From descriptive to reflective: Reading and interpreting generative AI images

punctum.gr

BY: **Seyedeh Maede Mirsonbol**

ABSTRACT

As generative AI expands from text to image creation, critical questions emerge about the nature of meaning-making in machine-generated visuals. This paper theoretically explores the shift from '*descriptive*' to '*reflective*' reading of generative AI images, drawing on linguistic and semiotic theories of Chomsky, Halliday, and Culioli. Chomsky emphasizes formal, rule-based structures, while Halliday and Culioli highlight meaning as contextual and inferential. Although studies of AI-generated images often focus on surface-level features such as style, coherence, and resemblance, this paper argues for a semiotic engagement that considers the underlying structures and contextual processes (or their simulation) at play. The central question of whether generative AI engages in the complex, contextual processes of human meaning-making is explored through an applied approach focusing on prompting and image inquiry. This is to construct and encourage humans in showing reflection and interpretation of visuals, firstly, by describing the relationship between AI and humans to form an interactive partnership; secondly, by moving to a reflective interpretation that requires a human viewer to supplement AI's syntactic fluency with socially, culturally, and cognitively grounded meaning-making.

ARTICLE INFO:

Volume: **12**

Issue: **02**

Winter **2025**

ISSN: **2459-2943**

DOI: **10.18680/hss.2025.0021**

Pages: **47-68**

KEYWORDS:

Generativity

Artificial Intelligence

Semiotics

Human-AI interaction

Images

Punctum. International Journal of Semiotics | **11:**02**:**2025
ISSN 2459-2943 | DOI: 10.18680/hss.2025.0021 | **punctum.gr**

1. Introduction

In the contemporary digital sphere, the generativity structure of visual media displays is increasingly intertwined with the production and circulation of images, for example, in education, health, marketing, environmental sustainability, etc. (Hobbs 2024). This process has been significantly intensified by advances in image-generating technologies, including those powered by Artificial Intelligence (AI). In the context of human interaction with AI, the generativity of its systems relies heavily on human evaluation, particularly when the focus is on the capacity to autonomously produce and manipulate visual content (Lin et al. 2024). This notion further highlights the human side as the stronger side of the scale, while AI-generated content is intentionally positioned on the weaker side and is considered a pivotal mediator of socio-cultural and communicative processes. This phenomenon aligns with a long-standing production of aesthetic images and is also central to the philosophical field of semiotics (the study of signs and meaning-making). From a semiotic perspective, this evolution in human-AI interaction constitutes a shift in sign-making practices, in which the generativity of AI contributes to the reconfiguration of visual signification and to the anthropological dimensions of how meaning is co-constructed between humans and machines (Leone 2024).

This paper aims to theoretically explore the concept of generativity through interaction with AI and to examine its visual outputs through a semiotic reading of signs, a form of analysis performed by human agents. This method can help demonstrate the impact of linguistic enunciation on a generation procedure (prompt engineering), which involves placing linguistic text within generative AI platforms to create visual representations. The human interaction with the visual signs and their interpretation will feature the stage that is considered as a catalyst for deep reflection and a semiotic scaffolding that results in meaning creation (Lacković 2020a: 16). The purpose is to reflect on whether AI can truly generate meaning visually in the same way humans do, or it is only a simulation that misses the essential human processes such as creativity, intuition, and imagination that are important to be theorized in these linguistic models. The study foregrounds and builds on influential linguistic and semiotic theories, including Chomsky's (1968) Universal Grammar (UG), Halliday's (1978) systemic functional linguistics and social semiotics, and Culioli's (1991-2000) enunciative theory. They all offer contrasting yet complementary frameworks for analysing the processes of meaning-making in language, including the concept of generativity, to support the observation and evaluation stages of prompting from multiple perspectives.

To answer the main question *Do generative AI images engage in the complex, contextual, and inferential processes involved in human meaning-making*, this study employs an applied approach to further investigate the act of prompting to understand how

visuals are enunciated and produced. Inspired by the work of D'Armenio, Deliège, and Dondero (2024), this approach highlights the reflective dimensions of image interpretation, in which meaning is shaped by the viewer's awareness of communicative purpose and situational context. This emphasizes the interpreter's agency, subjectivity, and mental operations in constructing meaning. In this context, Culioli's inferential model can enhance reflective engagement with AI-generated images by framing meaning as the result of cognitive and contextual operations, whereby the interpreter constructs it through acts of questioning, predication, and validation within specific enunciative settings. By applying this method to AI-generated images used for conceptual understanding, the paper argues that while such systems can simulate structural and stylistic elements, their capacity for deep, situated, and reflective meaning-making remains limited unless we move from descriptive to reflective interpretation of visual artifacts. This attempt requires a human viewer to supplement AI's syntactic fluency with socially and cognitively grounded frameworks of meaning. This theoretical and interdisciplinary idea invites critical reflection on how to approach images, while the goal is to read, use, and (maybe) trust AI-generated visual media in different contexts.

2. The concept of generativity

2.1. Generativity in linguistics and semiotics

According to Cambridge Dictionary, the term '*generativity*' is simply defined as "the quality of being able to produce or create something new"[1] and in social sciences it is expanded to have close connection to a person's interest while acting as a motivational force in life that impacts on work and engagement (Ward and King 2017: 67). Through the years, generativity in linguistics has been explored from the perspective of semiotics, encompassing textual and structural dimensions, sociocultural diversity, interactional processes, and the dynamics of influencing perception, creativity, and narrativity (Salthe 1999; Clarke 2009; Galofaro 2013). Similarly, generativity in images and visual displays has been examined with respect to symbolic representation, compositional structure, and the communicative potential of visual forms (Bertling 2019; Dondero 2025). If the aim is to employ generativity of images and emphasis on its language, then we need to first consider semiotics as "central to human communication toward [sign interpretation as] a more holistic view of semiosis" (Topalidou Laskaridou, Papadopoulos, and Koutsogiannis 2022: 123).

[1] Retrieved from Cambridge Dictionary Press (July 2025): https://dictionary.cambridge.org/dictionary/english/generativity

Punctum. International Journal of Semiotics | **11:02:2025**
ISSN 2459-2943 | DOI: 10.18680/hss.2025.0021 | **punctum.gr**

While it is difficult to view life from a single-dimensional perspective, it is also difficult to characterize language generativity as a rigid, rule-based, and analytic principle that operates along a one-way path, disconnected from other socio-cultural contexts. To explain generativity in linguistics and show its arc through time, it goes without saying to begin with Noam Chomsky, the father of modern linguistics. Chomsky's focus is on the linguistic aspect and grammatical structures of a language, and its generative behaviour follows rule-based strategies. In his 1960s research and papers, he articulated his views on the formalist view known as Universal Grammar (UG) (Chomsky 1968), which posits that all human languages share an innate set of grammatical principles and structures embedded in the human brain. According to this theory, the capacity for language acquisition and generating it into something more is hardwired, enabling children to learn any language rapidly and efficiently despite limited exposure. UG suggests that although languages may differ in certain respects, they all obey a universal set of rules governing sentence structure and syntax. This innate linguistic faculty enables children to generate and understand novel sentences that extend beyond the scope of their immediate experience (Chomsky 2000). UG challenges behaviorist notions of language based solely on imitation and reinforcement, and its main shortcoming is that context is eliminated in the study of language while it is examined in isolation from social use. Furthermore, meaning is often secondary in these circumstances, whereas structure holds the primary focus in language generativity (Chomsky 2002). Despite Chomsky's theory being revolutionary at its time, it has been criticized for its limited consideration of social, cultural, and communicative contexts. Michael Halliday, among other scholars, argued against UG by emphasizing the functional and social dimensions of language, proposing that the generativity of language is deeply influenced by interaction and the purpose language serves in communication rather than solely by innate grammatical structures.

Halliday's seminal work, *Language as Social Semiotic* (1978), introduced a paradigm shift by emphasizing language as a socially embedded semiotic system. The core of this perspective emphasizes the role of meaning as ideational (content), interpersonal (relationships), and textual (organization) as it is exchanged and interpreted in contextually generative communication shaped by socio-cultural and personal factors (Halliday 2013; Peluso 2021). Halliday (2005) argued that meaning is made through language, and its generative nature and usage in life cannot be fully understood without considering its social functions and the communicative purposes it serves within specific cultural contexts. In this sense, he sought to expand the scope of linguistic understanding and generativity with elements that could influence its production and implementation. His systemic functional linguistics proposed that language's generativity is not purely innate but is also shaped and constrained by

socio-cultural norms and human experience, which are "the total context of the interaction between an individual and his human [or non-human] environment" (Halliday 1978: 9; brackets added). This perspective moved linguistic research beyond Chomsky's cognitive and innate framework toward a more comprehensive view that incorporates social interaction and cultural environment by establishing the necessity of social semiotic in the theory of language generativity with elements like: "text, situation, register[tion], code, the linguistic system, social structure" (Halliday 1978: 108; brackets added). Therefore, broadening the generativity ideology to include external influences alongside biological components is essential if the goal is to create a constellation of meanings. Halliday's approach thus laid the groundwork for the sociolinguistic era, where the interplay between innate capacities and social contexts became central to understanding language development and use.

Both theories introduced by Chomsky and Halliday are powerful and serve different aims: Chomsky explains how language is possible, while Halliday explains how language is used. However, this study adopts an interdisciplinary approach that extends beyond them to complete the language generativity arc by building on the personal dimension of language described by linguist Antonio Culioli. Culioli's (1995) enunciative theory further advanced linguistic inquiry by focusing on the act of enunciation itself, being the answer to how a speaker (enunciator) produces meaning in context (situation) through choices and interactions (Groussier 2000: 161). Culioli's (1991-2000) collection of volumes complements and extends Halliday's theory by emphasizing on a more dynamic point of view and situational nature of language use, instead of viewing language as a static system, highlighting this important perspective that "we cannot countenance considering languages to be codes of a stable, universal reality" (Culioli 1995: 27). Regarding this notion, a text is not predefined as well and meaning emerges not only from social norms but also from the speaker's subjective position and interactional intentions, thus enriching the understanding of language and its generative aspect as a cognitive phenomenon (Culioli 2000).

To sum up this section, Halliday and Culioli are complementary in their rejection of Chomsky's formalism, and both focus on meaning in context, speaker agency, and the dynamic nature of language (Table 1). While Halliday centers on social function, Culioli digs into the mental procedures that underlie those functions. Both offer valuable perspectives for understanding how meaning is constructed, especially relevant when examining AI prompting resulting in image generation and human-machine interaction. Culioli's contextual and inferential approach, which this study adopts as its fundamental method for analysing AI's generativity, emphasises the speaker's subjectivity and mental operations in constructing meaning, which are central to understanding the relationship between human and AI.

Punctum. International Journal of Semiotics | **11:02:2025**
ISSN 2459-2943 | DOI: 10.18680/hss.2025.0021 | **punctum.gr**

Table 1.
Key differences between linguistic and semiotic theories of generativity by Chomsky, Halliday, and Culioli

Scholar	**Key Work(s)**	**Core Contribution to Generativity**	**Linguistic / Semiotic Principles**
Chomsky	*Syntactic Structures – UG Theory*	Proposed Universal Grammar (UG), an innate cognitive system enabling humans to generate infinite sentences from finite rules	Emphasized innate syntactic structures and cognitive mechanisms underlying language acquisition (linguistic)
Halliday	*Language as Social Semiotic*	Shifted focus to language and social semiotic system where generativity arises from social functions and cultural contexts, not solely innate structures	Developed systemic functional linguistics emphasizing language's social and communicative functions (semiotic)
Culioli	*Cognition and Representation in Linguistic: Enunciative Theory*	Introduces enunciative theory, focusing on language as dynamic acts of meaning production shaped by speaker's interactional choices. Extended generativity to include situational and subjective aspects.	Highlighted the role of enunciation and speaker's agency in meaning-making (linguistic and semiotic integration for generativity)

2.2. Generativity in the language of images

Images function as a unique language that communicates through shapes, colors, composition, and symbolism rather than words (Mitchell 1994). This language is inherently generative because, semiotically speaking, each viewer brings their own experiences, memories, and perceptions to the acts of semiosis, which can transform and expand the image's significance and generate new meanings (McAdams and de St. Aubin 1992). The great role that images have in today's scientific world is that they "invite us to revisit our conception of enunciation, to relate it more explicitly to a structure of experience, and to give it a place and a meaning in the exploration of our world" (Dondero and Fontanille 2014: 16). In this context, Lacković asserts:

> Indeed, in today's world we still need more images that show diverse human experiences, capabilities and identities, including non-human beings and things, and that is one of many reasons why we need to talk about images […]. Arguably, images can show exclusive representations. (Lacković 2020a: 6)

Within this visual domain, generativity refers to the ability to create and produce novel and meaningful forms, expressions, or shapes; especially pictures that explicitly convey messages capable of being expressed, interpreted, and transmitted within the operation of the semiotic transaction of visual communication (Louvel 2013: 13-14). Generated images to be used in this setting are also a means to be entangled with the action of signs and mind, because "mind translates differences, similarities, objective, subjective, wrong or right in a picture with semiotic systems" (Semetsky 2019: 9). According to this, the concept of generativity will undergo an evolution, extending from its foundational roots in psychology and sociology to encompass broader domains shaped by meaning-making processes in today's image-saturated world. When applied to images, generativity follows from Halliday's and Culioli's theories, which help us conclude that signs are not static or fixed. Rather, they are dynamic, open-ended, and capable of eliciting multiple layers of meaning, subjective interpretation, and emotional response within a broader sociocultural context that we can imagine as the sphere of semiosis.

The generation of new images offers a valuable means to understand society through the lens of pictorial signs. With respect to the digital environment, an image, "if not considered *art pour l'art* without any reference and an artifice on its own, is a reference to something else and as a means of reference is not isolated from a context" (Haase 2022: 158). The capacity to comprehend and translate the language of images into meaningful contexts facilitates the investigation of the transposition of the theory of enunciation in the digital environment and the realm of visual discourse (Dondero 2020: 2). Image generation involves an interdisciplinary dialogue between linguistics and semiotics, which informs the development of specific methodologies for reading pictorial signs within a text or reflecting on those interpreted through a network of connections between the visible elements and the external world beyond the image. This approach aligns closely with the theory of 'visual rhetoric' [2] (Group μ 1992 cited in Dondero 2020: 33). Ultimately, as Boling et al. (2014: 28) argues, "viewers of images must interpret them" because the interpretation of the generated contents inherent in the language of image highlights a dynamic interplay among the creator, the image, and the audience, shaping a fluid process of enunciation wherein meaning remains unfixed and is continually redefined through interpretation and creativity.

[2] Visual rhetoric theory is a critical study of the function of imagery and based on this theory "to understand an image, it is necessary to account not only for what is offered to vision, but also for what lies beyond the frame of the image, that is, what has been excluded" (Dondero 2020: 8). For more information about this theory see F. Andersson's 'The Visual Semiotics and Rhetoric of Groupe μ,' available at: https://research.abo.fi/en/publications/the-visual-semiotics-and-rhetoric-of-groupe-%CE%BC-opening-a-dialogue-

Punctum. International Journal of Semiotics | **11:**02**:**2025
ISSN 2459-2943 | DOI: 10.18680/hss.2025.0021 | **punctum.gr**

Sless (1986) points out that "creators of the texts – including images as texts – have to imagine the readers of those texts who, in return, must imagine the creators of the texts as part of the process of interpretation" (cited in Boling et al. 2014: 31). If this kind of attention to the creator and the audience of the visuals is emphasized more, interpretation of the generated visuals becomes focused on users' empowerment and relies on their creativity and intuition without centralized control that has an important role in the process of critical thinking. Therefore, the functional role of images is closely linked to this interpretive act and to how the viewer observes the text within a specific illustrative context. It can be understood from these that images can play a key role in many areas to impact on humans' reflection and inquiry-based knowledge construction, for instance, in the educational research and practice, "visual representations like photographs, diagrams, and models enable students to engage in scientific practices by allowing them to interact with complex phenomena" (Okulu 2025: 2).

2.3. **Generativity in the age of artificial intelligence**

This section begins by examining the term *generativity*, distinguishing its traditional meaning from its recent transformation amid its redefinition in the age of AI. Generativity, which has acquired new meanings in the context of AI, now involves co-creation between humans and AI in tasks such as text-to-image generation, which leverages AI to convert human-generated ideas into visual representations (Silvennoinen et al. 2024). This act is also known professionally as *prompting*. Prompt engineering is defined by Hobbs (2024: 21) as "the skillful design of language that can be used to enable generative AI to produce high-quality and desirable results." In fact, AI's manipulability through algorithms enables the automatic generation of images, a mechanism that takes input and makes a pivotal shift from human-authored creation to computational generation of pre-existing visual data as output. This process led scholars to ethically question the nature of generativity in AI (Lin et al. 2024), currently viewed as a matter of "more ethical than technical" (Leone 2023: 4). Moreover, in this way, AI's generativity is viewed as a serious concern challenging the idea of "human uniqueness" (Dondero 2025: 133).

Semiotically speaking, D'Armenio, Deliège, and Dondero (2024) take ethical objectives into consideration and pivot their exploration toward a more systematic approach to understanding the degree of human control over the generativity of AI. They evaluate the generated images in a three-layered structure: first, individually; second, with respect to their relevance to the input prompts; and finally, as a whole. In their research, they divided the dimensions of visual composition and evaluation of the text-to-image ability of AI into three categories called: "plastic

category, multimodal translation of actions, and enunciation" (D'Armenio, Deliège, and Dondero 2024: 6). These categories are applied while the viewer intends to semiotically study and describe the accuracy of image generation performed by AI. It begins with the plastic surface, which is a formal reading of the image, then moves to the next dimension to engage more deeply with the multimodal translation of signs into visual representations and their meanings. At this stage, viewers compare the linguistic instructions used for prompting with the generated image. Finally, the third layer of the study evaluates the enunciative articulation presented in the visual images to understand each sign and element within the image, as well as their connections to one another and to other signs invisible in the picture but existing in the surrounding sphere, which can influence the overall meaning of the image (D'Armenio, Deliège, and Dondero 2024: 14).

Having these three-layered structures to describe generativity in AI does not fully address the linguistic component of text creation; rather, it focuses on the second act of prompting, which concerns the image itself. From a linguistic perspective, I draw on the well-known sentence coined by Chomsky (1957) – '*Colorless green ideas sleep furiously*'– to illustrate how the use of generative AI's text-to-image prompting should be approached to get a hold on the linguistic composition for prompt activity. Additionally, the structure proposed by D'Armenio and colleagues is to set the stage for examples that demonstrate the human-AI interaction, based on a "comprehensive exploration of their impact on authority and visual rhetoric in the digital age" (Leone 2024: 432). Chomsky's example reveals that a phrase can be grammatically correct yet semantically nonsensical. This highlights the plastic surface, where verbal language is the enunciator's primary focus, acknowledging that proper sentence structure alone is insufficient and does not guarantee meaningful communication. In essence, while the organization of words provides a necessary framework for comprehension, it does not inherently convey the intended message. Therefore, a better approach is needed to promote generative AI visual simulations that transcend basic instruction by creating new layers of meaning that emerge through viewers' evaluation and interpretation (Dai, Suzuki, and Chen 2024). The intended meaning of an AI-generated image depends heavily on contextual factors such as prior knowledge, cultural cues, and the viewer's interpretive framework, much like Halliday and Culioli noted that sentence structure alone is insufficient to grasp the full meaning without context. The challenge of interpreting AI-generated images is therefore amplified: such visuals may incorporate symbolic or surreal elements, and their meanings are not immediately evident without additional explanation; their compositions may yield multiple, potentially conflicting interpretations that require systematic exploration.

Punctum. International Journal of Semiotics | **11:**02**:**2025
ISSN 2459-2943 | DOI: 10.18680/hss.2025.0021 | **punctum.gr**

On the one hand, a similar semiotic dynamic operates in the interpretation of AI images; on the other hand, visual arrangement and the inquiry into an image in isolation do not ensure that viewers will grasp the creator's intentional idea in the process of semiosis (the first step to enunciation). Evidently, the production of an image is necessarily followed by the significance of building its contents, and the next step is to guide the consumption of the image according to the context in which it will be used (Lacković 2020b: 450). Following this view, for effective recognition and consumption of AI-generated products and their use across a semiotic scope, it is essential to refine current methodologies and emphasize the generative potential of AI without neglecting the roles of human creativity, interpretation, and competence in controlling and guiding AI outputs. Thus, just as a grammatically correct sentence can be ambiguous or misunderstood without appropriate contextualization, a well-composed AI-generated image can be open to diverse interpretations that diverge from the enunciator's original intent. In summary, both linguistic and visual structures provide scaffolding for meaning but do not guarantee it. The first stage of scaffolding occurs when the enunciator creates a linguistic structure for the prompt to generate an image, while in the following, second stage, the viewer takes the product and uses various factors to consume and interpret what is seen, thereby creating a meaningful illustration. Ultimately, human agency and subjectivity foreground fruitful meaning and generate perspectives through interactive reflection on AI-generated content. This point brings us back to the beginning, where the scale of human-AI interaction and co-creation was explained by positioning humans as the ones who control the input and give instructions to the machine, while AI does the work of generating content. The division of labor between humans and machines can be represented as two sides of a scale, with one side exceeding the other (in this case, AI is doing more *work* in generating new visual content). However, when the question of interpretation and inquiry into what has been created arises, the human scale absolutely becomes *stronger* than the machine's. Particularly in the case of AI images, meaning-making is a dynamic, dialogic process involving the image itself, the viewer's perspective, and contextual clues, which are the very added substance that extends beyond mere linguistic structural elements and introduces us to a constellation of meanings. What should be the object of study forward is the extension beyond a simple descriptive analysis to a level of reading the visuals as signs can be illustrated in a polysemic[3] manner for "analysis of all socio-cultural practices involving generative AI" (Dondero, Alonso Aldama, and Leone 2025: 2).

[3] Also referred to as Polysemous: having multiple meanings. Retrieved from: https://www.merriam-webster.com/dictionary/polysemous (July 2025)

3. Going from 'describing' an image to 'reflecting' on it

The trajectory of Chomsky's UG theory parallels contemporary debates about AI across scholarly and practical fields. This parallel highlights a tension between descriptive frameworks and inquiry-based methods that require critical reflection on the subject matter. In this light, contents which are understood as signs can be developed, and the purpose of generativity is expanded within socially and culturally situated contexts. Crucially, such reflection and meaning-making processes remain grounded in human agency, while AI functions as a supportive tool. As mentioned before, humans occupy the dominant role in interpreting and shaping AI-generated content, and their place "is important in defining the meaning and intentions 'behind' image uses" (Lacković 2020b: 450). Now that humans are surrounded by the digital sphere, there is a strong emphasis on their role as users and on their competence in reading and understanding digital text. This ability is described as a reflective act, known as "the textual and reading experience" (Fadeev 2022: 87), which is essential in contemporary life, intertwined with AI algorithms. Reflection on immersive AI-generated images is essential because these images are created with human enunciation, yet no introductory course or method is available to analyse or reflect on them; nevertheless, they still carry aesthetic and ideological weight.

Dewey defines reflection as "the kind of thinking that consists in turning a subject over in the mind and giving it serious and consecutive consideration" (1934: 3). This meaning aligns with the goal of a deeper analytical approach to image interpretation. Reading images reflectively allows us to uncover the layered meanings embedded in visual culture, helping us understand the world more critically. Reflection, in this context, explores the power relations, dynamics, and ideologies that are often subtly embedded within images. For instance, a beauty advertisement may reflect dominant norms around femininity and whiteness, illustrating how visual media can reinforce societal ideologies. Kress and van Leeuwen (2021) support this approach by proposing a grammar to read visual signs with attention to Halliday's social semiotic model, which, as previously explained, suggests that meanings arise by the action of interpretation, inquiry, and reflection on signs in a social context. The goal of this grammar is to understand "how depictal elements –people, places, and things – combine in visual 'statements' of greater or lesser complexity and extension" (Kress and van Leeuwen 2021: 1). This reflective reading approach requires a shift from mere observation to deeper interpretation. This involves recognizing that visual signs are not neutral or passive exhibits but are active participants in semiosis. As Bertling (2019: 30-31) notes, this reflective process highlights the socio-cultural and relational nature of knowledge construction, where meaning emerges through collaborative and continuous reflection.

Punctum. International Journal of Semiotics | **11:**02**:**2025
ISSN 2459-2943 | DOI: 10.18680/hss.2025.0021 | **punctum.gr**

To apply the reflective grammar of reading to AI-generated images, this study adopts a comprehensive framework introduced by Gillian Rose (2022) (Figure 1). Her approach encourages moving beyond surface-level description to a deeper critique that includes contextual, ideological, and discursive inquiry. Rose describes this method as a rich pivot in visual analysis, where reflection serves as a bridge between seeing and knowing. By engaging in deliberate contemplation, observers move beyond mere perception of visual stimuli to a deeper understanding of their significance and context. Reflection transforms the passive act of viewing into an active process of interpretation, enabling insights that connect sensory experience with cognitive awareness. Consequently, this approach fosters a more nuanced comprehension of visual material, enriching both the analytical process and the knowledge derived from it. The four sites of her visual methodology highlight the importance of semiosis as the primary act for critically investigating different layers of an image and finding meanings according to the subjective interpretation of signs. Furthermore, she explains:

> Interpretations of visual images broadly concur that there are four sites at which the meanings of an image are made: the site(s) of the production of an image; the site of the image itself; the site(s) of its circulation; and the site(s) where it is seen by various audiences and users. I also want to suggest that each of these sites has three different aspects. These different aspects I will call 'modalities,' and I suggest that there are three of these that can contribute to a critical understanding of images. (Rose 2022: 47)

Although Rose's method was not originally developed for the study and interpretation of AI-generated images, this research demonstrates that her approach is highly effective as a semiotic-based methodology for investigating AI text-to-image outputs to better understand their contextual significance. It is essential to be equipped with the right method to engage with images reflectively, comprehending not only what is depicted but also how and why the image conveys its meaning. Reflection asks not just what is shown, but what is missing, who or what is left out, and why that matters. This shift is even more important in the age of AI, as it reframes the investigation of prompts from description to reflection and deeper understanding. The outcome of such a visual methodology in the study of generative AI images is a record of the visual elements presented, their critical evaluation, and an attempt to explain how and why the image works culturally in a polysemous manner. In this way, it would be easier to answer the main research question underlying that: *Do generative AI images engage in the complex, contextual, and inferential processes involved in human meaning-making, or do they simply reproduce surface-level visual patterns without reflective or semantic depth?*

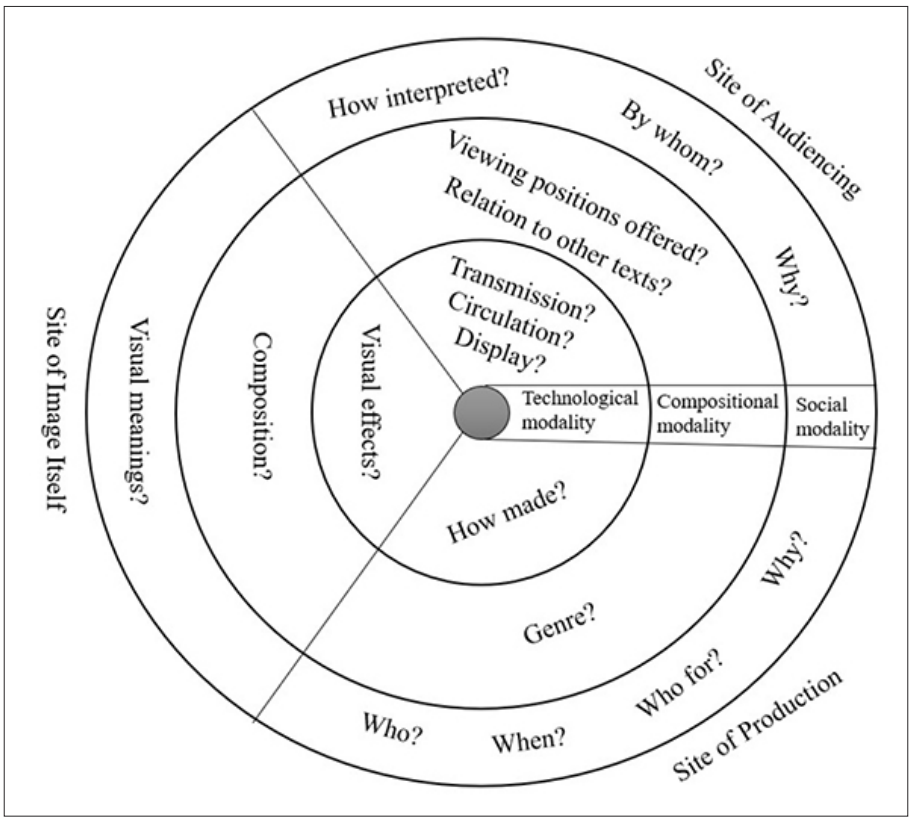

Figure 1. The sites and modalities for interpreting visual materials (Rose 2022)

4. AI-generated images: prompt examples and discussion

This section shows examples of images that were created by the generative AI platform Midjourney and explains how the inquiry was done according to Rose's (2022) visual methodology – separated from her work on semiology as a method– for reflection on the linguistic enunciation, its visual content output, and the socio-cultural and subjective interpretation that affect the context of the images. This approach is chosen to demonstrate the contrast between descriptive and reflective observation of viewers encountering AI-generated prompt content. This study specifically selects the generative platform Midjourney because it is "incapable of reasoning in a meta-semiotic way, namely, to operate a negation of an element expressed in the prompt" (Dondero 2025: 137). Another reason to use Midjourney iit offers artistic choice for non-artistic users who "want to explore their creative ideas but lack a strong grounding in traditional art approaches" (Hanafy 2023: 6). In this case, there will be no restriction in selecting who wants to perform in human-AI case studies that involve an interactive partnership with AI as a buddy. AI generative models, such as Midjourney, operate by integrating probabilistic patterns derived from large-scale training datasets to generate visually coherent images, without necessarily depicting a real-world reference.

Punctum. International Journal of Semiotics | **11:**02**:**2025
ISSN 2459-2943 | DOI: 10.18680/hss.2025.0021 | **punctum.gr**

In Figures 2 and 3, there are examples of text-to-image prompts created by Midjourney to demonstrate an illustration of the theories discussed; hence, to establish a comparison between the linguistic enunciation and the visual reading of the signs that is guided by Rose's methodology to understand the content (the visible signs existing in the image) and the context (the consumption of invisible signs and elements in relation to what can be seen). For this purpose, it is simply a matter of questioning, reflecting, and interpreting acts which are all part of humans' interactive role in connection with AI-generated images. Examples of questions about the content of the images, to first describe, then interpret, and finally reflect on the context in relation to real life, are listed below.

Figure 2. Prompt (Concept): *'a child sitting on a bench in a countryside and is looking away from the camera'*

- The first stage is the description of the foreground content of the images (production and image sites in Rose's methodology). The focus is on the visible signs and their meanings, which are more descriptive in nature, to ask critical questions about image production:

1) Who do you see in the image? Small child sitting on the bench? Is the child young? Who are the child's parents? The age of the child?

2) When is it? Sometime outside of late spring/autumn? Does it look like a nice day? The time of the year? Old or new image?
3) Where is the location? Somewhere rural/in the countryside?
4) What is the child doing? Is the child looking away from/at the camera? What are the outfits and colors that the child is wearing?
5) Why is the child alone? Is he/she waiting for someone (parents-friends) to return or play games? Why is the child smiling/sad? Why is the child outside on the road/in the forest?
6) How does the eye-level viewpoint position the viewer in relation to the figure?

• The second stage is reflection on the context of the images (circulation and audience sites in Rose's methodology). The focus is on invisible signs and guessing their connection to the elements in the images, following a deeper semiosis inquiry to understand the hidden layers of linguistic enunciation to interpret the contexts within the AI-generated images:

7) Where does the image come from?
8) How does the bench organize spatial relationships?
9) Why was it created and with what purpose? What was on the mind of the creator? What was the affective intention behind its production?
10) How does seasonal color function in this image?
11) Who consumes the image? Who are the intended audiences of this image?
12) What was the sentence used to create this prompt?
13) Is this a digital image? Have you seen such an image before? Do you think you can find it in Google stock? In what kind of media platform can it be used?
14) What aesthetic norms does the image reproduce as being created by AI?
15) Who is the source of information? Who is standing behind the camera? What do you think was it said to the child to sit in this position?
16) How does the visual style align with widely accepted ideas of "good" or "pleasing" imagery?
17) What kinds of visual diversity seem absent or minimized?
18) What would be other's reaction and consumption of this image? For example, what would a police office, teacher, or security guard think about this child in the image?
19) How does knowing the image is AI-generated change how you read its visual language?
20) What kinds of decisions are visible as machine decisions rather than human ones?

Punctum. International Journal of Semiotics | **11:**02**:**2025
ISSN 2459-2943 | DOI: 10.18680/hss.2025.0021 | **punctum.gr**

Figure 3. Prompt (Concept): *'a sustainable classroom to enhance the quality of education'*

The same process can be repeated for images in Figure 3, or any other prompt outputs to question the meaning and intention behind their creation by the enunciator and how generative AI is trying to represent them. Ultimately, a stronger, more interactive bond between humans and AI is possible, and by engaging with AI's generative platforms to imagine future scenarios, humans learn to envision different interpretations and to address semiosis in relation to real-world events and outcomes. A balanced approach to AI integration is essential, and it involves fostering a thoughtful human-AI partnership aligned with ethical principles and critical thinking. Human roles should be guided to view AI as a partner and to question its role in shaping knowledge, with reflection on issues such as bias, accuracy, and implications for real-world contexts. Changing the formation of reading images from descriptive to inquiry-based, which requires a true reflection on what is being observed, evaluation, and ultimately personal interpretation to generate meaning, would cultivate a critical approach that can increase awareness of visual discourse among humans across various contexts. By instilling the theoretical perspective advanced in this article, human-AI interaction moves closer to being reflective rather than descriptive, while using AI responsibly, empowering it as both a consumer and a creator in an ethically sound manner. This study emphasizes that meaning in AI-generated images does not reside only in what is visible but also in how and why certain visual configurations emerge. Inquiry – characterized by careful observation, analytical questioning, and contextual interpretation – enables viewers to examine the latent mechanisms underlying AI image production, including prompt engineering, model biases, dataset composition, and probabilistic generation. This reflective approach directly supports the theoretical claim that understanding AI-generated imagery requires more than technical knowledge; it necessitates cultivating interpretive skills that enable humans to critically engage

with AI-generated visuals across contexts. By learning to question how images are constructed, what assumptions they reproduce, and what visual norms they reinforce or disrupt, individuals can better understand AI agency and prepare humans to make informed decisions in a technology-driven, image-saturated world (Silvennoinen et al. 2024).

Culioli's view that meaning arises from contextual, inferential operations was presented in Rose's visual methodology for reading and critically questioning the concept of generativity in AI. Additionally, this method prompted the study to ask whether AI-generated images convey meaning beyond statistical mimicry. The purpose of our study is to advance future research on AI-generated image production through a reflective, multi-layered inquiry that yields a constellation of meanings, in which images, like narratives, unfold from abstract structures to concrete representations. This framework becomes essential when AI-generated images are inevitable, and there are no choices other than interpretations that evoke mood, narrative, or symbolic and iconic depth. By comparing these linguistic models with the processes behind AI image generation, the paper highlights a tension: AI can simulate visual syntax and style but struggles with the enunciative and epistemic dimensions of meaning. The transition from descriptive viewing (what is seen) to reflective interpretation (what is meant, implied, or felt) requires readers to actively negotiate meaning, which AI itself does not do; however, its outputs provoke such negotiation in human viewers. This calls for acknowledging the layered, semiotic nature of generative images and interrogates their aesthetic, cognitive, and communicative status in a human-AI ecology. In this way, development towards socio-cultural and semiotic awareness of generative AI and its role as a mediator and sign creator will be established.

The grammar of reflective reading of AI-generated images can be used in structures that draw on readers' experiences to highlight the necessity of an apt reading grounded in human agency (subjectivity). More precisely, regarding visual images, their generative style, and the interpretation of the signs at play is a critical approach that would entail the (meta)awareness of the possibility of different viewings and interpretations of the world, because "any genuine sign that *closes* on itself is simultaneously *open* to further interpretation and meaning-making through the paradoxical logic behind the self-referential dynamics of semiosis" (Semetsky 2019: 15). This view helps users of AI's prompts to be socially, culturally, and textually situated on the heavier side of the scale compared to AI in expressing their emotions and understandings in layers of semiotic reflections.

Punctum. International Journal of Semiotics | **11**:02:2025
ISSN 2459-2943 | DOI: 10.18680/hss.2025.0021 | **punctum.gr**

Conclusion

The aim of this study was to examine how the visual generativity of AI (text-to-image) is perceived and what can be critically done to advance reflective, deeper approaches to meaning-making in machine-generated images, focusing on a shift from descriptive to reflective interpretation. Drawing on linguistic and semiotic theories from Chomsky, Halliday, and Culioli, it explores how AI's image generation differs from human meaning-making. While Chomsky's formalist model privileges internal, rule-governed, and rigid structures, it aligns with the computational capacity and algorithmic foundations of generative AI, which similarly rely on statistical and rule-based operations to produce coherent visual outputs. However, such outputs often remain at the descriptive level, without reflecting on the level and interpretation of resemblance and semantic depth. In contrast, Halliday's socially grounded model, combined with semiotics, and Culioli's contextual and inferential approach, challenge the claim that AI-generated images convey meaning beyond statistical mimicry, emphasizing subjectivity, cognition, and context in the construction of meaning. To investigate, the study uses text-to-image prompting examples to explore how human-AI interaction partnerships produce visual content, questioning the concept of generativity and assessing whether AI images involve complex, contextual meaning-making. Findings suggest that meaningful knowledge construction and interpretation require shaping linguistic and enunciative processes in accordance with the language of images in AI-human interaction. Moving from descriptive to reflective interpretation requires that human viewers' key role in evaluating and interpreting generated content be more fully recognized and guided, to supplement AI's syntactic fluency with socially and cognitively grounded frameworks. This interdisciplinary perspective encourages critical reflection on how we read, use, and trust AI-generated visuals in knowledge creation.

AI text-to-image generated copyright statement by author

All AI text-to-image generated images in this article were created using Midjourney. To create images with Midjourney, you will need a subscription plan; the corresponding author purchased one on March 7, 2025. According to Midjourney's content rights, the creator of the images owns all created assets (Midjourney, n.d.-c.).
https://docs.midjourney.com/hc/en-us/articles/32083055291277-Terms-of-Service

References

Bertling, Joy G. 2019. Layered Collaborative Visual Reflection: Sharing Lived Experiences and Making Meaning. *Art Education* 72 (3): 28-38, https://doi.org/10.1080/00043125.2019.1578022

Boling, E., C. M. Gray, M. G. Modell, A. Altuwaijri, and J. Jung 2014. Learners Interpreting Instructional Images: Meaning-Making and Decision-Making Strategies. *Journal of Visual Literacy* 33 (2): 27-52, https://doi.org/10.1080/23796529.2014.11674716

Chapman, Siobhan and Christopher Routledge 2009. *Key ideas in linguistics and the philosophy of language.* Edinburgh: Edinburgh University Press.

Chomsky, Noam 1957. *Syntactic Structures*. The Hague: Mouton.

Chomsky, Noam 1968. *Language and Mind*. New York: Harper & Row.

Chomsky, Noam 2000. *New Horizons in the Studies of Language and Mind.* Cambridge: Cambridge University Press.

Chomsky, Noam 2002. *On Nature and Language*. Cambridge: Cambridge University Press.

Clarke, Eric F. 2009[1995]. Expression in performance: generativity, perception, and semiosis. In: John Rink (ed.) *The Practice of Performance: Studies in Musical Interpretation.* Cambridge: Cambridge University Press, 21-54.

Culioli, Antoine 1995. *Cognition and Representation in Linguistic Theory.* Amsterdam: John Benjamins.

Culioli, Antoine 2000[1989]. Representation, Referential Processes, and Regulation: Language activity as form production and Recognition. In: Antoine Culioli (ed.) *Pour une linguistique de l'énonciation Vol. 1, Opérations et représentations*. Paris: Ophrys, 177-213.

D'Armenio, E., A. Deliège and M. G. Dondero 2024. A semiotic methodology for assessing the compositional effectiveness of generative text-to-image models (Midjourney and DALL· E). In: Del Bue, A. C., Canton, J. Pont-Tuset, and T. Tommasi (eds.) *European Conference on Computer Vision*. Cham: Springer Nature Switzerland, 112-127, https://doi.org/10.1007/978-3-031-92089-9_8

Dai, D., S. Suzuki and G. Chen 2024. Generative AI for professional communication training in intercultural contexts: where are we now and where are we heading? *Applied Linguistics Review* 16 (2): 763-774, https://doi.org/10.1515/applirev-2024-0184

Punctum. International Journal of Semiotics | **11**:02:2025
ISSN 2459-2943 | DOI: 10.18680/hss.2025.0021 | **punctum.gr**

Dewey, John 1934. *Art as Experience*. New York, NY: Minton, Balch, and Company.

Dondero, Maria Giulia 2020. *The language of Images*. Cham: Springer Nature.

Dondero, Maria Giulia 2025. Semiotics of artificial intelligence: enunciative praxis in image analysis and generation. *Semiotica* 262: 111-146, https://doi.org/10.1515/sem-2024-0195

Dondero, M. G., J. Alonso Aldama and M. Leone 2025. Aspects of AI semiotics: enunciation, agency, and creativity. *Semiotica* 262: 1-3, https://doi.org/10.1515/sem-2025-0014

Dondero, Maria Giulia and Jacques Fontanille 2014. *The Semiotic Challenge of Scientific Images. A Test Case for Visual Meaning*. Ottawa, Canada: Legas Publishing,

Fadeev, Aleksandr 2022. Concept formation and the text in digital culture. *Punctum. International Journal of Semiotics* 8 (2): 83-99, http://dx.doi.org/10.18680/hss.2022.0017

Galofaro, Francesco 2013. Formalizing narrative structures: Glossematics, generativity, and transformational rules. *Signata: Annales des sémiotiques/Annals of Semiotics* 4:227-246, https://doi.org/10.4000/signata.757

Groussier, Marie-Line 2000. On Antoine Culioli's theory of enunciative operations. *Lingua* 110(3): 157-182, https://doi.org/10.1016/S0024-3841(99)00035-2

Haase, Fee-Alexandra 2023. Beyond Mona Lisa's Smile: A theoretical approach to the persuasion of likeness in the digital image. *Digital Age in Semiotics and Communication* 1: 157-173, https://doi.org/10.33919/dasc.22.5.8

Halliday, Michael A. K. 1978. *Language as Social Semiotic: The Social Interpretation of Language and Meaning*. London: Edward Arnold.

Halliday, Michael A. K. 2005. On matter and meaning: The two realms of human experience. *Linguistics and the Human Sciences* 1 (1): 59-82, http://dx.doi.org/10.1558/lhs.2005.1.1.59

Halliday, Michael A. K. 2013. Meaning as choice. In: Fontaine, L., T. Bartlett and G. O'Grady (eds.) *Systemic Functional Linguistics: Exploring choice*. Cambridge University Press, 15-36.

Hanafy, Nervana Osama 2023. Artificial intelligence's effects on design process creativity: A study on used AI text-to-image in architecture. *Journal of Building Engineering* 80, 107999, https://doi.org/10.1016/j.jobe.2023.107999

Hobbs, Renee 2024. *Media Literacy in Action: Questioning the Media*. London: Rowman & Littlefield.

Kress, Gunther and Theo van Leeuwen 2021[1996]. *Reading Images: The Grammar of Visual Design*. London: Routledge.

Lacković, Nataša 2020a. *Inquiry Graphics in Higher Education: New Approaches to Knowledge, Learning and Methods with Images*. Cham: Springer Nature.

Lacković, Nataša 2020b. Thinking with digital images in the post-truth era: A method in critical media literacy. *Postdigital Science and Education* 2: 442-462, https://doi.org/10.1007/s42438-019-00099-y

Leone, Massimo 2023. The main tasks of a semiotics of artificial intelligence. *Language and Semiotic Studies* 9 (1): 1-13, https://doi.org/10.1515/lass-2022-0006

Leone, Massimo 2024. Semiotics of the black box: on the rhetorics of algorithmic images. *Visual Communication* 23 (3): 426-451, https://doi.org/10.1177/14703572241247120

Lin, Francis Y. 2017. A refutation of universal grammar. *Lingua* 193:1-22, https://doi.org/10.1016/j.lingua.2017.04.003

Lin, Z., D. Pathak, B. Li, J. Li, X. Xia, G. Neubig and D. Ramanan 2024. Evaluating text-to-visual generation with image-to-text generation. In: Del Bue, A. C. Canton, J. Pont-Tuset, and T. Tommasi (eds.), *European Conference on Computer Vision*. Cham: Springer Nature, 366-384, https://doi.org/10.1007/978-3-031-72673-6_20

Louvel, Liliane 2013. From intersemiotic to intermedial transposition: Ex-changing image into word/word into image. In: Nancy Pedri and Laurence Petit (eds.), *Picturing the Language of Images*. Newcastle upon Tyne: Cambridge Scholars Publishing, 13-31.

McAdams, Dan P. and E. D. de St Aubin 1992. A theory of generativity and its assessment through self-report, behavioral acts, and narrative themes in autobiography. *Journal of Personality and Social Psychology* 62 (6): 1003-1015, https://doi.org/10.1037/0022-3514.62.6.1003

Mitchell, W. J. Thomas 1994. *Picture Theory: Essays on Verbal and Visual Representation.* Chicago: University of Chicago Press.

Okulu, Hasan Zuhtu 2025. Creating and evaluating instructional materials with generative artificial intelligence: visual representations in astronomy education. *Education and Information Technologies*: 1-20, https://doi.org/10.1007/s10639-025-13580-y

Peluso, Marina 2021. The functional approach, semiotics, and professional discourse. *Training, Language and Culture* 5 (2): 62-72, https://doi.org/10.22363/2521-442X-2021-5-2-62-72

Rose, Gillian 2022. *Visual Methodologies: An Introduction to the Interpretation of Visual Materials* (5th ed.). London: Sage.

Salthe, Stanley N. 1999. A semiotic attempt to corral creativity via generativity. *Semiotica* 127 (1-4): 481-496, https://doi.org/10.1515/semi.1999.127.1-4.481

Punctum. International Journal of Semiotics | **11**:02:2025
ISSN 2459-2943 | DOI: 10.18680/hss.2025.0021 | **punctum.gr**

Semetsky, Inna 2019. Visual semiotics and real events: Exploring the bricolage of images. *Video Journal of Education and Pedagogy* 4 (2): 90-110, https://doi.org/10.1163/23644583-00401006

Silvennoinen, J. M., K. Fenyvesi, T. Yada, M. Fisk and J. Oppenlaender 2024. Visualizing futures: children's co-created sustainability solutions with text-to-image generative AI. Available at SSRN 5014965. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5014965

Sless, David 1986. *In Search of Semiotics*. Totowa, NJ: Barnes and Noble.

Topalidou Laskaridou, A., N. Papadopoulos and D. Koutsogiannis 2022. Contemplating post-digital narrativity: Co-active, multimodal meaning-making on Instagram and its implications on learning. *Punctum. International Journal of Semiotics* 8 (2): 119-142, http://dx.doi.org/10.18680/hss.2022.0019

Ward, Sarah J. and Laura A. King 2017. Work and the good life: How work contributes to meaning in life. *Research in organizational behavior* 37: 59-82, https://doi.org/10.1016/j.riob.2017.10.001

AUTHOR

Seyedeh Maede Mirsonbol, PhD candidate in Semiotics and Media, University of Turin, Italy.