

A Multimodal Metaphor Analysis of WeChat Emoticon From the Perspective of Visual Grammar

Fangfang Wang

Linfen Vocational and Technical College, Linfen, 041000, China;
SEGi University, Kota Damansara, Petaling Jaya, Selangor, 47810, Malaysia

Nalini Arumugam

SEGi University, Kota Damansara, Petaling Jaya, Selangor, 47810, Malaysia

Danfeng Qiao

Linfen Vocational and Technical College, Linfen, 041000, China

Abstract—The WeChat emoticon in China is a multimodal discourse comprising words, pictures, and other symbols. It figuratively represents communication purposes using static symbols, such as words and images and dynamic symbols, such as animation, in synchronisation with both parties' perceptive abilities. This paper analyses the construction of multimodal metaphors with WeChat emoticons as the corpus, based on the three meta-functional theories of visual grammar and the concept of cognitive linguistics, to improve communicators' metaphorical understanding of emoticons and thus promote communication.

Index Terms—WeChat, emoticon, multimodal metaphor, visual grammar

I. INTRODUCTION

Saussure (1916) noticed that linguistic symbols connect auditory pictures and conceptions rather than names and things. The sound image, for example, relates to the medium of linguistic symbols. The conventional language, as the most essential mode of communication, has altered as the internet and science and technology have grown in popularity. As paralanguages, symbols such as images, sounds, and actions engage in our daily lives as media and help generate meaning alongside linguistics and characters (Barthers, 1964). Since then, the world has entered a multimedia, multi-perspective, and multi-mode pluralistic period (Cai et al., 2023). Kress and Leeuwen (1996) think that understanding meaning in today's society expresses the same meaning as words in images or texts. For example, what is expressed in language by choosing between different parts of speech and clause structures can be expressed in visual communication by choosing between different colours or constituent structures. Hu Zhuanglin, a famous linguist in China, also thinks that the meaning of written language and various senses of vision, hearing, smell, taste, touch, and even abstraction should all participate in the construction of meaning (Hu, 2007). Therefore, for the interpretation of images, it is necessary to combine language mode and visual mode for multimodal discourse analysis to interpret and analyse the best meaning of images (Liu, 2022). Among them, as one of the most influential communication media in China's internet age, WeChat emoticons have enriched our social meaning expression. Visual-oriented WeChat emoticon is a nonverbal symbol, the most instinctive expression and emotional demand of human beings. Its entertainment, ease of use and concealment are worthy of discussing its symbolic value and implied metaphorical information (Liu, 2022).

Kress and Leeuwen (2006) put forward visual grammar, so WeChat emoticons, mainly visual metaphors, conform to this theoretical framework. Given the need for more research on visual, multimodal metaphors in WeChat emoticons, this paper makes a multimodal discourse analysis of WeChat emoticons from the perspective of cognitive linguistics and systemic functional social symbols. At the same time, it tries to deconstruct and analyse the multimodal metaphor of WeChat emoticons briefly, expecting users to understand and interpret emoticons more deeply, use and even create emojis more scientifically and reasonably, and promote communication.

II. VISUAL GRAMMAR AND MULTIMODAL METAPHOR

A. Visual Grammar

Halliday (2008), the founder of social semiotics and a contemporary British linguist, put forward the view of social semiotics in systemic functional linguistics, arguing that the meta-function of language can be divided into conceptual, interpersonal, and textual meanings. Based on Halliday's meta-functional theory, British semiotist Gunter Kress and

Dutch semiotician Van Leeuwen applied it to the visual mode level to create a visual grammar. They think that visual grammar, like language, has clauses, sentences, and text rules; each element of visual grammar has its own grammatical rules through different arrangements, thus expressing the meaning of images (Kress & Van, 1996). Any semiotic model, as human beings can experience, must be able to represent all aspects of the world. It must be able to represent objects and their relationships in the world of the representation system. The concept of social sign language lies in image, that is, image making. It regards image as a symbol, a motivational connection of meaning. In symbolic linguistics, motivation has nothing to do with image making, but it is defined as the original relationship between the symbol and the symbolised object. Kress and van Leeuwen believe that visual design, like all symbolic language patterns, has three main functions. First, the conceptual function is to reproduce the relationship between objective things and the outside world. The second is the interactive function, which makes social interaction a function of social relations. The third is the textual function; images' reproduction meaning and interactive meaning are integrated into a meaningful whole (Kress & Van, 1996).

B. Multimodal Metaphor

Modality refers to the way that human beings interact with the outside world. The interaction with a single sense is called monomodal, the interaction with two senses is called bimodal, and the interaction with three or more senses is called multimodal. Barthes (1964) is one of the earliest researchers in multimodal discourse analysis. In his paper *Rhetoric of the Image*, published in 1964, he discussed the interaction between images and language in an expressive sense. However, it did not attract enough attention at that time. It was in the 1990s that the *New London Group* first put forward the concept of multimodal, and then multimodal research developed rapidly (Liu, 2022).

Kress and Van (2001) advanced the viewpoint of multimodal construction of symbol conversion behaviour; Royce (2002) investigated the complementarity of different symbols in multimodal discourse and multimodal collaboration in second language classroom teaching; and Kress explained the theory of multiple writing abilities in detail. Halloran (2005) specialises in multimodal phenomena in mathematical discourse and the theoretical development of multimodal. Language is merely the surface embodiment of conceptual metaphor, a way of thinking and acting. Metaphor occurs daily, and we use them to guide our thinking and behaviours. The core of metaphor is understanding and experiencing the present via another object. In the late 1990s, Forceville believed that metaphor should be expressed in language and other communication modes (such as pictures, music, sounds, and gestures). These patterns at least include (1) pattern symbols, (2) handwriting, (3) Spoken symbols, (4) gestures, (5) sound, (6) music, (7) smell, (8) taste, and (9) touch. The cognitive system of metaphor is multimodal, including metaphors in the same modality or different modalities of the source and target domains (Forceville, 1996).

III. THE DEVELOPMENT OF WECHAT EMOTICONS

WeChat, China's most popular app, is an instant messaging software developed by Tencent and debuted in January 2011. According to official data, the average number of registered WeChat users from January to June 2023 was 1.319 billion, with 1.26 billion users on May 11, 2023 (Wang, 2022). Its emoticon package was first available on WeChat 5.0 ISO in 2013. The WeChat emoticon package updated the content of the current emoticons and added technological methods to create its unique emoticon.

The first stage of the expression pack is the character expression (ASCII), such as (smile) to the Japanese *kaomoji* such as (〒 _ 〒) (tears), and then to the emoji expression famous all over the world, such as:



In the second stage, the cartoon emoticon package is



represented by *Tusky*, after which WeChat specially launched emoticon stores, such as *Happy Rabbit* ; The third stage is the custom emoticon package, which is made by netizen themselves and added to the custom emoticon bar for use in chatting or added to the custom emoticon collection sent by friends during chatting, as shown in *Figure 2*. The appearance of cartoon emoticons and custom emoticons gives the public more tools to screen. Also, it enriches the original dry text communication into timely, funny, humorous, entertaining, and warm effects.

IV. MULTIMODAL INTERPRETATION AND ANALYSIS OF WECHAT EMOTICONS UNDER VISUAL GRAMMAR

WeChat emoticons are used in this study to analyse multimodal metaphors from three perspectives: reproduction, interaction, and composition. It further examines how multimodal metaphors are formed and the applications of WeChat emoticons in interpersonal communication.

A. Reproducing Meaning Analysis and Metaphor Construction

The reproduction meaning of visual grammar can be divided into narrative reproduction analysis and conceptual reproduction (Cai et al., 2023). Narrative reproduction refers to the interaction between interactive or reproduction participants through vectors. The participation of vectors is a remarkable feature of the narrative process. The concept of reproduction does not involve vectors. Narrative reproduction can be divided into action, reflection, speech, and

psychological processes (Peng, 2022). For the metaphorical interpretation of the meaning of reproduction, the author uses the perspective of social symbols in systemic functional linguistics and conceptual metaphor in cognitive linguistics to analyse the speech and psychological process in the function of reproduction. Among them, conceptual metaphor mainly discusses the relationship and interaction among source domain, target domain, and mapping items based on the metaphor model of *A IS B* (Zhang & Guo, 2013). *Figure 1* is a mixed expression of images and words between various modes. The target domain and the original domain of such expressions appear, with the image as the target domain and the text as the original domain. In *Figure 1*, the expression of a cute bear, using the conceptual metaphor *bear is a human being*, belongs to the anthropomorphic metaphor of the existence chain from the high-level source domain to the low-level target domain, giving *bear* a *human* behaviour, and his body is shaking from side to side as if waiting for the arrival of his beloved. The rose in his hand symbolises love. The smiling expression in the image is a metaphor for the inner happiness and sweetness of the bear, and the words *waiting for one* (*Deng ni lai*) are the target domain, which complements the expression (Fernández, 2021). The expression background is the bridge between the yellow moon and the orange, and the whole picture belongs to the warm colour system. Bright colours construct a passionate and energetic emotion, which gives people familiarity, narrows the distance between readers, and symbolises warm love. Using this expression, we can see that the sweetness, happiness, and warmth between the interactive people are beyond words.

B. Interactive Functional Analysis and Metaphor Construction

Any model must be able to represent the specific social relations between the participants, observers, and the objects represented. Images involve two kinds of participants: those who reproduce (people, places, and things depicted in the images) and those who are active with each other (people who communicate with each other through images, producers and observers of images) (Kress & Van, 1996), and three kinds of relationships: (1) the relationships between the participants; (2) The relationship between interactive participants and reproduction participants (the attitude of interactive participants towards reproduction participants); (3) The relationship between interactive participants (what interactive participants do through images) in which Kress & van Leeuwen proposed that the meaning of interaction includes contact, social distance, viewpoint and modality (Kress & Van, 1996).

Contact is an imaginary relationship between participants and image viewers through their eyes (Zhong & An, 2023). In the interactive sense, if the participants make eye contact with the viewers, they want to ask for specific information and enter an imaginary interaction for the *asking* images. If there is no eye contact between the participants and the viewers, it is a *providing* image, which means providing certain information or services. According to the analysis of contact elements in *Figure 2*, there is no eye contact between participants and viewers, and it *provides* expression. Dad's eyes are mapped on his daughter's body, effective interaction is formed between father and daughter, and a relatively close relationship is established. The viewer participates in the picture and feels his father's kindness.

Social distance includes close-up, middle, and long-distance (Zhong & An, 2023). Different distances can imply different relationships between the reproducer and the audience. Feng Dezheng (2011) thinks that social distance in the interactive sense can be a metaphor for intimacy or not. The distance between the lens of the image can indicate the social distance between the characters and the viewers in the picture. The expression in *Figure 3* comes from the close-up of the picture. The heart shape, lips, thumb and fist are specific body parts belonging to close-up and intimate distance. This implies the participants' intimate relationship of *love, friendship, praise and admiration*.

Attitude depends on the perspective of image shooting. In visual grammar, perspective reflects the subjective and objective attitudes of image participants, with the subjective referring to people and the objective referring to things (Zhong & An, 2023). There are three perspectives: high perspective, horizontal perspective and low perspective - the high-angle metaphor of the viewer's strength, authority and superiority. From the horizontal perspective, the observer of the head-up metaphor and the participant of reproduction are in the same position, and the positive metaphor is *empathetic and integrated*. In contrast, the lateral metaphor is detached and indifferent. From a low perspective, the metaphor viewer is disadvantaged or in a lower position and authority than the participant (Kress & Van, 1996).

Figure 2 shows a picture of a rainy day in which a father holds an umbrella for his daughter, and the daughter lets him eat a lollipop. The interactive meaning of the image formed by the long-distance, positive and horizontal perspectives reflects the father's image. Long-distance as a participant in reproduction, the relationship between father and daughter and the viewer is alienated. A positive horizontal perspective conveys an objective message to the observer and resonates with the reader. Father's love is like an umbrella, always supported on rainy days. Father's love is like an umbrella, always held up on rainy days. This picture implies that the father is tall and warm, and the daughter is weak and cute and needs care and love. Let users understand the image of a tall father with a mountain of love and a cute and petite daughter expressed in the emoji package.

C. Composition Function Analysis and Metaphor Construction

The author adopts Chinese famous scholar Feng Dezheng's spatial metaphor and multimodal meaning construction. He explains how spatial elements construct meaning in multimodal discourse from the cognitive perspective of experience (Feng, 2011). The metaphor of image composition is based on information value. Different areas of the image: left, right, top and bottom, center and edge, represent given and new, ideal and reality, and core and subsidiary

information values, respectively. Based on philosophical experience, three sub-mappings are formed between spatial position and information value (Fernández & J., 2021).

(a). *Left and Right Composition Metaphor*

According to our experience of reading habits, we usually read articles from left to right. The elements placed on the left are presented in a known form, and those on the right are taken as new elements. For what is known, it means that it is presented as what the audience already knows, and it is familiar information. As a new thing, it means that it is presented as something unknown, or something that the audience has not agreed to attract readers' attention (Kress, 2006). *Figure 4* is a left-right composition, which shows a picture of a baby grabbing his mother's face and waking up. On the left is a picture of a child as a known element, and on the right is a picture of a mother sleeping. It is new information, which implies the mother's hard work, silent efforts, and fatigue after a day with the baby. It highlights the difficulty of raising children and the selfless greatness of maternal love. The spatial metaphor makes the image of the baby's cuteness and the mother's happy life more prominent, leaving a deep and warm picture of the mother and daughter's affection in viewers' hearts.

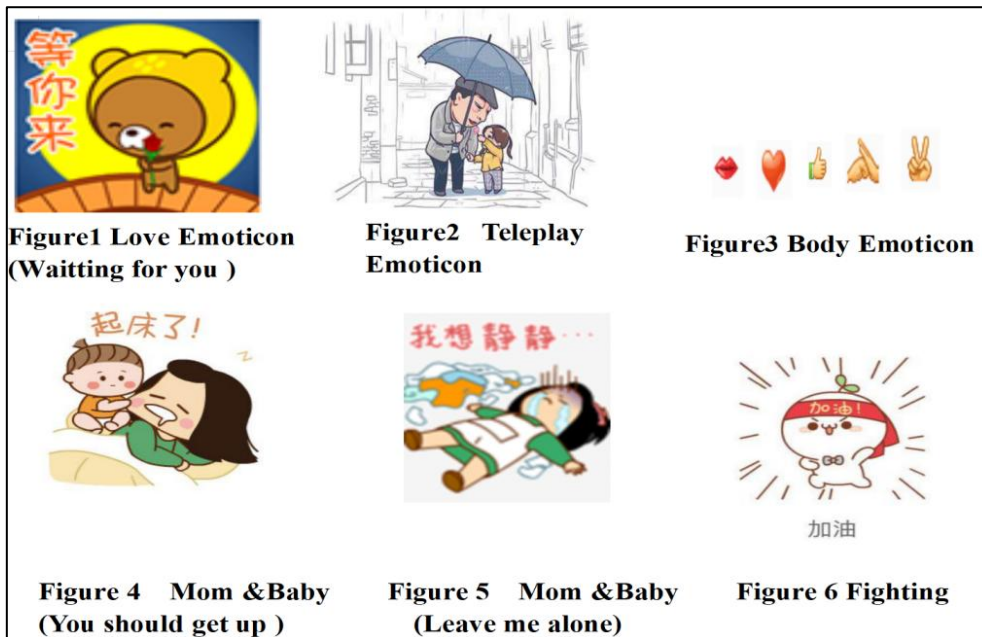


Figure 1-6. WeChat Emoticon Figures

(b). *The Metaphor of Up and Down Composition*

If, in a visual image, some constituent elements are placed in the upper part, and other different elements are placed in the lower part of the picture, then the content placed at the top will be presented as the most prominent part of the idealised or generalised nature. At the same time, the thing placed at the bottom will be presented as accurate (Feng, 2011). As far as up and down directions are concerned, we borrow the theory of conceptual metaphor. Lakoff and Johnson (1980) believe that the metaphor of the *upper* in the orientation is all positive, healthy and valuable, while the metaphor of the *lower* is mapped to negative, sad, few and worthless. Then, based on people's practical experience and social cognition, the upper orientation represents an ideal, good and desired goal in visual images. At the same time, the lower orientation conveys a practical and realistic situation, thus forming a metaphorical mapping (Cai et al., 2023). In *Figure 5*, the text *Let Me Alone* stands for *good, valuable and idealised*, symbolising the mother's ideal. The mother hopes to have her own independent space and life. Below is a pile of clothes and her mother's crying expression lying on the ground in an apron, which stands for *evil and harmful*, symbolising her inner confusion, the hard work of raising children and much housework to do. It can make users of WeChat emoticons understand the difficulties of their mothers and face their mothers and partners with grateful hearts.

(c). *Center-Edge Composition Metaphor*

Suppose visual composition uses a lot of centers, putting one element in the middle and other elements in the surroundings. In that case, we call the center element the center and the surrounding elements the edge. The central element of metaphor is the core of information and is the most important, and all other elements are subordinate information in a sense (Feng, 2011). Marginal elements are metaphorically regarded as subordinate, dependent and less essential elements. In the middle of *Figure 6*, the cartoon man raises his eyebrows, opens his mouth, moves forward, and wears a headrope to refuel. The overall image of the cartoon character is the most essential information, which means that people are ready for something, full of passion, and work hard to do something well. The edge is a line, the

second most crucial information, representing sunshine (Zhang et al., 2023). It is a metaphor for the image of the whole person working hard, full of positive energy and full of sunshine. Users of WeChat emoticons can encourage and support themselves or others and give others incentives or motivation.

V. CONCLUSION

This paper analyses multimodal metaphors in visual grammar, using WeChat emoticons as an example. It discusses the construction mechanism and meaning representation of multimodal metaphors employs conceptual metaphors to analyse reproduction meaning in visual grammar, and examines the construction of multimodal metaphors from the perspective of system function under the interactive function and composition function of WeChat emoticons. It facilitates the viewer's comprehension of the hidden meaning of WeChat emoticons. Make the propagation and growth of WeChat emoticons more beneficial.

There are many forms of WeChat emoticons, and dynamic graphics are one of them. Due to the limitation of technology and analysis methods, this paper does not analyse the multimodal metaphor of dynamic emoticons but analyses them as static pictures. Moreover, due to the complexity and diversity of WeChat emoticons, the emoticons analysed in this paper are selected based on their own knowledge and cultural background, which is subjective and needs further improvement and development.

ACKNOWLEDGEMENTS

The authors thank Linfen Vocational and Technical College and SEGi University for their help. This work was supported in part by a grant from the Linfen Vocational and Technical College's Fund Project: the 2023 New Era Foreign Language Education Research Topic Research on the Construction and Implementation Path of Cultural Confidence in Higher Vocational English Teaching under the Cross-Cultural Background in China (Project No.: 202341).

REFERENCES

- [1] Barthers, R. (1964). Rhetoric of the image. *Communication*, (4), 33-40.
- [2] Belgrimet, S., & Rababah, G. (2021). A multimodal discourse analysis of English posters in violence awareness campaigns against women. *Theory and Practice in Language Studies*, 11(2), 191-201. doi: <https://doi.org/10.17507/tpls.1102.12>
- [3] Cai S., Deng X., & Du, N. (2023). The image expression of China's national image propaganda film from the perspective of visual grammar. *News Sentinel*, 6, 4.
- [4] De Saussure, F. (1916). Nature of the linguistic sign. *Course in general linguistics*, 1, 65-70.
- [5] Forceville, C. (2009). Metonymy in Visual and Audiovisual Discourse In: Ventola E., Moya, A. (Eds.), *The World Told and the World Shown: Multisemiotic Issues*, New York: Palgrave Macmillan.
- [6] Forceville, C., & Urios-Aparisi, E. (Eds.). (2009). *Multimodal metaphor* (Vol. 11). Walter de Gruyter.
- [7] Forceville, C. (2008). Metaphor in pictures and multimodal representations. *The Cambridge handbook of metaphor and thought*, 462-482.
- [8] Feng D., & Xing, C. (2011). Spatial Metaphor and Multimodal Meaning Construction-Taking Automobile Advertising as an Example, *Foreign Languages*, (3), 24-28.
- [9] Fernández-Vázquez, J. (2021). Measuring environmental website communications in Latin American universities: Multimodal and ecolinguistic benchmarking. *International Journal of Sustainability in Higher Education*, 22(3), 599-614. doi:<https://doi.org/10.1108/IJSHE-07-2020-0248>.
- [10] Halliday, M.K. (2008). *An Introduction to Functional Grammar*. London: Edward Arnold.
- [11] Hu, Z. (2007). *Multimodality in Social Semiotics, Language Teaching and Research*, (1), 1-10.
- [12] Kress, G., & Van Leeuwen, T. (1996). *Reading Images: The Grammar of Visual Design*. London: Routledge: 5,124.
- [13] Kress, G., & Van Leeuwen, T. (2001). *Multimodal Discourse: The Modes and Media of Contemporary Communication*. London: Arnold: 20.
- [14] Kress, G., & Van Leeuwen, T. (2006). *Reading Images: The Grammar of Visual Design*, London: Routledge.
- [15] Kress, G., & Van Leeuwen, T. (2002). Colour as a semiotic mode: notes for a grammar of colour. *Visual communication*, 1(3), 343-368.
- [16] Lakoff, G., & Johnson, M. (1980). *Metaphors We Live By*. Chicago: The University of Chicago Press.
- [17] Lakoff, G., & Johnson, M. (2000). *Metaphors We Live By*. London: The University of Chicago Press.
- [18] Liu, X. (2002). Research on the graphic relationship of multimodal narrative discourse based on visual grammar. *Journal of Bengbu University*, 11, 4:7.
- [19] O'halloran, K. L. (2005). *Mathematical Discourse: Language, Symbolism and Visual Images*. London: Continuum.
- [20] Peng, Z. (2022). A multimodal discourse analysis of movie posters from the perspective of visual grammar - A case study of "hi, mom". *Theory and Practice in Language Studies*, 12(3), 605-609. doi:<https://doi.org/10.17507/tpls.1203.22>.
- [21] Royce, T. (2002). Multimodality in the TESOL classroom: Exploring visual-verbal synergy. *Tesol Quarterly*, 36(2), 191-205.
- [22] Teng, D., & Miao, X. (2018). Meaning construction of multimodal metaphors in picture books from the perspective of visual grammar. *Journal of Foreign Languages*, (5), 33-36.
- [23] Wang, G. (2022). Multimodal metaphor construction and cognitive analysis in educational cartoons. *Theory and Practice in Language Studies*, 12(3), 543-550. doi:<https://doi.org/10.17507/tpls.1203.14>.

- [24] Zhong, W., & An, X. (2023). From concrete symbols to abstract concepts—A multimodal metaphor analysis of Beijing Winter Olympics posters. *Journal of hubei university of arts and science*, 44.1(2023):8-15.
- [25] Zhao, S., Jiang, H., Tao, H., Zha, R., Zhang, K., Xu, T., & Chen, E. (2023). PEDM: A Multi-task Learning Model for Persona-aware Emoji-embedded Dialogue Generation. *ACM Transactions on Multimedia Computing, Communications and Applications*, 19(3s), 1-21.

Fangfang Wang was born in Taiyuan, Shanxi, China, in 1990 and is a lecturer at Linfen Vocational and Technical College.

She is currently studying for a Ph.D. at SEGi University in Malaysia, majoring in Education. Her research interests are language teaching and linguistics. She has published articles in some peer-reviewed Chinese journals.

Nalini Arumugam was born at Seremban, Malaysia on 17 September 1962. Ph.D. is obtained from the Faculty of Modern Languages and Communication (English Linguistics), Universiti Putra Malaysia, Serdang, Selangor, Malaysia. 2010.

She was a visiting professor at Universitas Islam Sumatera Utara, Indonesia in 2020, Universitas Islam Negeri (UIN) Imam Bonjol Padang, Indonesia in 2021 and Punjab College in 2021. Currently, she is affiliated with SEGi University, Kota Damansara, Malaysia. She has published in national and international journals, such as *Studies in English Language and Education*, *Higher Education and Oriental Studies (HEOS)*. She has also authored textbooks and reference books for tertiary. Her research interests revolve around classroom management and pedagogy, Learning styles, Humanising Education, and ESP.

Associate Professor Dr. Nalini Arumugam. She is a member of Asia TEFL and EDUCOOP, Malaysia. She is an active member of the Sathya Sai International Organisation.

Danfeng Qiao was born in Linfen, Shanxi, China, in 1989 and is a lecturer at Linfen Vocational and Technical College.