

## READING IMAGES - THE GRAMMAR OF VISUAL DESIGN (THIRD EDITION)

Authors: Gunther Kress and Theo van Leeuwen

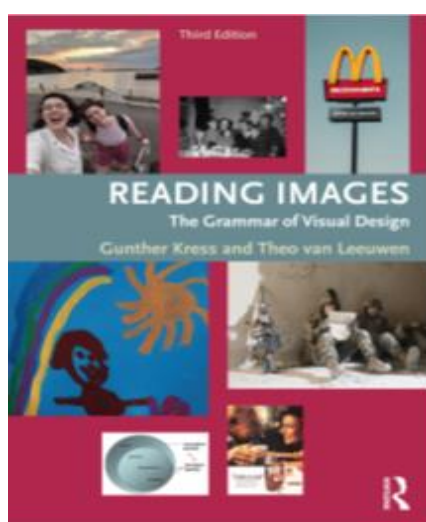
Publisher: Routledge, 2021

Nguyen Thi Thuy Linh\*

VNU University of Languages and International Studies, Pham Van Dong, Cau Giay, Hanoi, Vietnam

Received 22 February 2023

Revised 2 August 2023; Accepted 30 August 2023



Since the first edition was published in 1996, Gunter Kress and Theo van Leeuwen's *Reading Images: The Grammar of Visual Design* has received wide acknowledgment from the research community with over 18,000 citations on Google Scholar. The book has offered readers a systematic toolkit for the analysis of visual images from a social semiotic perspective. After nearly 30 years, the development of technologies has changed the ways of visual social practice, particularly in production and distribution of digital images. These deep changes in the social, the technological and hence the semiotic world led to the question whether the Visual Grammar framework is still applicable to the new semiotic resources of digital media. The third edition seeks answer to this question and indicates that with some refinements, the framework proposed in earlier versions remains valid and stands the test of time. Therefore, the major contributions of this new edition lie in the authors' expansion of the framework and its application to various contemporary images. This review will highlight these two issues by comparing the updated version with the previous ones and also discuss how the current edition deals with the critiques of previous editions. The information of the previous editions could be found in previous reviews like Forceville (1999) for the first edition and Thuy (2017) for the second edition, thus it is not within the scope of this present review.

### 1. A Comparison With the Old Editions

The third edition of *Reading Images: The Grammar of Visual Design* still keeps the original multidisciplinary principle with the "use of the *semiotic foundation* of a specific linguistic theory and aspects of its descriptions" (p. 41). It takes Halliday's social semiotic approach to language, with its three metafunctions as a starting point for the interpretation of visual images. According to Kress and van Leeuwen, similar to verbal language, visual texts

---

\* Corresponding author.

Email address: [linhnguyen804@gmail.com](mailto:linhnguyen804@gmail.com)

also fulfil three communicative functions, including representational, interpersonal and compositional ones.

Therefore, the structure of the book remains the same with eight chapters, following the three SFL metafunctions. The first chapter sets the theoretical background with the semiotic landscape. Chapters 2 and 3 deal with ideational meaning in the images by discussing visual resources for the representation of narrative and conceptual relations. Chapters 4 and 5 provide tools for designing the viewer's positioning and models of reality through a set of interpersonal design choices and validity markers. The last metafunction of compositional meaning is discussed in chapter 6, which examines how different resources are combined into coherent wholes and encodes intended reading paths. The last two chapters expand the framework to other modes like materiality, colors or dimensionality.

Although the original principle is still maintained, three major adjustments have been made, which shows the development of the theory over the time. Regarding the representational meaning, while the realization of processes and participants is still the same, a change can be found in the analysis of Circumstances. Earlier versions of the book strictly drew on Halliday's system with three equivalent subcategories of Settings, Means and Accompaniment. Meanwhile, in this third edition, no further classification is provided. Circumstances are simply defined as "participants that could be left out without affecting the basic proposition realized by the narrative structure" with the typical example of Setting (p. 70). Kress and van Leeuwen admitted the 'foreignness' of visual representation, and thus the irrelevance of linguistic framework in this case. The original Accompaniment and Means can either be "fused" into the main participants or classified as "Symbolic Attribute" to build up the identity of the participants (p. 70). This is not only a change in the analysis of Circumstances but also reflects the flexibility in the analysis of processes, which is further highlighted in the later section of Complexity in page 105. Different processes, which may include both narrative and conceptual processes, can be found in a single visual image and their relation is similar to that of clauses with coordination and subordination (p. 107).

Another important modification can be found in chapter 5 with the employment of the new term Validity instead of Modality in the previous editions. Modality was proposed by Halliday to refer to "the truth value or credibility of linguistically realized statements about the world" (Kress and Van Leeuwen, 2021, p. 150). This idea can be applied to visual images. However, while linguistic truth is based on probability and frequency, visual truth is decided by the idea of realism. The new term of Validity is used to highlight the different realizations of the truth in different semiotic modes. This subtle but important shift distinguishes Validity from Modality, which, in the social semiotic core idea, is based on "the values, beliefs and social needs of social group" (p. 154). The chapter then goes on with the discussion of eight validity markers, which, in fact, are similar to eight modality markers in the second edition. Therefore, this shift of terms mainly reflects the refinement of theoretical concepts rather than the actual analytical framework.

The meaning of composition also sees some adjustments in the subcategories, although the three elements of Information Value, Framing and Salience remain the same. Particularly in the last element of salience, a new issue of Reading path, which is designed by salience hierarchies, is introduced and discussed in detail.

In addition to the refinement of the framework, the accommodation to the changing visual landscape is another plus point of the third edition. The two previous versions were widely recognized for the variety of visual samples including photographs, paintings, cartoons,

charts and scientific diagrams, films, television, folk art, and even children's scrawlings. While many of these examples are still used to illustrate the theory, there is a shift to digital media with many updated images like websites, social media instances, cellphone user interfaces, and gaming platforms. This change is also reflected in the theory with its application to digital media data like the additional discussion on the validity in the digital age in chapter 5 or the analysis of visual composition of interfaces in chapter 6. These examples, interspersed with more traditional ones taken from the previous editions, show the practicality of Kress and van Leeuwen's social-semiotic theory in analyzing a range of both material and digital objects.

## 2. Critical Responses

The problem of categorization, one of the main critiques of the previous editions, has been addressed in this updated version. According to Forceville (1999), this framework provides readers with a practical toolkit for the analysis of visual images with hierarchical categories, which is useful to chart a new field, but at the same time, may also lead to the problem of categorization. "Categories are seldom clear-cut; many categories are fuzzy, and describe a continuum between extremes rather than a binary opposition with an either/or structure" (Forceville, 1999, p. 168); therefore, Kress and van Leeuwen was suggested to adopt a different view of categorization. Hence, the new edition has shown more flexible application of many categories in all three metafunctions. Many categories are analyzed according to the degree (more or less) of the feature, e.g. salience or subjectivity in attitude. Moreover, different categories or subcategories can exist in the same visual image. For example, one image can include both narrative and conceptual processes or can be designed in both left-right and center-margin in terms of information value. This flexibility may increase the accuracy of the interpretation of images.

However, this flexibility, in turn, may lead to the second problem of being "discursive" and "difficult to verify or disprove" (Bateman, 2008, p.46). Symbolic attribute can be a good example. Although the features of Symbolic Attributes are stated clearly in page 102, analysts may still feel confused in distinguishing Symbolic attributes from Circumstances. The question is how salient is enough to make a participant a symbolic attribute and separate it from the surrounding circumstance? For example, in Figure 3.27, the apple is interpreted as a symbol due to its "unnaturally large" size and "the Christ child holds it in a way that does not suggest he is about to eat it" (p. 103). However, with the given religious background knowledge, can analysts realize the symbolic meaning of the apple even without its unnaturally large size (i.e., the salience). This decision, in many cases, is subjective and bases on many socio-cultural and contextual factors.



Figure 3.27 (p.103)

Another evidence of this ambiguity can be found in the interpretation of representational meaning of, for example, a movie poster with the image of a couple hugging or kissing each other. According to Kress and van Leeuwen, advertising images (including movie posters) make frequent use of Symbolic attributive structures, in which human participants "usually pose in meaningful ways, rather than being shown as involved in some action" (p. 104). This

explanation is reasonable as it interprets the image as posing in front of the camera and thus regards this as a conceptual structure. However, the movie poster can also be a snapshot of a scene in the movie, thus, it can represent an action at that time and belongs to narrative structure. In this case, both interpretations seem acceptable and both are difficult to verify or disprove. The important point here is there is no specific system of indications stated in the book to help analysts make the right decision in such ambiguous cases. In this specific example of a movie poster, the context of production may account for the interpretation but unfortunately this information is unknown to ordinary analysts.

Furthermore, flexibility may also cause another problem of generalization, particularly in dealing with a larger corpus. The analysis, then, seems too complex to make a comparison among a larger number of samples and find prominent patterns in the corpus. This explains why the framework is often applied to reach the thorough and detailed description of a small group of visual images while research with corpus approach tends to seek for alternative or adapted version of framework (e.g. Bateman (2008)).

The last problem is the overgeneralization of visual resource choices in different genres. There is a good attempt throughout the book to cover many different genres with a wide choice of visual images as samples for analysis. However, some statements are made from the authors' observation without support from the literature. For example, they claim that the layout of textbooks increasingly employs the Center - Margin structure (p. 201) without any evidence from the analysis of any textbook or previous studies. This may lead to the similar problems of the previous versions when they claimed the rare use of Center - Margin structure and then have to correct the information in this third edition: "In earlier editions of this book, we wrote that centre-margin compositions are relatively uncommon. However, many of the examples we have shown, especially in Chapter 3, show that this may have changed, and that centred composition is, once more, becoming increasingly common" (p. 200). Given that multimodality is a new area of research, this problem is quite understandable, but it indicates the needs for more empirical studies to validate the features of multimodal genres.

In short, the book provides a systematic approach to analysis of visual elements with well-grounded theory, a practical toolkit with hierarchical categories and fruitful discussion of various samples from both material and digital objects. Despite the refined theory, the actual application of the framework to the diverse and complex structure of visual images may encounter difficulties in terms of categorization, thus becoming subjective and interpretative. Furthermore, the framework is particularly challenging in applying to a large corpus while there is an urgent need for further studies on multimodal genres.

### Acknowledgement

I'd like to express my gratitude to Gunther Rolf Kress and Theo van Leeuwen for their wonderful work in all three editions of this book, particularly to Kress, who sadly passed away in 2019. Their pioneering book has inspired and shaped my ideas on my PhD in the field.

### References

- Bateman, J.A. (2008). *Multimodality and Genre: A Foundation for the Systematic Analysis of Multimodal Documents*. Palgrave Macmillan. <https://doi.org/10.1057/9780230582323>
- Forceville, C. (1999). Educating the eye? Kress and Van Leeuwen's Reading Images: The Grammar of Visual Design (1996). *Language and Literature*, 8(2), 163-78.
- Thuy, T. T. (2017). Reading images - the grammar of visual design. *VNU Journal of Foreign Studies*, 33(6), 164-168. <https://doi.org/10.25073/2525-2445/vnuufs.4217>