

## Intersemiotic expansion in AI-generated images: An SFL-informed analysis

Dennis Lindenberg

Faculty of Science and Engineering, Waseda University, Tokyo, Japan

Email: lindenbergaoni.waseda.jp

Recent advances in text-to-image generation have transformed AI-generated images into multimodal texts that increasingly circulate across educational, creative, and professional contexts. Although these systems are commonly evaluated in terms of prompt alignment, creativity, or technical performance, comparatively little attention has been devoted to the question of how meanings encoded in language are reconstrued during visual realization. From the perspective of Systemic Functional Linguistics (SFL), language constitutes one semiotic resource among many, and meaning is not transferred intact across modes but reconstrued through intersemiotic relations between language and image (Halliday & Matthiessen, 2014; O'Halloran et al., 2019). Text-to-image generation therefore provides a productive empirical environment for investigating how linguistic meanings are systematically transformed through the affordances of the visual mode.

Rather than functioning as direct visual translations, AI-generated images emerge through interactions between prompt semantics, the semiotic affordances of the visual mode, and culturally sedimented patterns learned from large-scale multimodal training data. From an SFL perspective, this process can be understood as *intersemiotic transduction* (Kress, 2003), in which linguistic meanings are selectively expanded, reorganized, and reinterpreted across semiotic modes. Building on an SFL-informed framework for analysing visual-verbal relations in AI-generated images (Lindenberg, 2026), this study examines how minimal linguistic prompts are systematically expanded during image generation. It addresses two research questions:

1. How are meanings encoded in minimal linguistic prompts expanded and reconfigured through intersemiotic relations during visual realization?
2. How do different participant types influence these patterns of semantic expansion?

To investigate these questions, a purpose-built corpus of 120 AI-generated images was created using OpenAI's image generation system through ChatGPT. Four participant types differing in animacy and conventional agency ("person," "fish," "robot," and "spoon") were combined with six predicates representing material, mental, verbal, and projective process types. Each participant-process configuration was generated five times using a standardized two-line prompt, allowing systematic comparison across contrasting participant-process configurations.

The images were analysed using an SFL-informed coding framework grounded in multimodal social semiotics. Drawing on work on intersemiotic logico-semantic relations (Bateman, 2014; Lindenberg, 2025), expansion was operationalized as *specification*, *addition*, and *enhancement*, while *projection* captured visually represented speech and thought. Interpersonal meaning was examined through a three-level affect scale informed by the appraisal framework of Martin and White (2005). Coding reliability was established through independent coding by a trained second coder, with disagreements resolved through discussion.

The analysis demonstrates that AI-generated images systematically reconstrue

rather than merely reproduce meanings encoded in linguistic prompts. Material processes strongly license specification and experiential addition, whereas mental and interpersonal processes rely more heavily on projection and affective intensification to realize meanings that cannot be represented directly through observable action. Participant type further conditions these strategies. Human participants, functioning as highly schema-compatible entities, require comparatively little semantic compensation. Conversely, biologically implausible or non-agentive participants trigger systematic shifts toward anthropomorphism, environmental enhancement, projection, and intensified affect, enabling the visual mode to realize ideational and interpersonal meanings that remain underspecified in the original prompt.

From an SFL perspective, these findings suggest that intersemiotic expansion constitutes a systematic semantic process rather than an arbitrary by-product of image generation. The recurrent deployment of specification, addition, enhancement, projection, and affect indicates that visual realizations instantiate logico-semantic relations through which linguistic meanings are reconstrued according to the meaning potential of the visual mode. Rather than representing random additions or simple prompt alignment, these expansions constitute motivated semantic transformations that maintain interpretive coherence by drawing on culturally sedimented representational schemas.

The study extends SFL-informed accounts of intersemiotic meaning-making into the domain of generative AI. By demonstrating how participant-process configurations encoded in minimal linguistic prompts are systematically reconstrued during visual realization, it contributes to broader linguistic theories of semantic realization, intersemiotic transduction, and logico-semantic relations across semiotic modes while illustrating how AI-generated images provide a valuable empirical site for investigating multimodal meaning-making.

## References

- Bateman, J. A. (2014). *Text and image: A critical introduction to the visual/verbal divide*. Routledge.
- Halliday, M. A. K., & Matthiessen, C. M. I. M. (2014). *Halliday's introduction to functional grammar* (4th ed.). Routledge.
- Kress, G. (2003). *Literacy in the new media age*. Routledge.
- Lindenberg, D. (2025). Logico-semantic relations between spoken text and slides' visual elements in student presentations conducted online in the English as a Foreign Language context. *Visual Communication*, 24(2), 319–345. <https://doi.org/10.1177/14703572231180215>
- Lindenberg, D. (2026). How a text-to-image model adds meaning beyond minimal prompts: An SFL-informed intersemiotic analysis of visual-verbal relations. *Multimodality & Society*. Advance online publication. <https://doi.org/10.1177/26349795261459896>
- Martin, J. R., & White, P. R. R. (2005). *The language of evaluation: Appraisal in English*. Palgrave Macmillan.
- O'Halloran, K. L., Tan, S., & Wignell, P. (2019). SFL and multimodal discourse analysis. In G. Thompson, W. L. Bowcher, L. Fontaine, & D. Schönthal (Eds.), *The Cambridge handbook of Systemic Functional Linguistics* (pp. 433–461). Cambridge University Press.