

Doing Multimedia Analysis of Visual and Verbal Data: A Guide

Jay L. Lemke
University of Michigan

Increasingly, researchers find ourselves needing to analyze more than just textual (i.e. verbal) data. Verbal text is mixed in with visual representations such as photographs, drawings, diagrams, charts, graphs, etc. In video transcripts, we have parallel transcriptions of speech and still images or video clips of what was visible during the time of that speech. Visual and verbal representations are co-deployed to make meanings that go beyond what either would make alone. It is this synergy across semiotic modalities that we need to analyze.

Obviously such an analysis needs to build on what we already know about discourse analysis for text and speech, and on the foundations of visual semiotic analysis of still images. We are still only at the beginning of knowing how to analyze the dynamic flow of representations in video and film, but for this, too there are useful sources.

Here I provide first a set of Focal Questions to direct the attention of the analyst to key perspectives and issues which can help get us started. And secondly, I will list a number of my own works and a few by other people that are likely to be particularly helpful for the multimedia analyst. The focal questions are meant primarily for non-dynamic media (i.e. not for video or film), but can be applied, with caution and some creativity to these media as well. In the list of select references, I include more work aimed at dynamic multimedia and hypertext (such as website analysis and traversals across links).

Focal Questions for Initial Analysis

The following questions are designed to focus the analyst's attention on a variety of meaning types and functions during the analysis of multimodal semiotic artifacts. The initial questions apply to (nearly?) all media; some of the more advanced ones only in more specialized cases (indicated). Not all meaning types are equally salient in all genres. And not all questions provided here are equally relevant for all analytical purposes or research questions. But they are a good starting point.

I imagine the typical semiotic artifact to which these questions are addressed to be a graphical-textual display such as a page from a print publication or a webpage. Typically included are text headers and labels, cohesive text, and visual images and design elements. Many questions also apply, *mutatis mutandis*, to plain text and its typography and page layout, and to more complex artifacts such as hypertexts, hypermedia, multi-page spreads, etc. Different meanings are made in more extended texts and multimedia than in shorter ones or constituent units. Interactivity adds additional meaning dimensions, as do dynamic media such as audio and video/film or animations. Concrete objects, such as machines

and toys, require special consideration on some points, as do immersive environments, whether material (architecture, streetscapes) or virtual.

The guiding questions are arranged according to the analysis of three primary types of meaning presented in Lemke (1998a, revised 2011) and Lemke (1998b). Briefly, these are: **Organizational meaning**, the dimension of all meaning-making that tells us what goes with what, what smaller units belong to what larger units, and how meaning themes are distributed across a whole work. **Presentational meaning**, the dimension of all meaning-making that creates and/or references a state-of-affair, who is doing what to whom and with what means, what is happening, what is related to what and how. **Orientalional meaning**, the dimension of all meaning-making that indicates the relation being created between creator and interpreter (e.g. telling, asking, offering, etc.) and the attitudinal and/or evaluative stance being taken by the creator towards their own Presentational content (e.g. serious/humorous, valuing/disparaging, etc.).

I recommend beginning with the Organizational aspect of meaning because it helps to define units of analysis for dealing with the other two aspects. Ultimately, all three aspects intimately and complexly interact with one another, shifting each other's effects. So it is necessary to iterate or repeat the analysis a few times across each aspect to capture these cross-functional re-contextualizations. See references below for details and examples.

Organizational Functions and Meanings

Basic and General:

What constitutes the whole text or multimedia artifact under analysis? how do you know what is and what is not a part of it? Does it have a definite boundary or frame in space? in time? Is it part of a larger-scale semiotic entity? Which internal features indicate its role in the larger unit?

What are the largest-scale visual elements or entities within the artifact? the longest-distance and longest-time separated elements that appear to have a definite link or relationship? What is the basis of the link/relationship? what type of meaning relation does it create (e.g. part-part within a sub-maximal whole, similarity or category membership; cf. co-meronymic and co-hyponymic semantic relationships)? What are the longest chains of elements? on the basis of similarity? of visual linkage within a larger structure (spatial-geometrical relationship, graphical linkage)?

What is the most salient visual element? Where is your eye drawn to first? What visual pathways or vectors lead to, and from this element? What other salient elements lies along these pathways? How is salience realized graphically? How are the vectors visually defined in terms of path and directionality?

What are the largest-scale regions? what defines their (fuzzy) boundaries? what are their geometrical relationships to one another? what other meaning-

relationships does this global geometry support? For each region, what are its largest-scale constituents? (Iterate the questions until a minimal visual unit is reached.)

What are the visual roles of text in the artifact? How is text associated with other visual elements? Do visual patterns create relationships between distinct text units? Do textual-cohesion relationships create relationships between other distinct visual units? What are the semantics and meaning-functions of these relationships?

Within textual units on each text-scale, how is cohesion achieved? What are the structural units and their semantic relationships to one another at the scales of phrase or group, clause, clause-complex or sentence, rhetorical structure, and genre structure? How does cohesion operate within and across these structural units?

Advanced or Specialized:

What are the typical sequences of events or actions? what are the functional meaning relations of neighboring events? of more distant events along a chain or traversal? What are the options for alternative actions? What differences in types of sequences do different action-choices produce? on shorter sequence-scales? longer?

How do the material affordances of the object constrain organizational relationships of included elements? what are the available dimensions of spatial relationships? of sequential actions/events?

What kinds of meanings are made over longer timescales or space-scales that are not made on shorter ones?

How are organizational relationships created by presentational meanings? (e.g. similar or category-contrasting actors, processes, etc. distributed in different parts of the same text)

How are organizational relationships created by orientational meanings (e.g. similar or contrasting evaluations? source-user stances?)

Orientalional Meanings

Basic and General:

Who is the intended user/viewer/reader of this text/artifact? What internal features index anticipated qualities of the user? Is there a ranking of primary, secondary, etc. expected user-types? What indications are there of anticipated audiences or 'overhearers' who are not directly 'addressed' by the text?

What qualities of the source/author of the text/artifact are indexed by internal features? Are there multiple? ranked? sources or authors, or social voices?

How does the text position the source relative to the user? in relations of power? dominance? intimacy? formality? and to what degrees?

What does the text/artifact offer the user by way of information? services? options for action? What does the text/artifact request or demand of the user or constrain in action? How?

How does the text/artifact work to change the relationship of the user and the source? What is it doing to or for the user?

How does the text/artifact index the stance of the source (or any voice it projects) toward the text/artifact itself? toward the source/voice itself? toward the user? toward its own presentational content? with respect to implied or explicit evaluations of warrantability (probability, possibility, doubt, truth, realism), usuality/surprise/salience, normativity (required/fordidden), un / desirability (of all kinds), importance/significance, seriousness/irony/humorousness?

How does the text/source evaluate its own abilities? the abilities of the user? of others?

How does the text/artifact/source evaluatively appraise persons, things, phenomena, events, circumstances?

Advanced or Specialized:

How do indexed or implied qualities of the source/voices change over the course of the text, or from one visual region of the artifact to another? how do those of the anticipated/implied user and audiences change? how do the presupposed and created relationships between them change?

How do evaluative stances shift with shifts in presentational content? How do stances shift toward the same content at different points in the text or across different visual regions of the artifact? at different dynamical moments?

What sorts of changes take place on shorter time/space-scales? on longer ones?

Presentational Meanings

Basic and General:

What recognizable actors or participants in processes and relationships are presented? persons? other animate or quasi-animate actants? concrete things? abstract ideas, qualities, concepts, processes and relationships as things?

What quantities, qualities, types or category memberships, and degrees of these are attributed to participants, etc.?

What relationships are presented as existing (past, present, future, conditional, contrafactual) among these participants? In what common or shared processes, activities, actions, events, happenings are they presented?

What are the locations, settings, conditions for, causes of, modes and manners of, temporal locations for, durations of, and other circumstances attending or modifying these relationships and processes, as presented, represented, implied, created or presupposed?

Are the processes, circumstances and relationships presented as permanent, changing, new/recent, long-standing, sudden, stable, persistent, habitual, repetitive? and how?

How are processes, circumstances, and relationships synchronized or sequenced in time? on what timescales are they presented as typically enduring or changing?

What logical relationships among processes and relationships of participants are presented? How does one presuppose or enable another? add to, deviate from, restate, expand upon, or modify its meaning?

What categories of similar, contrasting, or identical meaning are created by the variety of participants and processes in the text/artifact? how are members of the same or contrasting categories distributed spatially, textually, and temporally in/ through the text/artifact?

If there are chains or patterns of distribution of members of such categories, what typical or recurring relationships between their members (including co-participation in the same process or relationship) are presented in the text/artifact? what relationships are constructed between larger or whole patterns?

Advanced or Specialized:

How are the presentational meanings of visual-graphical elements modified by their organizational relationships to textual elements? and vice-versa? (e.g. disambiguated, more delicately specified, made polysemic, etc.)

What information is presented redundantly between visual-graphical elements and textual elements? what is presented only in one mode or the other? what is presented only jointly between both modes?

RECOMMENDED SOURCES

Lemke, Jay L.

Analysing Verbal Data: Principles, Methods, and Problems. In K. Tobin & B. Fraser, (Eds). *International Handbook of Science Education*. London: Kluwer Academic Publishers. pp. 1175-1189. 1998a. [see also the revised chapter in the 2nd edition, 2011]

Multiplying Meaning: Visual and Verbal Semiotics in Scientific Text. In J.R. Martin & R. Veal, Eds., *Reading Science*. London: Routledge. (pp.87-113). 1998b.

Multimedia demands of the scientific curriculum. *Linguistics and Education* 10 (3): 247-272. 1998c.

Typological and Topological Meaning in Diagnostic Discourse. *Discourse Processes* 27(2), 173-185. 1999.

Multimedia genres for science education and scientific literacy. In M. Schleppegrell & M.C. Colombi, Eds. *Developing Advanced Literacy in First and Second Languages*. Mahwah, NJ: Erlbaum. pp. 21-44. 2002.

Travels in Hypermodality. *Visual Communication* 1(3): 299-325. 2002.

Mathematics in the Middle: Measure, Picture, Gesture, Sign, and Word. In Anderson, M., Saenz-Ludlow, A., Zellweger, S. & Cifarelli, V., (Eds.). *Educational Perspectives on Mathematics as Semiosis: From Thinking to Interpreting to Knowing*. pp. 215-234. Ottawa: Legas Publishing. 2002.

Video Epistemology In-and-Outside the Box: Traversing Attentional Spaces. In Goldman, R., Pea, R., Barron, B. & Derry, S.J. (Eds.), *Video Research in the Learning Sciences*. Mahwah, NJ: Erlbaum. pp. 39-52. 2007.

Conducting video research in the learning sciences. *Journal of the Learning Sciences* 19(1): 3-53. 2010. [with S. Derry, R. Pea, B. Barron, R. Engle, R. Hall, F. Erickson, R. Goldman, M.G. & B.L. Sherin, & T. Koschmann]

Other Authors

Arnheim, Rudolf. 1956. *Art and Visual Perception*. London: Faber.

Goldman, R., Pea, R., Barron, B. & Derry, S.J. (Eds.). 2007. *Video Research in the Learning Sciences*. Mahwah, NJ: Erlbaum.

Kress, G. & van Leeuwen, T. 1996. *Reading images*. London: Routledge.

N.B. This paper was originally created for my website and is not published.