

English Ambiguity in Multimodal Contexts: A Cognitive Communication Study Based on Internet Memes

Linping Chen^{1*}

¹ Guilin University of Technology, Guilin, China

*Correspondence: Linping Chen

Received: 21 April 2026; Accepted: 14 May 2026; Published: 15 June 2026

Abstract:

This study focuses on how semantic ambiguity drives the generation, variation, and diffusion of internet memes within multimodal contexts. From syntactic, semantic, and pragmatic perspectives, and by integrating the participatory culture of memes with visual “family resemblance” variation characteristics, it reveals the reconstruction mechanisms and communication patterns of ambiguity (isomorphic heterogeneity, heteromorphic homogeneity, heteromorphic heterogeneity). Cognitively, working memory and context resolve ambiguity through dual-process theory, while virtual environments may lead to cognitive contamination. The findings provide a basis for cognitive optimization and risk prevention in meme communication.

Keywords:

Multimodal, Semantic Ambiguity, Internet Meme, Cognitive Communication

1. Introduction

In the digital communication ecosystem, multimodal symbols—images, text, voice, and video—collectively shape online expression. Among these, internet memes have emerged as significant cultural carriers due to their participatory, mutable, and viral nature. Notably, many viral memes derive their appeal not from precision but from the deliberate creation of ambiguous meanings, transforming semantic ambiguity from a traditional “problem” into a key driver of online cultural production and dissemination.

Given the fragmentation of existing disciplinary perspectives, this study adopts an interdisciplinary approach. It integrates Zhang Xiaohong’s tripartite framework of English ambiguity (grammatical, semantic, pragmatic), Jia Xiang tong’s analysis of memes through “virtuality and materiality,” Feng Jianming’s discussion of “family resemblance” variation in visual memes, and cognitive communication perspectives such as dual-process theory. Grammatical ambiguity arises from structural vagueness (e.g., “Flying planes can be dangerous”); semantic ambiguity from lexical polysemy or selectional restriction failures (e.g., “bank”); and pragmatic ambiguity from context-dependent use and cooperative principle violations (e.g., “He is a real dragon”).

Regarding meme transmission, Jia Xiang tong emphasizes the unity of virtuality and reality: memes exist as digital symbols in virtual space yet exert tangible effects

on real-world cognition and behavior. Participatory culture further drives meme dynamics through user co-creation, secondary creations, emotional mobilization, and cross-circle dissemination, with algorithmic recommendation systems amplifying the “Matthew effect” in meme diffusion. Together, these foundations provide an integrated framework for analyzing how semantic ambiguity operates as both a cognitive and communicative mechanism in the multimodal dissemination of internet memes.

2. Generation Mechanism of English Ambiguity in Multimodal Contexts

2.1. The Impact of Multimodal Characteristics of Meme Carriers on Ambiguity

The uncertainty of the relationship between images and texts serves as the surface structure and direct driving force for the generation of ambiguity. The multimodal carriers of internet memes, particularly the combined relationships between images and texts, create semantic inconsistencies, abnormal juxtapositions, and high context dependency, which constitute the surface structure and direct driving force for ambiguity generation. This ambiguity does not stem from the vagueness of a single modality but arises from the “gaps” between multimodal symbols. Specifically, when an image depicting a scene with universal emotions (such as a person with a confused expression) is paired with highly abstract, decontextualized, or multi-referential text (such as a philosophical aphorism or a trending term related to current events), an open “interpretive framework” is formed. Audiences must actively draw upon their own cultural knowledge bases, subcultural backgrounds, and immediate social contexts to establish connections between the two. Take the widely disseminated “chaos” gesture meme as an example. The gesture image itself has no fixed meaning, but when combined with different text labels (such as those representing different ideologies, academic theories, or internet memes), it can instantly point to vastly different stances and objects of satire. This templated production mechanism allows the same image matrix to be infinitely “customized.” Its original context is gradually eroded through repeated copying, adaptation, and dissemination, with each recombination of image and text representing a new battle for meaning and the release of ambiguity potential.

Therefore, the combination of images and texts is essentially an invitational rhetoric. Through the incomplete matching between modalities, it cedes part of the interpretive authority to the audience, making ambiguity the starting point for their participation in interpretation and recreation.

The openness of cross-modal mapping serves as the deep cognitive mechanism and inherent instability underlying ambiguity generation. At the cognitive level, the deep-seated root of ambiguity lies in the inherent incompleteness and openness of the mapping process when memes construct “cross-modal metaphors.” According to Conceptual Blending Theory, a complete metaphor requires feature mapping from a “source domain” to a “target domain.” In multimodal memes, the key information of these two domains is often split and concealed across different modalities. For instance, the target domain (such as “a complex life”) may be explicitly stated in text, while the source domain (such as “a tangled mess”) is subtly hinted at through non-verbal imagery. These splitting forces the receiver to actively complete cross-modal conceptual blending in their mind, and the way of blending varies from person to person. More crucially, metaphors in memes are rarely simple; they are usually tightly

intertwined with metonymy, forming a “metaphor-metonymy interaction chain.” For example, in a political meme, a specific expression of a political figure (metonymy: using the individual to represent the entire policy style) may be employed to metaphorically imply “stubbornness” or “absurdity” (metaphor). However, the choice of which expression to use as metonymy and which target trait to map it onto is filled with ideological and subjective judgments. Take the classic “Brain Upgrade Diagram” meme as an example. The images at each level (from simple lines to complex halos) are visual metonymies representing “levels of thinking.” But when used to compare different viewpoints, theories, or groups, they constitute metaphors with strong value judgments. Every link in this chain—the selection of metonymy, the direction of metaphor, and the extraction of traits—is not definitive but leaves vast cognitive space for diverse or even opposing interpretations.

Therefore, cross-modal metaphors are not solid crystals of meaning but dynamic, meaning sketches that rely on cognitive filling. Their inherent instability makes ambiguity an inherent cognitive gene of multimodal memes.

2.2. Typical Case Analysis

2.2.1. Grammatical Ambiguity in Multimodal Memes: The Dislocation Game between Textual Structure and Image Context

In cases of grammatical ambiguity, the influence of multimodal features is primarily manifested in a dynamic dislocation game between the inherent syntactic or lexical ambiguity at the textual level and the “hints” or “interferences” provided by the image level. The core of this phenomenon is not that images directly resolve textual ambiguity. Instead, images exacerbate selectivity and even create new ambiguous spaces by offering a possible interpretive context. For example, consider a typical online homophonic meme where the text reads, “This position requires years of ‘gōng dú’ experience.” The ambiguity arises from the homophonic nature of “gōng dú,” which can be interpreted as either “gōng dú” (meaning “to pursue academic studies”) or a literal reference to “attacking viruses.” If the accompanying image is a thick academic book, it will guide the audience towards the interpretation of “pursuing academic studies.” Conversely, if the image depicts a laboratory scene of virus research, it will reinforce the literal meaning of “attacking viruses.” However, the complexity of multimodality lies in the fact that images can also provide conflicting or irrelevant information. For instance, if the meme is paired with an image of a leisurely vacation, the contrast between the text and the image will suspend the ambiguity, throwing the interpretive focus entirely back onto the audience’s contextual knowledge (such as whether they are familiar with recruitment jargon or specific industry memes).

So, multimodal carriers do not simplify grammatical ambiguity. Instead, they expand a single textual ambiguity into a more open cognitive multiple-choice question regulated by the relationship between text and image by adding variables (the directionality of images). The ultimate resolution of such ambiguity highly depends on the audience’s tacit understanding of subcultural symbolic rules.

2.2.2. Pragmatic Ambiguity in Multimodal Memes: The Fuzzy Consensus between Metaphor Construction and Ideology

In cases of pragmatic ambiguity, the role of multimodal features is even more profound. They collaboratively construct a complete metaphorical or satirical

framework through images and text, yet the ultimate intention and stance of this framework remain open-or even deliberately ambiguous-across different contexts. Take common political memes featuring leaders as an example. By juxtaposing an exaggerated facial expression of a leader (metonymically representing their overall policy style) with the text “moving forward steadily,” a multimodal metaphor is formed. Supporters may interpret the combination as “firm and resolute” (with the image serving as proof of determination), while opponents may see it as “blindly reckless” (with the image embodying rashness). Here, the ambiguity no longer stems from grammatical structure but from the ideological stances wrapped in a veneer of humor or irony. Research indicates that Finnish far-right memes excel at using such “bitter humor” to convey exclusionary messages in a playful manner while maintaining ambiguity towards outsiders. Images and text together delineate a field of meaning but refuse to provide a definitive answer. The true pragmatic intention-whether it is criticism, mobilization, or entertainment-needs to be confirmed within specific community consensus and context. Thus, multimodal carriers serve as an ideal tool for creating “deniable ambiguity” in this context. Creators can claim it is merely a joke, while receivers can complete ideological identification and dissemination through tacit understanding.

In summary, in grammatical ambiguity, multimodality exacerbates the complexity of choice; in pragmatic ambiguity, it constructs the fuzziness of stance. Together, they demonstrate the core mechanism by which online memes produce ambiguity and drive dissemination through the collaboration of multiple carriers.

3. The Dissemination Mechanism and Cognitive Evolution of English Ambiguity in a Multimodal Context

3.1. Mechanism of Ambiguity Variation in the Dissemination of Internet Memes

The dissemination of internet memes is not a simple process of replication and diffusion, but rather a dynamic process of “recontextualization”. During this process, the original ambiguous structures undergo three typical forms of variation through users’ secondary creative acts such as parody, collage, and reversal.

On the level of isomorphic yet heterogeneous nature: The same meme template retains its formal structure but is filled with different semantic content, thereby reactivating the original ambiguity within a new context. In such variations, the visual framework, layout, or image combinations of the meme remain stable, while the textual layer or the referent objects are replaced. Consequently, the “point of conflict” of the ambiguity is shifted to different domains. A typical example is the “Is this a pigeon?” meme. In the original image, the character mistakenly identifies a butterfly as a pigeon, creating a “referential ambiguity.” Users preserve the image structure and the “Is this a X?” sentence pattern, substituting X with various concepts (such as “a functional program,” “a healthy relationship”), causing the referential ambiguity at the grammatical level to repeatedly emerge in new contexts, with the form remaining unchanged but the referents varying. Such variations demonstrate that the transferability of ambiguity is key to the cross-domain dissemination of memes.

Regarding heterogeneous yet homogeneous nature: Multimodal memes of different forms tend to converge in semantic function, forming functionally equivalent “ambiguity families.” Although these memes vary in terms of images, text, and layout, they share the same semantic conflict pattern, relying on “saying something other than

what is meant” at the pragmatic level to construct meaning tension. A typical example is various ironic memes. The “That is the joke” meme directly points out through text the pragmatic situation where “the other party fails to understand irony”; the “I am something of a scientist myself” meme constructs implicit irony through a character’s self-aggrandizement in an inept situation, highlighting a mismatch between ability and context. Despite their vastly different forms, both rely on pragmatic ambiguity—the deviation between literal meaning and communicative intent—to drive meaning generation. This phenomenon of “heterogeneous yet homogeneous” nature corroborates Jia Xiangtong’s assertion that the virtual symbolic activities of memes can transcend specific media forms and achieve meaning aggregation at the functional level.

Finally, there is the heterogeneous and non-homogeneous type. In this case, memes undergo reconstruction in both form and semantics, with ambiguity no longer confined to the original structure but becoming a resource for users’ creative expression. In such variations, the “family resemblance” characteristic of the original meme is retained, but specific images, text, and combinations are replaced, with the point of ambiguity shifting accordingly. A typical example is the “Woman Yelling at Cat” meme. The original material is a collage of two images: an emotionally agitated woman and a perplexed cat sitting at a dining table. This combination itself constitutes a visual ambiguity of “mismatched emotions and responses,” creating a space for conflicting interpretations between the woman’s emotional expression and the cat’s calm response. Users continuously derive new conflicting semantics by replacing the text overlay (e.g., “Me explaining something vs. My brain processing it”), without retaining the original text content each time, and only preserving the ambiguous schema of “conflicting confrontation.” The characteristic of “family resemblance” in visual memes proposed by Feng Jianming is particularly evident here: ambiguity is not carried by a single meme but forms a recognizable “ambiguity family” among meme variants, sharing a certain cognitive conflict pattern rather than a fixed form.

The three variation mechanisms are not mutually exclusive; they often alternate or overlap during the dissemination process. Their common feature lies in the transformation of ambiguity from a “cognitive obstacle that needs resolution” into a “creative resource that can be utilized,” enabling users to shift from being passive recipients of meaning to active manipulators of meaning.

3.2. The Cognitive Dual-Process Mechanism of Ambiguity Resolution

The journey of ambiguity from its generation to comprehension is, at its core, a cognitive processing endeavor. According to the dual-process theory, when interpreting multimodal memes, users rely on two cognitive mechanisms that work in tandem.

The first is intuitive processing, which is characterized by its speed, automaticity, experience-based nature, and low cognitive load. When confronted with a meme, users initially rely on visual features, familiar contexts, and pre-existing knowledge to form a preliminary interpretation. Intuitive processing depends on pre-existing “semantic schemas” in working memory—that is, multimodal meaning-matching patterns accumulated by users through past exposure. For instance, when encountering the “Distracted Boyfriend” meme, users can instantly grasp the core semantic schema of “attention diversion” without the need for grammatical analysis,

because the visual composition of the meme (the man turning his head, the woman staring) highly matches the social script of “betrayal or preference shift.”

The second is controlled processing, which refers to being slow, rational, requiring cognitive resources, and allowing for error correction. When intuitive processing fails to uniquely determine meaning—such as in cases of text-image conflicts, the possibility of multiple interpretations, or a lack of background knowledge on the part of the user—the user invokes contextual information from working memory (such as comments, tags, platform atmosphere, and the identity of the poster) to reason and resolve ambiguity. Controlled processing involves mechanisms of “selective activation” and “suppression”: users select one interpretation from multiple possibilities and suppress the others. A typical example is the “This is fine” meme. In the image, a dog is drinking coffee in a burning room with a calm expression. Intuitive processing may lead to two conflicting interpretations: “the character is unaware of the danger” or “the character is self-consoling.” Users need to combine the meme’s dissemination history (originating from the comic “Gun show”) and its common usage context (expressing “pretending to be calm in a difficult situation”) to infer, through controlled processing, its pragmatic meaning of “ironic self-mockery.” During this process, users also need to suppress the literal interpretation (“the room being on fire is a good thing”) to resolve the ambiguity.

It is noteworthy that the interaction between intuitive processing and controlled processing is modulated by multiple factors. The first is contextual richness. In multimodal contexts, text, images, platform tags, and other elements constitute “redundant information” that can reduce reliance on controlled processing; conversely, fragmented contexts (such as images without accompanying text or being detached from their original platform) force users to engage in controlled processing. The second is cognitive load. When users are under high cognitive load (such as when skimming quickly or multitasking), they tend to rely more on the system and are more likely to accept the “most readily available” interpretation, even if it may not be accurate. Additionally, individual differences must be considered. Factors such as a user’s familiarity with specific memes, media literacy level, and cultural background influence the proportion of reliance on the two processing systems.

In virtual environments such as the metaverse and cross-platform dissemination, contextual information often exhibits characteristics such as fragmentation, decontextualization, and multimodal conflicts, which may undermine the effectiveness of controlled processing and result in unresolved ambiguities or even their erroneous consolidation. Jia Xiangtong’s discussion on the dual dimensions of “virtualization and reality” echoes this point: while symbolic activities in virtual spaces can enable meaning reconstruction, they may also lead to uncontrolled ambiguities and the formation of “cognitive pollution” due to their detachment from physical contexts.

3.3. The Hierarchical Transformation from Individual Cognition to Social Communication

The dissemination of ambiguity is not a linear process; rather, it undergoes hierarchical transformation between individuals and groups. This transformation process can be divided into three levels: micro, meso, and macro.

At the micro level, the primary focus is on individual cognition. Users identify and resolve ambiguities within their working memory, forming an initial understanding.

The core mechanism at this level is dual-process processing, with its outcomes falling into three states: (1) successful resolution of ambiguity, where users arrive at a stable interpretation; (2) unresolved ambiguity, where users retain multiple possible interpretations; and (3) erroneous resolution of ambiguity, where users form interpretations that deviate from the intended meaning of the communication. All three states serve as the cognitive foundation for subsequent communicative behaviors.

Next comes the meso level of group interaction. After individual interpretations enter social platforms, through interactive behaviors such as commenting, reposting, and secondary creation, a “negotiated consensus” or “divergent diffusion” emerges at the group level. The core mechanism at this level is “interactive contextual reconstruction”: users supplement context through replies (such as explaining the origin of a meme), adjust the ambiguity structure through secondary creation (such as modifying the text layer to reinforce a particular interpretation), and filter mainstream interpretations through likes or controversial voting. If the ambiguity structure of a meme is highly malleable (such as memes like “They did surgery on a grape,” whose meaning entirely relies on user input), it is more likely to stimulate diverse interpretations and facilitate cross-circle dissemination. Conversely, if the ambiguity structure is relatively closed (such as grammatical ambiguities that rely less on context), group negotiation may quickly converge to a stable interpretation. The outcomes of meso-level interactions, in turn, regulate individual cognition: users revise or solidify their interpretations based on group feedback.

Finally, there is the macro level of cultural dissemination. As memes spread widely and transcend different communities, platforms, and even cultural circles, some ambiguities become “conventionalized,” evolving into stable meaning resources within specific communities or cultures. Meanwhile, other ambiguities fade out of circulation due to contextual loss or excessive mutation. A typical example is the “Sad Frog” meme (Pepe the Frog). Initially a comic character with no obvious ambiguity, it gradually became solidly associated with meanings such as “grief culture” (a culture of resignation or despair), “loneliness,” and “failure,” forming a stable cultural symbol. This conventionalization process relied not only on sustained user engagement (accumulated at the meso level) but also on platform algorithms reinforcing “emotional tags” (mediated by macro-level technology). At the same time, the meme was also co-opted with negative connotations such as racism, leading to uncontrolled ambiguity. The original creator attempted to “reclaim” the image but achieved limited success. This case demonstrates that while conventionalization of ambiguity at the macro level grants memes communicative stability, it can also result in uncontrolled meaning shifts and the risk of alienation.

The hierarchical transformation mechanism refers to the bidirectional interactions among the three levels. Individual interpretations at the micro level provide material for meso-level group interactions; group negotiations at the meso level influence individual cognition through the mechanism of “social proof”; and cultural conventions at the macro level constitute the “background context” for micro-level cognition, pre-setting users’ expectational frameworks for memes. Conversely, innovative individual interpretations may spread through meso-level interactions and ultimately rewrite macro-level conventions. This hierarchical transformation process reveals the underlying logic of ambiguity-driven meme dissemination: ambiguity is neither a completely open “free-for-all interpretation” nor a completely closed “fixed encoding,” but rather exists in a dynamic range between structural constraints and user creativity.

4. Risks and Countermeasures in the Dissemination of Internet Memes

4.1. Risk of Alienation

The risk of alienation in internet memes first manifests as a systemic cognitive pollution. The core mechanism lies in the fact that memes, with their highly emotional, fragmented, and viral dissemination characteristics, are reshaping the public's information processing patterns and continuously weakening the social foundation for rational thinking and in-depth contemplation

Firstly, there is a “dimensional reduction” of thinking and “cognitive rot.” Memes strive for “instant comprehension,” compelling complex social issues to be compressed into simple binary oppositions, emotional symbols, or stereotypes. Long-term exposure to such “short, flat, and fast” content reinforces the brain's tendency toward “cognitive laziness,” leading individuals to rely on intuitive reflexes instead of logical verification. This is the crisis revealed by the phenomenon of “cognitive rot”—excessive indulgence in trivial, low-challenge online content may result in the degradation of in-depth thinking abilities and creativity. Then, there is the construction of “information cocoons” and “echo chambers.” The algorithmic recommendation mechanisms of social platforms continuously push similar memes that align with users' preferences, forming cognitive closed loops. Repeated exposure to homogeneous emotional information leads individuals to equate the “pseudo-environment” created by algorithms with the real world, rejecting heterogeneous viewpoints and thereby exacerbating group polarization. Harmful memes (such as those containing fake news or hate speech) resonate frequently within circles, continuously self-reinforcing and ultimately distorting the common cognitive foundation of members.

Finally, there is the deconstruction of serious discourse and the dissemination of “post-truth.” Digital populism adeptly utilizes the entertaining facade of memes to undermine the seriousness of elite discourse and mainstream narratives. Through playful and ironic means, memes can bypass rational debate and directly construct an emotional “truth,” prioritizing emotional resonance over fact-checking. When the impulse to “play with memes” overrides the pursuit of truth, public discussions lose their common benchmark, making it difficult to form social consensus.

The prevalence of cognitive pollution is deeply driven by a severely distorted attention economy. Under the underlying logic of “traffic is king,” the creation and dissemination of memes are no longer centered on information transmission or creative expression but have alienated into a brutal competition for users' limited attention. This distortion is reflected at three levels:

Alienation of production motives. “Traffic fishing” and the sale of opposition. To quickly capture attention amidst an information deluge, content creators (and those manipulating them from behind the scenes) use memes as efficient “fishing tools.” They identify, prepare, and cast their “bait” seamlessly: recognizing sensitive issues such as gender, region, and generation, they simplify them into highly conflicting labels (e.g., “Apple people” vs. “Android people”) and create easily disseminated memes. Their purpose is not to provoke thought but to deliberately evoke high-intensity emotions such as anger and opposition, thereby leveraging platform interaction algorithms to achieve traffic monetization. From “kitschy imitations”

(deliberately highlighting physical characteristics) to malicious labeling statements that create social divisions, all are products of this logic.

Alienation of distribution mechanisms. Algorithmic amplification of extreme emotions. Social platform algorithms are essentially “attention allocation machines,” with their core metric being user interaction rates (likes, comments, shares). Since emotional and conflicting memes are more likely to stimulate instant user interactions, algorithms naturally prioritize recommending such content, creating a vicious cycle where “extreme content gains more exposure, and more exposure breeds more extreme content.” This mechanism systematically rewards inciters and suppresses rational, complex voices.

Alienation of consumption behavior. From in-depth participation to “reflexive” likes and shares. In the information environment shaped by algorithms and the supply of content seeking instant gratification, users' attention becomes highly fragmented. Faced with a barrage of memes, in-depth reading and thinking become difficult, and political participation and social discussions are reduced to instinctive likes or shares of a meme. This “reflexive” participation inadvertently turns users into “data fuel” and emotional amplifiers for the dissemination of alienated content.

4.2. Optimization Paths

The core of optimizing meme design standards lies in establishing a set of non-mandatory “meaning guardrails” and “value orientations” for spontaneous and disorderly meme creation, thereby suppressing cognitive pollution at the production source. The key is not rigid rules but guiding memes toward more responsible and constructive directions by advocating best practices and transparent creation. Specific implementation encompasses three levels: Firstly, promoting “meta-discourse anchoring” at the symbolic level. To address cross-modal metaphor ambiguities, creators are encouraged to actively use tags (e.g., satire, educational), borders, or additional textual explanations to self-declare the intent and context of memes. For instance, a study on educational Instagram accounts demonstrated that using clear meta-discourse strategies to explain the relationship between images and text significantly enhances the understanding of complex information and reduces misinterpretations. Secondly, establishing a “responsible adaptation convention” at the behavioral level. To counter the malicious appropriation of classic templates (e.g., using harmless meme images for hateful expressions), communities can be encouraged to reach a consensus requiring secondary creators to indicate the purpose of adaptation and differences from the original context through watermarks or fixed text positions, similar to “Creative Commons” agreements, thereby forming cultural-level creative ethics. Lastly, defining and promoting the paradigm of “constructive memes” at the value level. Platforms and institutions can systematically support meme creations dedicated to scientific dissemination, social care (e.g., mental health support), and promoting cross-group understanding.

By offering traffic incentives and official certifications, new benchmarks for imitation can be established. An Austrian study has experimentally proven that constructively adapting political memes (e.g., adding hints of solutions) can effectively improve audiences' political efficacy and attitudes. Therefore, the ultimate goal of design standards is to partially transform memes' creativity from a “deconstructive force” that creates opposition and chaos into a “constructive force” that enhances understanding and consensus.

The core of optimizing user cognitive training lies in systematically cultivating a “critical slow-thinking” immune system at individuals’ information processing terminals that can counter the immediate emotional stimulation of memes. This is not simply about informing users of risks but reshaping their cognitive habits in decoding multimodal information through structured training, transforming them from passive “emotional reactors” into active “meaning auditors.” Training should focus on three core habits: First of all, developing the “muscle memory” of “multimodal disassembly.” Through repeated practice, users should instinctively separate elements such as images, text, and layout when encountering memes and question their relationships: Does this stereotypical image reinforce or contradict the textual label? Is its purpose to simplify or distort facts? Research indicates that fact-checking training for multimodal disinformation can significantly enhance citizens’ discernment abilities. Secondly, cultivating a “detective mindset” for “intent tracing.” Users should be trained to establish a reflexive chain of questioning: Where did this meme originally come from within a subculture? What is the historical stance of the publishing account? Is it trying to anger me, amuse me, or persuade me? What specific behavior (like, confrontational comment, share) is it attempting to induce? This mindset effectively deconstructs pragmatic ambiguities and reveals hidden ideological or commercial motives. Thirdly, acquiring the metacognitive ability of “emotion labeling.” This is crucial for countering the distortions of the attention economy. Users should be trained to pause and self-alert when feeling instantly angered or elated by a meme: “My intense emotion at this moment is being triggered by this symbol deliberately designed with a ‘us vs. them’ framework.”

Finnish media literacy education projects prioritize identifying “anger baits” and emotional manipulation. When a large number of users possess such “cognitive antibodies,” malicious memes aimed at polluting cognition and capturing traffic will naturally lose their intended propagation fuel and market. Therefore, the ultimate goal of cognitive training is to systematically reduce the “toxicity” carrying capacity of the entire information ecosystem by elevating the collective level of media literacy.

5. Conclusions

The multimodal characteristics of memes—text-image integration and cross-modal metaphor—shape ambiguity generation and dissemination through modal synergy, cultural presupposition, and technological mediation. Ambiguity from text-image integration arises when unclear logical relations or unequal modal status yield multiple interpretations due to shifting information focus or cultural differences. Ambiguity in cross-modal metaphor stems from open symbolic mapping and divisive cultural presuppositions, enabling the same meme to acquire opposite meanings across contexts. Technological mediation, notably algorithmic recommendation, amplifies this ambiguity by filtering user groups or reconstructing meaning frameworks, creating high uncertainty in dissemination effects. These findings reveal a core contradiction: efficient diffusion coexists with loss of meaning control, fostering cultural innovation while challenging information authenticity.

Future research should address three dimensions: (1) Technological ethics—developing ambiguity warning models and optimizing feedback mechanisms to balance efficiency and accuracy; (2) Cross-cultural communication—constructing a cultural adaptability framework for cross-modal metaphor to analyze divergent interpretations across cultural groups; (3) Cognitive science—using eye-tracking and brain imaging to uncover neural mechanisms of multimodal ambiguity processing.

Additionally, with the rise of generative AI (e.g., DALL·E, ChatGPT), future work should examine the ambiguous characteristics of AI-generated memes and their social impacts, offering new perspectives for digital media literacy education.

Conflicts of Interest

The author declares that there is no conflict of interest regarding the publication of this article.

Funding

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

References

- [1] Polli, C.; Sindoni, M.G. Multimodal computation or interpretation? Automatic vs. critical understanding of text-image relations in racist memes in English. *Discourse, Context & Media*, 2024, 57, 100755.
- [2] Shar, S.A.; Amur, S.F.; Ujan, N. Language Change in the Digital Age: The Influence of Social Media on Contemporary English. *ASSAJ*, 2025, 4(02), 3145-3155.
- [3] Yus, F. Incongruity-resolution humorous strategies in image macro memes. *Internet Pragmatics*, 2021, 4(1), 131-149. DOI: 10.1075/ip.00058.yus.
- [4] Udom, C.G. Impact of Social Media on Language Evolution and Effective Communication. *LALICO Journal of Languages, Literature, and Communication*, 2025, 3(2), 85-92.
- [5] Nayak, A.; Agrawal, A. Detection of hate speech in social media memes: A comparative Analysis. In 2022 Third International Conference on Intelligent Computing Instrumentation and Control Technologies (ICICICT). IEEE. 2022, pp. 1179-1185. DOI: 10.1109/ICICICT54557.2022.9917633.
- [6] Su, J.H.; Wu, W.N. The effect of Chinese learners' working memory span on the online processing of English temporary ambiguous sentences. *Journal of Xi'an International Studies University*, 2023, 31(04), 69-74. DOI: 10.16362/j.cnki.cn61-1457/h.2023.04.016.
- [7] Zhang, X.H. Classifying English ambiguity from syntactic, semantic and pragmatic perspectives. *Journal of Dalian University of Technology (Social Sciences Edition)*, 2008, 01, 83-86. DOI: 10.19525/j.issn1008-407x.2008.01.016.
- [8] Zhang, S.J. The internal rationale and international trends of the formation and development of cognitive pragmatics. *Modern Foreign Languages*, 2025, 1-10. DOI: 10.20071/j.cnki.xdwy.20251219.007.
- [9] Feng, J.M. A study on the dissemination mechanism of network visual memes. *Foreign Language and Literature*, 2023, 39(04), 112-120.
- [10] Jia, X.T.; Hu, Y. On the dual dimensions and alienation of information meme dissemination in the digital age. *Guangxi Social Sciences*, 2023, 02, 87-93.



© 2026 by the author(s); licensee International Technology and Science Publications (ITS), this work for open access publication is under the Creative Commons Attribution International License (CC BY 4.0). (<http://creativecommons.org/licenses/by/4.0/>)