

# “The murderer is him ✓”

## Multimodal humor in *danmu* video comments

Leticia Tian Zhang and Daniel Cassany

Universitat Pompeu Fabra

This paper analyzes humorous comments created through a popular viewing-and-commenting system used in China and Japan, known as *danmu* (or *danmaku*). This system enables its users to superimpose anonymous comments on the video frame, which are displayed in subsequent viewing. We collected 327 user-selected ‘funniest’ screenshots of comments from *danmu* video sharing sites. Using content and discourse analysis, we re-contextualized the comments and identified main mechanisms of humor. Results show that speakers make fun of the plot, characters and of each other, relating to the video frame, Chinese culture and Japanese fandom. They rely on non-aggressive but rather playful teasing, allusions and retorts, and apply multimodal resources such as color, layout, and symbols to enhance the humorous effect. Our study contributes to the emerging research focus on multimodal humor (Yus 2016), social semiotics and a discursive approach to *danmu*-mediated communication.

**Keywords:** *danmaku*, affordances, incongruity, semiotic modes, video sharing, Bilibili

### 1. Introduction

A little more than a decade ago, the Chinese video-sharing website AcFun (an abbreviation of anime, comic and fun) launched a viewing system known as *danmu*, which marked the beginning of a new era for online video sharing in China. Inspired by the Japanese prototype created by Nico Nico Douga (“Smiley Smiley Video”), *danmu* displays viewers’ comments on the video content, scrolling from right to left on the screen as the video displays. This phenomenon seems to resemble a “bullet curtain,” which is the literal meaning of the term *danmu* in Chinese (Figure 1).

*Danmu* authors can select the text size, color and movement; audience can also personalize their viewing experience (adjust the font, transparency, speed,



Figure 1. A screenshot of a *danmu* video from Bilibili: World Cup 2018 official song Команда<sup>1</sup>

etc.), filter certain comments, or deactivate *danmu*. This interface rapidly gained popularity in China and is now supported by both fan-based platforms and many mainstream video streaming sites. The most popular *danmu* server, bilibili.com, claims around 90 million visitors per month, and more than 4 million monthly paying users (Bilibili 2019; see Nakajima 2019 for a comparison of Nico Nico with Bilibili).

Regarding its popularity, a survey of 248 participants (Chen, Gao and Rau 2017) reveals that people view *danmu* to obtain: (1) information (e.g., background story, music, actors); (2) entertainment (e.g., complaints, parodies); and (3) social connectedness, i.e., a “pseudo-synchronous” feeling of shared viewing (Johnson 2013). Meanwhile, infrequent users complain about the visual clutter (e.g., ugly styles or fonts), the excess of information, and information ‘pollution,’ in other words, irrelevant and redundant comments, for instance, personal emotions, quarrels between fans, or spoilers. These could destroy the aesthetic effect of the original video, hinder viewers’ understanding, and thus bring inconvenience to the viewing experience.

Nevertheless, Chinese internet users embrace *danmu*, and in some cases the comments seem to garner more attention than the video itself. As Hsiao (2015: 113) observes, “for viewers, the programs are not the only main points; rather, the other viewers’ comments that scroll across the screen in real time are more popular.” It provides a habitat for the Japanese fandom (also known as ACG, an acronym for anime, comic and game) in China, and fosters a complex and dynamic participatory culture (Zheng 2016; Chen 2021); in the wider social context, it serves as an alternative platform for free speech and potential democratic participation (Yin and Fung 2017).

1. See <https://bit.ly/2RTtlfl> (accessed 11 December 2018).

This phenomenon has drawn scholarly interest from different disciplines, including information and communication studies, cultural and film studies (Nozawa 2012; Johnson 2013; Xu 2016; Zheng 2016; Dwyer 2017; Steinberg 2017; Yin and Fung 2017; Díaz-Cintas 2018; Nakajima 2019; Chen 2021), and applied linguistics (Hsiao 2015; Zhang 2017; Yang 2020, 2021; Zhang and Cassany 2019a, 2019b). However, *danmu* remains a largely unknown practice to both international academic audience and the public.

Our study is situated in this emerging body of work and explores an unstudied aspect of *danmu*. As can be readily seen from Figure 1, *danmu* comments are increasingly colorful and multidirectional. The texts are written in diverse scripts (e.g., Chinese characters, Latin alphabet, Cyrillic script) with non-standard semiotic resources (icons, smileys, etc.). The interaction between software design and its use is a crucial point in multimodal studies, in particular, studies on semiotic technologies (Djonov and Van Leeuwen 2013; Zhao, Djonov and Van Leeuwen 2014; Zhao and Zappavigna 2018). Our study expands earlier research that has mainly centered on Western multimedia platforms such as YouTube (e.g., Adami 2009; Sindoni 2013; Benson 2017), in order to explore original modes of interaction in the Asian context.

This paper emerges from the first author's PhD research, which focuses on the role of *danmu* in language and intercultural learning (Zhang and Cassany 2019a, 2019b). We draw upon data from a popular post titled 'What are the funny *danmu*?' on the Chinese question-and-answer website Zhihu. With the most liked screenshots selected by fans across years, platforms and genres, we aim to answer the following research questions:

RQ1: What humor mechanisms are employed in funny *danmu* comments?

RQ2: How do multimodal resources contribute (or not) to the humorous effect?

## 2. Humor and the *danmu* language

Our data deal with what is funny, amusing or laughable, or the study of humor in academic terms. Scholars from many fields like psychology, sociology, literature, philosophy and computer science have analyzed the concept of humor extensively. From our linguistic approach, the *incongruity-resolution* (IR) model is among the most popular theories that explain the strategies underlying humorous communication, in particular, the (canned) joke (e.g., Suls 1983; Raskin 1985; Attardo and Raskin 1991; Ritchie 2004; Dynel 2009a; Yus 2016; Attardo 2017).

The premise of the IR model is that people interpret jokes in a dual-phase pattern: (1) the hearer comes across certain incongruity or *cognitive dissonance* (Yus

1997); and (2) the hearer finds a resolution that reconciles the incongruity and obtains a humorous outcome. Yus (2016) differentiates between *discourse-centered incongruity*, which requires the hearer to adopt inferential strategies to process to joke for relevance (Sperber and Wilson 1995), and *frame-based incongruity*, which requires the hearer to construct an appropriate mental situation to make sense of the joke. In either case, the expectation of a non-serious conversation, the pleasure in solving incongruities, and shared social values and attitudes with the joker all contribute to a humorous outcome.

Beyond (canned) jokes, many researchers have explored *conversational humor* (Dynel 2009b, 2011) or *humor in interaction* (Norrick and Chiaro 2010). This work has shed light on a range of semantic and pragmatic devices that construct humorous effects in dialogic interactions. The first set encompasses devices that are frequent but not particular to conversational texts (Dynel 2009b), such as humorous lexemes and phrasemes; punning; allusions to existing texts with distortions and quotations; stylistic figures (comparison, irony and sarcasm); and register clash (mixing informal and formal discourse items).

The second set concerns mechanisms that are inherently interactive and interwoven into a conversation exchange, where humor is often co-constructed, constituting *conjoint humor* (Holmes 2006). Many scholars (e.g., Norrick 1993; Boxer and Cortés-Conde 1997; Hay 2000; Dynel 2008, 2009b) have analyzed, among others, teasing (jocular utterances performing playful and non-aggressive functions such as mock challenges or imitation), retorts (a quick and witty response to a preceding turn with which it forms an *adjacency pair*), putdowns (abusive and disparaging remarks), and banter (a rapid exchange of humorous lines for mutual entertainment).

Focusing on *danmu*-mediated interactions, some scholars have foregrounded the humorous feature of the *danmu* language. Hsiao (2015: 113) identifies a distinctive verbal art, *tucao*, which consists of “commenting on someone or something by uncovering the truth about it in a sarcastic, harsh, and humorous tone.” *Tucao* is realized through repetition, rhetorical questions and internet slang. Successful *tucao* acts “evoke laughter, elicit empathy, or threaten someone’s face” (Hsiao 2015: 124). They are not taken too seriously and echoed by other users. As a result, they facilitate a group rapport in the *danmu* community, which is also proven an important function of conversational humor (Hay 2000; Holmes 2005; Coates 2007).

Zhang (2017) relates *danmu*-mediated writing practices to Bakhtin’s (1984) notion of the “carnavalesque” (cf. Díaz-Cintas 2018). The concept is defined as a subversive literacy mode against the assumed dominant style, and uses humor and chaos for free expression. For example, users employ vulgar expressions to transliterate English used in the video, e.g., “Shǐ dà kē” (literal meaning: “shit big

chunk”) for Stark (the Marvel Comics superhero Iron Man). In choosing certain characters resembling the original English pronunciation, speakers manifest their plurilingual competence, a good sense of humor, and the “carnival” or “grotesque” orientations of the *danmu* language.

Unlike verbal humor or humor conveyed by language, multimodal humor remains a largely under-researched area for linguists. A few exceptions include analyzing comedy footage to identify multimodal stimuli like prosody and gesture (e.g., Attardo et al. 2003), examining cartoons (Tsakona 2009) and postcards (Francesconi 2011) through a combination of cognitive and semiotic approaches, and testing the IR model on advertisements and cartoons (Yus 2016) and image macros (Dyner 2016), i.e., user-generated online memes where a humorous text (a caption) is overlaid on a visual image.

Our study departs from previous research that conceptualizes multimodal humor via two semiotic modes, the verbal and the visual. Instead, we take into account modes made available through the *danmu* system, such as color and layout, to identify patterns or mechanisms to convey humor. This leads us to the field of multimodality and, specifically, the meaning potentials of semiotic modes.

### 3. Multimodality and semiotic modes of *danmu*

Multimodality refers to the combination of semiotic modes in the meaning making process (Kress and Van Leeuwen 2001). Multimodal discourse analysis (MDA) studies of social media have explored emerging meaning-making forms in online discourse, thanks to the advancement of sophisticated Web 2.0 technologies. Zappavigna and Zhao (2017) consider selfies as a visual genre, and analyze how they represent the everyday experiences of motherhood on Instagram. Zappavigna (2018) focuses on social tagging or hashtags, which enable metacommentary or metadiscourse to be embedded in the social media communication, with diverse linguistic functions from taxonomic classification to evaluation. Georgalou (2017) builds a comprehensive dataset of Facebook posts and interviews, to document how users discursively construct and negotiate their identity.

Many MDA studies have adopted a social semiotic perspective (Halliday 1978; Kress and Van Leeuwen 2006). This approach is based on Systemic Functional Linguistics (SFL), which asserts that language performs simultaneously three *metafunctions* in any communicative event: an *experiential* function of enacting experience, an *interpersonal* function of negotiating relationships, and a *textual* function of organizing information (Halliday and Matthiessen 2004).

From an SF perspective, we can draw some general considerations on two main semiotic modes to which *danmu* provides access, i.e., color and movement. First, as discussed by Kress and Van Leeuwen (2002: 343), color can be discussed as a semiotic resource or a mode that “is multifunctional in its uses in the culturally located making of signs.” Color can realize the ideational function to denote specific or classes of people, places and things, like colors of flags, corporations or universities; it can convey interpersonal meaning to influence others, such as orange signs of obstruction; on a textual level, it can distinguish differences as well as create coherence, creating a *color coordination* based on the same degree of brightness, and/or saturation. Moreover, color has an *associative value*, or *color symbolism* (Van Leeuwen 2011), which relates to its cultural and historic provenance, and should be carefully interpreted in the given socio-cultural context.

Second, *danmu* enables users to select the movement or the *dynamic layout* of comments. Layout is considered an integrative semiotic resource, which enables images, words and other spatially co-present elements to be combined to form cohesive and coherent multimodal texts (Kress and Van Leeuwen 2006). In semiotic software such as PowerPoint, layout has been critically examined as a primary meaning-making resource that can both benefit or constrain users’ awareness and experience in the design of slideshows (Djonov and Van Leeuwen 2013; Zhao, Djonov and Van Leeuwen 2014). In other words, while the software has considerable capacity to impose certain norms on the use of layout, users also actively go beyond those principles for their own objectives and within particular culture.

As Zhao and Zappavigna (2018) suggest, the meaning-making potential of social media genres is shaped by the constantly evolving technologies. In the case of *danmu*, the platform makes available *affordances* which guide the commenting practice, i.e., the embedded comment section and its typographic design options (See also Yang 2020 for a detailed analysis on *danmu* from a semiotic technology perspective). However, users do not take up all the resources or make semiotic choices correspondingly. As we will see in the examples, sometimes they go against the indications of the platform and produce creative and even stylized ways for humorous meaning making.

#### 4. Data collection and analysis

To collect examples of interaction in *danmu* commentary, we ran a search on Zhihu, the biggest Chinese social question answering site, similar to Quora in the Western world. Using the keyword *danmu*, we found popular posts such as “What are the funny *danmu* (comments)?”, “What *danmu* made you shed tears?” and “What *danmu* you sent went viral?”. The first question, created in January 2016

under the original title “What are the funniest *danmu* you have ever seen?”<sup>2</sup> was by far the most popular one. The answers usually contained one or more screenshots, a short explanation made by the author, and several pages of other users’ comments. The question attracted nearly 5,000 answers and over 68,000 followers. Given its highly relevant content and popularity, we chose this post as our data source.

To construct the dataset, we focused on answers with over 1,000 likes, with the most liked one hitting 17,000 likes. We excluded a few text-only answers, since it was difficult to retrieve the original viewing context. The rest of the answers presented screenshots taken from different platforms, including *danmu*-themed websites Bilibili and AcFun, official video streaming and sharing sites like Souhu, Tudou, iQiyi, etc. and the Japanese video-sharing site Nicovideo. They were selected by experts in the fan community, i.e., fans of media products and frequent users of *danmu*, and valued by the public. In other words, they constituted a representative sample from the platform for our analysis.

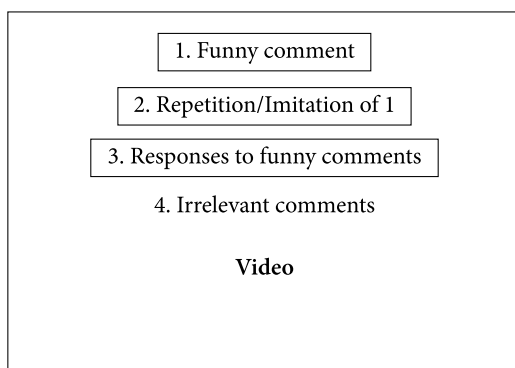
We collected 327 screenshots dated 16–19 July 2018. To facilitate the analysis, we first cleaned the data in two steps: (1) discarding repetitive screenshots; and (2) merging a sequence of continuous screenshots into one item for analysis (e.g., Figure 3), incorporating enough data to explicate the mechanism for humor. This resulted in 134 screenshots for analysis.

We based our analysis on the whole screen capture, which represents a communicative situation triggered by the funny comment (Figure 2). Since *danmu* comments are shown without authorship, it is impossible to identify the number or identity of individuals behind them. However, we consider *danmu* a social action, where certain comments tend to induce others to follow them, as our examples will illustrate. Many copied or similar messages serve to follow on from or amplify earlier messages, creating nonsensical memes (Zheng 2016). We also included responses to funny comments, as long as they were complete utterances in the capture; references to other aspects in the video were considered irrelevant and excluded.

We examined the data focusing on how commenters used the *danmu* technology to make fun of the video. Since humor was produced within the particular fan context (*danmu* culture and series/movies), the first author used the following references for contextualization and cross-check: the original post, including the author’s explanation and other fans’ comments, the original video whenever available, and Moegirlpedia (an encyclopedia of terms in Chinese and Japanese fandom) for specialized search.

---

2. <https://www.zhihu.com/question/19778466> (accessed 17 November 2018)



**Figure 2.** Basic analytical unit

Finally, to understand the humor from an insider’s perspective, the first author consulted a Chinese fandom expert. She has used Bilibili for nearly five years with a membership of the second highest level. She enjoys *danmu* viewing and is knowledgeable about East Asian subculture. The first author shares a similar background, and thus the discussion between the two further contextualized our observation and enriched the analysis.

We now discuss the examples in terms of whether fun is made of the plot, characters, or each other, and analyze in each case whether, and if so how, color, layout, etc. contribute to humor. Note that we did not consider alterations in font size (Figure 1), since its visual presentation depends on the device used for viewing and the lack of information would mislead the analysis.

## 5. Constructing humor through multimodal resources

### 5.1 Making fun of the plot

*Danmu* users make fun of the plot in three ways. To begin with, through the embedded comments, speakers can point out some detail in the video frame that affects the plot development, such as a spoiler. Figure 3 illustrates the use of downwards arrow to indicate the criminal in the Japanese anime series *Detective Conan*.

*Detective Conan* is one of the most broadcast animations in China, and many fans’ favorite activity is guessing the murderer in each episode. With *danmu*, some discovered an original way to enjoy the series, i.e., inserting deictic emojis at the exact time and place to reveal the culprit beforehand. In the aligned central comments, various speakers (in different colors) also add a check mark in their affirmation, signaling that the mission of this episode, i.e., searching for the murderer, is completed. This practice is reminiscent of giving spoilers prior to the digital era,





Extract 1. Screenshot 1 (with symbols)

‘The murderer is him (check mark).’

‘Criminal (downwards arrow).’

Figure 3. A spoiler

when comic readers drew a circle on the killer’s head upon his appearance. In both cases, the spoiler remains anonymous, which probably motivated participation.

Despite being voted as one of the funniest *danmu*, Bilibili actually forbids spoilers in its netiquette, since the result can be unwelcomed in the fan community. By examining the open access chat history, users can make an official report to the platform to mute or block the related account. From this perspective, the humorous element emerges not in the spoiler itself, but through how it is achieved collectively and at the right moment.

The multidirectional arrow, such as ←, →, ↑ and ↓, plays an important role in Figure 3 and is also among the most preferred symbols in the dataset. Zhang (2017) identifies three main functions of deictic emojis in *danmu* screening: (1) to pinpoint and correct/edit one’s own comments; (2) to pinpoint the video content; and (3) to mark one’s reply to another user. In a complex environment where messages are gradually layered onto the video, arrows act as markers of turn-taking and direct cohesive conversations. However, due to their ability to indicate certain detail in the video frame, arrows also facilitate humor which is specific to the plot.

Moreover, speakers choose different colors and a specific position to collectively repeat the punchline: ‘The murderer is him (check mark).’ While color seems to be selected based on personal taste, the layout of the phrase is designed as top *danmu*, appearing at the top center of the frame for a few seconds. According to the user guide of Bilibili, scrolling *danmu* are the default setting, and hence, multifunctional, whereas top *danmu* overlay the scrolling text and often serve as annotation. Being designed to pop up and remain static before disappearing, top *danmu* are more readily apparent and observed in many cases, as we are about to see in Figure 4.

Secondly, speakers can also interpret the plot in a humorous way, which often involves more than one video frames. For instance, Figure 4 belongs to a series of 20 screenshots, where 16 of them feature the yellow text ridiculing the narrative.



Extract 1. Screenshot 1 (in yellow)

‘An African bumps into a Japanese attempting to take over the golden oil.’

‘Then the African blames him angrily that how he dared taking them since Africa is less developed in agriculture.’

Extract 2. Screenshot 2 (in yellow)

‘Naturally the Japanese wouldn’t surrender.’

‘He would laugh.’

‘And say that their natural resources are too scarce.’

**Figure 4.** Rewriting the narrative

The original scene is taken from *Tiny Times*, a Chinese romantic drama film series, in which four young women quarrel over an alleged scandal. Argument between women is a recurring theme of the film, but also the target of criticism

and mockery, such as the yellow comment in Figure 4. Using the central position and a striking color, the commenter marks a difference from other speakers in white, suggesting an unconventional move: proposing an alternative story based on the meme of *Jinkela*.

As one of the most famous and long-lasting Chinese internet memes, *Jinkela*, or “golden oil,” is a brand of fertilizer best known for its advertisements. In one of its most popular clips, a Japanese and an African argued fiercely to take over the product for their own countries, until an American intervened and ruled in favor of the African. The exaggerating plot and the overacting of foreigners speaking Chinese motivated numerous remixes and parodies.

The humor in Extract 1 and 2 is achieved through a script opposition (SO) (Raskin 1985), enhanced by a consistent use of color and layout. The yellow text potentially suits two unrelated and incompatible scripts: the meme and the video scene, which causes incongruity and evokes laughter. To maximize the humorous effect, the speaker inserts the rewritten argument accurately according to the protagonists’ acting in each frame. The text persists in the same color and position of top *danmu* in consecutive frames, which reminds others of the parallel narrative. This captivated many viewers, with responses like: ‘What on earth did the yellow font say?’, ‘Good job the yellow font!’; and ‘OMG, I have to pause in every frame, otherwise I would miss the yellow font.’

The third way of making fun of the plot consists of a trans-modal practice. The texts acquire a graphic quality and are perceived as part of the frame. Figure 5 shows an elaborate example where speakers combine star-like symbols to *illuminate* the night.

Figure 5 is taken from the Japanese romance anime *My Little Monster*. In this screenshot, the boy is ashamed since he fails to catch fireflies for the girl. This is the series finale, and they are still not a couple officially, which makes many fans impatient. To compensate for the loss, viewers draw fireflies, signalling an alternative script with possibility for romance. From the incongruity perspective, the symbols also challenge the original script. However, unlike previous cases, the outcome does not produce a plot twist, but contributes to the plot development in a bona fide manner.

This practice exploits the pictorial register of *danmu* and plays with the visual design of the video. Users employ a variety of symbols including punctuation marks, geometric shapes and other special symbols, using different colors. Together they recreate a night of glittering fireflies and win the other audience members’ praise (Extract 1, Figure 5). As Zheng (2016: 337) suggests, “these comments have transcended the function of verbal communication, turning into a collective performance and spectacle.” The next section introduces more examples of



Extract 1. Screenshot 1

‘..... They are really lighting up!’

‘You guys are too cute (enamored face)...’

‘Fuck, I almost thought a bunch of fireflies showed up.’

Figure 5. Completing the scenario

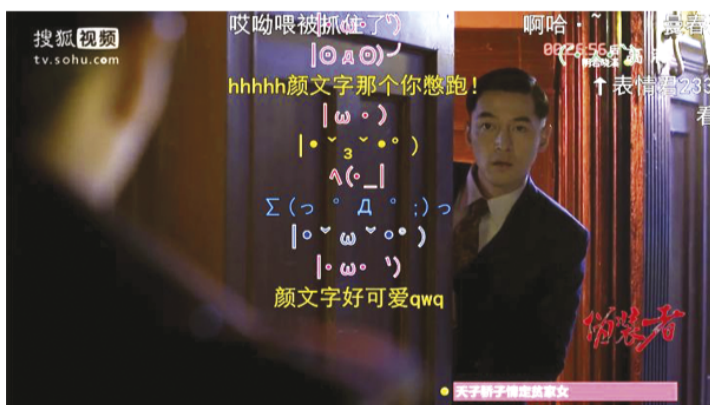
the imagine-oriented writing and the display of other resources, such as color and position.

## 5.2 Making fun of the characters

In addition to the plot, commenters also make fun of the foregrounded characters directly. A common strategy is using *kaomoji*, a Japanese emoticon style to represent the facial expression and even actions. *Kaomoji* is written in the Japanese writing systems (adopted Chinese characters or kanji and syllabaries of hiragana and katakana) and is read horizontally, e.g., (\* ^ ω ^) for joy. Figure 6 shows a straightforward example where speakers employ *kaomoji* to mimic the character and provoke laughter.

Figure 6 is captured from the Chinese espionage series *The Disguiser*. The protagonist is a spoiled young master from a rich household trained to become a spy. In this screenshot, he is caught by surprise during a mission. Triggered by the innocent and astonished face, many speakers insert *kaomoji* to tease him. Similar to Figure 5, the degree of aggression in teasing is nearly non-existent. In other words, viewers or fans do not mean to be genuinely offensive towards the character, but rather to challenge the latter jocularly (Dynel 2009b: 1293).

The writing of *kaomoji* also manifests commenters’ plurilingual repertoire. The emoticons are composed of Japanese characters and punctuation marks, but



Extract 1. Screenshot 1 (top *danmu*)

‘Hahaha to that *kaomoji*, don’t you leave!’

‘*Kaomoji* is so cute (crying face).’

Figure 6. Imitating characters

also Latin letters, e.g., *qwq* (eyes with falling tears and a quivering mouth). Other than the facial expression, *kaomoji* can draw contextualized body language. For instance,  $|\omega\circ)$  and  $\wedge(\circ\_|$  in Figure 6 represent hiding and resemble the man peeping from behind the wall.

While *kaomoji* seems popular among speakers, other genres of emojis, such as smileys and objects, are not present in our data. Our informant suggested that they seem too “universal” and “middle-aged” to be adopted by *danmu* users, who are mostly young people and fans of Japanese subculture. As a substitute, they prefer emoticons and especially *kaomoji*, which evoke a feeling of cuteness, distinguish them from ordinary users, and represent their identity. According to Zhang (2017), in *danmu* comments, emoticons and *kaomoji* are employed four times more than emojis.

A more sophisticated way to make fun of the characters is alluding to idioms and proverbs, which are recognizable to recipients with sufficient cultural knowledge. As Dynel (2009b: 1290) asserts, “their humorous force stems primarily from the language user’s acknowledgement of the pre-existing text and the quote’s relevance to the situation.” Figure 7 exemplifies how speakers exploit green, a color with idiosyncratic meanings in Chinese culture, to tease the character.

Figure 7, belonging to a series of five screenshots, is taken from the Chinese classical television series *Romance of the Three Kingdoms*. It captures the moment when the commandant Lü Bu finds out that Diaochan, while betrothed to him, is also a concubine of Dong Zhuo, his superior and foster father. In response to this



Extract 1. Screenshot 1 (in green)

‘I feel pity for Lü Bu.’

‘What? My *danmu* just turned green automatically.’

‘Green cloth (homophone for Lü Bu).’

Figure 7. Culture-related humor

plot development, many green squares appear on Lü Bu’s head at the top of the screen. Although the use of symbols and position also stand out, on this occasion it is the color that plays a crucial role in building the humor.

Traditionally in Chinese, ‘wearing a green hat’ refers to being cuckolded. This derives from earlier Chinese history, when green was considered an ‘in-between color’, used for inferior quality clothing (Xing 2009); since the 13th century, green headwraps were required for family members of prostitutes, which further evolved into ‘green hats’ for husbands whose wives commit adultery. Nowadays its implication is so widespread that it would be familiar to all the audience and create a humorous effect.

To reference this traditional association, speakers combine color with other semiotic resources: they employ square emojis to form a hat, and place them exactly at the center of the frame, covering the protagonist’s head. Others turn their texts green, expressing pity for him, acting surprisingly, and making a pun on his name (Extract 1). This prompts many laughs, acknowledgement and followers (e.g., ‘Testing testing (in green).’). In a total of four cases, we found speakers creating a green hat for husband made a cuckold, which points to a common humorous practice through *danmu* (cf. Xu 2016; Teng 2018).

Another original way to tease the characters is making up conversations for them through bottom *danmu*. Located at the bottom center, bottom *danmu* normally occupy the position of the subtitles with the same animation. This is conve-



nient if a video has no subtitles, e.g., the Russia 2018 World Cup song (Figure 1), because volunteers, or ‘wild subtitlers’ as they are known informally, provide lyrics and translation using bottom *danmu* (see Yang 2021 for an exploration on *danmaku* subtitling). However, the content can also be altered or manipulated to create a comical effect.

Figure 8 is taken from the Korean television series *My Love from the Star*. It is a romantic fantasy story between a girl and an alien with human appearance. In this intense scene, the injured alien refuses to be taken to hospital but cannot explain why, which leaves a silence in the last screenshot. Appropriating the blank space for subtitles, a speaker inserts bottom *danmu* as his response.



Extract 1. Screenshots 1–4

(Subtitle) ‘Go to the hospital.’

(Subtitle) ‘I can’t.’

(Subtitle) ‘Why can’t you?’

(*Danmu*) ‘Because... I... don’t have... health insurance.’

Figure 8. Creating subtitles

The reference to health insurance is hilarious; it spoils the mood and runs counter to the male character’s cold and distant persona. In terms of conversational humor, the response acts as a retort, intended to amuse the hearer unexpectedly, especially towards the indirect addressee or the third party, i.e., the audience of the series. The intention is succeeded, since given the minimal difference between the font used for original subtitles and *danmu*, many people do not realize the trick on first inspection.

Finally, Figure 9 represents perhaps the most extreme case where the main objective is not to convey humor, but to showcase the multimodal skills of the commenter. The screenshot captures an image-oriented writing that is far more complex than Figures 5, 6 and 7, signaling deep worship of the character, created solely and vividly by *danmu*. It shows a portrait of Rem, from the anime *Re:Zero – Starting Life in Another World*. She won the most popular Japanese anime/manga character awards on Bilibili in 2016, and inspired numerous creations of fan art. In this screenshot, she appears in the form of advanced *danmu*, involving complex typographic alterations of the text. Upon seeing the image, many fans express their excitement and admiration for the unknown creator (Extract 1, Figure 9). They also use blue font, which is Rem’s hair color.



Extract 1. Screenshot 1 (In blue)  
 ‘Amazing, this is actually *danmu*.’  
 ‘I love you Rem.’

**Figure 9.** Comment art

Different from previous cases, pixel-like images such as Figure 9 are usually created by one person, known as a *comment artist*. Johnson (2013: 308) reports that on the Japanese video-sharing site Nicovideo, they fulfill a role that “commands a kind of prestige.” Although comments are not attributed to individuals, artists enjoy more attention or praise, sometimes more than the content of the video. On Bilibili, skillful users upload tutorial videos on how to send advanced *danmu*, while the site also organizes competitions to promote *danmu* art.



### 5.3 Making fun of each other

The last category deals with commenters making fun of each other. Similar to Figure 8, this practice also relies on bottom *danmu*, but instead of referring to the narrative, speakers interact with potential viewers. In real-life conversations, formal phrases can be repeated verbatim in certain contexts (e.g., in court or army discourses) to create humor (Dynel 2009b: 1291). Similarly, Figure 10 shows a bottom comment faking a system notification, which triggers different reactions besides laughter. As a user explained in the comment section of the post, ‘this is to scare those who watch videos on their mobile phones through Wi-Fi.’ Indeed, the author of the post wrote ‘This scares the hell out of me’, although afterwards he posted the screenshot as one of the funniest *danmu*, and got more than 11,000 likes. According to other comments, this kind of trick is widely practiced in *danmu*. They admit having being tricked by ‘low battery alert,’ ‘computer shut off’ when the screen becomes temporarily black, and ‘End’ or ‘After 10 years’ in unexpected moments.



Extract 1. Screenshot 1

‘You have switched to 2G/3G/4G network.’

Figure 10. A fake system notification

On the other hand, participants also complain that they are so familiar with these tricks that they are caught unprepared when the messages are real: ‘I saw it once too, and I didn’t pay attention since I thought it was a joke. Not long after several

episodes there was an SMS reminding me that my data would exceed the limit... that was the hardest month I have ever lived...!

## 6. Discussion and conclusions

Regarding RQ 1, our analysis sheds light on a variety of mechanisms to create humor in *danmu*. Speakers interact with and make fun of the plot, the characters and each other. From an incongruity-resolution perspective, speakers challenge the original script and put forward an alternative scenario where their participation makes sense. To “get the joke,” the audience has to contextualize the funny comment in the incongruous frame (Yus 2016), using shared encyclopedic and specific knowledge, such as: (1) traditional culture (Chinese history and values); (2) popular culture and, in particular, elements of Japanese fandom (meme, idols, *kaomoji*); (3) the audiovisual product (conventions of the genre); and (4) the viewing situation (display device and network). Meanwhile, speakers resort to conversational humor mechanisms, especially when they make fun of the character and each other. Teasing, allusions and retorts are employed with a non-aggressive but rather playful intention to elicit laughter. The lightheartedness is also evident in the absence of more offensive humorous forms, such as putdowns.

Regarding RQ 2, the analysis shows that employing multimodal resources enhance the humorous effect. *Danmu* users not only draw upon the affordances of the system to make changes in color and layout, but also actively incorporate special symbols to create shapes or graphical effects. The semiotic resources contribute to humorous meaning making through textual, ideational and interpersonal level. On one hand, the resources help organize the comments, differentiating and highlighting certain messages; on the other hand, some resources represent associative values and play a key role in amusing others, e.g., green as a derogatory color for men in Chinese culture, *kaomoji* as a sign of youth and subculture. To maximize the humorous outcome, speakers even break the community’s norm, e.g., dropping major spoilers or intentionally misplacing some comments to simulate “official voices” (Bakhtin 1984). Speakers also frequently combine multimodal resources, which indicates their rich experience in practicing different usages, despite the relatively new system.

The study both corroborates and suggests new directions for related fields. First, the analysis proves that classic theories on humor such as the IR model (or other similar theories that emphasize a contrast between incompatible scripts/interpretations), can be used to identify humorous strategies in multimodal communication. Furthermore, different from cartoons (Tsakona 2009) or image macros (Dynel 2016; Yus 2019), which convey humor through either visual or ver-

bal modes or a juxtaposition of the two modes, *danmu* users employ contemporary and innovative modes (color, movement, emoji) to boost humor. In a few cases, the funniness does not originate in common humorous mechanisms, but rather *surprisingly sophisticated skills of comment making*. In other words, expert users take their liberty of creation to extremes, which negotiates and broadens audience's dynamic understanding of humor. Finally, viewers acknowledge the humor through common support strategies similar to those in face-to-face interactions (Hay 2001), such as contributing more humor, playing along with the gag, and offering sympathy to the object of teasing.

This paper contributes to several emerging research focuses: multimodal humor (Tsakona 2009; Francesconi 2011; Yus 2016); multimodal meaning making on social media (Georgalou 2017; Zappavigna and Zhao 2017), in particular, the humorous and evaluative functions of *danmu* as metadiscourse (Zappavigna 2018) targeting aspects of the primary text or the context; and the study of *danmu* from a discourse analytic perspective (Hsiao 2015; Zhang 2017; Zhang and Cassany 2019a, 2019b). Further studies may intend to deepen the discussion on discourse features of the *danmu* language, such as interaction patterns and topical development. While it is crucial to understand how *danmu* mediates, facilitates or problematizes online polylogues, scholars may also be interested in exploring more usages of the cutting-edge technology, for other purposes, like education or advertisement, and for different audiences, such as its potential application in the Western context.

## Acknowledgements

The authors thank Michael Salmon and Willa Wei Wang for their generous help in proofreading the article. This research 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-I00; 2019–2021; Ministry of Science, Innovation and Universities).

## References

- Adami, Elisabetta. 2009. "'We/YouTube': Exploring sign-making in video-interaction." *Visual Communication* 8(4): 379–399. <https://doi.org/10.1177/1470357209343357>
- Attardo, Salvatore (ed.). 2017. *The Routledge Handbook of Language and Humor*. New York: Routledge. <https://doi.org/10.4324/9781315731162>
- Attardo, Salvatore, Jodi Eisterhold, Jennifer Hay, and Isabella Poggi. 2003. "Multimodal markers of irony and sarcasm." *Humor* 16(2): 243–260. <https://doi.org/10.1515/humr.2003.012>
- Attardo, Salvatore, and Victor Raskin. 1991. "Script theory revis(it)ed: Joke similarity and joke representation model." *Humor* 4: 293–348. <https://doi.org/10.1515/humr.1991.4.3-4.293>

- Bakhtin, Mikhail. 1984. *Rabelais and His World*, trans. by Hélène Iswolsky. Bloomington: Indiana University Press.
- Benson, Phil. 2017. *The Discourse of YouTube: Multimodal Text in a Global Context*. New York: Routledge.
- Bilibili. 2019. “Bilibili Inc. announces fourth quarter 2018 and fiscal year 2018 financial results.” <http://ir.bilibili.com/news-releases/news-release-details/bilibili-inc-announces-fourth-quarter-2018-and-fiscal-year-2018> (accessed 3 March 2019).
- Boxer, Diana, and Florencia Cortés-Conde. 1997. “From bonding and biting: Conversational joking and identity display.” *Journal of Pragmatics* 27: 275–295. [https://doi.org/10.1016/S0378-2166\(96\)00031-8](https://doi.org/10.1016/S0378-2166(96)00031-8)
- Chen, Yue, Qin Gao, and Pei Luen Patrick Rau. 2017. “Watching a movie alone yet together: Understanding reasons for watching Danmaku videos.” *International Journal of Human-Computer Interaction* 54(4): 731–43. <https://doi.org/10.1080/10447318.2017.1282187>
- Chen, Zhen Troy. 2021. “Poetic prosumption 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* 21(2): 257–277. <https://doi.org/10.1177/1469540518787574>
- Coates, Jennifer. 2007. “Talk in a play frame: More on laughter and intimacy.” *Journal of Pragmatics* 39: 29–49. <https://doi.org/10.1016/j.pragma.2006.05.003>
- Díaz-Cintas, Jorge. 2018. “Subtitling’s a carnival: New practices in cyberspace.” *The Journal of Specialised Translation* 30: 127–149.
- Djonov, Emilia, and Theo van Leeuwen. 2013. “Between the grid and composition: Layout in PowerPoint’s design and use.” *Semiotica* 197: 1–34. <https://doi.org/10.1515/sem-2013-0078>
- Dwyer, Tessa. 2017. “Hecklevision, barrage cinema and bullet screens: An intercultural analysis.” *Participations: Journal of Audience & Reception Studies* 14(2): 571–589.
- Dynel, Marta. 2008. “No aggression, only teasing.” *Lodz Paper in Pragmatics* 4(2): 241–261. <https://doi.org/10.2478/v10016-008-0001-7>
- Dynel, Marta. 2009a. *Humorous Garden-Paths: A Cognitive-Pragmatic Study*. Newcastle upon Tyne: Cambridge Scholars Publishing.
- Dynel, Marta. 2009b. “Beyond a joke: Types of conversational humour.” *Language and Linguistics Compass* 3(5): 1284–1299. <https://doi.org/10.1111/j.1749-818X.2009.00152.x>
- Dynel, Marta. 2011. “Joker in the pack: Towards determining the status of humorous framing in conversations.” In *The Pragmatics of Humour across Discourse Domains*, ed. by Marta Dynel, 217–241. Amsterdam: John Benjamins. <https://doi.org/10.1075/pbns.210.15dyn>
- Dynel, Marta. 2016. “‘I has seen Image Macros!’ Advice Animals memes as visual-verbal jokes.” *International Journal of Communication* 10: 1–29.
- Francesconi, Sabrina. 2011. “Multimodally expressed humour shaping Scottishness in tourist postcards.” *Journal of Tourism and Cultural Change* 9(1): 1–17. <https://doi.org/10.1080/14766825.2010.521561>
- Georgalou, Mariza. 2017. *Discourse and Identity on Facebook*. London: Bloomsbury.
- Halliday, Michael Alexander Kirkwood. 1978. *Language as Social Semiotic*. London: Arnold.
- Halliday, Michael Alexander Kirkwood, and Christian Matthiessen. 2004. *An Introduction to Functional Grammar*. London: Routledge.
- Hay, Jennifer. 2000. “Functions of humor in the conversations of men and women.” *Journal of Pragmatics* 32: 709–742. [https://doi.org/10.1016/S0378-2166\(99\)00069-7](https://doi.org/10.1016/S0378-2166(99)00069-7)
- Hay, Jennifer. 2001. “The pragmatics of humor support.” *Humor* 14(1): 55–82. <https://doi.org/10.1515/humr.14.1.55>

- Holmes, Janet. 2005. "Politeness, humor and gender in the workplace: negotiating norms and identifying contestation." *Journal of Politeness Research* 1(1): 121–149.  
<https://doi.org/10.1515/jplr.2005.1.1.121>
- Holmes, Janet. 2006. "Sharing a laugh: pragmatic aspects of humour and gender in the workplace." *Journal of Pragmatics* 38: 26–50. <https://doi.org/10.1016/j.pragma.2005.06.007>
- Hsiao, Chi-Hua. 2015. "The verbal art of *tuciao* and face-threatening acts in *danmu* screening." *Chinese Language and Discourse* 6(2): 109–132. <https://doi.org/10.1075/cld.6.2.01hsi>
- Johnson, Daniel. 2013. "Polyphonic/pseudo-synchronic: Animated writing in the comment feed of Nicovideo." *Japanese Studies* 33(3): 297–313.  
<https://doi.org/10.1080/10371397.2013.859982>
- Kress, Gunther, and Theo Van Leeuwen. 2001. *Multimodal Discourses: The Modes and Media of Contemporary Communication*. London: Arnold.
- Kress, Gunther, and Theo Van Leeuwen. 2002. "Colour as a semiotic mode: Notes for a grammar of colour." *Visual Communication* 1(3): 343–368.  
<https://doi.org/10.1177/147035720200100306>
- Kress, Gunther, and Theo Van Leeuwen. 2006. *Reading Images: The Grammar of Visual Design* (2nd edn.). London: Routledge. <https://doi.org/10.4324/9780203619728>
- Nakajima, Seio. 2019. "The sociability of Millennials in cyberspace: A comparative analysis of barrage subtitling in Nico Nico Douga and Bilibili." In *China's Youth Cultures and Collective Spaces: Creativity, Sociality, Identity and Resistance*, ed. by Vanessa Frangville, and Gwennaël Gaffric, 98–115, Abingdon: Routledge.  
<https://doi.org/10.4324/9780429056093-6>
- Norrick, Neal R. 1993. *Conversational Joking: Humor in Everyday Talk*. Bloomington: Indiana University Press.
- Norrick, Neal R., and Delia Chiaro (eds.). 2010. *Humor in Interaction*. Amsterdam: John Benjamins.
- Nozawa, Shunsuke. 2012. "The gross face and virtual fame: Semiotic mediation in Japanese virtual communication." *First Monday* 17(3): 1–23. <https://doi.org/10.5210/fm.v17i3.3535>
- Raskin, Victor. 1985. *Semantic Mechanisms of Humor*. Dordrecht: Reidel.
- Ritchie, Graeme. 2004. *The Linguistic Analysis of jokes*. Abingdon: Routledge.  
<https://doi.org/10.4324/9780203406953>
- Sindoni, Maria Grazia. 2013. *Spoken and Written Discourse in Online Interactions: A Multimodal Approach*. New York: Routledge.
- Sperber, Dan, and Deirdre Wilson. 1995. *Relevance: Communication and Cognition* (2nd edn.). Oxford: Blackwell.
- Steinberg, Marc. 2017. "Converging contents and platforms: Niconico video and Japan's Media Mix Ecology." In *Asian Video Cultures: in the Penumbra of the Global*, ed. by Bhaskar Sarkar, and Joshua Neves, 91–112. Durham: Duke University Press.
- Suls, Jerry M. 1983. "Cognitive processes in humor appreciation." In *Handbook of Humor Research Vol. 1: Basic Issues*, ed. by Paul E. McGhee, and Jeffrey H. Goldstein, 39–57. New York: Springer-Verlag. [https://doi.org/10.1007/978-1-4612-5572-7\\_3](https://doi.org/10.1007/978-1-4612-5572-7_3)
- Teng, Man. 2018. "Colors in Dan Mu: Emerging meaning-making norms and community online." Paper presented at the 9th International Conference on Multimodality: Multimodality, University of Southern Denmark, 15–17 August 2018.
- Tsakona, Villy. 2009. "Language and image interaction in cartoons: Towards a multimodal theory of humor." *Journal of Pragmatics* 41: 1171–1188.  
<https://doi.org/10.1016/j.pragma.2008.12.003>

- Van Leeuwen, Theo. 2011. *The Language of Colour: An Introduction*. London: Routledge.
- Xing, Janet Zhiquan. 2009. “Semantics and pragmatics of color terms in Chinese.” In *Studies of Chinese Linguistics: Functional Approaches*, ed. by Janet Zhiquan Xing, 87–102. Hong Kong: Hong Kong University Press.
- Xu, Yizhou. 2016. “The postmodern aesthetic of Chinese online comment cultures.” *Communication and the Public* 1(4): 436–451. <https://doi.org/10.1177/2057047316677839>
- Yang, Yuhong. 2020. “The *danmaku* interface on Bilibili and the recontextualised translation practice: a semiotic technology perspective.” *Social Semiotics* 30(2): 254–273. <https://doi.org/10.1080/10350330.2019.1630962>
- Yang, Yuhong. 2021. “*Danmaku* subtitling: An exploratory study of a new grassroots translation practice on Chinese video-sharing websites.” *Translation Studies* 14(1): 1–17. <https://doi.org/10.1080/14781700.2019.1664318>
- Yin, Yiyi, and Anthony Fung. 2017. “Youth online cultural participation and Bilibili: An alternative form of democracy in China?” In *Digital Media Integration for Participatory Democracy*, ed. by Rocci Luppichini, and Rachel Baarda, 130–154. Hershey, PA: IGI Global. <https://doi.org/10.4018/978-1-5225-2463-2.ch007>
- Yus, Francisco. 1997. “La teoría de la relevancia y la estrategia humorística de la incongruencia-resolución [Relevance theory and the humorous strategy of incongruence-resolution].” *Pragmalingüística* 3–4: 497–508.
- Yus, Francisco. 2016. *Humour and Relevance*. Amsterdam: John Benjamins. <https://doi.org/10.1075/thr.4>
- Yus, Francisco. 2019. “Multimodality in memes: A cyberpragmatic approach.” In *Analyzing Digital Discourse: New Insights and Future Directions*, ed. by Patricia Bou-Franch, and Pilar Garcés-Conejos Blitvich, 105–131. London: Palgrave Macmillan. <https://doi.org/10.1007/978-3-319-92663-6>
- Zappavigna, Michele. 2018. *Searchable Talk: Hashtags and Social Media Metadiscourse*. London: Bloomsbury.
- Zappavigna, Michele, and Sumin Zhao. 2017. “Selfies in ‘mommyblogging’: An emerging visual genre.” *Discourse, Context & Media* 20: 239–247. <https://doi.org/10.1016/j.dcm.2017.05.005>
- Zhang, Leticia Tian, and Daniel Cassany. 2019a. “Is it always so fast?: Chinese perceptions of Spanish through *danmu* video comments.” *Spanish in Context* 16(2): 217–242. <https://doi.org/10.1075/sic.00035.zha>
- Zhang, Leticia Tian, and Daniel Cassany. 2019b. “El fenómeno «danmu» y la participación mediática: Comprensión intercultural y aprendizaje de lenguas a través de «El Ministerio del Tiempo» [The ‘danmu’ phenomenon and media participation: Intercultural understanding and language learning through *The Ministry of Time*].” *Comunicar* 58: 19–29. <https://doi.org/10.3916/C58-2019-02>
- Zhang, Yi. 2017. “Heteroglossic Chinese online literacy practices on micro-blogging and video-sharing sites.” PhD dissertation, University of South Florida.
- Zhao, Sumin, Emilia Djonov, and Theo van Leeuwen. 2014. “Semiotic technology and practice: A multimodal social semiotic approach to PowerPoint.” *Text & Talk* 34(3): 349–375. <https://doi.org/10.1515/text-2014-0005>
- Zhao, Sumin, and Michele Zappavigna. 2018. “The interplay of (semiotic) technologies and genre: The case of the selfie.” *Social Semiotics* 28(5): 665–682. <https://doi.org/10.1080/10350330.2018.1504724>

Zheng, Xiqing. 2016. "Borderless fandom and contemporary popular cultural scene in Chinese cyberspace." PhD thesis. University of Washington. <https://digital.lib.washington.edu/researchworks/handle/1773/36544>

## Address for correspondence

Leticia Tian Zhang  
Universitat Pompeu Fabra  
Roc Boronat, 138, Office 53.212  
08018 Barcelona  
Spain  
[leticiatian.zhang@upf.edu](mailto:leticiatian.zhang@upf.edu)

## Biographical notes

**Leticia Tian Zhang** is a PhD candidate in Translation and Language Sciences at Universitat Pompeu Fabra (Barcelona, Spain). Her thesis investigates *danmu* and its potential for language and intercultural understanding. Her research interests include computer-mediated discourse analysis, amateur translation (fansubbing and fandubbing), and multimodality. She has published in journals like *Comunicar*, *Spanish in Context*, and *Language Learning & Technology*. For more information, please visit <https://leticiatianzhang.wordpress.com>.

**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, also in English. Website: [https://portal.upf.edu/web/daniel\\_cassany/](https://portal.upf.edu/web/daniel_cassany/).

## Publication history

Published online: 16 October 2019

Corrected: 28 September 2021

References and their in-text citations have been updated since the original Online-First version of this article was published on 16 October 2019.