

An Exploratory Approach for Modeling Image-Enabled Discourse

Jaime Snyder

School of Information Studies, Syracuse University

221 Hinds Hall, Syracuse NY, 13203 USA

jasnyd01@syr.edu

Abstract

A proposed methodology for the study of image-enabled discourse is introduced, focusing on a conceptual framework that builds on theoretical contributions of related research and provides working definitions for key concepts, and a brief description of future empirical research. The goal of this methodology is to produce an analytic model of image-enabled conversations that could be applied to a range of multimodal information problems including system building for information access and extraction. Drawing is used as an example of an image-centered discourse practice.

Introduction

We might tend to think of words as being the primary building blocks from which we construct conversations. However, a number of other tactics are available to us when we try to communicate with each other. We can gesture, growl, frown or draw a picture. Of particular interest to this research is this last option: the ability to exchange meaning by constructing images. In spite of the ubiquity of visual information in our daily communications, both face-to-face and online, the various disciplines involved in image-centered research have not yet coalesced around integrated research regarding the role images play in communicative practices. As academic researchers and system builders direct increasing attention to managing multimedia information, the nascent nature of our understanding of the role visual information plays in communicative practices has been thrown into sharp relief.

Problem definition

The evolving framework described focuses on a particular type of spontaneously created images, so-called “napkin drawings.” This is a ubiquitous form of ad hoc visualization consisting of marks made on an available surface during the flow of a conversation, sometimes kept, sometimes abandoned, and notoriously cryptic for those not involved in the discussion. These drawings often mark moments of great insight or idea generation, and thus are

clearly an important component in the process of communication between individuals. So-called napkin drawings can anchor, bridge, and facilitate the flow of information at crucial moments. Rarely seen as aesthetic objects of great admiration for their own sake, these images answer to a different set of requirements than other constructed images (such as paintings and photographs).

On a certain level, we exhibit great expertise at deploying images just when we need them. Bringing this ability into a more conscious and deliberate domain will allow us to better exploit this seemingly innate human communication practice when building systems seeking human-like qualities. Insight provided by a functional understanding of the role of images within conversation has great potential to enhance traditional information models by incorporating notions of context and functionality into the overall meaning of an image.

This research takes inspiration from information access research that relies heavily on sound theories for the categorization of complex meaning in text, such as Fillmore’s FrameNet (Fillmore, Johnson, & Petruck, 2003) and Wiebe et al’s work on opinion and sentiment (Wiebe, Wilson, & Cardie, 2005; Wilson, Wiebe, & Hwa, 2004), to name just two examples.

Research questions

What is it about ad hoc visualizations that make them such good conveyors of meaning at certain points in a conversation? How can we capture the vital role of the image within the context of the conversation? And why, in situations where we have great need for clarity and precision, do we forgo words, which have the benefit of dictionary definitions and specific rules of grammar, and turn to the seemingly nebulous, ambiguous realm of images in order to be more clear? Image-enabled discourse is the term introduced here to refer to this phenomenon. This concept is built on the notion that investigating the communicative activity of *creating* an image will contribute to our understanding of the nature of visual information and allow us to integrate models of visual meaning with existing models of linguistic meaning.

Investigating image-enabled exchanges

Conceptual framework

In contrast to much research conducted in the area of visual studies, image-enabled discourse focuses on the context of the creation of an image and the role that the image plays in multimodal communicative practices, de-emphasizing analysis of the image as an artifact.

Functional aspects of images. Traditionally, image-oriented fields of study (e.g., art and art history, design, architecture and advertising communications), have focused almost exclusively on the image artifact: paintings, photographs, ads, video, etc. Art historian James Elkins has argued for a more inclusive study of images, art and non-art, opening that discipline to new perspectives. Elkins also notes that art historians and cultural scholars have paid little attention to the process of image-making (Elkins, 1999). In his survey of the study of visual culture, Elkins notes the lack of integration between studio art programs and art history departments as evidence of this inattention and points to this as a potentially rich area for future research (Elkins, 2003).

In his well-known series of books about the visual design of information, Edward Tufte also advocates for an expanded perspective on the function of images, describing various types of graphical representations as conveyors of complex information. He addresses issues related to context of use by advocating more thoughtful engagement with the information conveyed by graphics such as charts and graphs, however his focus is firmly on heuristics, prescribing best practices for creating more effective information graphics (Tufte, 1983, 1990, 1997). He is less concerned with the context of making and is more attentive to interpretation of the information being presented.

Discussions of visual literacy are found in the education literature, for example: (Cook, 2006; Harrison & Treagust, 2000; Pettersson, 1989). This research tends to focus on pedagogy, refraining from delving more deeply into the specific role visual information plays in individual exchanges of meaning at an interaction level.

Discussions of drawing as a quasi-functional activity can also be found in research discussing the use of drawing as a psychological diagnostic tool rather than a communicative act in and of itself (Freedman, 1994). For example, drawing has been used diagnostically as an indicator of an underlying cognitive pathology, as with autistic children who are able to draw hyper detailed and realistic seeming images very early in their development (Jolliffe & Baron-Cohen, 1997). Research related to cognitive load and mental processing, such as (Sweller, 1988), could inform the identification of potential contextual circumstances that may contribute to the decision to use images during conversations.

Context of language use. Discourse linguistics is the study of the influence of context on language. It encompasses research that looks at linguistic characteristics of passages of text, analyzes language use at and above the sentence level and looks at the linguistic structures of exchanges between individuals in a variety of contexts (Cutting, 2002; Littlejohn, 1996; Searle, 1969, 1992; van Dijk, 1979, 2000). It should be noted that the term discourse can be used in a number of ways: a discourse could be a self-contained passage of text (discourse linguistics) it could refer to the political beliefs held by a speaker (discourse analysis), or it could refer to a series of coherent utterances that include “conversations” (conversation analysis). Discourse analysis, while sharing common ancestry with the more structural linguistic approach, takes a more socio-cultural approach by using text to identify underlying set of beliefs and assumptions held by the speaker that indicate a world view subtext (Littlejohn, 1996; Miniard, Bhatla, Lord, Dickson, & Unnava, 1991; Philips & Jørgensen, 2002). Although none of these approaches to the study of discourse specifically invite theories to be extended into the realm of images, there is an emerging subfield, multimodal discourse analysis, that seeks to identify the influence of mode and context on meaning, focusing on co-occurrence and interaction between image and text (Royce & Bowcher, 2007). Researchers in this area have looked to the study of semiotics in order to extend discourse theories and methods to include visual information (Matthiessen, 2006; Royce, 2007). Rich discussions of the relationship between signs and meaning can be found in readings on the history and philosophy of semiotics, providing a philosophical context for extending the discussions of language into the realm of multimodal (and image-enabled) discourse (Sebeok, 2001).

Researchers in the area of multimodal discourse analysis have noted the strong connection between technology and multimodal communication, describing the increasingly multimodal nature of the complex communication environments in which we operate (Levine & Scollon, 2004). Researchers in the area of natural language processing (Carberry, Elzer, & Demir, 2006), human computer interaction (HCI) and computer supported collaborative work (CSCW) (Kraut, Fussell, & Siegel, 2003; Kraut, Gergle, & Fussell, 2002; Smith & Fiore, 2001; Wahlster, 1991) have also been working to try to better understand the interdependent dynamics of multiple modes of communication within a single interface from a discourse-oriented perspective.

Drawing as a communication practice. Viewing drawing as a communicative practice aligns image-enabled discourse with other theories of socially constructed meaning and interactive, dynamic language-based perspectives. Combining social theory with functional analysis of language use, researchers in sociolinguistics have revealed the effects of social interactions on

language, citing changes in structural linguistic indicators as evidence of meaning being created and changed over time (R. Wardhaugh, 2006). Similarly, science historians Knorr-Cetina and Amann have investigated the role images play in scientific inquiry and the creation of new knowledge within the physics research community (Knorr-Cetina & Amann, 1990). Extending through to view drawing as a generalized communicative practice (rather than a more specialized artistic mode of expression) is not a perspective commonly taken in visual studies research. However, evolving approaches to the investigation of language as an interactive and constructed phenomenon can be extended to support the notion that drawing is a functional meaning-making activity serving a specific role within multimodal communication (G. Kress & Van Leeuwen, 2001; G. R. Kress & Van Leeuwen, 2006).

Definitions

In order to examine the dynamics of image-centered communication practices, it is necessary to define the boundaries of the phenomenon. Image-enabled discourse refers to any interaction (face-to-face or virtual) between individuals that relies at some point on the interjection of an image in order to meet the goals of the exchange. Drawing is a type of image-enabled strategy and is the initial focus of this research because of its accessibility, ubiquity and ability to be created with very primitive tools.

Drawing is defined as a persistent and visible mark. For example, a hand gesture is not a drawing but a series of scratches in the dirt is a drawing. Drawings can contain alpha-numeric characters but do not have a strict left-to-right orientation. For example, a phone number written on a napkin is not a drawing because it needs to be “read” from left to right to make sense; a diagram showing numbered measurements for the construction of a box would be drawing because it does not need to be “read” from only one orientation.

People involved in conversations have certain interactions goals. We know from conversation analysis research that these goals can take a variety of forms, and awareness of them can be conscious or unconscious. People employ a range of communicative strategies in response to these goals, and as a result of the conditions in which the interactions take place (Littlejohn, 1996; Wardhaugh, 1985, 2006). The goal of the conversation determines which conditions are most hospitable for interactions to take place. The absence of salient conditions can lead to a weakening of or break in the bi-directionality of interactions. In order to create (or re-establish) hospitable conditions, alternative communication strategies can be employed. Because of specific interactive affordances, drawing enables people to employ special strategies in order to effectively create necessary conditions for reaching communicative goals.

A model of image-enabled discourse will seek to account for the ways in which drawing gives people the ability to create or re-establish salient, hospitable conditions for reaching goals within face to face interactions. Viewing image-enabled exchanges as bi-directional communicative interactions is informed by linguistic theories of conversation and discourse strategies such as code-switching and contextuality (Gumperz, 1982). The notion of co-occurrence of multiple modes of communication (or simultaneous semiotic systems) is supported by research in the area of multimodal discourse analysis as well as being aligned with some research in multilingual communication (Auer, 1998; Gumperz, 1982; G. Kress & Van Leeuwen, 2001; Myers-Scotton, 1993).

This framework also makes use of the concept of affordances (Gibson, 1977; Norman, 1999). This term refers to those aspects of the nature of visual communication that make it uniquely suited to provide the conditions needed to employ certain communicative strategies. For example, drawing is by nature persistent, tangible and visible. These are affordances of visual communication. The concept of affordances disavows the adage that “a picture is worth a thousand words.” The existence of visual affordances means that at times, no number of words can replace a visual representation, because there are concepts (or information) in the world that are inherently visual.

Initial phase of exploratory empirical research

In order to begin to identify the affects of image creation on the overall sense of a multimodal exchange, short semi-structured interviews focusing on descriptions of face-to-face conversations that have included the creation of one or more drawings will be conducted. The goal of semi-structured interviews is to gather rich, systematic descriptions from conversation participants that will reveal patterns of image-enabled communication practices that can be used to contextualize the meaning represented in multi-modal artifacts. Much like relying on human analysts to establish a gold standard for system evaluation, this approach begins with human insight in order to establish a framework for machine-level models of meaning. While an ideal initial data set would include multiple rich descriptions of people drawing in natural settings, the spontaneous nature of such exchanges makes it very challenging to collect a large number of such observation within controlled settings.

Interview protocol. In lieu of “in the wild” observations, two research techniques commonly used in information science research have been selected to guide the design of semi-structured interview questions: critical incident technique (CIT) (Flanagan, 1954) and sense-making methodology (Dervin, 1999). Like CIT, sense-making seeks to situate respondents in a specific moment related to a phenomenon of interest.

Type of practice	Goal of interaction	Salient condition for interaction	Enabling affordance
Bridge	Achieve functionally similar understanding	Ability to synchronize disparate vocabularies	Drawing allows people to build isomorphic bridges by enabling them to map their domain of knowledge to another's through the use of a shared visual language.
Sandbox	Create shared space for collaboration	Ability to create new concepts in cooperation with others	Drawing allows people to establish a mutual and tangible point of reference.
Transformation	Verify the content of a message	Ability to separate information content from form by systematically transposing to a different modality	Drawing allows people to transpose or transform the content of a message in order to verify the meaning being conveyed, separately from the mode of conveyance.
Show	Share visual information	Ability to accurately convey information by maintaining its native format.	Drawing allows people to express visual information in a visual mode.

Table 1. Example categories of image-enabled communicative practices

Pilot results. A pilot study involved administering the interview protocol via eleven face-to-face interviews. The main study will involve gathering responses to a very similar set of questions via an online interface. Inductive analysis of the eleven pilot interviews yielded a preliminary set of categories of image-enabled communication practices that will be refined during the main study. These preliminary categories (Table 1) are provided as an example of the scheme that is expected to be generated by the qualitative data collection described above.

Future research

Because of the nascent nature of this research, it falls short of offering an actionable, applicable model of image-enabled exchanges at this point. During future phases of this research, increasingly robust methods, such as experimental protocols, will be developed to further test the reliability and validity of categories of image-enabled communication practices identified during preliminary phases. The goal of this research is to significantly contribute to a highly robust analytic model that can be extended to a range of applications, including applied visual analysis, system building and heuristic evaluations. This model will also be able to inform and be integrated with existing linguistic-based models, which will allow them to be extended for use with multimodal data.

References

Auer, P. 1998. *Code-Switching in Conversation: Language, Interaction and Identity*. New York: Routledge.

Carberry, S., Elzer, S., & Demir, S. 2006. Information Graphics: An Untapped Resource for Digital Libraries. Paper presented at the SIGIR'06, August 6-11, Seattle, Washington.

Cook, M. P. 2006. Visual Representations in Science Education: The Influence of Prior Knowledge and Cognitive Load Theory on Instructional Design Principles. *Science Education*, Published online in Wiley InterScience www.interscience.wiley.com(DOI 10.1002/sce.20164).

Cutting, J. 2002. *Pragmatics and Discourse: A Resource Book for Students*. New York: Routledge.

Dervin, B. 1999. On studying information seeking methodologically: the implications of connecting metatheory to method. *Information Processing and Management* 35(6): 727-750.

Elkins, J. 1999. *The Domain of Images*. Ithaca, NY: Cornell University Press.

Elkins, J. 2003. *Visual Studies: A Skeptical Introduction*. New York: Routledge.

Fillmore, C. J., Johnson, C. R., & Petruck, M. R. L. 2003. Background to FrameNet. *International Journal of Lexicography* 16(3): 235-250.

Flanagan, J.C. 1954. The Critical Incident Technique. *Psychological Bulletin* 51(4): 327-358.

Freedman, M. 1994. *Clock Drawing: A Neuropsychological Analysis*. Oxford, UK: Oxford University Press.

Gibson, J. J. 1986. *The Ecological Approach to Visual Perception*. Mahwah, NJ : Lawrence Erlbaum Associates.

Gumperz, J. J. 1982. *Discourse Strategies*. New York: Cambridge University Press.

Harrison, A. G., & Treagust, D. F. 2000. Learning about Atoms, Molecules, and Chemical Bonds: A Case Study of Multiple-Model Use in Grade 11 Chemistry. *Science Education* 84(3): 352-381.

Jolliffe, T., & Baron-Cohen, S. 1997. Are People with Autism and Asperger Syndrome Faster Than Normal on the Embedded Figures Test? *Journal of Child Psychology and Psychiatry* 38(5): 527-534.

Knorr-Cetina, K., & Amann, K. 1990. Image Dissection in Natural Scientific Inquiry. *Science, Technology & Human Values* 15(3): 259.

- Kraut, R. E., Fussell, S. R., & Siegel, J. 2003. Visual Information as a Conversation Resource in Collaborative Physical Tasks. *Human-Computer Interaction* 18: 13-49.
- Kraut, R. E., Gergle, D., & Fussell, S. R. 2002. The Use of Visual Information in Shared Visual Spaces: Informing the Development of Virtual Co-Presence. Paper presented at the CSCW '02, New Orleans, LA.
- Kress, G., & Van Leeuwen, T. 2001. *Multimodal Discourse: The Modes and Media of Contemporary Communication*. London, UK: Arnold.
- Kress, G. R., & Van Leeuwen, T. 2006. *Reading Images: The Grammar of Visual Design*. New York: Routledge.
- Levine, P., & Scollon, R. 2004. *Discourse and Technology: Multimodal Discourse Analysis*. Washington, D.C.: Georgetown University Press.
- Littlejohn, S. W. 1996. *Theories of Human Communication* (5th ed.). New York: Wadsworth Publishing Company.
- Matthiessen, C. M. I. M. 2006. *The Multimodal Page: A Systemic Functional Exploration*. In T. D. Royce & W. L. Bowcher, eds. *New Directions in the Analysis of Multimodal Discourse*, pp. 1-62. Mahwah, NJ: Lawrence Erlbaum Associates.
- Miniard, P. W., Bhatla, S., Lord, K. R., Dickson, P. R., & Unnava, H. R. 1991. Picture-based Persuasion Processes and the Moderating Role of Involvement. *Journal of Consumer Research* 18: 92-106.
- Myers-Scotton, C. 1993. *Social Motivations for Codeswitching: Evidence from Africa*. Oxford, UK: Clarendon Pr.
- Norman, D. A. 1999. Affordance, conventions, and design. *interactions* 6(3): 38-41.
- Pettersson, R. 1989. *Visuals for Information: Research and Practice*. Englewood Cliffs, NJ: Educational Technology Publications.
- Philips, L., & Jørgensen, M. W. 2002. *Discourse Analysis as Theory and Method*. London, UK: Sage Publications.
- Royce, T. D. 2007. *Intersemiotic Complementarity: A Framework for Multimodal Discourse Analysis*. In T. D. Royce & W. L. Bowcher, eds., *New Directions in the Analysis of Multimodal Discourse*. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.
- Royce, T. D., & Bowcher, W. L. eds. 2007. *New Directions in the Analysis of Multimodal Discourse*. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.
- Searle, J. 1969. *Speech Acts: An Essay in the Philosophy of Language*. Cambridge, UK: Cambridge University Press.
- Searle, J. 1992. *(On) Searle on Conversation*. Philadelphia, PA: John Benjamins Publishing Company.
- Sebeok, T. A. 2001. *Signs: An Introduction to Semiotics*, 2nd ed. Toronto, Ontario: University of Toronto Press.
- Smith, M. A., & Fiore, A. T. 2001. Visualization Components for persistent conversations. Paper presented at the SIGCHI conference on Human Factors in Computing Systems, Seattle, WA.
- Sweller, J. 1988. Cognitive Load During Problem Solving: Effects on Learning. *Cognitive Science* 12: 257-285.
- Tufte, E. R. 1983. *The Visual Display of Quantitative Information*. Cheshire, Connecticut: Graphics Press.
- Tufte, E. R. 1990. *Envisioning Information*. Cheshire, Connecticut: Graphics Press.
- Tufte, E. R. 1997. *Visual Explanations*. Cheshire, Connecticut: Graphics Press.
- van Dijk, T. A. 1979. Relevance Assignment in Discourse Comprehension. *Discourse Processes* 2:113-126.
- van Dijk, T. A. 2000. *Cognitive Discourse Analysis. An introduction*. At <http://www.discourses.org/UnpublishedArticles/cogn-dis-anal.htm>: University of Amsterdam, Universitat Pompeu Fabra, Barcelona.
- Wahlster, W. 1991. *User and Discourse Models for Multimodal Communication*. In J. W. Sullivan & S. W. Tyler, eds. *Intelligent User Interfaces*, pp. 45-67. New York: ACM Press.
- Wardhaugh, R. 1985. *How Conversation Works*. New York: Blackwell Basil.
- Wardhaugh, R. 2006. *An Introduction to Sociolinguistics*. Malden, MA: Blackwell Publishing.
- Wiebe, J., Wilson, T., & Cardie, C. 2005. Annotating Expressions of Opinions and Emotions in Language. *Language Resources and Evaluation* 39(2): 165-210.
- Wilson, T., Wiebe, J., & Hwa, R. 2004. Just how mad are you? Finding strong and weak opinion clauses. Paper presented at the Proceedings of 19th Annual National Conference on Artificial Intelligence.