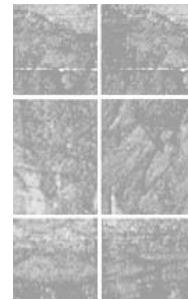


visual communication



ARTICLE

Travels in hypermodality

JAY L. LEMKE
City University of New York

ABSTRACT

The power of visual communication is multiplied when it is co-deployed with language in multimodal texts. In hypertexts, such as websites, the interaction of these two semiotic resources affords new forms of informational and design complexity and presents us with some new political choices. This article offers a semiotic scheme for the analysis of composite verbal–visual meanings and some discussions of the semiotic politics of visual communication in hypermedia design.

KEY WORDS

hypertext • language • multimedia design • multimodality • semiotics • website design

HYPERMODALITY

‘Travels in hypermodality’ playfully echoes Umberto Eco’s title *Travels in Hyperreality* (1990), which examined our fascination with imitation realities that are somehow more appealing than the everyday world. Eco wondered why anyone would prefer a Disneyland village to a real one, as we might wonder why a child prefers a doll to a playmate or a cartoon to a home movie. Some scholars today still wonder why anyone would prefer an illustrated text to a purely verbal one, or an interactive hypertext to a printed page.

Scholars are professionally obtuse. Designers, on the other hand, know very well that simplicity gives us welcome respite from the demanding complexity of everyday life and the less we’re constrained by realism, the more engaged our imaginations become. Both verbal text and visual images can be built to be more constraining of the meanings a reader makes or more enabling of the reader as a co-conspirator. Semiotic products can be designed to be more passive objects of contemplation or more active resources for the creation of further meaning. They can invite us to follow, or they can invite

us to lead. Good design builds in both functions, in varying combinations, depending on the known purposes of designer and client and the imagined, or fantasized, purposes of a prospective user (Kress and Van Leeuwen, 2001).

Hypermodality is one way to name the new interactions of word-, image-, and sound-based meanings in hypermedia, i.e. in semiotic artifacts in which signifiers on different scales of syntagmatic organization are linked in complex networks or webs. I propose here that one useful way to understand the design resources afforded by hypermodality is to consider multiplicative combinations of the *presentational*, *orientational*, and *organizational* resources of each semiotic mode (language, depiction/ imagery/graphics, and soundforms).

Hypermodality is more than multimodality in just the way that hypertext is more than plain text. It is not simply that we juxtapose image, text, and sound; we design multiple interconnections among them, both potential and explicit. In the simplest form of hypertext, we might have a web of 'pages' (or paragraphs, sentences, or even single words) in which the whole or some part of the page was linked to the whole or some part of another page (or even another part of the same page) in some way other than by the default sequential convention of ordinary reading. The links might be invisible, discoverable by exploring the technology that actuates them. They might be partially explicit (e.g. a unit marked visually as the source of a linking vector), but the target of the link, the nature of the meaning relation between source and target, and whether the link is reversible might not be explicit. Links make hypertexts multisequential (Aarseth, 1997). There are many possible trajectories, or *traversals*, through the web of a hypertext. Meaning on a time- and text-scale long compared to the typical scale of linked units (e.g. a paragraph or page) becomes a creation of the user/reader that is far less predictable to the designer than in the case of a printed book whose narrative or argument has a single conventional sequence.

It's important to be clear from the beginning in what sense hypertextuality differs from textuality. The difference is, first of all, one of medium. Typical meaning differences then arise because people exploit the affordances of one medium differently from those of another. A printed text is not itself truly linear or sequential as a medium in the sense that, say, spoken monologue is. Imagine a technology in which written words were presented to us visually one at a time. That would be a linear or singly sequenced written textual medium. We don't do it that way. We present written texts almost always, whether in handwriting, print, or on the computer screen, as a two-dimensional array, and we exploit that visual medium in many ways. We distinguish headers and sidebars from main text. We read clusters of words at least in a single horizontal line, and to varying degrees also up and down a page. We are visually aware of paragraphing and sectioning.

And our eyes wander. There are many sources of visual salience on a page, and just as our eyes traverse a painting or diagram according to salient features and vectors linking them (Arnheim, 1956), we may look away from

nearby words on a page to more distant words that are salient because of typeface (italic, bold, small caps), or by their recognizability (e.g. proper names or key words of interest to us), or because they happen to sit in a header or sidebar, or are the initial or final words of a paragraph or section. So do our interests wander, too. We do not always start a printed text at the title page, or the first paragraph of the main text. We may leaf through a book, glancing at this page or that; we may turn to an index, we may follow the page numbers in a table of contents, we may look from a line to a footnote, to a bibliography of references, to an author index, and back to another page. This is a traversal in the print medium, using the technology of the book, both materially (turning pages) and by means of its genre elements (page numbers, index, etc.). But it differs from the hypertext medium not simply because the technology is different – one could use the technology of hypertext to simulate a book in all these respects – but because the web of connectivity of a hypertext activates our expectations that there will be links out from any present text unit and that there will be no single default reading sequence of a main text to return to, or against which we should be reading the content of an excursus. In hypertext there is only excursus – trajectories and loops on different scales without a single unifying narrative or sequential development of a thesis.

Hypermodality is the conflation of multimodality and hyper-textuality. Not only do we have linkages among text units of various scales, but we have linkages among text units, visual elements, and sound units. And these go beyond the default conventions of traditional multimodal genres. Even on a single printed page of a magazine, newspaper, or scholarly article in the sciences, we know to connect certain graphical images with certain verbal units (via labels, captions, explanatory text) and vice versa (illustrations of narrative events, figures cited in the text). Organizational devices such as bounding boxes and nearness or juxtaposition combine with semantic content to indicate to us what goes with what across the modal divide between text and image. In hypermedia, there are more kinds of connection than those provided for in print genres.

A guided sequence

I have tried in the previous section to give a rough idea of what I mean by traversals and hypermodality. I have no more than suggested a number of technical distinctions and terminological niceties that a more systematic analysis would provide in detail. See Lemke (2002) for some distinctions among: the medium of hypertext, the technologies for implementing that medium, and the informational content of a particular hypertext web; the sequence of signifiers that constitute a trajectory through a hypertext web and the meanings made with those signifiers that constitute the traversal as such; the various scales or units of signifiers and meaning-making practices both extensionally (size) and in time.

The sequence of topics I have planned, if you read what follows in its default order, will first distinguish three kinds of meaning made by every semiotic act (presentational, orientational, and organizational), and then consider how meanings based on signifiers in different sign systems (language, depiction, music, etc.) can be combined or integrated to produce more specific and new kinds of meanings not otherwise available. I illustrate these principles with analyses of some website pages and then conclude with some issues of the politics of hypermedia design.

MULTIPLYING MODALITIES: PRESENTATIONAL, ORIENTATIONAL, AND ORGANIZATIONAL MEANING

If we are concerned with the kinds of meaning that can be made with hypermedia, we need to examine two kinds of resources that extend beyond the affordances of plain text. One of these is the semantics of hypertextuality, which will be considered in the next section. The other is the semiotics of multimedia, particularly the integration of verbal and visual resources for meaning.

I take the position that, fundamentally, all semiosis is multimodal (Kress and Van Leeuwen, 1996; Mitchell, 1994): you cannot make meaning that is construable through only one analytically distinguishable semiotic resource system. Even if for many purposes we analytically distinguish the linguistic semiotic system from that of depiction or visual-graphic presentations, and both from others such as the music-sound system or the behavioral-action system, the fact that all signifiers are material phenomena means that their signifying potential cannot be exhausted by any one system of contrasting features for making and analyzing meaning.

If I speak aloud, you may interpret the acoustical sounds I make through the linguistic system as presentations of lexical items, organized according to a linguistic grammar, etc. But you may also interpret them as indexical signs of my personal identity, social category memberships, state of health and emotional condition. Interpretation of my speech-sound-stream through the terms of the linguistic system alone does not and cannot exhaust its possible meanings in the community. If I choose to write down my words, eliminating those affordances of vocal speech that give rise to this supra-linguistic meaning potential, I must still create material signs, which now and again afford other ways of meaning: in handwriting there are many indexical nuances of meaning, in print there are choices of typefaces and font, page layout, headers and footers, headings and sidebars, etc. Each of these conveys additional kinds of meaning about the historical provenance of the text, its individual authorship, the state of the author (in the case of handwriting), the conventions of the printer, which parts of the text are to be seen as more salient, how the text is to be seen as organized logically, etc. – all through non-linguistic features of the visible text.

Furthermore, we can hardly help interpreting word-pictures with

pictorial imaginations, visualizing what we hear or read, whether as image, technical or abstract diagram, graph, table, etc. And conversely, when we see a visual-graphical image, whether a recognizable scene or an abstract representation of logical-mathematical relationships, we cannot help in most cases also interpreting it verbally. Language and visual representation have co-evolved culturally and historically to complement and supplement one another, to be co-ordinated and integrated (Lemke, in press a). Only purists and puristic genres insist on separation or monomodality. In normal human meaning-making practice, they are inseparably integrated on most occasions.

What kinds of meanings can be made by combining verbal, visual, and other signs from other semiotic resource systems? How do the meanings of multimodal complexes differ from the default meanings of their monomodal components in isolation? How do we construe the meanings of components in multimodal complexes and of whole complexes as such?

My basic thesis is that the meaning potential, the meaning-resource capacity, of multimodal constructs is the logical product, in a multiplicative sense, of the capacities of the constituent semiotic resource systems. When we combine text and images, each specific imagetext (Mitchell, 1994) is now one possible selection from the universe of all possible imagetexts, and that universe is the multiplicative product of the set of all possible linguistic texts and the set of all possible images. Accordingly, the specificity and precision which is possible with an imagetext is vastly greater than what is possible with text alone or with image alone.

That said, there are a few important qualifications. First, the existence of cultural traditions means that the probabilities for all possible combinations of textual items (on any text scale) with all possible visual features (on any image scale) are never equal. So the total information in any imagetext is always a great deal less than the maximum information possible if all combinations occurred with equal probability. It would only be in a culture in which language and image were entirely redundant, where there was one and only one picture that could be associated with each text, and one and only one text that could be associated with each picture, that this multiplicative model would not apply.

Consider also the issue of crossmodal translations. Even though a culture may create conventions about how, say, a painting is to be described in words, or commented on in scholarly fashion, or how a mathematical equation is to be graphically represented, text, image, and other semiotic forms are *sui generis*. No text *is* an image. No text has the exact same set of meaning-affordances as any image. No image or visual representation means in all and only the same ways that some text can mean. It is this essential incommensurability that enables genuine new meanings to be made from the combinations of modalities.

For meaning to actually multiply usefully across semiotic modalities, there must be some common denominators. At what level of abstraction can

we say that images and texts and other kinds of semiotic productions make meaning in the same way?

All semiosis, I believe, on every occasion, and in the interpretation of every sign, makes meaning in three simultaneous ways. These are the generalizations across modalities of what Halliday (1978) first demonstrated for linguistic signs, when considered functionally as resources for making meanings. Every text and image makes meaning presentationally, orientationally, and organizationally. These three generalized semiotic functions are the common denominator by which multimodal semiosis makes potentially multiplicative hybrid meanings.

Presentational meanings are those which present some state of affairs. We construe a state of affairs principally from the ideational content of texts, what they say about processes, relations, events, participants, and circumstances. For images, one could apply the same terms, recognizing what is shown or portrayed, whether figural or abstract (Kress and Van Leeuwen, 1996). It is this aspect of meaning which allows us to interpret the child's unfamiliar scrawl on paper through his use of the word 'cat', or his indecipherable speech through his pantomime of eating.

Orientational meanings are more deeply presupposed; they are those which indicate to us what is happening in the communicative relationship and what stance its participants may have to each other and to the presentational content. These are the meanings by which we orient to each other in action and feeling, and to our community in terms of point of view, attitudes, and values. In text, we orient to the communication situation primarily in terms of speech acts and exchanges: are we being offered something, or is something being demanded of us? Are we being treated intimately or distantly, respectfully, or disdainfully? We assess point of view in terms of how states of affairs are evaluated and which rhetorics and discourses are being deployed. The actual signs range from the mood of a clause (interrogative, imperative) to its modality (uncertainty, insistence), from markers of formality to the lexis of peer-status, from sentence adverbials (unfortunately, surprisingly) to explicit evaluations (it's terrible that ...). Visually, there is also a presumptive communicative or rhetorical relationship in which the image mediates between creators and viewers and projects a stance or point of view both toward the viewer and toward the content presented in the image.

Organizational meanings are largely instrumental and backgrounded; they enable the other two kinds of meaning to achieve greater degrees of complexity and precision. Most fundamentally, organizational resources for meaning enable us to make and tell which other signs go together into larger units. These may be structural units, which are contiguous in text or image-space, and usually contain elements which are differentiated in function (subject/predicate in the clause; foreground/background in image composition). Or they may be cohesive or catenative chains, which may be distributed rather than contiguous, and in which similarity and contrast-

within-similarity of features tie together longer stretches of text or greater extent of image as a unity or whole (repetition of words and synonyms; unity of palette).

In multimodal semiosis, we make crossmodal presentational (orientational, organizational) meaning by integrating the contributions to the net or total presentational (orientational, organizational) meaning from the presentational meanings of each contributing modality (Lemke, 1998b). Indeed, in many multimodal genres (and in all multimedia productions to some extent), the presentational (orientational, organizational) aspect of the meaning of a multimodal unit is underdetermined if we consider only the contribution from one modality. It may be ambiguous or unidentifiable or simply too vague and imprecise to be useful in the context of the next larger whole.

In the complexity of real meaning-making, there are further complications to this basic principle. First, *within* a semiotic modality, presentational, orientational, and organizational meanings are not by any means totally independent of one another. The possible combinations do not all occur with equal probability, and functionally each one helps us to interpret the others, especially in short, ambiguous, or unfamiliar texts or images. Human semiotic interpretation is both gestalt and iterative. That is, we recognize patterns by parallel processing of information of different kinds from different sources, where we are not aware of any sequential logic, and we refine our perceptions and interpretations as we notice and integrate new information into prior patterns in ways that depend in part on our having already constructed those prior, now provisional, patterns. It is well known in the case of reading a text of some length, that we form expectations about text-to-come and we revise our interpretation of text-already-read in relation both to new text we read and to the expectations we had already formed before reading it.

The viewing of images proceeds in a somewhat different fashion, but still undergoes similar processes through time. We may see a certain gestalt of a whole image, but if the image is complex enough in its details, if there are many scales of visual organization embedded within one another in its composition, then we will not have taken in all the details at first glance, nor will we have become aware of the many kinds of relationships, contiguous and at-a-distance, in and across the total image. We examine relationships within different scales of organization, and we move our attention along different pathways through the image until we have exhausted these possibilities and made provisional interpretations, which then lead us to examine still more details at various scales, through the iterative process which may, as with text, converge on some overall interpretation, or diverge into many possibilities, or simply be unstable.

I believe that it is customary in our culture to pay conscious attention primarily to presentational meanings, to orientational ones only in special circumstances, and to organizational ones only if you are a professional user

of the medium. We rely on familiarity with genre conventions to automate our use of organizational and orientational cues and to allow us to proceed directly to presentational information, at least in institutionalized use of media, where we are taught that it is only the presentational content which is important for institutional purposes. Such approaches, of course, are highly uncritical. They ignore power relationships, presupposing institutional roles. They ignore the limitations of genre conventions on possible new meanings. They increase a certain narrow kind of efficiency, and minimize the ongoing threat to the social status quo. As professional analysts and designers, we concern ourselves very much with organizational meaning in an instrumental sense: as means to orientational and presentational ends. We also pay attention to orientational meaning, but again very selectively. We may design rhetorical strategies, but we may not question our own role or imagine alternative possible relationships to the users of what we design. We are likely to adopt a particular evaluative stance toward our presentational content (desirable, likely, surprising, obligatory, usual), but we may not consider where that stance positions us in the social universe of discourses about these matters or in our social relations to others and their interests.

HYPertext SEMANTICS

By the semantics of hypertext, I mean the affordances of the hypertext medium for constructing meaning-relationships along traversals. This is the analogue of what might be called longer-scale text semantics in the case of more conventional verbal media. Just as we make meanings across many paragraphs or chapters that we do not make within a single paragraph or chapter, so we can make meanings in hypertext along long traversals (across say 10, 30, 100 or more units or *lexias*) that are not made in any one *lexia* or even across links between two *lexias*. We know relatively little about long-scale conventional text semantics (Lemke, in press b). I believe that on intermediate scales (dozens of paragraphs, say) we know two basic things: (1) that meanings are made through the nested embedding of structures on different scales, particularly genre structures and rhetorical-argumentative structures; and (2) that meanings are also made through extended cohesion chains (Halliday and Hasan, 1976) and by their intersections in structures at much shorter text scales (Lemke, 1995).

In hypertext webs that are rich in interconnection, it is possible to make use of genre-like structures, but difficult to hierarchically organize them sequentially. Cohesion chains, on the other hand, which are based on relations of similarity of units across extended text, work equally well in hypertext. It is possible in hypertexts for readers to gradually cumulate the wealth of details and the model of a richly drawn universe which we also get over long stretches of ordinary narrative or novelistic text. It is less easy to create compelling logical arguments (Kolb, 1997) or coerce the reader to agreement on matters of analysis and interpretation, but it is still possible to

raise for the user of a (branching, multisequential) hypertext web a great many interrelated questions that perhaps better reflect the complexity of real issues. Hypertexts can offer multiple perspectives, and even accommodate multiple authorship of a web.

I am not going to attempt here the task of developing a general, long-scale hypertext semantics, for either textual or multimodal hypermedia. I want only to sketch out some foundational notions and then see what happens in the webpage traversal examples I will provide. We need to begin, however, with at least an inventory of the shortest-scale crosslink meaning relations.

What kinds of meaning relations are typically construed across a single hypertext link? Such a link provides what is essentially an intertextual meaning relation, and we know (Lemke, 1995) that the kinds of meaning relations made between texts include those made within texts on longer scales. We also know that within texts there is a certain scale-invariance of meaning relations from the clause complex (Halliday, 1994) to the rhetorical formation (Mann and Thompson, 1986) and beyond. These basic relations are all specializations of the general semantic relations of Expansion and Projection (Halliday, 1994: ch. 7). In Halliday's model, one clause expands on another when, for instance, it re-states it, specifies it further, gives an example of it, adds related information, or qualifies some circumstantial conditions for it or consequences of it. A clause projects another when it introduces it as something said or meant. The relationships of expansion are the minimal ones for rhetorical structure; they are semantic generalizations of many relationships coded in the grammar by conjunctions and by those ubiquitous, if today somewhat obsolescent, abbreviations *e.g.*, *i.e.*, and *viz.* Projection was long ago recognized as being central to narrative structure by Bakhtin and his colleague Voloshinov (1986[1929]) in discussing the many varieties of direct and indirect discourse in Russian. What is useful about Halliday's model for our purposes is that these relations are all essentially *binary*, they can make local linkages of meaning without depending on the existence of larger structures (e.g. genre structures or extended arguments) that conflict with the openness of hypertext to alternative traversals.

I believe that as a first approximation the kinds of relations construed between consecutive webpages or hypertext lexias are these:

Presentational

- Links which tie one topic-specific set of semantic relationships to another in the same way such sets are internally connected (cf. thematic formations, Lemke, 1983, 1995); e.g. activity-to-actors, object-to-qualities, event-to-manner, activities linked by common actors and vice versa, etc.
- Logical relations of expansion and projection: restatement, specification, exemplification, commentary; addition, exception, alternative;

conditionality, causality, contextualization; quotation, opinion (Halliday, 1994: ch. 7)

- Rhetorical relations which further specify the logical relations, such as concession, opposition, disjunction, problem–solution, cause–consequence, proposal–evidence, events–generalization, etc. (Mann and Thompson, 1986)

Orientalional

- Offer & response (accept, consider, demur, decline, reject, counter-offer); demand & response (comply, refuse, etc.); more generally offer/demand-information, -action & response; degrees of offer/demand: entice, suggest, propose, insist, etc. (Halliday, 1994: ch. 4)
- State-of-affairs & evaluation (warrant, desire, importance, normativity, usuality, comprehensibility, humor); evaluative propagation chain elements; heteroglossic alliance/opposition (Lemke, 1998a)

Organizational

- Functional relations among the elements of such structures as: nominal group, clause, complex, rhetorical formation, genre (Halliday, 1994; Martin, 1992)
- Covariate chain element: similarity chain, co-hyponymic chain, co-meronymic chain; based on presentational or orientational features (Halliday and Hasan, 1976; Lemke, 1995)

As we proceed from lexia to lexia along a traversal, each linked pair may have many of the above relations. For each of the general classes of semantic connections listed above, there are corresponding visual principles and forms. In fact, many of the verbal relations can be conceptualized as visual metaphors (e.g. chains, multislot structures, narrative scenes, viewpoints). Nonverbal visual works of comparable complexity and scale show all these features, as for example in rich, traversable visual environments (e.g. architectural spaces, artist-exhibition spaces, designed landscapes, online gaming worlds), dynamic visual displays (e.g. silent films, animations, auto-scrolling displays, theatrical and dance performances), and extended static visual series or sequences (cartoon books, graphic novels, ukiyo-e print collections, long scroll paintings).

I now illustrate some of the principles I have put forward here regarding multimodality and hypertext semantics at least for the short-range scale of single pages, page-pairs across single links, and very short traversals of webpages. I draw on my experience of having analyzed these instances in the context of much longer traversals (e.g. Lemke, 2002b).

Let's begin with a fairly rich and complex page from the domain of

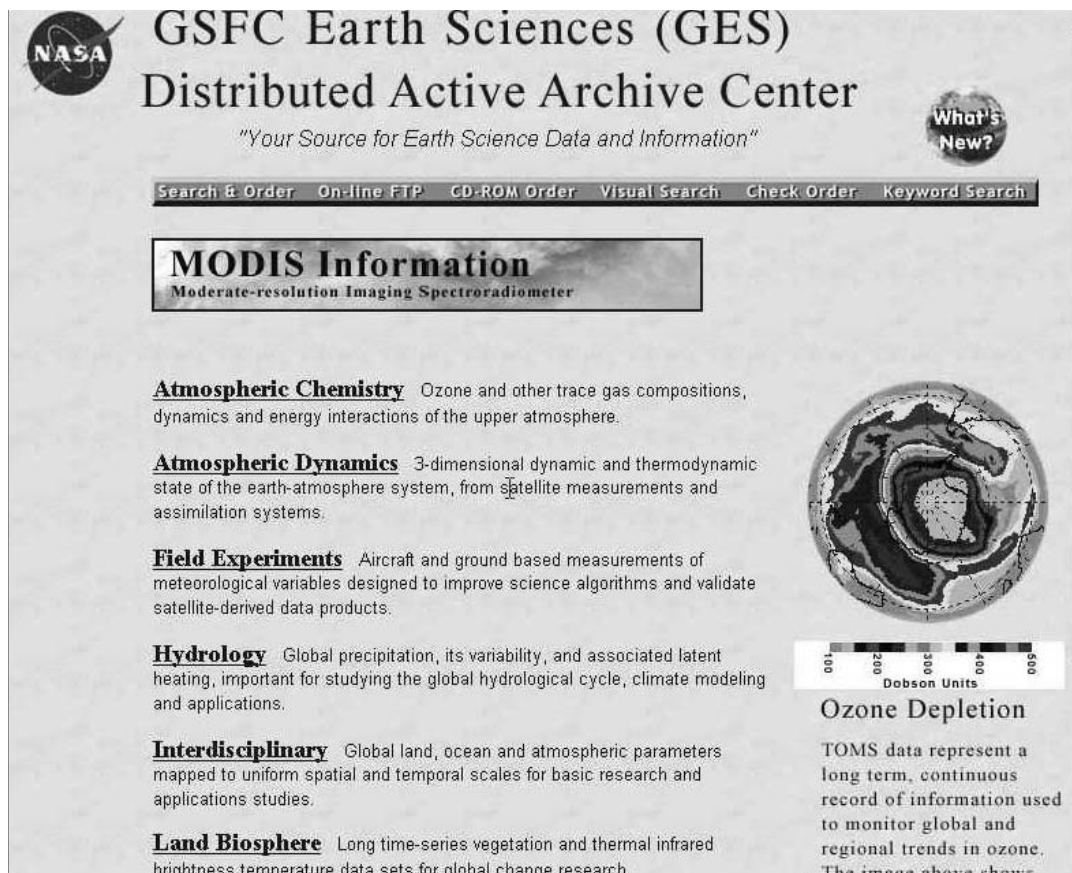


Figure 1 NASA Goddard Space Flight Center Earth Sciences Distributed Active Archive homepage. *Source:* GSFC Earth Sciences (GES) Distributed Active Archive Center. [http://daac.gsfc.nasa.gov/DAAC_DOCS/gdaac_home.html]

the U.S. National Aeronautics and Space administration (NASA) in a website at the Goddard Space Flight Center (GSFC), see Figure 1.

This page is a gateway to data collected by NASA satellites which record the spectrum of radiation reflected from the earth. These data can be interpreted as giving measures of the temperature, density of atmospheric ozone, degree of vegetation cover, etc. all over the earth, as seen from space. The primary content of the page is a menu of seven links to such datasets (see Figure 2) with brief textual descriptions and associated color images showing one important application of each type of data. This is a dynamic webpage: when the mouse is moved over the links (but not clicked), a new color image and caption text appear for each link.

The first link anchor is:

Atmospheric chemistry: ozone and other trace gas compositions, dynamics and energy interactions of the upper atmosphere.

The accompanying image is a polar-view globe map with a few dotted

longitude and latitude lines and is colored to represent data values. Below is a spectrum key, with color spectrum labeled in hundreds of Dobson units. The caption text has a header, Ozone Depletion, and reads:

TOMS data represent a long term, continuous record of information used to monitor global and regional trends in ozone. The image above shows the October 1996 Antarctic ozone hole as viewed by the TOMS ADEOS instrument.

The second link anchor in the menu is:

Atmospheric dynamics: 3-dimensional dynamic and thermodynamic state of the earth–atmosphere system, from satellite measurements and assimilation systems.

The image in this case shows flow patterns of atmospheric and oceanic ‘currents’ against a land–ocean topography. Flow lines are represented by arrows forming a vector field (Figure 2).

Let’s consider the page as a whole and the dynamic images and captions for these two links. I present a brief analysis of the textual–linguistic and visual–semiotic features of the text and their integrative interrelationships in terms of the three generalized semiotic functions common to both: presentational, orientational, and organizational.

It is often useful to begin a multimodal analysis with the visual–organizational composition of the whole page, because, as Arnheim (1956) noted, this is the salient structure which guides the eye in its traversals across the page.

There is a top frame with the NASA logo on the left, the title in large serif type, below the title a motto in small quoted italics, on the right an iconlink for What’s New. Below these is a single thin blue navigation bar dividing the top matter from the page content. Below the bar is a graphical link for highlighted information on a new satellite instrument (MODIS). It is a rectangle with a bluesky-and-clouds background overprinted with large and smaller type.

The main content of the page, all of which has a uniform sky-blue background, consists of the menu of textual link anchors noted above, each of which is both a link and an active mouse-over site, and each of which is followed by a descriptive gloss. Moving the mouse over a textlink causes a complex graphic image to appear in a frame at the far right, and below it, as a caption, a short explanatory text. These imagetext pairs illustrate the kinds of data to be found under the header link.

The overall organization here is very much like a typical printed page of text: horizontally aligned paragraph-like units of text, stacked vertically down the virtual page. The deviations from this are the two globe-shaped icons at top left and right, and the inset dynamic images with their caption

text below them in a separate 'column' to the right of the menu of textual links. You cannot normally see the whole length of the page at once, but you do initially see the first of the dynamic images. The scrollbar on the right indicates that there is more content on the page, below. It defines a vertical unoriented vector (up or down); this as well as the left-aligned edges of the textlinks, menu bars, and MODIS rectangle, and the conventional reading orientation (down the page, especially for the caption text), impose a vertical ordering on a page dominated visually by discrete horizontal elements.

The page has a compositional unity deriving from this horizontal-and-vertical-grid structure, and this unity is reinforced by the uniform sky-blue background and the blue color tones of the text, menu bars, NASA logo, What's New globe icon, and MODIS rectangle. This uniformity throws into salient relief the non-blue colors of the dynamic images. The first, titled Ozone Depletion, shows yellow as its primary color, with secondary green, central pink and white, a lesser area of dark red and a tiny bit of sky blue. All these colors also appear in a color-bar below the image, but in extremely small tokens. This color contrast sets off the dynamic image from everything else on the page, including its own caption which is similar in color but slightly larger in typeface. The eye is strongly drawn to the dynamic images, and as the user discovers that the images change with movement of the mouse down the textlink menu, there is set up another dynamic intersection of horizontal shifts of attention back and forth between the textlinks and the images with the vertical up and down exploration of the page's affordances and returns to previously seen images.

Each of the other dynamic images also includes an element of sky-blue, and there is both a sort of tying together (visual cohesive tie) and an alternation of the eye back and forth between these blue elements and the dominant blue of the rest of the page. The blue elements in the images take on a special prominence because of the color context of the rest of the page.

The use of color to produce a cohesive tie which has a definite semiotic significance is clear in the Monsoon Dynamics image (associated with the second textlink, see Figure 2). Here, apart from the use of blue to indicate two of the most thematically important elements (the blue ribbons for air mass drift and the blue arrows for wind movement), there are two yellow ribbons which are not otherwise visually connected to one another, but which are meant to represent movements of air masses with some similar (not noted) property, in contrast to those represented in blue.

The twin organizing principles of both text and visual images are (1) structural and (2) cohesive. The former is the familiar principle of wholes and parts (e.g. lists, figure-and-caption); the latter that of similarities of salient features (e.g. color, or typeface).

The textual units here are structurally organized into headers or titles or link-names and explanatory extended nominal groups. The structural relation is Halliday's Expansion:Elaboration, here realized paratactically (Halliday, 1994). The cohesive organization of the list is based on co-

Moderate-resolution Imaging Spectroradiometer

Atmospheric Chemistry Ozone and other trace gas compositions, dynamics and energy interactions of the upper atmosphere.

Atmospheric Dynamics 3-dimensional dynamic and thermodynamic state of the earth-atmosphere system, from satellite measurements and assimilation systems.

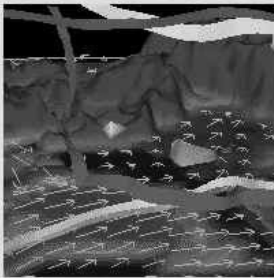
Field Experiments Aircraft and ground based measurements of meteorological variables designed to improve science algorithms and validate satellite-derived data products.

Hydrology Global precipitation, its variability, and associated latent heating, important for studying the global hydrological cycle, climate modeling and applications.

Interdisciplinary Global land, ocean and atmospheric parameters mapped to uniform spatial and temporal scales for basic research and applications studies.

Land Biosphere Long time-series vegetation and thermal infrared brightness temperature data sets for global change research.

Ocean Color Remote sensing ocean color data used to investigate ocean productivity, marine optical properties, and the interaction of winds and currents with ocean biology.



Monsoon Dynamics
3-D perspective of flow patterns during the Indian Monsoon, as computed by the GEOS-1 Data Assimilation System at GSFC. Air parcel trajectories (yellow and blue ribbons) and low level wind vectors are shown for June 1998.

Tools Documents Links Education Information

Goddard DAAC User Services: 301-614-5224 or 1-800-794-3147 -- daacuso@daac.gsfc.nasa.gov
 Web Curator -- web-curator@daac.gsfc.nasa.gov
 NASA Official: Steve Kempler, DAAC Manager -- kempler@daac.gsfc.nasa.gov
 Last modified: Mon Dec 20 16:00:34 EST 1999

Figure 2 NASA GSFC DAAC homepage, Atmospheric Dynamics link, Displaying Monsoon Dynamics. *Source:* GSFC Earth Sciences (GES) Distributed Active Archive Center. [http://daac.gsfc.nasa.gov/DAAC_DOCS/gdaac_home.html]

hyponymy (Halliday and Hasan, 1976), or more generally co-thematic collocations, with an intertextual basis (knowing the earth system division into *lithosphere*, *hydrosphere*, and *atmosphere*, supplemented by the cross-cutting *biosphere*).

There are also other very direct ties between principal elements. For example, the first two textlinks are 'Atmospheric Chemistry' and 'Atmospheric Dynamics', corresponding closely to the thematic complementarity of chemistry and physics. The former is cohesive with 'ozone' in its image's title, the latter with 'dynamics' repeated in its image's title. These lexical semantic ties reinforce and are reinforced by the visual ties noted above, and by the interactive connection of the dynamic change of images with movement of the mouse over the vertically stacked textlinks.

The cohesive linguistic ties, whose function here is organizational, are of course also functioning presentationally (thematically) to help construct the semantic relationship between the words of the textlink and its elaboration on the one hand and the thematic meaning of the image caption

on the other. There is a very high degree of lexical cohesion of the strongest type (word repetition, identity chaining, reference chaining) between these, e.g. in the case of the 'Monsoon Dynamics'. The overall semantic relationship is one of re-statement, with some specification or subcategorization, and particularly the semantic relationship of Exemplification (dynamics: monsoon:India).

How is the presentational semiotic link built between the textlink phrases and the visual image itself? First, there is the supporting organizational connection, implying that there is a presentational or thematic relationship. Thus the lexical identity from the linktext to the caption title ('dynamics') combines with the visual organizational convention of the genre (structural:caption-to-figure), to connect to the image as such. When we interpret the image thematically, we do so in relation to the caption text. We identify 'flow patterns', 'trajectories ... yellow and blue ribbons' and 'wind vectors' visually in the image. The other element in the image, inter-pictorially recognizable as mountains or topographic relief of some sort, cohesively links back to the elaboration of the textlink at left: 'dynamic ... state of the earth-atmosphere system', in which the other elements are 'atmosphere' and the mountains 'earth'.

But what of the orientational meanings here? How do the verbal and visual signs combine to produce (a) a stance to the viewer/user, and (b) a stance toward the thematic content presented?

The page as a whole has an explicit orientation: 'Your Source for Earth Science Data and Information' in italics just below the page title (see Figure 1). The second-person address here defines a Source-User relationship and makes a proposal, an offer (Halliday, 1994). But the page also makes demands, most clearly that you use the mouse to reveal the images and captions, and also that you click on one of the links to proceed to those and only those pages the designer has provided links to (apart from the default Back operation, or options afforded by the browser rather than the page). The sudden, and unexpected, appearance of the Monsoon image at right clearly makes a demand on the user's attention. These are demands for action, not for information. The caption text, when it refers to visual elements in the image ('yellow ... ribbons'), also makes an implicit demand that we match these to visual elements.

Semantically, the textlink and its elaboration and the caption text present pure offers of information. As such, we should consider the putative source's stance toward this information: its warrantability (high, there is no qualification of probability), its desirability (unmarked, except that monsoons may be considered dangerous, and so information about them useful and desirable), its normativity (no proposal, so no issue of obligation), its usuality (monsoons are expectable, the conventions of scientific imaging of the data presented also as of unmarked usuality), importance (monsoons are important events, information about them presumably also), comprehensibility (high for the source, the data and image help us

understand monsoons, but may be confusing or mysterious for the user, depending on expertise and familiarity with these terms and conventions), and seriousness (unmarked, high for this genre and register).

Visually, we may ask how the image itself expresses such stances.

Warrantability: this is not a photorealistic image, it does not have high visual resolution or realistic detail. It is an abstract representation, which we infer (from the thematic content of the caption and intertextual knowledge) is based on satellite data and faithful to those data, which would make the warrantability high, even though purely by visual conventions it might not seem so (Kress and Van Leeuwen, 1996). We can make comparisons with the other dynamic images. In the Ozone Depletion image, also abstract, there is a quantitative coding that adds a high degree of warrantability, and this coding is presented visually and numerically (for a similar instance, see the NDVI page analysis below). The image represents point for point real scientific data. It is not just an imagined scenario. In the Monsoon image we might still have some doubt about the exact contours of the landscape shown, because conventionally such images abstract and simplify. The text-warrant in the caption is only for the accuracy of the 'flow patterns', leaving this issue open.

Desirability: nothing in the image itself shows a 'monsoon' as a dangerous storm or flood-causing rain, except for the choice of a *black* background for the sky. This choice is presumably purely conventional, not dictated by the conventions of scientific representation of data on air flows. I think we can plausibly interpret it as a marker of the dangerous or serious or undesirable nature of the monsoon, and so indirectly also as reinforcing the sense of the importance of having NASA data to understand it.

Normativity: there does not seem to be an 'ought to' element in this image, though there is a demand of sorts, that the eye follow the arrows (there are so many of them, almost all aligned in the same direction). The other orientational stance to the viewer here is signaled by the overhead view: that we see this as from above (from the satellite, but much closer up), with a 'God's eye' view, which is culturally associated both with warrantability (from on high you see more of what is) and with science (an objective view from above the fray, of what is true). So the Source–User relationship is being visually constructed here as one of a giver of information to one who needs this information, one who is in a position to know, to one who is not but needs the knowledge. One who can guide understanding to one who needs to follow that guidance (e.g. the image's arrows, but also the links themselves). So there is a kind of normativity here, that the user ought to trust and use the knowledge the source offers.

Usuality: this is not a very conventional image for most viewers, though it may be for specialists, and it is not so unusual for anyone with scientific training. Within the image itself, there are non-naturalistic features in contrast with the naturalistic ones, mainly the ribbons and arrows, which are the features explained in the caption. The fact that they need explanation would seem to indicate that the source regards them as at least to some

degree unusual. In this respect, usuality is linked to Comprehensibility: what is unusual is also often less comprehensible. The explanation makes these features more comprehensible, but cannot make them more usual. The degree of comprehensibility also tends to define the implied user: someone for whom the explanation given is adequate.

Finally, Importance and Seriousness: the color choice in the Monsoon image is a palette that might be associated in other genres with a cartoon image, and so not fully serious. This is clearly not the case here, and seriousness is indicated in part by the many arrows, suggesting a systematic pattern, not a playful arrangement. How does the image show the importance of its content? Not, it seems, by attention to photorealistic detail, and if we compare with the other images, not by the use of numerical codes either. Importance here seems mainly to be conveyed verbally through the connection to 'monsoon', as it is also in an image that is titled 'Ozone depletion' (intertextually marked as important in the community).

I now pass on to a series of three webpages from a different website within the large NASA domain, the Earth Observatory site. These pages are intended for a less technically expert or scientifically knowledgeable public, as we can see by comparing its much-reduced use of technical language and complex visual displays with the pages we have just analyzed.

The Earth Observatory is a website with many, presumably deliberate, parallels to the GSFC site. Where the GSFC site is intended to allow technically expert users to access large scientific datasets of satellite images, the Earth Observatory site provides a similar but greatly simplified service for students, teachers, and the general public. The page displayed here (see Figure 3) is the dataset portal page, very similar in content and design to the GSFC page.

This time there is a set of instructions (not visible in Figure 3) on how to use the mouse to activate dynamic images and lists of links to various datasets. Passing the mouse cursor over each of the five globe icons and their brief titles on the right displays a larger more detailed globe image, presumably false-colored with data-based information, and a set of links to datasets. These are very simple images with no captions. They are also tied cohesively to the introductory note for each set of displayed links by a fairly trivial repetition of lexical items. For example, the Life on Earth menu item displays text which reads:

Life on Earth

Parameters measuring life on Earth include (follow the links to view the datasets):

[Biosphere](#)

[Landcover Classification](#)

[Vegetation](#)

[Chlorophyll](#)



View data holdings: The data chart shows all the environmental measurements displayed by the Earth Observatory, and the months they are available.

UPDATED DATASETS:
[4km² TRMM Fire](#) (November 1999)

Parameters measuring life on Earth include (follow the links to view the datasets):

[Biosphere](#)
[Landcover Classification](#)
[Vegetation](#)
[Chlorophyll](#)

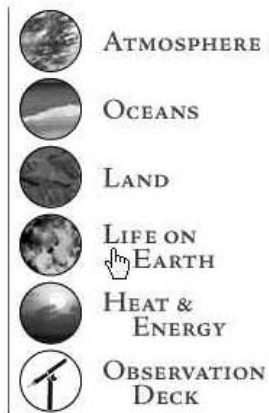
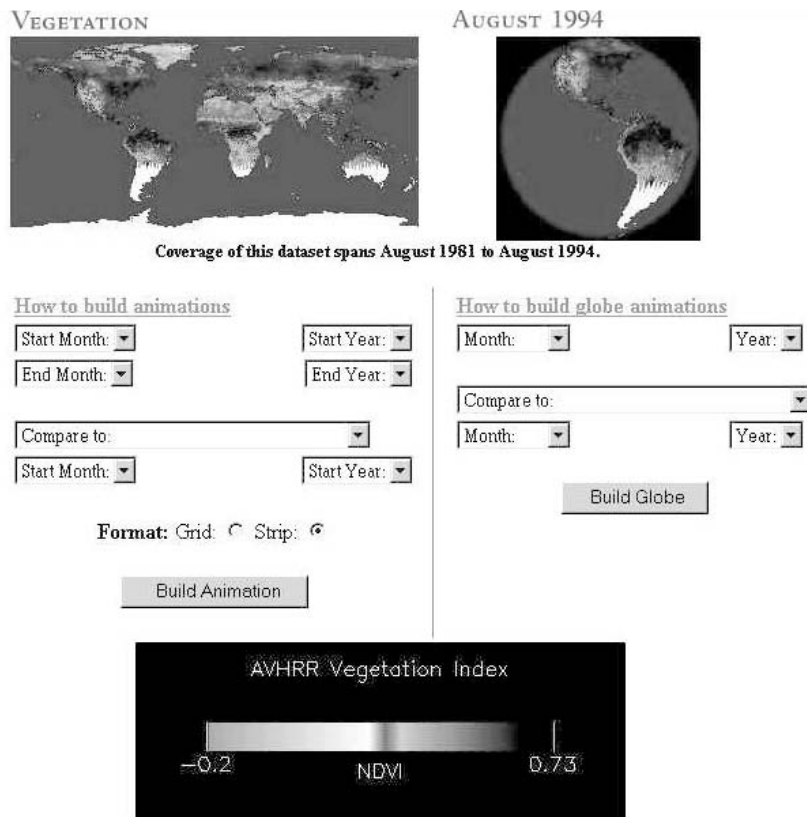


Figure 3 NASA Earth Observatory Observation Deck webpage.
 Source: [http://earthobservatory.nasa.gov/Observatory/]

They are cohesively tied to the displayed links by thematic collocation ('biosphere', synonymy; 'vegetation', meronymy; 'chlorophyll', collocation), in the same way, verbally, that the displayed introduction and links are cohesive with each other. The menu icon and the displayed globe image are cohesive visually only by the similarity of shape (round) and dominant color (green). But there is a crossmodal cohesive link between the word 'earth' in the menu and image of a globe of the earth, which one might call a crossmodal collocation, since this association is a conventional one produced by their frequent co-occurrence in other contexts.

Structurally, we have a similar but simplified structure to that of the GSFC page. The displayed introductions and links have a similar internal structure, with a title, a specification, and the list of links. In addition there is the co-displayed globe, organizationally suggesting a thematic tie to the items in the list as well as the title. With this prompt, we can interpret the Life on Earth displayed globe as showing vegetation, green with chlorophyll, covering the landforms of Africa in an uneven way. We now read regions of the globe image as jungles and deserts, in part *because* the dynamic visual



Product description: Normalized Differential Vegetation Index data shows the 'greenness' of plants. For example, evergreens and deciduous trees both have high

Figure 4 NASA Earth Observatory: Life on Earth, Vegetation Dataset webpage. *Source:* [<http://earthobservatory.nasa.gov/Observatory/Datasets/ndvi.avhrr.html>] (since updated as [<http://earthobservatory.nasa.gov/Observatory/Datasets/ndvi.fasir.html>])

and interactive organization of the page implies a thematic relation to the verbal items in the links list.

Orientationally, this page addresses the user verbally through imperatives (Proposal:demand:action). The default lefthand display gives instructions: 'Click to view ...', 'Pass your cursor ...' and the one for Life on Earth, like the others, says 'Follow the links ...'. This constructs a Guide:Visitor relationship, with the author/source as more knowledgeable, not just about the earth or NASA data, but also about the use of the webpage itself. Otherwise we also have a (Proposal:offer:information) orientation to the user, as with the GSFC page. In terms of stance to content, there is very little content as such, so most of the evaluative dimensions are unmarked. The images show a fairly realistic, but not high-resolution image for the telescope (high warrantability, but no importance) and a globe (for Life on Earth) that has the same warrantability issues discussed above for the Monsoon Dynamics image.

I am more interested at this point in passing from the analysis of multimodality, which I think is sufficiently illustrated already, to hypertextuality. In a sense we have already seen in these dynamic pages a measure of hypertextuality within the page frame, insofar as the mouse-over display of images and text represents an interactive potential for multisequential display to the user of lexias or meaning units. The meanings along such traversals, beyond those of individual displays, are mainly cohesive: each element is an instance of some general category, and therefore with some thematic and/or visual similarities to the others, and as we catenate them we are cumulating toward an exhaustive exploration of the category. This is a fairly common traversal principle in websites. We are presented with a general theme or topic and links to subcategorized elements of it. We can examine these in any order, but as we examine more and more of them, we construct a traversal which is more than the sum of its parts, primarily insofar as the whole also exhausts the subcategorization structure and we find we have all the parts or pieces.

If, however, we now accept the page's offer of information and accede to the demand to click on a link, such as Vegetation under Life on Earth, we can consider the meaning relations between this portal page and what appears next (see Figure 4).

In terms of Organizational meaning relations, we must first recognize that the act of clicking and following the link, i.e. calling up the display of the page, is an actional-structural tie. This is similar to the crosstextual tie we find when we turn to an indicated page from one part of a book to another. Semantically, this is a two-part relation, in which the first part (anchor) in this case specifies the kind of information (dataset) to be found in the second part (target). We can consider this akin to the verbal semantic relationship of Expansion:Elaboration (between clauses). There are also of course cohesive ties, as we would expect. Verbally there is 'vegetation' and 'dataset' repeated on the new page. Visually, there is another globe, and a map of the earth, both with green coloring distributed unevenly.

This new page has a very different compositional structure, though it can still be read as a vertical stack of information units: website banner and menu, page title, sample visual data displays, input form to specify a dataset, visual-numerical color key, 'product description'. This is an invariant pattern for all the dataset access pages following from the links on the original portal page.

The page constructs a complex source: user relationship. In addition to its implied offers of further information and demands for action (click to follow link), it also now presents a [Proposal:demand:information] in the guise of the input forms. This is the interactive equivalent of a verbal question. In order to respond to this demand, we must also accede to procedural demands for action (e.g. use a drop-down list, click on a button to submit the data form). But we are now given a large number of choices, in effect a combinatorial freedom, which is typical of the open web hypertext medium.

This type of author:user relationship in hypertext gives rise to a sense of interactive dialogue. The page asks us to fill in the form, but by filling it in and clicking one of the 'Build' buttons, it is we who are now demanding information and action, and in general the programs and databases behind the website will accede to these demands. One could say that this page offers the opportunity to make a demand.

The 'instructional' links 'How to build ...' are also interestingly complex in terms of the semantics of exchange relations. On the surface they offer information, 'here is how to build ...', but they demand an action (click the link) to get that information. When the user clicks, she or he is not just acceding to the demand for action, but also demanding the information, i.e. asking the question 'How do I build ...' or demanding 'Tell me how to build ...'

If we fill in the data for building a 'strip' animation, and click *Build*, the result is a new page, not permanently on the webserver but created just for us (Figure 5). Below the usual website header and site title, there is a cohesive page title or image label 'Vegetation/NDVI', continuing the identity and similarity cohesion chains from the original portal page (vegetation) and the product description and color key image elements of the previous (anchor) page. Below this we see a Quicktime strip image in which, by dragging with the mouse along the slider bar at the bottom, we shift time backwards and forwards, with time labels beneath the image changing to index this. The rightmost and leftmost positions correspond to the input-form data we supplied for start month, end month, and year(s). What we see is the same world map, but the colorations change with the seasons. Below the animation strip-map is text which is instructional in nature; i.e. it is about the procedures for viewing the animated map, not about the thematic content of vegetation that changes with the seasons.

What shall we say about the crosslink meaning relation of these two pages? It is again a kind of exchange relationship (orientationally): this is the server's answer to our question, its response to our demand, which was itself an acceptance of an offer made on the prior page. The dialogue genre is extended one turn.

Organizationally, the crosslink relationship is a more complex actional-structural one than that between the portal page and the input-form page. The action is not simply clicking the Build button, but also supplying the information in the form. There will not be a strip page (but rather an error page, or no response) otherwise. So the interactive first-part of the structural pair involves supplying information as well as demanding it. The information supplied through the input-form is verbal-numerical, but its internal organizational semantics is structured by a visual table, i.e. by a layout of relations among elements similar to that of a table in a print medium. (See Lemke, 1998b for a discussion of the historical and semantic relations of tables to connected text and to more graphical visual displays.)

Presentationally and thematically, the interpage relationship would

Vegetation/NDVI

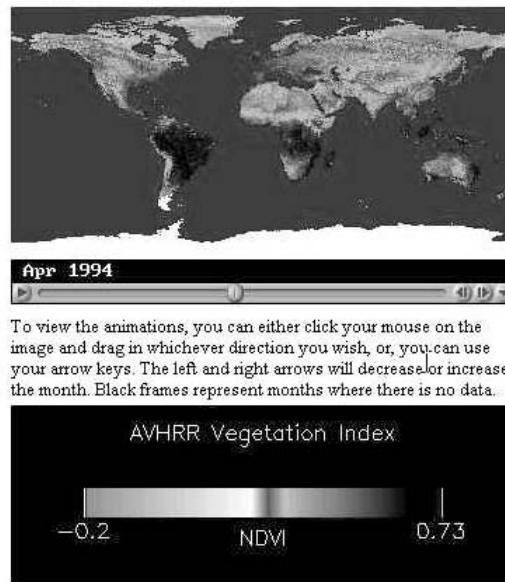


Figure 5 NASA Earth Observatory: Vegetation NDVI Data in Animated World Map.
Source: [<http://earthobservatory.nasa.gov/Observatory/showqt.php3>]

seem to be that of exemplification. The first-part page offers a template for a category of possible strip-maps, and the second-part displays one instance from this category. This is the primary meaning relationship, but secondarily one can also say that the new strip-map provides an Extension:additive to the information in the sample display. As if to say: in August 1994 the world looked like this, and in other months (e.g. April 1994) it looked like this.

In purely visual terms, we could compare the two strip-maps (sample and animation), to see a cohesive tie (the same continents and landforms) and a spatial-distributional difference in colors. There would be a very different pattern of salience and eye-movement across the two maps. How could we name their visual meaning-relation? It is somewhat akin to 'morphing', a constant general form with varying details (or the continuous passage from one image to another). It is also somewhat like the relation of theme-and-variations in music: a constant overall form, with varying lesser details. This assumes that we see, culturally, configurations of shapes and masses as more significant visually than relatively minor changes within a still relatively constant color palette. This visual relationship of the pages seems subordinated to the semantic relationships already discussed.

There is, of course, another visual relationship of the pages: their sequential display in the same visual space, like the frames of an animation or film. This is a dynamic visual relationship, in time more than in space, in the window of the browser.

THE POLITICS OF HYPERMODALITY

Is there a politics of multimodality as such? I have already argued that there are no media in which meanings are made purely with one semiotic resource system. So what we are really concerned with in seeking a politics of multimodality are the social conventions regarding the degree of importance assigned to different media and their combinations. I think there is no doubt that ‘logocentrism’ in modern European intellectual and academic culture represents a political ideology. To privilege linguistic meaning to the point of excluding or denigrating pictorial modes of representation must have a politics, it must favor some interests or modes of social control.

The conventional argument in these cases is that images have an inherent degree of ambiguity which makes them unsuitable to precise scholarly meanings and accurate reasoning. I do not think this premise is acceptable. The ambiguity of verbal text is very high, as anyone advocating the parallel ideological claims of mathematics or scientific notation as superior to verbal text could amply demonstrate. Linguistic registers and genres have evolved specialized rhetorical and textual strategies to reduce ambiguities of certain kinds (for reference, for implication) over long historical periods, but so have many visual genres, e.g. those employed in medical and botanical illustration, in scientific data visualizations such as those we’ve seen and mentioned earlier, etc. Moreover, visual representations can present meanings-by-degree (the shapes of clouds and mountains, degrees of brightness, shades of color, exact relative sizes, etc.) far more precisely than can the more gross categorial distinctions of verbal language. Mathematics was largely invented, beyond simple counting, to extend the semantics of natural language to this domain of meaning-by-degree (ratios and fractions, geometric relations, quantitative co-variation, etc.; cf. Lemke, in press a). The total amount of information in an image, and the total number of discernible contrasts which define its visual features are certainly comparable to those of natural language. Its two- and three-dimensional affordances for organizational relationships are also superior to what the syntagms of verbal text afford. Visual–diagrammatic logics are probably superior to both verbal reasoning and mathematical logic notations on many criteria.

So it seems much more likely that what is accounted as ‘ambiguity’ in visual images is nothing more than there not being a one-to-one correspondence of images to texts. If one takes language, a priori, to be the standard of precision in meaning, then anything that does not have a unique verbal reading is judged to be ‘ambiguous’. A high-resolution image of the

earth as seen by a satellite is in no way ambiguous; it is indeed more precise and accurate than any possible verbal description could be. A comparison of predicted and measured values in some experiment, shown on a data graph, is a far better basis for reasoned judgments than any verbal presentation of the comparison, or even any numerical or algebraic presentation.

So why the denigration of visual representations? My strong suspicion is that because text and image *mutually* contextualize one another, influencing our interpretations of each and both together, that it is the power of the image (and other semiotics) to subvert and undermine the authority of linguistic categories and categorical imperatives which is being politically suppressed by logocentrism and monomodal purism. Language affords a low-dimensional representation of experience and the complexity of social–natural realities. It reduces matters of degree to matters of kind, frequently to dichotomizing categories (masculine/feminine, gay/straight, capitalist/communist, heroes/terrorists) through which sentiments and allegiances can be more easily manipulated. Of course visual images can also be used in this way, but images inherently afford a much greater display of complexity and ‘shades of grey’, whether in unedited documentary footage from a war zone or in the daily gyrations of a stock price over months or years, or those of the earth’s average temperature in a debate on global warming. When we put images and text together, their very incommensurability, the fact that they cannot both present exactly the same message, casts doubt on the monological pretensions of either, but particularly those of language.

A more balanced multimodality is potentially more politically progressive, whether in the deliberate juxtaposition of texts and images that never quite tell the same story and force us to more critical analysis than either might do alone, or in the representation of issues of ‘race’, gender/sexuality, social class, culture, etc. in multidimensional ways as matters of degree and possibility rather than category and constraint.

Is there then also a politics of hypertextuality? I believe that hypertextuality invites and affords more complex dialogical (or pseudo-dialogical) chaining of offers and demands, choices and constraints between users and designers/sites (see Aarseth, 1997 and Lemke, 2002a to sort out these roles and agencies) than does text which is built with the strong expectation that readers will follow a default sequence.

As Kolb (1997) has also argued, this circumstance alters the affordances of the medium for making the traditional sorts of extended arguments (enthymemes) that are common in, say, modern academic philosophy and throughout most of the social sciences. Authors cannot count on readers staying within the grasp of their argumentation; it is harder to lead the reader rhetorically down the garden path to agreement with the author’s views. Instead, readers explore alternative pathways through a hypertext, or create their own traversals, particularly over longer text-scales. Authors may produce a consistent voice or viewpoint in all the lexias they

include in a hypertext and hope that their cumulative effect will naturalize their viewpoint for the reader. Or they may inscribe their viewpoint into the organizational structure of links and pathways (e.g. subcategorization schemes). Nonetheless, authors lose the power to make some traditional kinds of monological, coercive/cogent arguments.

What the hypertext medium affords the author or designer in a positive sense is the opportunity to escape monologism altogether, not in the trivial sense of creating a pseudo-dialogue with the user, but in the more profound Bakhtinian (1981[1935]) sense of including multiple social voices, giving the reader access to a field of heteroglossia, of discourse diversity and conflict. Why would authors do this? We might, for example, want to include in the same web both officially authoritative images or texts and counter-images and discrepant discourses. And link them together in such a way that the socially dominant viewpoint is constantly confronted with its Others, its usual monological voice constantly subverted by an implied dialogic opposition. We might also want to create for the user a space which affords and demands critical analysis by leaving the last word unsaid, by opening up possibilities and foregrounding alternatives and contradictions, inviting users to think for themselves.

When we combine the affordances of multimodality with those of hypertextuality, it is doubly possible to resist the monological voices of traditional genres (such as the one I am now writing). First, through the incommensurability of different semiotics and their differential affordances, particularly the opportunities for presenting 'non-essential' details, relational complexity, and meaning-by-degree provided by visual media; and second, through the cross-linking of diverse viewpoints, discourses, images, etc. in hypertext, which affords the user the opportunity to make meanings that are not the implied or explicit conclusions of the author/designer.

Visual communication is at its most powerful, not when it retreats into the splendid isolation of an imaginary semiotic autonomy, but when it confronts verbal language head-on and challenges its hegemony, when it takes its place as an equal (and equally often as the leading) partner in multimodal communication. The medium in which both confrontation and partnership, both subversion and empowerment, is most fully afforded today is that of hypertext. Travelling together in hypermodality, we can make meanings that will let people see and speak in new and more thoughtfully critical ways.

REFERENCES

- Aarseth, Espen (1997) *Cybertext*. Baltimore, MD: Johns Hopkins University Press.
- Arnheim, Rudolf (1956) *Art and Visual Perception*. London: Faber.
- Bakhtin, Mikhail M. (1981[1935]) 'Discourse in the Novel', in M. Holquist (ed.) *The Dialogic Imagination*. Austin: University of Texas Press.

- Eco, Umberto (1990) *Travels in Hyperreality*, ed. John Radziewicz, trans. William Weaver, 2nd edn. Fort Washington, PA: Harvest Books.
- Halliday, M.A.K. (1978) *Language as Social Semiotic*. London: Arnold.
- Halliday, M.A.K. (1994) *An Introduction to Functional Grammar*. 2nd edn. London: Arnold.
- Halliday, M.A.K. and Hasan, R. (1976) *Cohesion in English*. London: Longman.
- Kolb, David (1997) 'Scholarly Hypertext', *Hypertext '97: Proceedings of the ACM Conferences on Hypertext* (Association for Computing Machinery): 29–37.
- Kress, G. and Van Leeuwen, T. (1996) *Reading Images*. London: Routledge.
- Kress, G. and Van Leeuwen, T. (2001) *Multimodal Discourse*. London: Arnold.
- Lemke, J.L. (1983) 'Thematic Analysis: Systems, Structures, and Strategies', *Semiotic Inquiry* 3(2): 159–87.
- Lemke, J.L. (1988) 'Text Structure and Text Semantics', in R. Veltman and E. Steiner (eds) *Pragmatics, Discourse, and Text*, pp. 158–70. London: Pinter.
- Lemke, J.L. (1995) 'Intertextuality and Text Semantics', in M. Gregory and P. Fries (eds) *Discourse in Society: Functional Perspectives*, pp. 85–114. Norwood, NJ: Ablex Publishing.
- Lemke, J.L. (1998a) 'Resources for Attitudinal Meaning: Evaluative Orientations in Text Semantics', *Functions of Language* 5(1): 33–56.
- Lemke, J.L. (1998b) 'Multiplying Meaning: Visual and Verbal Semiotics in Scientific Text', in James R. Martin and Robert Veel (eds) *Reading Science*, pp. 87–113. London: Routledge.
- Lemke, J.L. (2002a) 'Notes on Multimedia and Hypertext', online at: [<http://academic.brooklyn.cuny.edu/education/jlemke/papers/hypermedia/appendix.html>]
- Lemke, J.L. (2002b) 'Multimedia Genres for Science Education and Scientific Literacy', in M. Schleppegrell and M.C. Colombi (eds) *Developing Advanced Literacy in First and Second Languages*. Mahwah, NJ: Erlbaum.
- Lemke, J.L. (in press a) 'Mathematics in the Middle: Measure, Picture, Gesture, Sign, and Word', in M. Anderson, V. Cifarelli, A. Saenz-Ludlow and A. Vile (eds) *Semiotic Perspectives on Mathematics Education*.
- Lemke, J.L. (in press b) 'The Role of Texts in the Technologies of Social Organization', in R. Wodak and G. Weiss (eds) *Theory and Interdisciplinarity in Critical Discourse Analysis*. London: Macmillan/Palgrave.
- Mann, William and Thompson, Sandra (1986) 'Relational Propositions in Discourse', *Discourse Processes* 9(1): 57–90.
- Martin, J.R. (1992) *English Text: System and Structure*. Philadelphia: John Benjamins.
- Mitchell, W.J.T. (1994) *Picture Theory*. Chicago: University of Chicago Press.

Voloshinov, V.N. (1986[1929]) *Marxism and the Philosophy of Language*, trans. L. Matejka and I.R. Titunik. Cambridge, MA: Harvard University Press.

BIOGRAPHICAL NOTE

JAY L. LEMKE is Professor of Educational Studies at the University of Michigan. He is the author of *Textual Politics: Discourse and Social Theory* (Taylor & Francis, 1995), *Talking Science: Language, Learning, and Values* (Ablex, 1990) and *Using Language in the Classroom* (Oxford, 1989). He is also co-editor of the international research journal *Linguistics and Education*. He earned his PhD in theoretical physics from the University of Chicago, and his current research interests include multimedia semiotics, multiscale dynamics in social–ecological systems, and the role of new information technologies in education.

Address: Department of Educational Studies, University of Michigan, 610 East University, Ann Arbor, MI 48109, USA. [email: jaylemke@umich.edu]