

(In press, 2024). In C. Vasquez & J. Chovanec (Eds.), *Experiencing digital discourses: Multimodality, engagement, activism*. Palgrave Macmillan.

Multimodality in Digital Discourse: Exploring Image-Text Interrelations in WeChat Sticker Use

Jing Ge-Stadnyk, University of California, Berkeley, U.S.A.

Susan C. Herring, Indiana University, Bloomington, U.S.A.

Lusha Sa, South-Central University for Nationalities, Wuhan, China

Abstract

This study examines stickers that combine images and text as a lens for exploring multimodality in digital discourse. It establishes a classification scheme of image-text relations that reflects the ways in which images and text interact, as well as the pragmatic functions served by image and text in relation to one another. Using computer-mediated discourse analysis, we analyzed and compared 544 standard (professionally designed) stickers and custom (user-created) stickers used in Chinese WeChat conversations. The findings show that multimodal custom sticker use predominates on WeChat; that the most common image-text relation is ‘image subordinated to text’ in both sticker groups; and that text most often conveys the main proposition of the sticker, with the image supporting the text in various ways, especially by adding cuteness and as illustration. The findings highlight user creativity and a preference for multimodal discourse on Chinese social media, while also suggesting potential future application areas.

1. Introduction

The study of multimodal digital discourse—specifically messages that combine text and images—has become a critical concern in disciplines such as linguistics, communication, and media studies (Bateman, 2014; Hardy-Vallée, 2016; Marsh & White, 2003). As Marsh and White (2003) argue, understanding the use of these two modes together is crucial for document producers to convey and process information effectively. As an online communication strategy, combining text and image allows internet users to convey difficult and complex ideas, initiate playful interactions, and reach out to and connect with other users. However, the interplay between text and images in digital discourse presents a number of analytical challenges (Marsh & White, 2003; Yus, 2019). First, image and text co-construction has become increasingly complex, while at the same time, increasingly flexible due to the availability of easily accessible editing tools in computer-mediated communication (CMC) (Bateman, 2014). Second, the production and reception of multimodal messages are shaped by culture-related choices made by their constructors, as well as by the semiotic resources available in any given situation (Royce, 2007).

Royce’s (2007) proposition is illustrated well by a particular form of graphical icons (or graphicons) that proliferates on Chinese social media, namely, stickers, or “character-driven illustrations or animations to which text is sometimes attached” (Konrad, Herring, & Choi, 2020, p. 217). Other

graphicons in digital discourse include emoticons (representations of facial expressions derived from a string of keyboard characters), emoji (small, colorful digital icons used to convey ideas, emotions, and action) (de Seta, 2018), GIFs (dynamic looping images drawn from popular media; Miltner & Highfield, 2017), and image macros (images with text overlaid at the top and/or the bottom). Image macros are a popular subtype of image memes (humorous images “intentionally created and transferred with an expectation for them to spread”; Yus, 2021). A memetic image macro consists of text superimposed on an image; its creators take one modality (text or image) and change components of it to express their own ideas, while maintaining a consistent resemblance to the memetic group (Nissenbaum & Shifman, 2018).

The complexity and particularities of Chinese stickers lend themselves well to multimodality analysis. On platforms such as Facebook, stickers are discrete, professionally designed cartoonish images that are larger than emojis but smaller than GIFs or image macros, and they are typically available in platform-arranged sets (Konrad et al., 2020). Chinese internet users also have access to stickers created by professional designers, but in addition, users can easily create and customize their own unique stickers that combine both text and image (de Seta, 2018; Tang et al., 2021). As a result, Chinese stickers vary in size and form; they can resemble either Western stickers or larger image macros. We use the term ‘standard stickers’ and ‘custom stickers’ to distinguish these two types. Standard stickers are offered as thematic sets in the interfaces of digital applications (e.g., WhatsApp, LINE, Facebook Messenger) and are organized in tabs (de Seta, 2018). Custom stickers are created by users themselves. While custom stickers that feature text superimposed on an image resemble memetic image macros, the custom stickers in our study are not memes, in that they do not exhibit the essential characteristics of internet memes. That is, they do not spread from one person to another through *copying* and *imitation*, while keeping *a consistent resemblance* to the memetic group (Nissenbaum & Shifman, 2018).

Custom stickers afford a surprising degree of freedom and adaptability on Chinese social media platforms; they encourage user creativity and democratize digital discourse (Zhou, Hentschel & Kumar, 2017). Democratization in this context refers to an increase in access to making and /or consuming media products or contents by ordinary social media users (Hesmondhalgh, 2019). For example, WeChat enables users to create custom stickers by using images from other sources and then editing them with simple editing tools (de Seta, 2018). Once these stickers are shared in WeChat conversations, other users can collect and save them in their local sticker gallery. The content and design of custom stickers is entirely under the control of the individual users, rather than the WeChat platform.

Although the phenomenon of multimodality has intrigued scholars for decades, it is becoming increasingly important nowadays due to the pervasiveness of multimodal discourse in CMC (Bateman, 2014; Vásquez, 2022). In this context, text-image relationships can be understood as ‘meaning multiplication’— in other words, the meaning derived from the combination of two

modes is more significant than meaning conveyed by a single mode (Lemke, 1998). For instance, pictures not only capture readers' attention, but they also add nuances to the accompanying text, while added text can clarify information conveyed by an image. Nevertheless, previous research on multimodal messages abounds with classification schemes proposing one-way relationships (that is, the functions of images in subordination to text or vice versa) (Bateman, 2014), whereas the interrelations of text and image (that is, the functions of images in relation to text *and* vice versa) remain understudied. Furthermore, Jewitt (2014) points out that the co-occurrence of text and images fulfills individuals' communicative needs. Yet much multimodal research has focused on materials developed by authorities and professionals, such as comic books and newspapers (e.g., Jewitt, 2016; Taboada & Habel, 2013; Wartenberg, 2012). Such a narrow focus is too limited to capture the complex and dynamic uses of multimodal discourse initiated by individuals, such as that produced by Chinese WeChat users, who tend to employ stickers that combine text and various forms of visuals to perform their identities and manage impressions (Wang et al., 2019).

To address these gaps, in this study we examine the use of multimodal custom and standard stickers in Chinese WeChat conversations, with a specific focus on the interrelations of images and text in these two sticker groups. Drawing on the existing literature (Konrad et al., 2020; Stökl, 2004), we define multimodal stickers as relatively complex, multi-layered, and intertextual combinations of illustrations (including animations of characters) and text. We understand instances of sticker use as "multimodal microevents in which all the signs present combine to determine its communicative intent" (van Leeuwen, 2004, p. 8).

This study expands the scope of previous investigations in several directions. First, it broadens the literature on internet stickers by adding a multimodality perspective and by shedding light on the democratized content creation facilitated by custom stickers. Second, it seeks to identify interactive two-way relations between text and image, that is, the functions of text in relation to image and vice versa. In so doing, it responds to the call made by several prominent language and communication researchers, who assert that scholars of digital discourse need to engage with the production of digital texts (Vásquez, 2022) and issues of multimodality (Page, 2022). Finally, our study addresses calls from digital media scholars to investigate graphicon use in languages other than English (e.g., Wang et al., 2019), and particularly, sticker use in non-Western contexts (Konrad et al., 2020).

We open by discussing and synthesizing two streams of research: sticker use in CMC and text-image relations in multimodal discourse. We then describe the main methods used to answer our research questions, which are presented at the end of the literature review: computer-mediated discourse analysis of a corpus of 544 multimodal stickers, comprising 56 standard and 488 custom instances. The analysis establishes and applies a classification scheme of image-text relations that reflects the ways in which images and text interact, as well as the functions served by image and text in relation to one another. Specifically, it focuses on interrelations of image and text on two

levels: one-way (asymmetrical) versus two-way (symmetrical) relations, and pragmatic functions of text in relation to image and vice versa. Overall, our analysis reveals the enormous popularity of custom (as compared to standard) stickers and the essential role of text (as compared to images) in conveying propositions, as well as a variety of functions conveyed by images. Among these functions, cuteness and humor are especially characteristic of Chinese digital culture. Finally, we present three overarching propositions in relation to user creativity and preferences in multimodal discourse, and point out the potential future applications of our two-tiered classification scheme.

2. Relevant Literature

2.1 Stickers as expressive repertoires in CMC

The production and reception of information expressed by multimodal stickers requires users to interrelate image and text both semantically and pragmatically (Stökl, 2004). Previous studies (e.g., Kato, Kato, & Ozawa, 2018; Konrad et al., 2020; Lim, 2015) have focused on standard stickers, and standard stickers such as those available on platforms like Facebook typically incorporate little text, if any, beyond formulaic expressions such as “good morning” and “happy birthday.” In contrast, it is not unusual for both standard and custom WeChat stickers to include longer text expressing more varied contents.

Most sticker studies have been anchored in a communication and social perspective, insofar as stickers fulfill a variety of communicative functions in digital space. For instance, in an early study of LINE (a popular mobile message app in Japan and South Korea), Lim (2015) claimed that the strategic and dynamic use of stickers allows users to attain communicative fluidity when expressing emotions, opinions, and intentions. Lim’s observations were supported by later studies that identified sticker functions such as rich and versatile expression, conveying nuances of meaning, changing the topic and flow of conversations (Kato et al., 2018), enhancing interpersonal intimacy (Wang, 2016), and performing user behaviors and actions (Zhou et al., 2017). A more recent study expanded the scope of sticker usage themes by adding the way stickers convey something specific, deliver humor and playfulness, and express personality (Konrad et al., 2020). In a comparison of stickers and emoji on Facebook Messenger, Konrad et al. (2020) observe that: “the elaborate and character-driven nature of stickers portrays body language and facial expressions more explicitly than emoji—they are demonstrative, rather than symbolic—and the rapid rate at which stickers are being generated offers users a great variety of ways to express their emotions, moods, and actions” (p. 222). While the above-mentioned studies paint an overall positive picture of sticker use, Cha et al. (2018) and Tang et al. (2021) found that image-only sticker use often results in ambiguity and misinterpretation. However, it is possible that a combination of image and text could enhance communication clarity. For example, accompanying text could help guide the receiver of a sticker to interpret the image as the sender intended.

A smaller strand of research has shed light on the attributes common to stickers on Chinese social media, many of which align with the above-mentioned usages. In examining memetic sticker competitions, where two or more users employ only stickers to show off the capacity of their sticker gallery as a form of competitive and playful interaction in group chat, Ge (2020) identified several features of stickers, such as expressing humor, representing mass media characters, and adding cuteness. The mass media characters category illustrates the role played by popular culture as discussed in Jiang and Vásquez (2020), in that the images come from mass media sources such as movies, television, and music videos. Moreover, cuteness is a salient motif in studies of Chinese stickers. Marcus et al. (2017) proposed a typology of cuteness in Chinese stickers: Children and animal cuteness fits with a traditional view of adorable features, whereas *Kawaii* style and *Moe* style are derived from Japanese culture. *Kawaii* style deploys structural features such as a round face, big head, and fat body in a visually exaggerated manner, and *Moe* style depicts sweet, vulnerable, innocent, and somewhat naïve girls and boys. Finally, *Baizou* (Rage) comic style or ugly cuteness underscores the peculiarity of Chinese digital culture (Jiang & Vásquez, 2020): *Baizou* stickers depict “a stick figure-style character” featuring a “rage face” to express rage or other emotions (Ma, 2016, p.18). Humor is sometimes embedded in *Baizou* (Rage) comic stickers, because of a parody effect generated by the combination of a vulgar, wacky appearance with anarchic wit (Ma, 2016).

Although the necessity to understand the interplay between images and text has often been acknowledged (e.g., Bateman, 2014; Stökl, 2004; Yus, 2019), studies of stickers thus far have mostly focused on stickers as images and have described their general functions. To date, systematic investigation of the relations between multimodal semiotic resources within stickers themselves has not been undertaken. The nearest parallels are studies of text-image relations in comics and multimodal documents, which are discussed in the following section.

2.2 *Text-picture co-construction and relations*

In the context of digital discourse, Bateman (2014) noted that the ways in which image and text work to formulate a message have become increasingly complex and flexible, and he pressed scholars to investigate how both modes interact with each other to fulfill different communication purposes. Barthes’s (1977) early typology provides scholars with a solid foundation to explore the role of image and text in conveying meaning. His categories include *anchorage* (text supports image by guiding readers’ interpretation of the image), *illustration* (image supports text by providing more details about a textual message), and *relay* (the two being equal). Scholars later applied and expanded Barthes’s classification scheme by examining professional documents in the offline domain. For instance, McCloud’s (1994) seminal study about comics dissected ‘relay’ into ‘duo-specific’ (image and text repeat the same proposition) and ‘interdependent’ (two modes co-construct a proposition that neither could express alone). Compared to Barthes’s work, ‘word-specific’ and ‘picture-specific’ in McCloud’s study are defined more explicitly: The first category means that images illustrate text without adding new meaning; the second means that images

convey a proposition and text does not add new meaning. While text and images fulfill their essential roles in these asymmetric relations, the assumption is that they are equally prominent (i.e., two modes must combine to convey an entire message meaning). This assumption is reinforced in later studies through characterizations such as ‘mutual determination’ (Noth, 1995) and ‘multimodal ensemble’ (Jewitt, 2016).

Other research examines one-way relationships (functions of image in relation to text or vice versa) in more tangentially-related discourse domains. For instance, in analyzing newspapers and scientific articles, Taboada and Habel (2013) found that figures often elaborate on the text, tables tend to provide evidence for claims in the text, and pictures provide background and motivation for the information in the text. Furthermore, Bednarek and Caple’s (2012) research on print advertising identified several key functions of visuals in relation to text, i.e., illustration, evaluation, and aesthetic. Wartenberg (2012) took another direction by identifying the potential role of text in comics books (e.g., the narration or commentary, pictorial element), rather than specific functions of text in relation to images.

In addition to this previous research that proposed classification schemes based on the analysis of professionally designed offline materials, a small number of CMC studies have classified image-text relations, mostly by analyzing professionally designed online materials. In examining various multimodal genres, including news websites and online galleries, Martinec and Salway (2005) identified both symmetric and asymmetric relations that align with previous studies. *Symmetric relations* involve interdependencies and complementarities between image and text, while *asymmetric relations* involve image subordinated to text and text subordinated to image. In a study of Instagram posts, Kruk et al. (2019) identified *divergent* relations (image and text convey different meanings), *additive* relations (one mode modifies the other), and *parallel* relations (two modes convey the same meaning). While these categories still completely reside in the well-established classification schemes proposed in earlier studies, a study of online advertisements by Zhang, Hwa, and Kovashka (2018) broadened this line of research by adding *opposite* relations (i.e., text and image alone imply opposite ideas than when used together) and *decorative* functions of image in relation to text (i.e., using aesthetic images only to beautify the text).

A few studies have gone further to propose alternate classification schemes. A taxonomy developed by Marsh and White (2003) based on an analysis of professional documents (e.g., dictionary and information design, children’s literature) identified 49 relationships, grouped into three broad categories for images: 1) functions having little relation to the text (e.g., image either decorates text or elicits emotion); 2) functions having a close relation to the text (e.g., images illustrate text or provide author/source); and 3) functions going beyond the text (e.g., images emphasize text). Moreover, Ge and Gretzel (2018) identified a variety of additional image functions in relation to the text in a study of Chinese marketer designed social media posts. Prominent categories include *elaboration* (image provides further details that relate to text),

restatement (repeating the information presented in the text), *evidence* (justification for the claim presented in the text), and *decoration* (visually appealing photos to beautify the text). However, only two studies of image-text relations that we are aware of have focused on user-generated multimodal online discourse. Ge and Herring (2018) analyzed messages combining Chinese text and emoji sequences, showing the prominence of emoji functioning as restatement and evaluation—in other words, users added emoji sequences to make subjective comments on ideas mentioned in the text. Yus’s (2019) study of multimodal Internet memes is perhaps the closest endeavor to the present study. It shows the essential role of text in the interpretation of image memes, with images tending to illustrate, amplify, elaborate, and alter the textual meaning.

In summary, previous studies of stickers have generally not focused on their multimodal properties—i.e., images and text—or how meaning is conveyed by the interplay of these two modes. Moreover, the literature on multimodality has largely outlined one-way image-text relations; potential two-way functions (that is, functions of text in relation to image and vice versa) remain largely unexplored. Against this backdrop, we address two questions: 1) How and why are customized stickers and standard stickers used in Chinese WeChat conversations? 2) What semiotic resources are exploited in each type of sticker, and how do the different modalities interrelate?

3. Methods

3.1 Data

We wanted to examine sticker use in naturally-occurring online conversations on the WeChat messaging platform—that is, asynchronous messages spontaneously exchanged between users for their own communicative purposes. Accordingly, we recruited 40 WeChat users, 20 females and 20 males (age: *mean*: 20.01; *SD*:1.79), using a combination of convenience and purposive sampling (Etikan, Musa, & Alkassim, 2016). All participants were experienced users who used both WeChat messaging and stickers daily. The third co-author invited participants to share screenshots of 10 recent WeChat conversations, with the specification that each of the screenshots had to contain at least three different stickers. Participants were asked to anonymize any usernames and profile photos. In total, 400 screenshots of WeChat conversations were collected. An initial review of the data led to the identification and removal of duplicate stickers, leaving a dataset comprising 1063 distinct stickers. Further, since the focus of this study is on multimodal stickers, text-only ($n=12$) and image-only ($n=507$) instances were removed from the data. Our final dataset for analysis consists of 544 multimodal stickers. This corpus size is considered sufficient for conducting detailed and fine-grained qualitative analysis to gain in-depth, nuanced, and contextualized insights (Vásquez, 2022).

3.2 Coding and analysis

We employed computer-mediated discourse analysis, or “language-focused content analysis” (Herring, 2004), to analyze our sample of sticker uses. We specifically focused on the pragmatic functions of image and text in relation to each other, while taking into account the discourse context surrounding each instance of sticker use. Our coding followed a grounded theory approach (Strauss & Corbin, 1997), according to which categories emerge from the data itself in a process of constant comparison and iterative coding. In line with later developments in this approach (see Kelle, 2007), we also took into account existing theoretical frameworks mentioned in the literature, including symmetric and asymmetric relations, structural features of cuteness, and linguistic characteristics of humor. In this way, we established a two-level coding scheme of image-text relations (Table 1) and pragmatic functions of image and text in relation to each other (Table 2), drawing on previous literature and allowing new categories to emerge from our corpus.

The multiple-phase coding process was led by the first author with the assistance of the two co-authors and two Chinese undergraduate students. The first phase aimed at identifying standard and custom stickers, stickers expressing humor, and stickers with cuteness. The students were first instructed to classify the stickers as standard or custom. It can be difficult for people who are not heavy WeChat sticker users to distinguish these two types of stickers; however, young WeChat stickers who are obsessed with sticker use (like our coders) can easily recognized the difference. The students were also tasked with determining the presence/absence of humor elements based on linguistic characteristics (e.g., plays on language to create non-literal meanings) (Berger, 1993; Weaver, 2010). The presence/absence of cute images was identified based on the typology of cuteness proposed by Gn (2018), focusing on appealing and endearing appearance. This includes *Kawaii* style (i.e., cute features such as round face, big head, and fat body), *Moe* style (i.e., sweet, vulnerable, innocent, somewhat naïve girls), and *Baozou* (or rage) comic style (i.e., a vulgar, wacky appearance combined with anarchic wit to achieve a parody effect) (Ma, 2016; Marcus et al., 2017). After one week of training, during which the first author and the students used 50 non-sample stickers to practice coding, the two students independently analyzed the entire data sample. Cases where agreement could not be reached were discussed with the first author.

In the second phase, we coded the sticker uses for image-text relations and pragmatic functions of image and text. First, we selected the first screenshot shared by each of the participants, resulting in 89 sticker uses. Next, the third author, a native Chinese speaker who is fluent in English, translated the text both embedded in and surrounding these stickers into English. The first two authors then analyzed the 89 instances together. Ambiguous cases (n=18) were resolved by discussing with the third author and the two students. The first author then conducted the third phase by coding the entire dataset, and the third author reviewed the coding results to ensure coding consistency and clarity. Ambiguous cases that caused disagreement (n=36) were resolved through discussion between the first and second author.

Our codebooks are presented in Tables 1 and 2.

Table 1: Types of image-text relations

Category	Definition
Image subordinated to text	The text conveys a complete proposition. The image modifies the text, without adding any new propositional meaning.
Text subordinated to image	The image conveys a complete proposition. The text modifies the image, without adding any new propositional meaning.
Image and text interdependent	A complete proposition is expressed by the co-occurrence of text and image. The text conveys part of the proposition, and the image conveys part of the proposition.
Image and text independent	The image and the text convey the same proposition in a redundant manner. A sticker receiver can get the meaning from either the text or the image.

Table 2: Pragmatic functions of image and text

Category	Definition
Background	The image has little relation to the text. A visually appealing photo or image serves as a decorative backdrop for the text. For example, the text says ‘good morning,’ and the accompanying image shows sunflowers.
Cuteness	The image presents one or more of the following structural features: 1) chubby children or animals; 2) <i>Kawaii</i> style (exaggerated features such as a round face, big head, and fat body); 3) <i>Moe</i> style (sweet, vulnerable, innocent, and somewhat naïve girls and boys); 4) <i>Baozou</i> (Rage) style (a stick figure-style character featuring a “rage face” to express rage or other emotions, e.g., ‘girl with a mushroom-shaped hairstyle’ and ‘panda man’).
Explanation	Information conveyed by the image provides a reason or justification for a proposition expressed by text.
Humor	Content described in the image or text has a non-literal meaning intended to be humorous. This includes wordplay, ridicule, irony, absurdity (nonsense or a situation that goes against all logical rules), caricature (a picture, description, or imitation of a person or thing in which certain striking characteristics are exaggerated to create a comic or grotesque effect), exaggeration, and comparison (a sudden unexpected visual change).
Illustration	The image adds an example or instance to the proposition conveyed by the text, making it clearer and more comprehensible. Content represented in the images closely relates to the text’s meaning but does not add any substantially new meaning. For example, the text says ‘(I’m) eating watermelon,’ and the image shows a boy who is eating a slice of watermelon.

Intensification	People or things described in the image strengthen the emotion or sentiment expressed by the text. For example, the text says, ‘don’t bother me,’ and the image shows a person with eyes wide open, indicating seriousness.
Restatement	The textual component, which is smaller and less salient than the pictorial component, redundantly describes the information conveyed by the image, and it does not add any substantial new meaning.
Source	The image depicts a mass media celebrity, popular online figure, cartoon character, etc. as the source of its message.
What character is saying	The text conveys what the character in the image is ostensibly saying and/or what sounds they are making, including laughter (e.g., ‘hahaha’) and expressions of contentment (e.g., ‘mmm’) and doubt (e.g., en?)

4. Findings

Overall, our initial dataset of 1063 stickers revealed a strong user preference for custom stickers (n=952) over standard stickers (n=111) in the Chinese WeChat conversations. This preference aligns with previous research showing the popularity of user-created stickers on Chinese social media (e.g., Ge, 2020; Ma, 2016; Wang et al., 2019; Zhou et al., 2017). Our data also show a trend of sticker uses that combine both text and image (n=544), as opposed to stickers that contain only one mode (n=519), that is, text-only (n=12) or image-only (n=507) stickers. Moreover, the final multimodal dataset of 544 stickers contains many more custom stickers (n=488) than standard stickers (n=56).

4.1 *Types of image-text relations*

The analysis of all the data in the final multimodal dataset uncovered both asymmetric and symmetric image-text relations, as identified in previous studies (e.g., Barthes, 1977; Martinec & Salway, 2005; McCloud, 1994). As Table 3 shows, the category of image subordinated to text accounts for the largest percentage by far in both standard and custom multimodal stickers, followed by image and text interdependent. WeChat users often used text to convey the essential point of the sticker or to express emotions and opinions, while the accompanying images modified the text in different ways (e.g., as illustration; by adding cuteness). In the cases where image and text are interdependent, the main message of the sticker was conveyed jointly by the text and its accompanying image, each mode contributing part of the overall proposition. Text subordinated to image and image and text independent relations occurred less frequently in both the standard and custom sticker groups. In the former category, images serve as the main source of information and text is supplementary, whereas the latter category illustrates the concept of equivalence in social semiotics—picture and words deliver the same content, resulting in some degree of redundancy in meaning (e.g., Yus, 2019).

Table 3: Image-text relations: Frequency distribution in standard and custom stickers

Categories	Number	Percent	Number	Percent
	Standard stickers		Custom stickers	
Image subordinated to text	47	83.9%	428	87.7%
Image and text complementary	5	8.9%	27	5.5%
Text subordinated to image	3	5.4%	21	4.3%
Image and text independent	1	1.8%	12	2.5%
Total	56	100%	488	100%

Each category in Table 3 is discussed and illustrated in the sections that follow.

4.2 Pragmatic functions of image and text for each image-text relation

4.2.1 Image subordinated to text

In the overwhelming majority (87.3%) of all the multimodal stickers in our corpus, text conveys the proposition or main idea of the sticker. In these stickers, images function in seven ways in relation to the text: adding cuteness, as illustration, as background, intensification of the text, indicating the source of the proposition (often a popular media figure), adding humor, and explanation of the text (see the definitions for these categories in Table 2.) Table 4 shows the frequency distribution of image functions when the image is subordinated to the text.

Table 4: Image subordinated to text: Frequency distribution of image functions¹

Categories	Number	Percent	Number	Percent
	Standard stickers		Custom stickers	
Cuteness	31	66%	210	49.1%
Illustration	16	34%	197	46%
Background	6	12.8%	3	0.7%
Intensification	6	12.8%	50	11.7%
Source	4	8.5%	156	36.4%
Humor	1	2.1%	115	26.9%
Explanation	1	2.1%	5	0.12%

As Table 4 shows, ‘cuteness’ in the image accounts for the largest overall percentage in both sticker types. ‘Cuteness’ is more common in the standard stickers, however, than in the custom stickers. As in the literature on cuteness in China (Marcus et al., 2017), we found that cuteness in our sticker data is associated with traditional cuteness (i.e., chubby children and animals), *Kawaii* and *Moe* style (i.e., Japanese animations and manga), and *Baozou* or rage comic characters (or

¹ The total percentage adds up to more than 100% because some categories are not mutually exclusive.

ugly cuteness peculiar to Chinese online culture, such as panda man and girl with mushroom-shaped hair). An interesting observation is the frequent occurrence of *Baozou* or rage comic characters in custom stickers, as in Figure 1. In contrast, cuteness is more commonly expressed through animals and cartoon characters in standard stickers, as in Figure 2.

In the custom sticker in Figure 1 below, the text conveys both the proposition and humor, while the image of a panda with the face of an ordinary man adds cuteness and indicates the ostensible source of the text. A free translation of the text is “If I don’t lick you two times today, you wouldn’t know what licking a dog means.” On a surface level, the text is a form of metaphor, implying that the addressee of the message is a dog. However, the cuteness expressed by the image suggests that this sticker is used to ridicule the addressee in a friendly way.

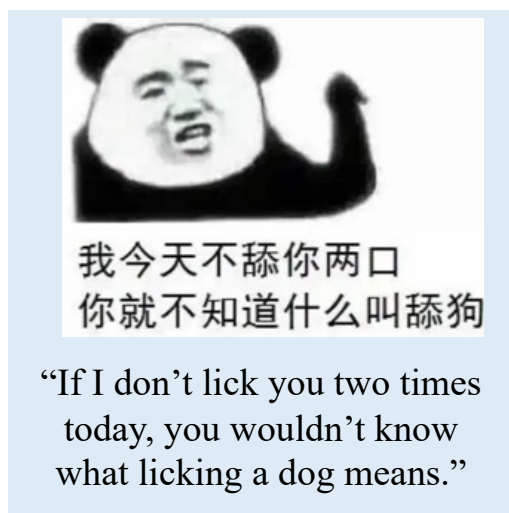


Figure 1 (Custom sticker). Image subordinated to text: text conveys proposition and humor, image adds cuteness and source.

Text expressing the proposition and simultaneously conveying humor was found in 10% of the custom stickers but was absent in standard stickers.

Figure 2 is another example in which the text conveys the proposition and the image functions to add cuteness, but in a standard sticker. A free translation of the text is “A pretty girl is speechless.” This sticker was used to mean ‘I (a pretty girl) have no idea how to respond to you.’



Figure 2 (Standard sticker). Image subordinated to text: text conveys proposition, image adds cuteness.

Illustration is the second-most prominent function in relation to text in both types of stickers. The image adds an example or instance to the proposition conveyed in the text, making it clearer and more comprehensible. This result is consistent with categories of image-text relations in the mainstream literature, such as McCloud's (1994) 'word specific' and Noth's (1995) 'pictorial exemplification.' Figure 3 is a custom sticker in which the text conveys the proposition, and the image illustrates it. A free translation of the text is "Even if this garlic can squeeze in, it's still not part of the orange." In Chinese culture, this saying means that some people do not belong to a certain social group due to their socioeconomic status. Even when they try their best to fit into that group, they still do not belong there.

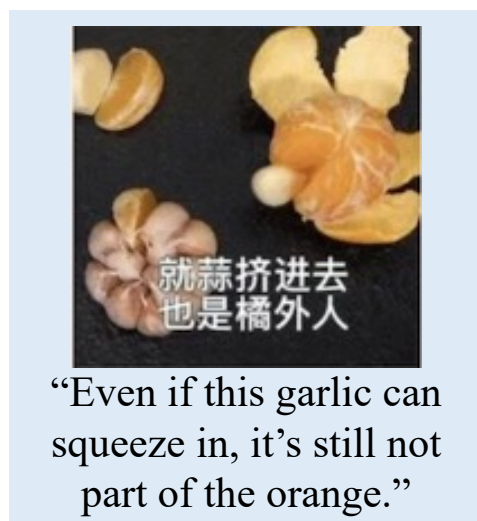


Figure 3 (Custom sticker). Image subordinated to text: text conveys proposition, image provides illustration.

Example 4 is a standard sticker in which the text conveys the proposition (“eating watermelon”) and the image both functions as illustration and adds cuteness. The image is often used by WeChat users to show that they are in a happy mood.



Figure 4 (Standard sticker). Image subordinated to text: text conveys proposition, image both functions as illustration and adds cuteness.

Images functioning as illustration of the text appear often in custom stickers (46%) and in standard stickers (34%).

While the above-mentioned functions are prominent in both data sets, the frequencies of other categories differ more across standard and custom sticker instances. In the standard sticker dataset, ‘background’ (i.e., image serving as decorative backdrop) and ‘intensification’ (i.e., image boosting an emotion or sentiment expressed by the text) are the third-most commonly used functions, followed by ‘source’ (i.e., appearance of mass media figures and Internet characters who are the ostensible source of the text). Figure 5 shows a standard sticker in which the text conveys the proposition, and the image provides background. This category aligns with ‘decoration,’ or the use of visually appealing images to beautify the text, as discussed in Ge and Gretzel (2018).



Figure 5 (Standard sticker). Image subordinated to text: text conveys proposition, image provides background.

In the standard sticker in Figure 6, a free translation of the text is “Don’t bother me.” It was used in context to indicate that the message sender was serious about the textual utterance. The serious expression on this Chinese actress’s face intensifies the meaning of the text.



Figure 6 (Standard sticker). Image subordinated to text: text conveys proposition, image functions as both intensification and source.

Indicating source in the image is relatively infrequent in the standard stickers. However, it is very popular in custom stickers. In particular, the WeChat users often create stickers that attribute textual sayings to *Baozou* (Rage) comic characters, as in Figure 1. This reflects the peculiar digital culture on Chinese social media (Jiang & Vásquez, 2020)—its ugly aesthetics can function as a vehicle to express complicated, out-of-control, and nuanced emotions (Ma, 2016). Chinese celebrities are another popular source, as in Figure 7, a custom sticker in which the text conveys the proposition, while the image functions as both source and humor. A free translation of the text is: “you are very coquettish.” The person shown in the sticker is Xiaosu Ling, an actor in the TV series *Home Temptation*.

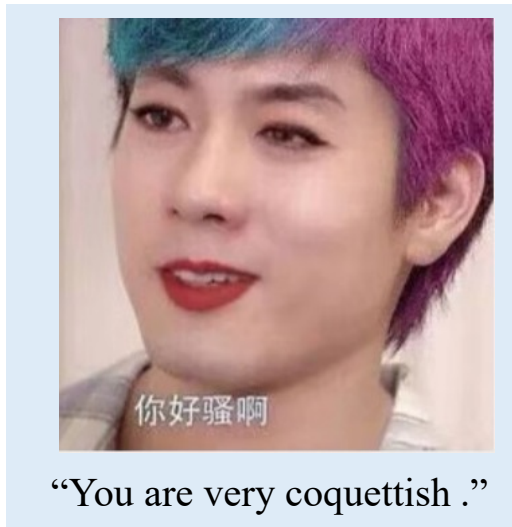


Figure 7 (Custom sticker). Image subordinated to text: text conveys proposition, image functions as both source and humor.

Another common characteristic of the image in custom stickers where image is subordinated to text is the robust use of humor, as illustrated in Figure 7. This aligns with previous sticker studies, which suggest that humor is a key feature of Chinese stickers (Ge, 2018) that users employ to amuse others (Konrad et al., 2020).

4.2 Image and text interdependent

Text and image co-construct the proposition in 5.9% of all the multimodal stickers. Each mode serves more functions in custom stickers than in standard stickers. For instance, in the custom stickers, text in a few instances also conveys humor, and the image also sometimes conveys humor and source. The frequency distribution of text and image functions when text and image are interdependent is shown in Table 5.

Table 5: Text and image interdependent: Frequency distribution of text and image functions

Mode	Categories	Number	Percent	Number	Percent
		Standard stickers		Custom stickers	
Text	Humor	N/A	N/A	2	7.4%
	What character is saying	N/A	N/A	2	7.4%
Image	Adding cuteness	5	100%	11	40.7%
	Humor	N/A	N/A	9	33.3%
	Source	N/A	N/A	7	25.9%

Figure 8 shows a custom sticker. A free translation of the text is: “This is my saving in recent years.” In this example, both modes contribute to the humor of the sticker: Here, the image resolves the meaning of the deictic *this*.



Figure 8 (Custom sticker). Text and image co-construct the proposition; both modes contribute to the humor of the sticker.

Meanwhile, quite a few of the custom images add cuteness (40.7%) (Figure 9), express humor (33.3%), and depict the source (25.9%). The text in Figure 9 says “wealthy women,” while the image of the cartoon character fishing with a heart as bait expresses the meaning “seduce.” The overall proposition is “seducing wealthy women,” presumably in hopes of marrying them or otherwise gaining access to their wealth.

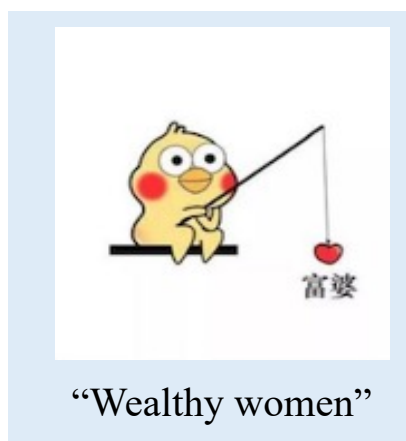


Figure 9 (Custom sticker). Text and image co-construct the proposition; image also adds cuteness.

In standard stickers, in contrast, text does not serve any other function apart from conveying part of the proposition, while images in all cases convey cuteness. This is illustrated in Figure 10; the text says “received,” and the image expresses the meaning “email.” The overall proposition is “email received.”



Figure 10 (Standard sticker). Text and image co-construct the proposition; image adds cuteness.

4.3 Text subordinated to image

In 4.4% of the multimodal stickers, it is the image that conveys the proposition or main idea; in these instances, the text is subordinated to the image. A comparative analysis of the two sticker types shows that when images convey the proposition, the accompanying text in custom stickers serves a greater variety of functions than the text in standard stickers, as summarized in Table 6.

Table 6: Text subordinated to image: Frequency distribution of text functions

Categories	Number	Percent	Number	Percent
	Standard stickers		Custom stickers	
Restatement	3	100%	12	57.1%
Humor	1	33.3%	4	19%
Intensification	N/A	N/A	5	23.8%
What character is saying	N/A	N/A	2	9.5%

Text functioning as ‘restatement’ accounts for the largest percentage in both data sets. While previous research found that images (emoji, photos/pictures) tend to repeat the textual information (Ge & Gretzel, 2018; Ge & Herring, 2018), this study shows that the WeChat users often used text to repeat the proposition conveyed by the image. This is illustrated in the next two examples. In these examples, the textual component, which is smaller and less salient than the pictorial component, redundantly describes the information conveyed by the image, and it does not add any substantial new meaning. i



Figure 11 (Standard sticker). Image conveys the proposition, and text (meaning “cheers”) restates the proposition in the image.

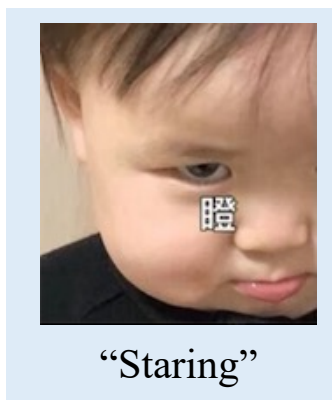


Figure 12 (Custom sticker). Image conveys the proposition (something like, “I stare at you”) and cuteness, text (meaning “stare”) restates the image.

‘Intensification’ was the second-most prominent function in custom stickers, although it was not observed in standard stickers. ‘Intensification’ differs from ‘restatement’ in that text in the latter simply describes the information conveyed in the image (Figures 11 & 12), whereas text in the former increases the strength of the proposition expressed in the image (Figure 13). In the custom sticker in Figure 13, the proposition conveyed through the image is that a person is eating sunflower seeds. The image also shows a source, a Chinese actress. The text repeats the words “crunch crunch” over and over to indicate the sound of cracking sunflower seeds, intensifying the proposition in the image.



Figure 13 (Custom sticker). Image conveys the proposition, text intensifies the proposition.

In addition to conveying propositions, some of the images serve other purposes. While only one image in a standard sticker was coded as adding humor, a higher percentage of images (61.9%) serving other functions was detected in custom stickers. These functions include ‘source’ (28.6%) (Figure 13), ‘cuteness’ (23.8%) (Figure 12), and ‘humor’ and ‘background’ (4.8% each).

4.4 *Image and text independent*

In 2.4% of all image-text relations, either text or image alone could convey the proposition. We identified only one case of this in the standard stickers, and in that case, the two modes do not have any other functions. In the custom stickers, although the text does not serve any other functions beyond conveying the proposition, some of the images depict the source (58.3%), express humor (25%), or convey cuteness (16.7%). Figure 14 is an example where the text and image convey the same proposition, laughter (the text says “hahahahaha”), and the image also depicts the source, a Chinese actress.



Figure 14 (Custom sticker). Text and image convey the same proposition, image also adds source.

5. Discussion

In this section, we present three overarching propositions about multimodal sticker use and the interplay between text and images, as revealed through our analysis.

First, multimodal custom sticker use emerged as a powerful trend in our WeChat data, evidence of a preference for bottom-up articulation on Chinese social media. Like other forms of user-generated content, custom stickers can be described as bottom-up cultural creations, moving from individuals to wider crowds. Individual users are given the freedom to choose images from other sources and then add text to these images using editing tools (de Seta, 2018). Once the individuals use these stickers to interact with their peers in WeChat conversations, their interlocutors can download and curate the stickers in the platform-afforded local sticker gallery for future use and sharing. The prevalence of custom stickers also illustrates ‘personalization,’ one of the key social affordances of CMC (Wellman et al., 2003). Compared to standard stickers, this form of expression is tailored to individual preferences and needs, advancing a more individualized way of interacting with others in the online space.

At the same time, custom stickers combining text and image were more prevalent in our corpus than stickers consisting of text alone or an image alone. This finding suggests that WeChat users appreciate creativity, flexibility, and clarity in multimodal discourse. Producing and customizing stickers combining two modes is a creative process—it requires at least a minimum level of graphic design and editing effort. This is in a sharp contrast to standard stickers—users can simply download them from the WeChat sticker gallery and then access them via their keyboard. Compared to custom stickers, we observed that both text and images embedded in standard stickers were less creative. For example, text consisted of mainly formulaic expressions such as “good morning” and “happy birthday,” and images tended to be generic and de-individualized, such as

of animals and flowers. In contrast, the text in custom stickers is user produced and hence an open set, while the images are specific and individualized, often representing concrete, recognizable individuals. Choosing to create customized multimodal stickers requires more effort, but it also allows users to express their creativity and communicate their complete meaning more effectively than using standard stickers or using stickers consisting of an image or text alone.

An obvious limitation of standard stickers is their lack of flexibility, in that users cannot edit their content. Perhaps consequently, standard stickers have been reported to lead to misunderstanding (Cha et al., 2018; Tang et al., 2021). The prominence of custom stickers combining images and text suggests that a wedding of these modes allows users to avoid ambiguity and misinterpretation caused by image-only stickers. Text may guide users to interpret images; images may illustrate text. However, the co-occurrence of both can go beyond clarifying the senders' intended meanings—it also permits them to perform pragmatic functions, such as exaggerating emotion, and achieve other purposes, such as creating intimacy.

Second, we found that text-image relations are representationally asymmetric: Text specializes in conveying propositions, whereas image functions are more versatile, modifying the text in various ways (e.g., illustration, adding cuteness, source). These findings expand on the well-established categories of 'illustration' (Barthes, 1977), 'word specific' (Yus, 2019), and 'exemplification' (Martinec & Salway, 2005) functions of images. As mentioned in the literature review section, these categories emphasize the essential role of text; images tend to merely help improve the receiver's comprehension by reinforcing information provided in text. In contrast, the variety of image functions derived from our study demonstrates that images play an important role as complements to text, consistent with the role of social media-based visuals in gaining attention, creating affiliation, and building engagement (Ge & Gretzel, 2018). By drawing on the social media engagement literature, one could argue, for example, that 'adding cuteness'—the most prominent function in both standard and custom stickers—allows users to evoke a sense of affection, prompting others to engage in WeChat conversations (Ge, 2020).

Specifically, our comparative analysis of functions of images in relation to text yielded a complex set of unique and shared dimensions of standard and custom stickers. Cuteness in the image accounts for the largest overall percentage in both sticker types. This result is consistent with Ge (2018), who found that cuteness is a common feature of Chinese stickers, while also aligning with sticker usage themes in Facebook Messenger as reported by Konrad et al. (2020) (i.e., adding cuteness to the message). Against this backdrop, one can assert that stickers lend themselves to expressing cuteness in general. Meanwhile, our coding, which was based on the cuteness literature (Marcus et al., 2017), revealed details about how Chinese users adapt their tendency to use cuteness as regards their sticker design and use, under the combined influence of Chinese traditions, cultural imports from Japan, and the emergent digital culture in China (de Seta, 2018). For instance, traditional cuteness (e.g., children, animals) is prevalent in standard stickers, while *Baozou* (Rage)

cuteness is salient in the custom sticker group. Marcus et al. (2017) characterize the *Baozou* sticker phenomenon in China as an ongoing grassroots creation movement. This ‘negative cuteness’ also reflects user needs and preferences on Chinese social media; that is, the use of ugly aesthetics is a means of self-deprecation and building mutual understanding. As example 1 illustrated, the message sender used the Panda Man as their sticker persona to humorously ridicule the addressee. Because of the vulgar and wacky appearance of the Panda Man, a common character in WeChat stickers, the addressee would likely interpret the text as a friendly joke.² Such shared understandings allow users to have freedom of choice in their digital discourse as regards what to say and how to say it.

Third, it seems that Chinese WeChat users tend to use custom stickers as a means to perform ‘networked individualism’ (Wellman et al., 2003). According to this concept, people are expected to fashion their unique identity and image, and by doing so construct their ‘selves.’ Meanwhile, they actively participate in the shaping of social networks, demonstrating an enduring human longing for communality. In our study, this is illustrated by users’ preference for adding source and expressing humor in images. Previous studies found that users often employ stickers to manage their online impression and implement self-presentation strategies (Wang et al., 2019; Zhou et al., 2017). Building on this, we suggest that the choice to depict different figures and characters (e.g., mass media celebrities, popular online figures) as the source of the stickers’ propositions allows users to express their desired online image, while creating affiliation with others who resonate with those sources. In a similar vein, users can use humor as a creative form of expression to fashion their identity and foster group cohesion with those who appreciate light-hearted and witty communication.

6. Conclusion

This article analyzed multimodality in digital discourse by classifying and coding the interrelations of image and text in WeChat sticker use. It focused particularly on comparing standard and custom stickers in order to shed light on user creativity and preferences in Chinese digital discourse, as well as to illustrate the democratized nature of content creation via stickers in a social media sphere. Our hope is that linguists, communication scholars, and digital media scholars might apply our classification scheme to advance research on multimodal digital discourse. The scheme can be used as a starting point for a grounded analysis of interrelations of image and text in various settings, such as graphicons (Herring & Dainas, 2017) used on different social media platforms and/or in different communication contexts. Such an investigation could also help identify text-image interrelations that are genre specific. A second area of inquiry that could be explored based on our taxonomy relates to users’ interpretations. For example, researchers could ask how different individuals and demographic groups respond to the range of functions served by text and images. Finally, by building a sticker-based classification scheme, our study helps define and describe the

² This statement is based on consultation with five Chinese sticker users.

domain of multimodal digital discourse. Future research could build on its findings by developing algorithms to support machine learning in this context.

Acknowledgements

We thank Prof. Charles Hirschkind for editing and offering constructive comments on the manuscript.

References

- Barthes, R. (1977). *Image, music, text*, translated by S. Heath. Hill & Wang.
- Bateman, J. (2014). *Text and image: A critical introduction to the visual/verbal divide*. Routledge.
- Bednarek, M., & Caple, H. (2012). *News discourse* (Vol. 46). A&C Black.
- Berger, A. A. (1993). *An anatomy of humor*. Transaction Publishers.
- Cha, Y., Kim, J., Park, S., Yi, M. Y., & Lee, U. (2018). Complex and ambiguous: Understanding sticker misinterpretations in instant messaging. *Proceedings of the ACM on Human-Computer Interaction*, 2(CSCW), 1-22.
- De Seta, G. (2018). Biaoqing: The circulation of emoticons, emoji, stickers, and custom images on Chinese digital media platforms. *First Monday*, 23(9). DOI: <https://doi.org/10.5210/fm.v23i9.9391>
- Etikan, I., Musa, S. A., & Alkassim, R. S. (2016). Comparison of convenience sampling and purposive sampling. *American Journal of Theoretical and Applied Statistics*, 5(1), 1-4.
- Ge, J. (2020, June). The anatomy of memetic stickers: An analysis of sticker competition on Chinese social media. In *Workshop Proceedings of the 14th International AAAI Conference on Web and Social Media: Workshop on Emoji Understanding and Applications in Social Media (Emoji2020)*. Atlanta, GA. DOI: 10.36190/2020.01
- Ge, J., & Gretzel, U. (2018). Social media-based visual strategies in tourism marketing. *International Journal of Semiotics and Visual Rhetoric (IJSVR)*, 2(2), 23-40.
- Ge, J., & Herring, S. C. (2018). Communicative functions of emoji sequences on Sina Weibo. *First Monday*, 23(11). DOI: <https://doi.org/10.5210/fm.v23i11.9413>
- Gn, J. (2018). Emoji as a 'language' of cuteness. *First Monday*, 23(9). <https://doi.org/10.5210/fm.v23i9.9396>.
- Hardy-Vallée, M. (2016). Text and image: A critical introduction to the visual/verbal divide by John A. Bateman. *Visual Studies*, 31(4), 366-368

- Herring, S. C. (2004). Computer-mediated discourse analysis: An approach to researching online behavior. In S. A. Barab, R. Kling, & J. H. Gray (Eds.), *Designing for virtual communities in the service of learning* (pp. 338–376). Cambridge University Press.
<https://doi.org/10.1017/CBO9780511805080.016>
- Herring, S. C., & Dainas, A. R. (2017). “Nice picture comment!” Graphicons in Facebook comment threads. In *Proceedings of the Fiftieth Hawaii International Conference on System Sciences*. IEEE Press. <http://hdl.handle.net/10125/41419>
- Hesmondhalgh, D. (2019). Have digital communication technologies democratized the media industries? *Media and Society*, 101-120.
- Jewitt, C. (2014). Multimodal approaches. In *Interactions, images and texts* (pp. 127-136). De Gruyter Mouton.
- Jewitt, C. (2016). Multimodal analysis. In A. Georgakopoulou & T. Spilioti (Eds.), *Handbook of language and digital communication* (pp. 69–84). Routledge.
- Jiang, Y., & Vásquez, C. (2020). Exploring local meaning-making resources: A case study of a popular Chinese internet meme (biaoqingbao). *Internet Pragmatics*, 3(2), 260-282.
- Kato, S., Kato, Y., & Ozawa, Y. (2018). Exploring potential factors in sticker use among Japanese young adults: Effects of gender and text messaging dependency. *International Journal of Virtual Communities and Social Networking (IJVCSN)*, 10(2), 1-23.
- Kelle, U. (2007). The development of categories: Different approaches in grounded theory. In A. Bryant & K. Charmaz (Eds.), *The Sage handbook of grounded theory* (pp. 191-213). Sage.
- Konrad, A., Herring, S. C., & Choi, D. (2020). Sticker and emoji use in Facebook Messenger: Implications for graphicon change. *Journal of Computer-Mediated Communication*, 25(3), 217-235.
- Kruk, J., Lubin, J., Sikka, K., Lin, X., Jurafsky, D., & A. Divakaran, A. (2019). Integrating text and image: Determining multimodal document intent in Instagram posts. In *Proceedings of EMNLP 2019*. arXiv preprint arXiv:1904.09073.
- Lemke, J. (1998). Multiplying meaning: Visual and verbal semiotics in scientific text. In J. Martin & R. Veel (Eds.), *Reading science: Critical and functional perspectives on scientific discourse* (pp. 87-113). Routledge.
- Lim, S. S. (2015). On stickers and communicative fluidity in social media. *Social Media+ Society*, 1(1), 2056305115578137.
- Ma, X. (2016). From internet memes to emoticon engineering: Insights from the Baozou comic phenomenon in China. In *Human-Computer Interaction. Novel User Experiences: 18th International Conference, HCI International 2016, Toronto, ON, Canada, July 17-22, 2016. Proceedings, Part III 18* (pp. 15-27). Springer International Publishing.

- Marcus, A., Kurosu, M., Ma, X., & Hashizume, A. (2017). *Cuteness engineering: Designing adorable products and services*. Springer.
- Martinec, R., & Salway, A. (2005). A system for image–text relations in new (and old) media. *Visual Communication*, 4(3), 337-371.
- Marsh, E. E., & White, M. D. (2003). A taxonomy of relationships between images and text. *Journal of Documentation*, 59(6), 647-672.
- McCloud, S. (1994). *Understanding comics: The invisible art*. William Morrow Paperbacks.
- Miltner, K. M., & Highfield, T. (2017). Never gonna GIF you up: Analyzing the cultural significance of the animated GIF. *Social Media+ Society*, 3(3), 2056305117725223.
- Nissenbaum, A., & Shifman, L. (2018). Meme templates as expressive repertoires in a globalizing world: A cross-linguistic study. *Journal of Computer-Mediated Communication*, 23(5), 294-310.
- Noth, W. (1995). *Handbook of semiotics*. Indiana University Press.
- Page, R. (2022). Analyzing multimodal interactions in social media contexts. In C. Vásquez (Ed.), *Research methods for digital discourse analysis* (pp. 159-176). Bloomsbury Publishing.
- Stöckl, H. (2004). In between modes: Language and image in printed media. *Perspectives on Multimodality*, 1, 9-30.
- Strauss, A., & Corbin, J. M. (1997). *Grounded theory in practice*. Sage.
- Taboada, M., & Habel, C. (2013). Rhetorical relations in multimodal documents. *Discourse Studies*, 15(1), 65–89.
- Tang, Y., Hew, K. F., Herring, S. C., & Chen, Q. (2021). (Mis) communication through stickers in online group discussions: A multiple-case study. *Discourse & Communication*, 15(5), 582-606.
- van Leeuwen, T. (2005). *Introducing social semiotics*. Routledge.
- Vásquez, C. (Ed.). (2022). *Research methods for digital discourse analysis*. Bloomsbury Publishing.
- Wang, S. S. (2016). More than words? The effect of line character sticker use on intimacy in the mobile communication environment. *Social Science Computer Review*, 34(4), 456-478.
- Wang, Y., Li, Y., Gui, X., Kou, Y., & Liu, F. (2019). Culturally-embedded visual literacy: A study of impression management via emoticon, emoji, sticker, and meme on social media in China. *Proceedings of the ACM Conference on Human-Computer Interaction*, 3(CSCW), 1-24.
- Wartenberg, T. E. (2012). Wordy pictures: Theorizing the relationship between image and text in comics. In *The art of comics: A philosophical approach* (pp. 85-104). Blackwell Publishing.

- Weaver, S. (2010). Developing a rhetorical analysis of racist humour: Examining anti-black jokes on the Internet. *Social Semiotics*, 20(5), 537–555.
- Wellman, B., Quan-Haase, A., Boase, J., Chen, W., Hampton, K., Díaz, I., & Miyata, K. (2003). The social affordances of the Internet for networked individualism. *Journal of Computer-Mediated Communication*, 8(3). DOI/10.1111/j.1083-6101.2003.tb00216.x
- Yus, F. (2019). Multimodality in memes: A cyberpragmatic approach. In Bou-Franch, P., Garcés-Conejos Blitvich, P. (eds), *Analyzing digital discourse: New insights and future directions* (pp. 105-131). Palgrave Macmillan, Cham.
- Yus, F. (2021). Incongruity-resolution humorous strategies in image macro memes. *Internet Pragmatics*, 4(1), 131-149.
- Zhang, M., Hwa, R., & Kovashka, A. (2018). Equal but not the same: Understanding the implicit relationship between persuasive images and text. *arXiv preprint arXiv:1807.08205*.
- Zhou, R., Hentschel, J., & Kumar, N. (2017, May). Goodbye text, hello emoji: Mobile communication on WeChat in China. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pp. 748-759.