understanding danmu: interaction, learning and multimodality in fan video comments*. PhD Thesis, Universitat Pompeu Fabra, Spain.
- Zhang L and Cassany D (2019a) The danmu phenomenon and media participation: intercultural understanding and language learning through The Ministry of Time, *Comunicar* 27(58): 19–29.
- Zhang L and Cassany D (2019b) 'Is it always so fast?': Chinese perceptions of Spanish through danmu video comments. *Spanish in Context* 16(2): 217–242.
- Zhang L and Cassany D (2019c) 'The murderer is him ✓': multimodal humor in danmu video comments. *Internet Pragmatics*. Epub ahead of print 16 October. DOI: 10.1075/ip.00038.zha.
- Zhang L and Cassany D (2019d). Strategies of audiovisual comprehension and Spanish to Chinese translation in a fansub community, *Revista Española de Lingüística Aplicada/Spanish Journal of Applied Linguistics* 32(2): 618–647.

- Zhang L and Cassany D (in press) Language play through multimodality: color, position, symbol, and shape in *danmu*-mediated communication. In Parodi G (ed) *Multimodality: From Corpus to Cognition*. London: Bloomsbury.
- Zhang Y (2017) *Heteroglossic Chinese online literacy practices on micro-blogging and video-sharing sites*. PhD Thesis, University of South Florida, USA.
- Zhang Y (2020) Adopting Japanese in a popular Chinese video-sharing website: heteroglossic and multilingual communication by online users of Bilibili.com, *International Multilingual Research Journal* 14(1): 20–40.
- Zhao S, Djonov E and Van Leeuwen T (2014) Semiotic technology and practice: a multimodal social semiotic approach to PowerPoint. *Text & Talk* 34(3): 349–375.
- Zheng X (2016) *Borderless fandom and contemporary popular cultural scene in Chinese cyberspace*. PhD Thesis, University of Washington, USA.

Author biographies

Leticia-Tian Zhang obtained her PhD in Translation and Language Sciences from Universitat Pompeu Fabra and will be a lecturer in Spanish Philology at Beijing Foreign Studies University since fall, 2020. She has published in areas including Chinese social media practices (e.g. *danmu/danmaku*), multimodality, informal language learning, and fansubbing.

Daniel Cassany is an Associate Professor in Discourse Analysis at Universitat Pompeu Fabra in Barcelona. He has published more than 15 books about written communication and language teaching in Catalan, Spanish and Portuguese, like *Describir el escribir* (1987) or *Laboratorio lector* (2019), and also more than 90 texts, among others, scientific articles in bulletins and reviews, in several languages. Website: https://portal.upf.edu/web/daniel_cassany/.