ELSEVIER

Moving multimodality beyond the binaries: A response to Gunther Kress' "Gains and Losses"

Paul Prior

University of Illinois, Urbana, IL 61801, USA

Abstract

This response to Kress agrees that it is vital to identify gains and losses afforded by changing media but questions his reliance on binaries and periodization. It suggests a return to James Gibson's affordances; more precise analysis of semiotic objects; and sociohistoric theories that link semiotic artifacts, individual development, and social practice.
© 2004 Elsevier Inc. All rights reserved.

Keywords: Affordances; Literacy; Multimodality; Practice theory; Remediation; Semiotics; Sociohistoric theory

1. Introduction

In a series of publications, Gunther Kress has persistently championed the view that logo-centric and text-centric views are no longer adequate for those interlocking fields that deal with communication and its links to sociohistoric developments. Kress and frequent co-author, Theo van Leeuwen, have worked not simply to forward the visual, the digital, or some other mode or medium but to specify a semiotic perspective in terms of multiple modes and media. Among his many works, my personal favorite remains *Before Writing* (1997), which explored the sign-making activity of children as they enter into social-semiotic life. When we read this work in graduate seminars, his reflections on multiple instances of children making cars, especially those made out of pillows and other elements in the bedroom, always evoke one or more graduate students to respond (both to him and to my own inclusion of it in a course on "writing") that children's representations of cars (drawn, cut out, or built out of materials at hand) are not just related to the serious business of writing. These reactions prompt important reflections on the unit of analysis for literacy studies (and on why our local program is named writing studies), especially as we later look at Peter Medway's (1996) research on architects,

---

* *Email address:* pprior@uiuc.edu.

whose disciplined professional practice routinely involved making drawings and small-scale models of buildings and landscapes.

In "Gains and Losses," Kress rehearses several key arguments that I believe are, in some form, central to the present task of advancing our understanding of literate activity and artifacts. First, he reiterates the call for recognition of the ubiquitous nature of multimodality. Second, he reiterates his and van Leeuwen's argument (2001) that multimodality should lead not to a series of compartmentalized modal disciplines but to a field where semiotic principles cross-cut modes. Framing (or chunking) is an example of such a principle. Whether music, written text, talk, film, mathematical equation, painting, or .swf file, signs are framed and chunked to achieve certain meaning-potentials (and suppress others)—they are organized to signify, evaluate, relate, relatively foreground or background, and so forth. The specific devices or resources that frame or chunk signs in specific modes, or even in particular genres within and/or across modes, may vary (for example, spatial relations in visual images, volume in talk, rhythm in music, visual montage in film, texture in painting), but all can be understood as achieving goals related to segmenting and marking an ensemble of signs. Third, as in *Literacy in the New Media Age* (2003), Kress argues here that different modes have different affordances (a notion that should be traced back to James Gibson, who sought to describe ways that objective properties of things structure fields of potential for perception and action) and that it is not only critical for us to understand and evaluate the affordances offered by modes and media but also to act pedagogically and politically in light of those evaluations. Kress presents this paper then as a framework for pursuing the project of tallying up the gains and losses that the affordances of particular modes and media offer. In broad strokes, I believe these key arguments are not only valuable but they represent a critical set of challenges for those interested in understanding and shaping literate, or perhaps more accurately semiotic, practices and artifacts.

However, as Kress translates this ambitious program into a series of particular claims and suggestions for how to enact it, especially when he proposes a series of strong binaries of mutually exclusive affordances—each of which is associated with a particular mode—I find myself in sharp disagreement not only in terms of basic questions of theory but also often in terms of readings of particular texts offered as illustrations. I would highlight three problems, each of which derives from an interpretive practice driven, it appears, by a theoretical imperative to uncover powerful binaries.

First, Kress presents a semiotic history defined by periods; his attention focuses particularly on what he describes as a long period (several centuries) dominated by one stable constellation, essayist literacy packaged in the highly conventionalized and coercive formats of book and page, a period that he suggests has been unraveling over the last twenty or so years, eclipsed by the nascent age, still in flux, but dominated by a new constellation, the image and the screen. However, the past Kress evokes is a very selective past. Consider the following image from J. Martin Miller's book *An Official History of the Japanese-Russian War* that was published in 1904 and, thus, was well ensconced in what Kress has identified as the settled age of print literacy presented in traditional book formats (Figure 1).

The two pages contain four photographs, captioned, each of which is discussed in the text plus an image of a Japanese text (presented as an official pass and interestingly reproduced sideways). The layout of these two pages is complex and includes typographic highlighting
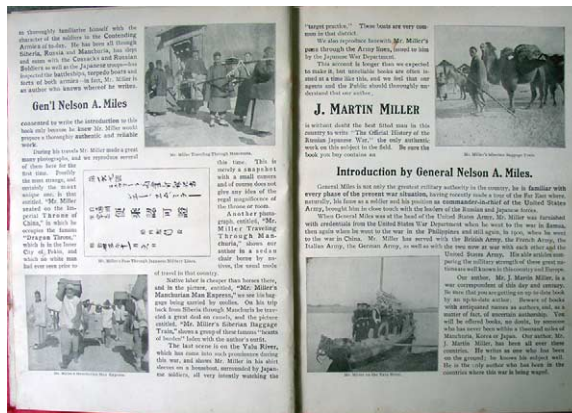
Fig. 1. From the introduction to *An Official History of the Japanese-Russian War*.

with centered, larger, bold typefaces, not only for a quasi-heading embedded in a sentence in the middle of the second page but also for two names, likewise embedded within sentences. The following page is no simpler, including facsimiles of letters (hand written and typed) from the United States Secretary of State and General Miles. The goal of this introductory part of the book is to establish the authority of the author by several means, first and foremost, by displaying that he was there for the events of the history and knew key players and second by authoritative endorsements of his credentials. This journalistic-historical version of being-there ethos is established multimodally. This book is not an especially remarkable artifact; it was among many books I inherited from a family friend's basement and it was the second old book I picked up in my basement (the first being a pictorial book on Ireland), yet it hardly appears to be a settled constellation of words on pages.

In *Graphic Design: A Concise History*, Richard Hollis (2001) presents a large number of texts from the first half of the twentieth century that could be read as a sustained counter-argument to Kress' claims that texts prior to the 1970s or 1980s were settled verbal packages. One particularly interesting image (p. 114) is of a 1944 training manual from the US Air Force on using a gun. The two-page spread involves a series of pictures (with shifting angles and perspectives) and text captions, linked across the pages by a meandering path of dotted lines and arrows.

Kress argues that written texts have long had a rigidly defined, conventional reading path and that the multiple entry points of, for example, the London Institute of Education webpage represent a revolutionary change, as visitors and/or readers have to make choices that would have already been made for them in a book. However, consider the following list of genres that have multiple entry points and typically self-organized reading paths: newspapers, magazines, dictionaries, encyclopedias, cook books, restaurant menus, quilts, hymnals, atlases, legal statues, university academic policy statements, product catalogues, tapestries, books of poetry, field guides to plants and animals, phone books, travel guides, the Bible, posters, product instructions, grammar books, and lists of all kinds (cf. Goody's, 1977 discussion of the significance of lists for understanding literacy). These counter examples, in short, are not quantitatively rare, socially restricted, or culturally trivial. They represent genres from religion

and science, politics and commerce, high art and popular culture, government and the home. It is clear that the settled age of print books written in certain linguistic registers (Kress stresses lexically and syntactically complex sentences) can only be sustained by sharply limiting the genres and/or texts sampled. None of this critique is meant to deny that we are in a period of rapid, even disorienting change, but simply to insist that we trace the past and project the future as precisely as possible, not falling back on prototypical representations that can only stand by obscuring so much of the semiotic landscape we set out to observe.

I would contrast Kress' views with those of Jay Bolter and Richard Grusin (1999), who offer up their own candidates for cross-modal logics, immediacy and hypermediacy, but who see these logics as ones that have been variously realized historically across media. Immediacy can be seen in linear perspective painting as well as virtual reality; hypermediacy appears in renaissance altarpieces and cabinets as well as hypermediated computer desktops. Like Kress, Bolter and Grusin also trace changes in newspapers. However, where Kress contrasts a text-only past (which requires careful selection of newspapers) to an increasingly visual present with modern newspapers apparently borrowing from computer screens, Bolter and Grusin argues for complex mutual effects, as computer screens borrow from texts and pages and texts and pages borrow from computer screens, and as television screens remediate ticker-tape news of the early twentieth century, and so on. In other words, where Kress sees a one-way sequence of unique semiotic objects, Bolter and Grusin see blurred, complex, and mutual relations. Seeking to define sociohistoric periods in terms of dominant communicational technologies is a project essentially identical to Walter Ong's (for example, Ong, 1982) and appears to recapitulate all the problems Ong's theories have offered. Periodization must erase (or discount) social and communicative hybridity to produce an image of homogeneity and, thus, tends to overdetermine modes of communication and their consequences around a small set of prototypical objects and scenes. The overall thrust of Kress' framework could easily lead us to a multimodal replay of the orality-literacy debates of the 1980s.

A second key problem I see relates to Kress' treatment of "affordances" as highly determinative, mutually exclusive, and binary. Gibson's (1979) basic notion of affordances was, in fact, intended to avoid turning objective properties of things into such hard categories. Gibson stressed that affordances are relational, ecological, and tendential (not determinative). For example, Gibson suggested that a post box "affords letter-mailing to a letter-mailing human in a community with a postal system" (p. 139), a formula that illustrates the way affordances of an object (the post box) are relational (the letter-mailing human user) and ecological (a community with a postal system). Gibson (1979) also stressed the Wittgensteinian fuzziness of categories, noting that objects such as hammers may afford—with relative ease or difficulty—a great many kinds of action. Kress, on the other hand, proposes a set of hard binary distinctions between words and images. Words in his account are finite, sequential, vague, conventional, authored, narrative and/or causal, and open to critique. Images are infinite, spatial, specific, natural and transparent, viewed, and available only for design.

Kress' attempt to describe modes in terms of mutually exclusive, binary affordances repeatedly leads to selective examples selectively read. Kress suggests, for example, that language is sequential (and associates sequence with narrative and causation) by comparing sentences like the following:

Bill hit Mary. vs. Mary hit Bill.
The sun rose and the mist cleared. vs. The mist cleared and the sun rose.

However, sequence is not alone responsible for the meaning effects. Consider further examples:

Bill hit Mary. Bill was hit by Mary. (No sequence change but a change in meaning.)
I put on my shoes and socks. I put on my socks and shoes. (Change in order but not meaning.)
I read my book, and Katie wrote in her diary. Katie wrote in her diary, and I read my book.
(Change in order of clauses, but no inference of causation or evoking of different worlds.)

Language is inflectional as well as sequential and the narrative causal interpretation of different sequences is part of a broader interpretative pattern (*post hoc* reasoning being one aspect, which can operate visually as well as verbally). Moreover, it is clear that images are also sequential. Film is one example, but print comics and process diagrams are other examples. In fact, in earlier work, Kress and van Leeuwen (1996) offered markedly more complex representations of images, distinguishing, for example, naturalistic images from diagrams. They noted that even single images may have preferred vectors for sequential processing, that images may be highly conventional (iconographies, international signage), and, most importantly, that many images evoke narrative and causal interpretations. Consider, for example, the narrative force of the recent Abu Ghraib images. The single image of a man, hooded in black, wired to pipes overhead, standing on a box, has, for many, renarrativized the US role in Iraq. Kress' binary oppositions ultimately threaten to undercut his broader project of finding multimodal principles or concepts variously achieved in particular instances of sign making, as it appears that, except for the principle of framing, word and image share no features.

In *Literacy in the New Media Age*, Kress (2003) offers an example of a page—a lesson on electrical circuits from a recent student science book (Fig. 9.7, p. 155)—intended to display how the affordances of modes are linked to historical semiotic periods. Kress suggests this page is emblematic of the screen age and image: though writing and image mix, his analysis argues that the modes are functionally specialized, that the words are relegated to the work of pedagogical framing while the images do the primary work of presenting the curricular content. In part, Kress states:

Image is used to represent that which is the issue, the core of the curricular issue here: what a circuit is, what the elements of a circuit are, how we think about circuits theoretically, and what circuits are like in practice. *That content does not appear in any part of the written text* [italics added]. (p. 155)

He is contrasting this example with pages from science books of earlier decades where, he argues, the primary curricular content was carried by the words. However, consider the following passages from the present-day textbook, found in the figure Kress chose:

Transistors and chips are examples of *semi-conductors*. They are made from special crystals like silicon. Transistors work because they conduct electricity in the right conditions. They are useful because they can turn on and off very fast, and they need very little electricity.. . . A transistor is a special semi-conductor. It has three connections: a base, a collector, and an emitter. When a small current is put on the base, it lets a much larger current flow between the collector and the emitter. So a tiny current can control a much larger one. (p. 155)

I submit that these passages, part of the written text in Fig. 9.7 that Kress described as carrying no part of the curricular content, do deal with what a circuit is, what its elements are, and other parts of the curricular issue. I agree that there is some functional specialization of modes and that exploring the contours of that specialization is vitally important, but I would argue that a finer-grained, more attentive analysis leads to a more complex rendering of what the words and images of this page are doing and of how their affordances are being harnessed for specific semiotic purposes.

Having argued in his paper that words and representations with words are tightly constrained by convention, whereas images and depiction are free and open-ended Kress (2005) sums up: "With depiction and images the situation is different: that which I wish to depict, I can depict. . . . I can draw whatever I like whenever I like to draw it. Unlike words, depictions are full of meaning; they are always specific". Again, I can only suggest that these remarkable claims are driven by an abstract theoretical commitment rather than by close and serious consideration of the actual semiotic landscape. I know of no evidence that those who work with images rather than words find the media and tools they work with so pliant, that they never need struggle with the recalcitrance of materials to their representational intents and desires. In short, I would suggest that the task of understanding multimodality and multimodal affordances is better pursued through more complex and less certain classifications and interpretive work (see, for example, Elkins, 1999; Rohan, in press; Wysocki, 1998, 2004).

Finally, near the conclusion, Kress (2005) states, "semiotics does not deal with learning; just as pedagogy or psychology does not deal with signs". Interestingly, this statement ignores the tradition of sociohistoric theory and research that traces back to Valentin Voloshinov and Lev Vygotsky, a tradition that, I would argue, offered a particularly rich set of resources for challenging another key binary for Kress, that of texts and practices. Voloshinov (1973) offers a semiotic theory centrally interested in concrete learning, of ways individuals appropriate and are appropriated by social and ideological systems of signs. Vygotsky's (1978) central insight was that human development, cognition, and action are all radically shaped by sign mediation. Work in this tradition, for example by William Hanks (1996) and James Wertsch (1991), offer a number of tools for pursuing not simply the multimodality of representational artifacts but also the nature of the multimodal practices that are as fundamental to understanding the production, reception, and distribution of semiotic objects as they are to understanding how multimodality is learned, changes, and shapes the sociogenetic (re)production of people and social formations. Recently, Jody Shipka and I (Prior & Shipka, 2003) have examined writing as a multimodal practice that involves structuring lives and environments as well as texts. Elinor Ochs, Sally Jacoby, and Patrick Gonzalez's (1994) analysis of physicists' multimodal work and Charles Goodwin's (1994) analysis of professional vision (in fieldwork and court room) points to the delicate ordering of text, talk, gesture, and image in situated practice. In "Gains and Losses," as in most of Kress' work, the focus on semiotic artifacts is matched by an almost total neglect of semiotic practices. (The notable exception to this neglect of practice seems to be in Kress' rich descriptions of the semiotic work of his own or other children around the home—which probably explains why *Before Writing* remains my favorite among his texts.) In any case, the hard binary of texts and practices in "Gains and Losses" and *Literacy in the New Media Age* means that distributed practices of production, interpretation and use tend to be relocated, as obligatory elements, inside texts themselves. I do not believe that we

can account for multimodality and affordances without a focus on the whole of practice—on artifacts, activity, and people alike.

While I have identified a number of concerns about Kress's current exposition of the gains and losses of multimodal developments in our semiotic landscape, I would conclude by returning to an area of relative agreement. Kress argues that it is critical to explore the affordances of different modes and media at this point because we are in a period of rapid and radical social, economic, political, cultural, and technological change, change reorganizing and realigning the uses and effects of modes and media. Without narrating these changes through a strict periodization, I believe that revolutionary change is afoot and that we must consider the complex, emerging affordances and consequences of semiotic practices, artifacts, and media carefully and precisely to understand and shape change. Indeed, it is because, as Vygotsky and Voloshinov recognized, semiotic practices so profoundly mediate human thinking and action, that, as Kress (2003) has argued, our investigations and interventions do need to be "entirely hard-headed and clear-sighted" (p. 175).

> **Paul Prior** is an Associate Professor of English and Associate Director of the Center for Writing Studies at the University of Illinois, Urbana-Champaign. His research explores connections among writing, reading, talk, learning, and disciplinarity. It has appeared in articles, chapters, and *Writing/Disciplinarity: A Sociohistoric Account of Literate Activity in the Academy* (Lawrence Erlbaum). Recently, he has co-edited with Charles Bazerman *What Writing Does and How It Does It: An Introduction to Analyzing Texts and Textual Practices* (Lawrence Erlbaum). Current projects include studies of the writing process as dispersed and embodied activity and of remediation of an online art object.

References

Bolter, Jay, & Grusin, Richard. (1999). *Remediation: Understanding new media*. Cambridge: MIT Press.

Hanks, William. (1996). *Language and communicative practices*. Boulder, CO: Westview Press.

Kress, Gunther. (1997). *Before writing: Rethinking the paths to literacy*. London: Routledge.

Kress, Gunther. (2003). *Literacy in the new media age*. London: Routledge.

Kress, Gunther (2005). *Gains and losses: New forms of texts, knowledge, and learning. Computers and Composition, 22* (1).

Kress, Gunther, & van Leeuwen, Theo. (1996). *Reading images: The grammar of visual design*. London: Routledge.

Kress, Gunther, & van Leeuwen, Theo. (2001). *Multimodal discourse: The modes and media of contemporary communication*. London: Arnold.

Elkins, James. (1999). *The domain of images*. Ithaca, NY: Cornell University Press.

Gibson, James. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.

Goodwin, Charles. (1994). Professional vision. *American Anthropologist, 96*, 606–633.

Goody, Jack. (1977). *The domestication of the savage mind*. Cambridge, UK: Cambridge University Press.

Hollis, Richard. (2001). *Graphic design: A concise history*. London: Thanes & Hudson.

Medway, Peter. (1996). Virtual and material buildings: Construction and constructivism in architecture and writing. *Written Communication, 13*, 473–514.

Miller, J. Martin. (1904). *An official history of the Japanese-Russian War*. n.p.

Ochs, Elinor, Jacoby, Sally, & Gonzales, Patrick. (1994). Interpretive journeys: How scientists talk and travel through graphic spaces. *Configurations*, *2*, 151–171.

Ong, Walter. (1982). *Orality and literacy: The technologizing of the word*. London: Metheun.

Prior, Paul, & Shipka, Jody. (2003). Chronotopic lamination: Tracing the contours of literate activity. In Charles Bazerman & David Russell (Eds.), Writing selves, writing societies: Research from activity perspectives (pp. 180–238). Fort Collins, CO: The WAC Clearinghouse and Mind, Culture, and Activity. Retrieved September 5, 2004, <http://wac.colostate.edu/books/selves_societies>.

Rohan, Elizabeth. (in press). I remember mamma: Material rhetoric, mnemonic activity, and one woman's turn-of-the-twentieth century quilt. *Rhetoric Review*.

Voloshinov, Valentin. (1973). *Marxisim and the philosophy of language* (Ladislaw Matejka & Irwin Titunik, Trans.). Cambridge, MA: Harvard University Press.

Vygotsky, Lev. (1978). *Mind in society: The development of higher psychological processes* (Michael Cole, Vera John-Steiner, Sylvia Scribner, & Ellen Souberman, Eds.). Cambridge, MA: Harvard University Press.

Wertsch, James. (1991). *Voices of the mind: A sociocultural approach to mediated action*. Cambridge, MA: Harvard University Press.

Wysocki, Anne. (1998). Monitoring order: Visual desire, the organization of web pages, and teaching the rules of design. *Kairos 3.2* (14 pages). Retrieved September 5, 2004, <http://english.ttu.edu/kairos/3.2/binder.html?features/wysocki.mOrder0.html>.

Wysocki, Anne. (2004). The multiple media of texts: How onscreen and paper texts incorporate words, images, and other media. In Charles Bazerman & Paul Prior (Eds.), *What writing does and how it does it: An introduction to analyzing texts and textual practices* (pp. 123–163). Mahwah, NJ: Lawrence Erlbaum.