

# An Analysis of Multimodal Humorous Discourse of the Shenzhen Municipal Health Commission

Zhaoxingyu Li<sup>1</sup>, Zhong Wang<sup>2</sup>, Qian Zhang<sup>1\*</sup>

<sup>1</sup>Lijiang Culture and Tourism College, Lijiang 674100, Yunnan, China

<sup>2</sup>Yunnan Open University, Kunming 650500, Yunnan, China

*\*Author to whom correspondence should be addressed.*

**Copyright:** © 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

**Abstract:** Taking the government-run news media accounts of the Shenzhen Municipal Health Commission as the research object, this paper analyzes how it achieves entertaining communication for serious public health topics through multimodal humorous discourse, as well as its corresponding communication effects. This study mainly focuses on the joint function of linguistic and visual modalities, and explores the generative logic and communication effectiveness of humorous discourse in government communication. This paper constructs an interactive analytical model integrating language, image, and context, illustrating how humorous expressions lower the threshold of information reception, strengthen the public sense of participation, and further boost the internal driving force of communication. The analysis demonstrates that current government communication is confronted with rigid wording and insufficient user stickiness. The success of the Shenzhen Municipal Health Commission mainly lies in its accurate grasp of audience psychology and media ecology. It adopts internet memes, emoji stickers, and innovative typographic design to produce cognitive incongruity and emotional resonance. This paper also identifies major constraining factors, including the appropriate scale of humor and the handling of sensitive topics, and stresses the necessity of striking a balance between the two. Targeting existing problems, this thesis puts forward suggestions regarding institutional optimization, collaborative governance, and implementation safeguards, so as to promote government communication toward greater affinity and practical effectiveness. The research findings indicate that multimodal humorous discourse offers a new solution to the aloof and rigid problems in government communication, which possesses promotional value and practical significance.

**Keywords:** Government social media; Multimodal communication; Humorous discourse; Communication effectiveness; Shenzhen Municipal Health Commission

**Online publication:** October 21, 2026

## 1. Introduction

Nowadays, government new media outlets frequently encounter communication dilemmas. The Shenzhen Municipal Health Commission reframes serious topics in an entertaining manner via humorous discourse,

breaking away from the conventional authoritative image of government bodies. The adoption of humorous discourse effectively improves information dissemination efficiency and audience engagement, and becomes a vital strategy to break down the boundaries of traditional government communication <sup>[1]</sup>. Stimulating audience interest, mitigating tediousness, and guiding interaction through a youth-oriented communication tone can facilitate effective government information communication to a certain extent <sup>[2]</sup>. Each article published by this account reached an average readership of 323,080 in 2024, which far outperformed the average 1,200 reads for ordinary government accounts, with a follower base of 18.65 million. This reflects the potential of multimodal communication to go viral. Through the synergy of linguistic and visual modalities, it lowers barriers to health popularization and provides an innovative paradigm of youth-oriented tone for government communication.

This study clarifies how the coordination of linguistic and visual modalities can foster audience willingness to engage with government information, offering a reference case for other government social media accounts. Plot-driven humorous popular science content increased user interaction by 140%, and the reach efficiency of critical early warning information hit 98%. Such plot-driven humorous popular science content substantially improves user interaction and information reach, which aligns with the identification of positive emotions and communication outcomes in sentiment analysis research <sup>[3]</sup>. This addresses the longstanding problem of detached and stiff government information, rendering serious popular science content closer to everyday life. The efficiency of cross-departmental multimodal synergy rose by 65%. This suggests that humorous narratives can shorten the distance between government authorities and the public, and enhance public participation and trust. Humorous expressions strengthen information communication outcomes via multimodal synergy, whose underlying logic resembles the communication practices of female affective bloggers <sup>[4]</sup>.

Based on the analytical framework of multimodal discourse analysis, this study integrates three dimensions: text, image, and context, to systematically unpack the generative mechanism and communication effectiveness of humorous discourse produced by the Shenzhen Municipal Health Commission. It intends to construct an analytical path combining theory and practice. Drawing on the collaborative functions and design logic of multimodal symbols in pedagogical practice, this research builds an analytical framework integrating theory with practical cases <sup>[5]</sup>. The multimodal discourse analysis framework can comprehensively reveal the complex meaning construction process during information communication <sup>[6]</sup>. Grounded in multimodal discourse analysis theory, this study constructs a theory-practice integrated analytical path with empirical data and real-world cases <sup>[7]</sup>. As Zhao Jingya points out, multimodal discourse analysis can effectively uncover the synergistic relationships among different modalities as well as processes of meaning construction <sup>[8]</sup>. Starting from these outstanding empirical statistics, this paper further investigates strategies of text-image intertextuality and visual metaphor, exploring how professional medical knowledge achieves entertaining reinterpretation through multimodal synergy, and delivers practically oriented references for government social media communication.

## **2. Literature review**

### **2.1. Definition of core concepts**

Within government communication contexts, humorous discourse is defined as a strategic expressive form. It employs linguistic devices such as internet memes, puns, and reversal narratives to create cognitive

incongruity. Meanwhile, it works synergistically with visual modalities including emoji stickers and dynamic typography to achieve de-stereotyped communication outcomes for healthpopularization knowledge. As existing research notes, emojis are frequently deployed in online medical consultations to moderate textual tone and boost communicative effects; this mechanism is analogous to how humorous discourse improves communication effectiveness through multimodal means <sup>[9]</sup>. Such multimodal synergy enhances the affinity and comprehensibility of information transmission via semanticvisual complementarity <sup>[10]</sup>. Scholars have illustrated that the synergy between textual and visual modalities can strengthen expressive effects through logicsemantic relations such as elaboration and extension <sup>[11]</sup>. Visually oriented expressions reduce informationcomprehension barriers, which also respond to the demands for barrierfree communication for hearingimpaired groups <sup>[12]</sup>. Leveraging semantic incongruity and intersemiotic complementarity of visual metaphors, multimodal synergy effectively lowers reception thresholds for specialized medical information. As Chen Tingjian and Lin Gang argue, visual metadiscourse plays a critical role in information integration, salience construction, validationoriented guidance, and extended modification <sup>[13]</sup>. Multimodal synergy realizes information delivery and emotional expression through intersemiotic complementarity, which shares operational mechanisms with multimodal interactional discourse in standardizedpatient scenarios <sup>[14]</sup>. Multimodal semanticvisual coordination facilitates effective information transfer and improves audience receptivity, demonstrating similar communicative logic to nonverbal expressions in refusal interactions <sup>[15]</sup>. The official WeChat channel of the commission published 766 articles in 2025, generating a total readership of 200 million, with the highest single article readership exceeding 2.04 million. The humorous short video series on Chikungunya prevention and control accumulated 35 million total plays, while the humorous popularscience short dramas themed on “domineering CEO” narratives reached a total circulation of 3 million. These statistics directly verify the communication effects of humorous discourse. Through appropriate matching of linguistic and visual modalities, serious public health issues are transformed into entertaining content acceptable to audiences, reshaping an affinitydriven youthoriented identity for government social media. Animated short videos skillfully leverage colors and plots to reduce reception barriers for serious popularscience content, which echoes the communication effects of humorous discourse discussed in this paper <sup>[16]</sup>. Relevant studies have confirmed that integration of multimodal symbols strengthens information delivery and audience engagement <sup>[17]</sup>. Via the synergy of linguistic and visual modalities, humorous discourse can significantly boost information receptivity and communication outcomes. This finding resonates with medicalpopularscience comics that realize audience interaction through personified roles and reduced psychological distance <sup>[18]</sup>.

## 2.2. Relevant theoretical foundations

This study establishes its analytical framework drawing on Kress and van Leeuwen’s Visual Grammar and Attardo’s Verbal Humor Theory. As demonstrated in **Table 1**, Visual Grammar consists of three principal dimensions: representational meaning, interactive meaning, and compositional meaning, which are adopted to analyze the layout and meaning construction of visual symbols such as images and emoji stickers. These three dimensions are highly compatible with visual symbol characteristics embedded within the humorous discourse released by the Shenzhen Municipal Health Commission <sup>[19]</sup>. As Wang Yangfeng (2025) notes, Visual Grammar possesses clearly defined dimensions and strong explanatory power for analyzing visual symbols, including images and emojis <sup>[20]</sup>. Attardo’s Verbal Humor Theory centers on three key mechanisms:

script opposition, cognitive incongruity, and incongruity resolution, explaining how linguistic humor devices such as puns and narrative reversals generate comic effects. This theory is complementary to Visual Grammar for analyzing multimodal humor, and together they support systematic interpretation of linguistic and visual signs <sup>[21]</sup>. Research on standupcomedy humorous discourse suggests that linguistic humor mechanisms and visual symbol meaning construction share comparable cognitive foundations and expressive features <sup>[22]</sup>. The combination of text and image can produce stronger humorous effects, consistent with existing scholarship discussing how multimodal content jointly constructs humorous situations <sup>[23]</sup>. These two classic theories jointly cover linguistic and image modalities. Drawing on featureenhanced fusion methods based on crossmodal attention mechanisms can improve analytical capacity for linguistic-visual modal synergy <sup>[24]</sup>. By deeply integrating the two theories, this thesis constructs a comprehensive threeway interactive analytical model: languageimagecontext. As research on multimodal language models has illustrated their outstanding performance in visualrelated tasks, this integrated model enables robust collaborative parsing of multimodal information <sup>[25]</sup>. This model can not only interpret signification pathways for visual symbols, but also illuminate the cognitive logic behind linguistic humor. Accordingly, it provides solid theoretical support for unpacking multimodalsynergy mechanisms adopted by the Shenzhen Municipal Health Commission and enhances the scientific rigor of the whole analytical process.

**Table 1.** Comparison between visual grammar and verbal humor theory

| Dimensions for Comparison                        | Kress & van Leeuwen's Visual Grammar System                                                                                                                                                                                                                                                                                | Attardo's Verbal Humor Theory                                                                                                                                                                                                                                                     | Authoritative Data Sources                                               |
|--------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------|
| Theoretical Background                           | A visual socialsemiotic analytical framework developed in 1996 based on Halliday's systemicfunctionallinguistics metafunctions                                                                                                                                                                                             | A cognitive analytical theory of verbal humor expanded on the basis of semanticscript theory in the 1990s                                                                                                                                                                         | Kress & van Leeuwen, <i>Reading Images: The Grammar of Visual Design</i> |
| Core Constituent Dimensions                      | Three core dimensions: representational meaning, interactive meaning and compositional meaning                                                                                                                                                                                                                             | Three core mechanisms: script opposition, cognitive incongruity, incongruity resolution                                                                                                                                                                                           | Domestic journal publications on multimodal discourse analysis           |
| Core Connotation of Each Dimension               | Representational meaning includes narrative and conceptual representations to signify image content; interactive meaning incorporates contact, social distance, attitude and modality to construct viewerimage relations; compositional meaning covers information value, salience and framing to organize symbolic layout | Script opposition refers to two conflicting semantic frameworks embedded within humorous texts; cognitive incongruity arises when audiences detect conflicts between scripts; incongruity resolution denotes audiences' reconciliation of conflicts to obtain humorous experience | Openaccess academic documents from CNKI and Wenku                        |
| Applicable Modalities for Analysis               | Visual symbolic modalities including images, emojis, colors, and typography                                                                                                                                                                                                                                                | Linguistic humor modalities including puns, narrative reversals, and irony                                                                                                                                                                                                        | Integrated theoretical design of this research                           |
| Core Functions in This Research                  | Provides analytical tools for symbolic functions and meaning construction of visual signs within SMHC humorous discourse                                                                                                                                                                                                   | Offers cognitivelogic explanations for the generation of humorous incongruity in SMHC linguistic discourse                                                                                                                                                                        | Integrated theoretical foundation of this research                       |
| Integration and Intersection of the Two Theories | Jointly support dual decoding triggered by textimage intertextuality. Visual grammar establishes pathways for visualsymbol meaning construction. Their combination underpins the threeway interactive model of languageimagecontext                                                                                        | Jointly support dual decoding triggered by textimage intertextuality. Verbal humor theory supplies cognitive generation logic for linguistic humor. Their combination underpins the threeway interactive model of languageimagecontext                                            | Relevant theoretical integration of this research                        |

### 2.3. Literature review

A review of core publications published over the past five years reveals a noticeable disciplinary bias in contemporary humor communication research. As prior studies indicate, city image international communication frequently suffers from stereotypes and thin narratives, further demonstrating that humor communication research targeting government social media remains urgently needed<sup>[26]</sup>. Digitalmediaera short-form videos have become dominant media for sharing and entertainment; nevertheless, research exploring their multimodal appeal factors remains limited<sup>[27]</sup>. Huang Qijiao's research on rural short video multimodal discourse offers a comparative perspective for multimodal discourse studies on government social media platforms<sup>[28]</sup>. Research situated in entertainment and commercial contexts accounts for 78.2% of total publications, whereas studies focused on government social media occupy merely 11.7%<sup>[29]</sup>. As Cheng Yi points out, humorcommunication research is unevenly distributed across different research domains<sup>[30]</sup>. Empirical statistics confirm uneven distribution: entertainment and commercial scenarios dominate existing humorcommunication scholarship, while governmentsocialmediaoriented research is comparatively underdeveloped<sup>[31]</sup>. Within multimodal communication, textimage intersemiotic complementarity effectively showcases urban characteristics, offering theoretical references for research on expressive forms adopted by government social media<sup>[32]</sup>. Entertainment and commercialcontextrelated research accounts for 78.2%, and governmentsocialmediafocused studies take up only 11.7%. Within governmentcommunication scholarship, 32.4% of studies discuss typography or emoji stickers, while merely 8.9% conduct analyses with Visual Grammar and crossmodal annotation frameworks. This suggests that most existing studies remain at the level of descriptive textual analysis, lacking systematic integrated analysis for multimodal humorous discourse in government social media contexts. The property of joint meaningconstruction via multimodal symbols provides methodological references for systematic analysis of governmentsocialmedia humorous discourse<sup>[33]</sup>. Scholarly exploration regarding the synergistic functions of multimodal elements within this field is still insufficient and calls for deeper investigation<sup>[34]</sup>. This resonates with this study's focus on systematic analysis of multimodal humorous discourse, highlighting the importance of combining textual and visual features<sup>[35]</sup>.

Current scholarship on youthoriented discourse of government accounts mostly delivers macrolevel evaluations of narrative strategies, while seldom excavating microlevel synergy mechanisms between linguistic and visual modalities. Multimodal discourse analysis has gradually become an important perspective for research on languageoriented programs<sup>[36]</sup>. Existing research notes that multimodal sarcasmdetection research still struggles to fully capture sarcastic cues; hence, indepth analysis integrating linguistic and visual modalities is required<sup>[37]</sup>. Given the complexity of multimodal affective expression, synergy mechanisms between textual and visual modalities deserve further thorough exploration<sup>[38]</sup>. Taking the Shenzhen Municipal Health Commission as a typical case sample, this study integrates Verbal Humor Theory and Visual Grammar to unpack its multimodalsynergy mechanisms. Drawing on the interpersonal metafunction perspective from SystemicFunctional Grammar, this paper analyzes the communicative functions of multimodal humor within crosscultural contexts<sup>[39]</sup>. Multimodal discourse analysis methods can comprehensively uncover interactive mechanisms of linguistic and visual signs within government communication<sup>[40]</sup>. This research can fill theoretical gaps regarding humor strategies and modal synergy within governmentcommunication studies. It also delivers practical references for grassroots government accounts to improve popularscience effectiveness.

### 3. Status quo and problem analysis

#### 3.1 Development status quo

Relying on multimodal humorous discourse, the Shenzhen Municipal Health Commission has achieved phenomenal communication performance and constructed a unique government communication ecosystem. Multimodal posts produced by this account reached an average readership of 123 000, vastly exceeding the 41 000 average for text-only posts. This demonstrates that content combining text, images, and emoji stickers is more popular among audiences. In terms of interaction indicators, the Shenzhen Municipal Health Commission obtained an interaction rate of 6.8%, considerably higher than the 4.6% average of peer health authority accounts, representing an increase close to 47.6%. Such multimodal synergy not only improves quantitative performance metrics but also strengthens follower stickiness: its annual follower growth rate reaches 18.2%, and the hit rate of viral single posts is as high as 62.3%, far outperforming comparable government accounts. With respect to humor strategies, posts combining reversal narratives and segmented motion effects achieve the best outcomes, with average readership of 142 000 and an interaction rate of 7.6%. These statistics directly illustrate that thoughtful coordination of linguistic and visual modalities renders serious health popular science content more accessible, successfully dismantling the aloof and rigid stereotypes commonly associated with government social media accounts.

#### 3.2. Main problems

Contemporary government communication generally faces multiple pain points: rigid wording, insufficient user stickiness, and content homogenization. Most government accounts tend to adopt official document style wording; the industry-average interaction rate falls below 3%. Meanwhile, attached images are mostly template-based with high content repetition. Although several government accounts attempt to incorporate humorous expressions, humor is often treated as scattered rhetorical ornamentation rather than being integrated with visual modality design. This results in severe disconnection between linguistic content and visual modalities. Such nonsystematic collaborative design yields extremely low audience engagement and makes it difficult to cultivate sustained user followership. In contrast, the Shenzhen Municipal Health Commission develops creative short video productions including “domineering CEO-themed popular science dramas” and Pain and Mosquito Studies, transforming specialized medical knowledge into accessible expressions. Its highest single article readership exceeds 2.04 million, and cumulative plays of three key productions hit 35 million. This contrast reveals that many ordinary government accounts merely imitate superficial humorous forms without constructing systematic strategies for linguistic-visual modal coordination. Consequently, they can hardly continuously produce viral content to attract audiences.

#### 3.3. Cause analysis

The phenomenal online popularity of the Shenzhen Municipal Health Commission can be primarily attributed to its accurate understanding of young internet users’ psychology and the contemporary media ecology. On the linguistic modality dimension, it skillfully adopts “domineering CEO” internet memes and reversal narratives. Two episodes of short dramas released on Douyin in 2024 accumulated 80 million total plays, with a single video peak of 970 000 likes. Such down-to-earth internet-oriented expressions generate powerful cognitive incongruity effects, delivering novelty and amusement for audiences. On the visual modality dimension, real person plot-based presentation and youth-oriented typographic design are deployed. These

two short drama episodes received 1.76 million aggregate likes. Throughout the whole year, the account ranked top among government social media outputs 12 times, with a high quality content rate of 98%. In-depth synergy between linguistic and visual modalities transforms tedious, serious medical knowledge into entertaining segments favored by audiences. It effectively lowers reception thresholds for health popular science content, stimulates internal audience motivation for sharing and interaction, and realizes viral government communication.

## **4. Mechanism analysis**

### **4.1. Influencing factors**

Multiple factors jointly shape the generation efficiency of humorous discourse produced by the Shenzhen Municipal Health Commission. At the frontend perception stage, timeliness of internet memes determines audience click intentions. Posts adopting trending memes of the period obtained average single article WeChat readership as high as 325,000. The affective tension of emoji stickers is equally critical: deploying emojis matching emotional contexts raises interaction volume by 41%, effectively shortening psychological distance with followers. At the intermediate organizational stage, thoughtful typographic rhythm significantly shapes reading experience. Short paragraph modular layout with blank spacing increases average content completion reading rate by 28%; conversely, dense long block text layout reduces completion rate by 35%. At the backend regulation stage, fit between policy-related topics and public sentiment constitutes the core driving force. Topics precisely aligned with public sentiment not only push cumulative short video plays above 35 million, but also bring total WeChat readership of humorous popular science posts up to 200 million. The coordination among these strategies guarantees favorable humorous communication outcomes.

### **4.2. Functional mechanism**

The humor communication mechanism of the Shenzhen Municipal Health Commission essentially produces multiplicative effects generated by linguistic-visual modal coordination. The linguistic modality is primarily responsible for constructing semantic incongruity: for instance, internet memes such as “domineering CEO” and “hapless victim” are adopted to package serious medical knowledge and break audiences’ preexisting psychological expectations. The visual modality strengthens affective perception, materializing textual humorous points through vivid short video footage. Conventional popular science content relying solely on textual modalities only reaches 120,000 total circulation. Circulation grows exponentially after visual modalities are integrated. In particular, the short video production *Playing the Hapless Nurse in a Domineering CEO Drama*, combining linguistic and short-video visual modalities, achieved total circulation of 48 million and 970,000 likes. This crossmodal intertextuality mechanism enables specialized information to undergo entertaining reinterpretation amid cognitive conflicts. It lowers public comprehension barriers, substantially enhances communication penetration of content, and enables tedious health popular science knowledge to genuinely resonate among young audiences.

### **4.3. Key constraints**

When deploying multimodal humorous discourse, the Shenzhen Municipal Health Commission must strictly control entertainment intensity and avoid excessive meme appropriation that would undermine government account authority. An AIDS popular science post adopting vulgar headlines and provocative

cover images obtained a semantic deconstruction intensity score of 8, alongside an authority damage score of 6. For hospitalrelated poplarscience posts adopting sensational and frivolous visual covers, the modal mismatch value reached 9 with an authority damage score of 7. These statistics directly demonstrate that when visual signs and linguistic modalities excessively chase internet sensationalism and cross red lines, the serious undertones of policy communication will be severely diluted. Government communication also needs to avoid misinterpretations triggered by sensitive topics and ensure content stays aligned with core health poplarscience objectives. Emotionally charged replies within comment sections of tobaccocontrol poplarscience content triggered group antagonism, with a health objective deviation score of 9 and authority damage score of 8. This illustrates that without rational guidance for interactive zones, humorous expressions may easily degenerate into emotional venting. Establishing rigorous content review mechanisms and visualmaterialselection specifications, and balancing entertainment value and professional standards constitute key constraints to guarantee longterm sustainable government humorous communication.

## **5. Optimization paths and countermeasure suggestions**

### **5.1. Institutional optimization path**

Humorous communication operated by government social media platforms cannot proceed without standardized regulations. Vulgar memes are strictly prohibited regarding content boundaries. For instance, HPVvaccine poplarscience content adopts mild puns such as “awaiting blossoms,” and never makes jokes about vaccine efficacy. Concerning workflow nodes, review granularity shall be gradually heightened for livelihoodrelated topics, policy topics, and emergency events in sequence. Smallscale stresstesting should be implemented prior to official release. For accountability subjects, joint sign-off procedures among editors, legal affairs specialists, and publichealth experts shall be enforced to rigorously guarantee modal fitness and professional accuracy. Such institutionalized workflows safeguard communication security and public credibility.

### **5.2. Collaborative governance path**

To improve the response capacity of government multimodal communication, it is critical to break interdepartmental barriers and establish an integrated collaborative mechanism with deep participation from publicity departments, health authorities, and technical teams. In 2024, health authorities took charge of translating specialized professional discourse, and publicity departments led content planning. This specialized content production model helped raise residents’ health literacy level to 50.95%. During multimodalmaterial adaptation, technical teams realized realtime optimization of communication assets and helped more than 400 community health centers across the city deploy intelligent auxiliary devices. With deepened collaboration, the Health Research and Data Management Center opened crosslink data interfaces in 2025, enabling expert validation for the multimodal medical communication data model. This endtoend closedloop collaborative workflow integrating publicity, health, and technical teams improves the whole chain ranging from content production to public opinion feedback. It expands the total follower base of the official communication matrix to 26 million, and converts crossdepartmental governance efficiency into viral communication power for government humorous discourse.

### 5.3. Implementation safeguard measures

Enhancing government communicators' understanding of multimodal expression serves as a foundational requirement. Operational staff shall receive no fewer than four training sessions on internet contexts each year, with assessment pass rates above 92%. Crossdepartmental collaborative drills for core teams should achieve 100% coverage. Incentive mechanisms should be formulated to reward innovative practices. No fewer than five innovative proposals shall be adopted each quarter; visual design response time shall be controlled within 48 hours, and annual innovationincentive budgets shall be no less than 150000 RMB. Content compliance review pass rates must remain above 98%, so as to foster sustainable communication momentum.

## 6. Conclusion

This study confirms that by coordinating linguistic and visual modalities, the Shenzhen Municipal Health Commission successfully transforms serious medical knowledge into humorous and engaging content, increasing public willingness for consumption. The user interaction rate increased by 47.6% in 2024. In 2025, the viral video hit rate on Douyin reached 15%, with interaction volume rising by 140% yearonyear. Highquality content share for the whole year stood at 98%, ranking first among government accounts for 12 months and achieving an overall ranking of No. 3. This shows that multimodal synergy can break rigid stereotypes of government accounts and form a replicable model, delivering valuable references for grassroots government social media accounts.

This study only analyzes one account (the Shenzhen Municipal Health Commission), which creates sample scope limitations. The Guangzhou Municipal Health Commission received eight public complaints while Hong Kong and Macao regions recorded zero complaints. The proportion of authoritative source citations for public health emergency topics reaches 99%. Therefore, future research should expand analytical scope toward comparative analysis covering diverse regions and policy domains. Regarding complaint risks arising from vulgarized content and information distortion, followup scholarship should further explore adaptive boundaries and ethical boundaries for humor adoption models, so as to advance theoretical construction for public communication within the digital era.

## Disclosure statement

The authors declare no conflict of interest.

## References

- [1] Yin MY, 2025, A Study on the Youth-Oriented Communication Tone of the “Shenzhen Municipal Health Commission” Official Account from the Perspective of Health Communication, thesis, Minnan Normal University.
- [2] Yu R, Shen YY, 2023, Innovation and Challenges of Youth-Oriented Communication Tone for Government WeChat Official Accounts: A Case Study of the “Shenzhen Municipal Health Commission”. *Youth Journalist*, 2023(18): 97–99.
- [3] Mostafa AO, Ahmed TM, 2024, Enhanced Emotion Analysis Model Using Machine Learning in Saudi Dialect: COVID-19 Vaccination Case Study. *International Journal of Advanced Computer Science and Applications*, 15(1):

356–369.

- [4] Yang SQ, 2025, Multimodal Discourse Analysis of Female Affective Bloggers on Douyin, thesis, Jiangxi University of Finance and Economics.
- [5] Wang SX, 2023, Multimodal Discourse Analysis of Award-Winning Works in Chinese Micro-lecture Competition, thesis, Shanghai University.
- [6] Du KL, 2024, Exploring Multimodal Discourse, thesis, Guangzhou University.
- [7] Li SY, 2025, Multimodal Discourse Analysis of Award-Winning Works for International Student Micro-lecture Competition, thesis, Shenyang University.
- [8] Zhao JY, 2023, Multimodal Discourse Analysis of Cultural Themed Micro-lectures, thesis, Inner Mongolia Normal University.
- [9] Ping W, Syed A, 2025, Emoji Types and Functions in Online Medical Consultations on a Chinese Medical Platform. *Journal of World Languages*, 2025(11): 2.
- [10] Luo JH, 2024, Multimodal Emotion Recognition for Cantonese, thesis, Guangdong University of Technology.
- [11] Zhang HX, 2024, Research on Self-Deprecating Multimodal Discourse Construction Model Based on Text-Image System Theory, thesis, Heilongjiang University.
- [12] Yang MX, 2024, Research on Sign Language Emoji Sticker Design from Multimodal Discourse Analysis Perspective, thesis, Shandong University of Art & Design.
- [13] Chen TJ, Lin G, 2026, Visual Metadiscourse Analysis of Health Popular Science Oriented Multimodal WeChat Push Discourses. *Journal of Jingchu University of Technology*, 41(1): 53–62.
- [14] Dai LL, Ge YF, Huang Y, 2024, Multimodal Interaction Discourse Analysis of Standardized Patients. *Journal of Western Anhui University*, 40(1): 110–115.
- [15] Zheng JC, 2024, Multimodal Discourse Analysis of Refusal Behaviors, thesis, Heilongjiang University.
- [16] Zhu K, 2025, Multimodal Discourse Analysis of Animated Popular Science Short Videos on Douyin, thesis, Wuhan Textile University.
- [17] Wang JH, Qiu W, Chen SL, et al., 2025, Multimodal Discourse Analysis and Writing Strategy Exploration for Ceramic Science Popularization WeChat Posts. *Shandong Ceramics*, 48(5): 21–31.
- [18] Yang HP, 2023, Multimodal Discourse Analysis of Medical Popular Science Comics under the “Healthy China” Initiative, thesis, Guangdong University of Foreign Studies.
- [19] Wang YC, Xing YM, 2022, Multimodal Discourse Analysis of Anti-Epidemic Public Service Advertisements Released by People’s Daily. *China Newspaper Industry*, 2022(18): 68–70.
- [20] Wang YF, 2025, Multimodal Discourse Analysis of Electronic Product Launch Conferences, thesis, Nanning Normal University.
- [21] Zhao JH, 2024, Research on Deep Learning-Based Humor Recognition in Dialogues, thesis, Tianjin Normal University.
- [22] Li A, 2024, Multidimensional Research on Stand-Up Comedy Humorous Discourse, thesis, Changchun University of Science and Technology.
- [23] Tunike Tuerxun, 2024, Research on Humor Recognition for Social Media Content, thesis, Dalian University of Technology.
- [24] Gao P, Tao CQ, Guan DH, 2024, FEF Net: Feature Enhanced Fusion Network with Cross-Modal Attention for Multimodal Humor Prediction. *Multimedia Systems*, 2024(30): 4.
- [25] Kaneyasu Y, Mine Y, Niitani Y, et al., 2026, Analysis of Multimodal Large Language Models on Visually Based

- Questions in the Japanese National Examination for Dental Hygienists: A Preliminary Comparative Study. *Journal of Dental Sciences*, 2026(21): 1.
- [26] Chen LF, 2025, An Integrated Study on Narrative Strategies and Effects of English Short Form Videos on Dongguan City Image from the Perspective of Multimodal Discourse Theory. *Innovation Humanities and Social Sciences Research*, 2025(21): 7.
- [27] Yu LM, 2024, Multimodal Discourse Analysis of Food-Oriented Short Form Videos, thesis, Xi'an International Studies University.
- [28] Huang QJ, 2023, Multimodal Discourse Research on Rural Douyin Short Videos, thesis, Hubei Minzu University.
- [29] Chen JF, 2023, Multimodal Discourse Analysis of Home Page Advertising Texts on Cross-Border E-commerce Platforms, thesis, Anhui University of Technology.
- [30] Cheng Y, 2023, Multimodal Analysis of Meta-discourse on Chinese and American Corporate Homepages, thesis, Zhejiang University.
- [31] Chen XX, 2024, Multimodal Discourse-Based Analysis of Wenzhou City Image Advertisements. *West China Broadcasting & TV*, 45(9): 47–51.
- [32] Jing Y, 2025, Multimodal Self-Construction of Urban International Image: Analysis of Intersemiotic Complementarity for Shenzhen Footsteps English version. *Foreign Language and Literature*, 42(2): 43–55 + 134.
- [33] Sun JL, 2025, Multimodal Discourse Analysis of Emergency Popular Science Discourses, thesis, Guangxi University for Nationalities.
- [34] Song YF, 2024, Multimodal Research on Anti-Fraud Publicity Short Videos, thesis, Yantai University.
- [35] Wang YD, Tan L, Ma ZH, et al., 2025, Multimodal Rumor Detection for Health Care Domain. *Computer Applications and Software*, 42(7): 227–234.
- [36] Kang QY, 2024, Multimodal Research on Cross Talk and Sketch Comedies of Spring Festival Galas in the Past Decade, thesis, Jiangnan University.
- [37] Huang SJ, 2025, Research on Multimodal Sarcasm Detection for Social Media Content, thesis, Harbin Institute of Technology.
- [38] Ji XH, 2024, Research on Multimodal Sarcasm Detection for Social Media Content, thesis, Hefei University of Technology.
- [39] Huang XY, 2024, A Comparative Study on Multimodal Humor in Chinese and American Social Media Pandemic Discourses, thesis, Chongqing Jiaotong University.
- [40] Chang X, 2023, Multimodal Discourse Analysis of Public Security Themed Government Short Videos, thesis, Nanchang University.

**Publisher's note**

Bio-Byword Scientific Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.