Visual Images Interpretive Strategies in Multimodal Texts[*]

Jing Liu

School of Foreign Languages, Qingdao University of Science and Technology, Qingdao, China

Abstract—**A wide range of semiotic resources has been used to construct meaning. Visual systems of meaning offer different resources and potentials for meaning making. Readers need to familiarize themselves with a variety of ways to make sense and read visual images. This paper, drawing on the theories of semiotics, art and visual communication grammar, presents a framework of interpretive strategies to approach, analyze and comprehend the visual images in contemporary multimodal texts, so as to expand the readers' interpretive repertoires and strengthen their capacity in constructing and interpreting multimodal texts.**

Index Terms—**visual images, multimodality text, interpretive strategies**

I. INTRODUCTION

With the rapid advance of science and technology, the ways of human communication have changed greatly. We live in an increasingly visual culture, images, color and other non-verbal resources are no longer used mainly to entertain and illustrate, rather, they are becoming significant in communicating and meaning-making. Today people are increasingly exposed to texts that contain elaborate visual images, unusual narrative structures, complex design elements and unique formats (Goldstone, 2004; Kress, 2003; Serafini, 2011). We communicate through a variety of modes in which language is only one, if not essential part. The use of various semiotic resources in public communication has shown that meaning is realized not only through language but also through the integrated use of a wide range of semiotic resources including static and dynamic ones. In addition, the pervasiveness of visual images in computer interface, children's picturebooks, graphic novels, textbooks, magazines and advertisements require readers to simultaneously employ written text, visual images, and design elements to make meaning. In fact, little contemporary books and articles are not accompanied by visual images, Paul Duncum (2004) states, "…there is no avoiding the multimodal nature of dominant and emerging cultural sites" (P259). Kress (2000) makes the argument that the combined use of different modes to make meaning has gone to the point that it is now possible that when making sense of a text, even of its linguistic parts alone, we have to have a clear idea that some other features might be contributing to the meaning of the text. But the reality is that little attention has been paid to the non-linguistic resources. Therefore in making sense of the multimodal texts, moving beyond the traditional cognitive strategies and enhancing the readers' interpretive abilities is an important part of reading comprehension instruction. Images and texts are being combined in unique ways, and readers in today's world need new skills and strategies for constructing meaning in transaction with these multimodal texts as they are encountered during the social practices of interpretation and representation (Serafini, 2009).

II. MULTIMODALITY AND MULTIMODALITY TEXTS

Multimodality is a term widely discussed by linguists and semioticians.It means "the combination of different semiotic modes — for example, language and music — in a communicative artifact or event."(Van Leeuwen, 2005, P28) It also refers to the diverse ways in which a number of distinct semiotic resource systems are both codeployed and co-contextualized in the making of a text-specific meaning. (Baldry & Thibault, 2006) Therefore, it describes the grammar of visual communication that is used by image designers. It is an analysis of the rules and principles that allows viewers to understand the meaning potential of relative placement of elements, framing, salience, proximity, color saturations, styles of typeface, etc. (Machin, 2007) So we can see that every semiotic mode is a meaning momentum or potential, and multimodality mainly focuses on the study of the interrelationships between various communicative modes, no matter whether they are visual or auditory, words or image. It is a complex combination of meaning making activities that have undergone rapid changes in the contemporary social, cultural, economic and technological context. Moreover, the concept of multimodality is a useful yardstick to measure and evaluate the diversity ways of meaning making.

Multimodal texts, which convey information by means of various modes such as visual images, written language, design elements and other semiotic resources, are more complex than written texts. According to Kress, different logics

---

govern the mode of written language and that of visual image: written text is governed by the logic of time or temporal sequence, whereas, visual image is governed by the logic of spatiality, organized arrangements, and simultaneity (Kress, 2003). That is, meaning is derived from position in the temporal sequence of written text, whereas meaning is made from the spatial relations or grammar of visual images (Kress & van Leeuwen, 1996). To understand the written language, temporal sequence or order in which words appear in a sentence is very important, for example, the meaning of "John killed Smith" is quite different from that of "Smith killed John". In visual images, the position, size, and composition of the contents of the image play a significant role in the meaning making.

Multimodal discourse analysis is to analyze how several or all of the different semiotic modes intertwine together to create a unified text or communicative event. The premise of multimodal discourse analysis is that in many domains of contemporary writing, textual structure is realized, not by linguistic means, but visually, through layout, color, and typography both at the level of the "clause" and at the level of "discourse". Actually multimodal discourse analysis has become a new trend in the studies of discourse analysis, for it focuses upon the complete communicative aspects of discourse that emerge within interaction.There are many ways to do multimodal discourses analysis, such as content analysis, conversation analysis, social semiotic analysis and so on (Van Leeuwen & Jewitt, 2001). And different perspectives can be taken to analyze them, for example, layout, modality, typography, color, genre, discourse, style and so on are the angles we can choose to do multimodal discourse analysis.

III. VISUAL COMMUNICATION GRAMMAR

Halliday holds that language is a semiotic mode and any semiotic mode has to represent three communicative broad metafunctions, namely the ideational metafunction, the interpersonal metafunction and the textual metafunction, which has a decisive influence on Kress and Van Leeuwen's social semiotic framework of visual communication grammar. In Kress and Van Leeuwen's opinion, the three metafunctions of linguistics can be extended to visual communication. In *Reading Images,* they see image as a resource for representation and thus will display culturally produced regularities. The meanings expressed by people are the first and foremost social meanings, so Halliday's three metafunctions for language can also be used as a starting point for their account of images because they assume the three metafunctions model works well as a source for thinking about all modes of representation (Kress & Van Leeuwen 2006).

Kress and Van Leeuwen (1996, 2006) assume that image, color, music, typography and other visual modes are similar to language and they can simultaneously fulfill and realize the three broad communicative metafunctions as language does. In their view, image and other visual modes can represent objects and their relations in a world outside the representational system, so there are many ideational choices available for visual sign-making in visual communication. They also think that image and other visual modes have the capacity to form texts, complexes of signs which internally cohere with each other and externally with the context in and for which they were produced. Besides, image and other visual modes are able to represent a particular social relation between the producer, the viewer and the object represented. And all semiotic systems are social semiotic systems which allow us to negotiate social and power relationships. They see images of whatever kinds as means for the articulation of ideological position (Kress & Van Leeuwen 2006). That's why in their *Reading Images,* they draw examples from many domains, such as text books, websites, advertisements, magazine articles and so on to express their meaning making in their social practices. The key notion in any semiotics is the "sign", or "sign making", so Kress and Van Leeuwen discuss forms (signifier) such as vector, modality, gaze, composition, perspective, line and color, as well as the way in which these forms are used to realize meanings (signified) in the making of signs.

Based on Halliday's theory, Kress and Van Leeuwen use a slightly different terminology in discussing the meaning of image in visual communication: representational instead of ideational; interactive instead of interpersonal; and compositional instead of textual. As for the representational meaning, they have distinguished two kinds of image in the light of the different characteristics of image: one is narrative images which involve four processes: action process, reactional process, speech and mental process, and conversation process, another is conceptual images which include three kinds: classificational process, analytical process and symbolic process. Kress and Van Leeuwen have suggested three ways to examine the interactive meaning of images from three aspects: contact (demand or offer), social distance (intimate, social, or impersonal), and attitude (involvement, detachment, viewer power, equality, representation power etc.). The compositional meaning of images is realized through three interrelated systems: information value (given or new, ideal or real, important or less), salience (achieved through size, color, tone, focus, perspective,overlap, repetition, etc.), and framing. From Kress and Van Leeuwen's visual grammar, we can see that images are made up of elements that can be decomposed when we analyzed the meaning of them. Just as language, images have meanings only when they are integrated together. That is, the meaning of visual images comes from the arrangement of different visual elements. The three metafunctions put forth by Kress and van Leeuwen (1996) are not direct relationships between semiotic resources and meaning. The concepts such as power, interaction, detachment and involvement and so on are not meanings hidden in the images, rather, they are meaning potentials, that is, they are the possible meanings which will be activated by the producers and viewers of images (Jewitt and Oyama,2001).

IV. VISUAL LITERACY

Nowadays human beings make meaning in a variety of ways, so the traditional notion of 'literacy" should be extended beyond the confines of just being able to read and write. A contemporary definition of "literacy" is " the ability to access, analyze, evaluate, and communicate messages in a variety of forms" (Hobbs, 1997). Then what is visual literacy? Is it interchangeable with another term media literacy?

Shepherd holds the ideas that consumers themselves interact with media messages in the process of making meaning, he is quoted as stating: "Media Literacy is an informed, critical understanding of the mass media. It involves an examination of the techniques, technologies and institutions that are involved in media production, the ability to critically analyze media messages and a recognition of the role that audiences play in making meaning from those messages" (www.media-awareness.ca. 2002.p6). So we can see that media literacy mainly focuses on three major aspects: mass media, how and for what purpose messages are constructed and consumed by the masses.

For textual literacy, reading and writing ability is essential, similarly, the capacity to manipulate and construct meaning with images is significant to visual literacy. Anthony Pennings (2002) gives a more specific definition of visual literacy: "Visual literacy is an emerging area of study which deals with what can be seen and how we interpret what is seen. It is approached from a range of disciplines that: 1) study the physical processes involved in visual perception. 2) use technology to represent visual imagery, and 3) develop intellectual strategies used to interpret and understand what is seen" Based on the definition of general literacy developed by Hobbs (1997), Chauvin (2003) states that "visual literacy is the ability to access, analyze, evaluate, and communicate information in any variety of form that engages the cognitive processing of a visual image". Here visual images may involve body language, motion, dance, two and three dimensional works of art, photographs and clipart, films and video, museum exhibits and dioramas, advertisements, illustrated written or verbal discourse, architecture, hypermedia and visual reality experiences, and so on.

There are even more definitions. From these discussions we can see that both media literacy and visual literacy concern message or meaning construction. Media literacy put emphasis on the mass media which may be visual in nature such as television, magazines, and may be outside the scope of visual experience such as radio and recordings; while visual literacy mainly focuses on the symbolic aspects upon which the meanings or messages are based.

Just as reading and writing are essential to conventional literacy, the ability to construct meaning with images is a core component of visual literacy. The advance of technology has made it possible that the ordinary people are becoming increasingly professional visual designers, for instance, using software photoshop to alter an image digitally, and thus given everyone the ability to join the artists and craftsmen in using visual images to express ideas and thoughts that have always been inside of us in visual forms.

One point should be noted that when defining intelligence, we have come to value linguistic competence very much, and are not aware that visual competence actually precedes linguistic competence and we may say that language evolved through its connection to visual ability, visual literacy may help us to achieve the goal of linguistic literacy. "From nearly our first experience of the world, we organize our needs and pleasures, preferences and fears, with great dependence on what we see (Donis, 1973:1)

V. VISUAL IMAGES INTERPRETIVE STRATEGIES

Through the integration of existing and updated knowledge and the appropriate use of strategies, an active image reader may promote, monitor, coordinate, and maintain the comprehensive comprehension of the image. In analyzing the meanings in Renaissance art, Art historian and critic Panofsky (1955) devised a method often referred to as iconography and iconology. He created an list of various components of a piece of art works, and then identified the conventional meanings and the underlying philosophical meanings and interpretations constructed within the given sociocultural context. He held that there are three levels of meaning: The first level, preiconographic, is the denotative level, mainly focusing on the interpretation of the elementary or natural meaning, which involves the identification of visual materials such as the objects known from daily experience. The second level, iconographic, is the connotative level, mainly focusing on the interpretation of secondary or traditional meanings, and during the interpretive process, the viewers need to move beyond the natural or surface meaning of the image to consider their knowledge and experiences. The third level, iconological, mainly focuses on the interpretation of ideological meanings of an image constructed in particular social, cultural political, and historical contexts. Base on Panofsky's three strata method, a framework of visual images interpretive strategies is introduced to make sense of multimodal texts.

A. Meta-interpretive Strategies

The first and foremost interpretive strategy is the readers' meta-interpretive strategies, that is, the readers must have the awareness of multimodal meaning constructing. They must bear in mind that visual images are no longer ornamental and subsidiary, and are increasingly employed to make sense of the world, often overshadow the once dominant mode of written language. In multimodal texts, the visual semiotic resources are different from the ones drawn upon to create printed texts, and therefore bring with them different meaning making potentials. Different modes which evoke different responses from the reader, writer, listener, and viewer convey meaning differently, but they are complementary and joint contributors to an overall meaning that is more than the meanings conveyed by the separate modes. This meta-interpretive strategy may help the readers to take a more critical reading position, challenging and interrogating the structures and components that authors, illustrators, and designers use to construct meanings (Zammit,

2007).

B. Perceptual Strategy

Then comes the perceptual strategy, that is, what has the readers noticed: the visual and design elements (e.g. pattern, line, shape, color, typography, texture) presented in the multimodal texts. The readers may create an inventory of visual representation, and devise a vocabulary for enumerating and depicting these various elements, which is an important aspect of visual meaning interpretation. Lewis (1990) explained, "When we name things, we call them into being. We permit them to enter our consciousness but only in the garb in which we have dressed them". (P139) Perceiving and noticing the visual elements of a multimodal text is an initial and primary aspect of the meaning comprehension process.

For example, the reader may ask himself/ herself some questions such as:

What can you see from this visual image, or what are the constructing elements of this visual image? What are the format and the dominant colors of the images? Where is the text positioned? Within or beside the image? Or is it separated by borders or white space? What fonts are used? Are there any unusual and abnormal elements (i.e., elements that stand out or seem out of place)? How about the framing of the images? Are there thick borders or faded edges?

The comprehension of visual images always begins with the perception of the visuals that artists, illustrators, and graphic designers use to render a story and communicate to readers (Arnheim, 1974; van Leeuwen, 2005; serafini, 2011). If readers don't notice these particular elements, they may not be able to extract meaning from them during their interpretive processes.

C. Analytical Strategy

After noticing what is in the visual images and the design elements of a multimodal text, what these objects and elements mean is an important and key aspect of the following comprehension process. The readers may interrogate the representations through profound considering what these representations allow them to see, and simultaneously what they prevent them from seeing. The visual grammar of Kress and van Leeuwen (1996) provides readers with various perspectives for attending to and interpreting visual images. For example, how various objects are organized and located in the visual image and how they interact and coordinate with other elements; relative to the objects and participants in the image, is the viewer positioned very close or far away? And so on. The readers may ask the following questions:

What first catches your eye in the image, that is, what is foregrounded, and what is backgrounded? How about the size and scale of the image? Why are certain elements larger than the others? Does the larger element add to the meaning of the image? Are there any dominant color and what does it mean to the readers? How about the white space used in the image? How is the image framed, and how does this position you as a viewer, face to face, above or below, what does that mean? Are there any recurring patterns? Through the integration of lines, lighting, colors, contrast and gestures, what does the designer try to lead you to look at? What are the relationship between images and written texts?

In the analytical process, making sense of the interrelationships among various visual elements and understanding the meanings associated with them in a certain culture is essential to the interpretation of multimodal texts. Readers of multimodal texts need to understand how the various elements in visual images interplay and coordinate to make meaning if they are going to move beyond the literal perception of images and multimodal texts.

D. Sociocultural Strategy

Scott (1994) argues that, "pictures are not merely analogues to visual perception but symbolic artifacts constructed from the conventions of a particular culture." So in order to interpret the meaning of images thoroughly, the readers need to take the related social meaning system into consideration, that is, they must look not only at the relationships within the images but beyond the image itself to the current social, cultural, scientific, ecological and political background. Focusing on the social, cultural, historical, and political contexts of the production and transmission of visual images in multimodal texts is an indispensable aspect of the interpretation process. How the images affect us as viewers hinges on the larger social, political and cultural contexts in which they are viewed. "To explore the meaning of images is to recognize that they are produced within dynamics of social power and ideology....images are an important means through which ideologies are produced and onto which ideologies are projected". (Sturken and Cartwright,2001, P21).For example, visual symbols, such as a rose or a cross signifying love or Christian values respectively, are constructed in sociocultural contexts and used by artists to convey meanings beyond the literal level.

VI. CONCLUDING REMARKS

The problem of the twenty-first century is the problem of the image, according to cultural theorist W.J.T. Mitchell (1995). Successfully making sense of visual images in multimodal texts may pose a challenge to today's readers, which requires them to get familiar with art, media, and semiotic theories. And the readers are supposed to consider alternative interpretive strategies that involve the descriptions of the world presented, the visual and design elements used, the semiotic resources with meaning momentum or potential in sociocultural contexts, and the ideological influences that may form and alter one' perceptions and interpretations. Moving beyond the traditional comprehension strategies will help expand the perspectives and strategies readers may draw on to be fully literate when shifting from a

word-dominated typographic era to a post-typographic era dominated by the visual images and multimodal texts.

REFERENCES

[1] Arnheim, R. (1974). Art and visual perception: A psychology of the creative eye. Berkeley and Los Angeles: University of California Press.
[2] Baldry, A & Thibault, P. J. (2006) Multimodal Transcription and Text Analysis. London: Equinox.
[3] Chauvin, B.A. (2003). Visual or media literacy? *Journal of Visual Literacy*, 23(2), 119–128.
[4] Donis, D. A. (1973). A primer of visual literacy. Cambridge, M A: The Massachusetts Institute of Technology Press.
[5] Duncum, P. (2004). Visual culture isn't just visual: Multiliteracy, multimodality and meaning. Studies in Art Education: A *Journal of Issues and Research in Art Education*, 45(3), 252–264.
[6] Goldstone, B.P. (2004). The postmodern picture book: A new subgenre. *Language Arts*, 81(3), 196–204.
[7] Hobbs, R. (1997). Literacy for the information age. In J. Flood, S. Brice Heath and D. Lapp (Eds.). *The Handbook of Reservation on Teaching Literacy Through the Communicative and Visual Arts* (p7-14). New York: Simon & Schuster Macmillan.
[8] Jewitt, C. & Kress, G. (Eds.). (2003). Multimodal literacy. New York: Peter Lang.
[9] Jewitt, Carey, & Oyama, Rumiko. (2001). Visual Meaning: A Social Semiotic Approach. In T. Van Leeuwen & C. Jewitt (Eds.), *Handbook of Visual Analysis*. (pp. 134–156). London: Sage.
[10] Kress, G. (2003). Literacy in the new media age. London: Routledge.
[11] Kress, G. (2000). Multimodality in Cope, B. & Kalantzis, M. (eds.). *Multiliteracies: Literacy Learning and the Design of Social Futures*. London: Routledge.
[12] Kress, G. & van Leeuwen, T. (1996). Reading images: The grammar of visual design. London: Routledge.
[13] Kress, G., & van Leeuwen, T. (2001). Multimodal discourse: The modes and media of contemporary communication. London: Edward Arnold.
[14] Kress, G., & van Leeuwen, T. (2006). Reading Images. London: Routledge.
[15] Lewis, D. (1990). The constructedness of texts: Picture books and the metafictive. *Signal: Approaches to Children's Books*, 62, 139.
[16] Mitchell, W. J. (1995). Picture Theory: Essays on Verbal and Visual Representation. Chicago: University of Chicago Press.
[17] Machin, D. 2007. Introduction to Multimodal Analysis. London: Hodder.
[18] Panofsky, E. (1955). Meaning in the visual arts. Garden City, NY: Doubleday.
[19] Pennings, A. (2002). What is visual Literacy? Retrieved November 9, 2002, from Marist College Website: http//www.marist.edu/pennings/viswhatis.htm.
[20] Scott, Linda M. (1994). Images in Advertising: The Need for a Theory of Visual Rhetoric. *Journal of Consumer Research*, 21, 252–273
[21] Serafini, F. (2009). Understanding visual images in picture books. In J. Evans (Ed.), *Talking beyond the page: Reading and responding to contemporary picture books* (pp.10–25). London: Routledge.
[22] Serafini, Frank. (2011). Expanding perspectives for comprehending visual images in multimodal texts. *Journal of Adolescent & Adult Literacy* 54(5) :342-350
[23] Sturken & Cartwright. (2001). Practices of Looking: An Introduction to Visual Culture. Oxford: Oxford University Press.
[24] Van Leeuwen, T. & Jewitt, C. (2001). Handbook of Visual Analysis. London: Sage
[25] Van Leeuwen, T. (2005). Introducing social semiotics. London: Routledge.
[26] Zammit, K. (2007). Popular culture in the classroom: Interpreting and creating multimodal texts. In R. Whittaker, M. O' Donnell, & A. McCabe (Eds.), *Advances in language and education*. (pp. 60–76). London: Continuum.

Jing Liu was born in Qingdao, China in 1974. He received his M. A. Degree in linguistics from Ocean University of China, in 2006.

He is currently an associate professor at the School of Foreign Languages, Qingdao University of Science and Technology, Qingdao, China. His research interests include applied linguistics and multimodality discourse analysis.