Studying Entextualization and Controversy: CDA, participant observation, computer-aided corpus analysis

Barbara Johnstone
Rhetoric in Detail

Discourse analyses of rhetorical talk and text

Edited by

Barbara Johnstone
Carnegie Mellon University

Christopher Eisenhart
University of Massachusetts at Dartmouth

John Benjamins Publishing Company
Amsterdam / Philadelphia
Studying entextualization and controversy

CDA, participant observation, computer-aided corpus analysis

Barbara Johnstone
Carnegie Mellon University

Like Chris Eisenhart’s and Susan Lawrence’s chapters in Section I, Craig Stewart’s chapter in this section enriches traditional sentence-grammar analysis in the Critical Discourse Analysis mode by adding rhetorical categories and concepts to the mix. Collectively, Eisenhart’s, Lawrence’s, and Stewart’s chapters make the point that the processes and categories linguists think of when they think about language—transitivity, modality, lexical choice, and the like—are not the only areas of choice in which subtle ideological work can occur. Eisenhart illustrates the ideological effects of genre; Lawrence uses patterns of figuration to drill into the details of the effects of naming and wording.

Stewart introduces two larger-scale variables: framing and argumentation style. The idea that human cognition is organized by means of structured sets of expectations, or frames, originated in the work of cognitive scientists like Schank and Abelson (1977). At the same time, sociologist Erving Goffman (1974) was suggesting that social interaction was organized and interpreted in terms of frames. The idea was taken up in linguistics in work by Deborah Tannen (1979, 1993) and others and in communication studies by scholars like Fan and Kosicki (1993).

Argumentation has long been a concern of philosophers and rhetoricians. Classical rhetoric offers Aristotle’s distinction between dialectic, in which arguments are based on shared sets of rules for arriving at truth, and rhetoric, the art of using arguments based on probability. In contemporary rhetoric and argumentation studies, vernacular modes of argumentation have been added to scholars’ purview, in addition to the formal, disciplinary modes that were traditionally the focus. At the intersection of rhetoric and linguistic pragmatics, “pragma-dialectics” (van Eemeren & Grootendorst 1994, 2004; Jacobs & Jackson 1989) explores how arguments arise in conversation and describes the inferential
processes people use (or should use) in everyday disputes. Stewart uses Stephen Toulmin's (1958) model of argument to describe one of the ways that scientific discourse about a controversial psychiatric study reflects popular-media discourse about it, despite the bright-line boundary scientists attempt to maintain between the scientific and the popular.

Susan Gilpin's chapter exemplifies some of the methods of participant observation. Participant observation is the primary research technique of ethnography, or the description of cultures. It was developed by cultural anthropologists interested in finding out from the perspective of natives what foreign cultures were like, particularly the small, isolated, traditional societies that have been the focus of anthropological research until fairly recently. Stocking (1983:7) describes the main features of traditional participant observation research this way: "Entering as a stranger into a small and culturally alien community, the investigator becomes for a time and in a way part of its system of face-to-face relationships, so that the data collected in some sense reflect the native's own point of view." The experience of participant observation fieldwork is a key element of anthropology's disciplinary self-image, in which such fieldwork is valued as an important formative experience for new scholars as well as because it reflects and encourages a holistic approach to describing culture and an egalitarian, relativistic ethic with respect to evaluating societies (Stocking 1983:8).

Beginning in the 1960s, anthropologists began to question the assumptions underlying participant observation, in part because worldwide decolonization began to make it clearer how deeply the approach was influenced by imperialistic attitudes about the relationships between "natives" and Western scholars. As a result, ethnographers began increasingly to talk about fieldwork and try out new ways of imagining whether and how participant observation can lead to knowledge. At the same time, scholars in other fields began to adopt participant observation research methods and join in the discussion about how to use the results.

What differentiates participant observation from casual looking around has to do with the nature of the researcher's participation in the group being studied and the care and systematicity with which records are made and analysis undertaken. Participant observers spend time developing roles for themselves in the groups they are interested in, and then more time as group members, filling one or more roles as insiders and simultaneously making systematic efforts to come to understand what is going on in the group from the perspective of other group members. Participant observers try to uncover and record the unspoken common sense of the group they are studying, the "immediate and local meanings" (Erikson 1986:119) in terms of which the local world hangs together for local people. Thus, as Spradley points out (1980:13–20), ethnography is a corrective to social theory that assumes that everyone's beliefs and behaviors can be explained in the same
terms, showing instead how social theory has to be "grounded" (Glaser & Strauss 1967; see the Introduction to Section II) in particular situations.

Every analysis of discourse starts with a corpus – a body of texts or transcripts which includes either everything the analyst wants to make claims about (all the letters Constantin Visoianu wrote to the U.S. authorities, all the government reports about the Waco incidents) or a systematic subset of it. Developing a corpus is thus an essential step in any study; any discourse analysis is in this sense a corpus analysis. However, the term corpus analysis has come to have a more specialized use both in linguistics and in rhetoric as a label for an analysis that is based on a large set of texts or transcripts (Sinclair 1991; Stubbs 1996; Biber 1988; Hart 1994; Collins, Kaufer, & Vlachos 2004). What "large" means in this context is highly variable. Among the most widely used corpora of text or talk in English are the Brown corpus of just over a million words of edited American English, the London-Lund corpus, which contains 435,000 words of spoken British English, the Lancaster-Oslo-Bergen corpus of one million words of British English, and the Colbault corpus of written and spoken English, which contains over 200 million words (Stubbs 1996:xvii-xviii). Most studies are based on a selection from one or more corpora.

Corpora like these began to be assembled at the same time computers began to be available to academic researchers. The Brown corpus, for example, was first published in 1964 “for use with digital computers,” according to its manual (Francis & Kucera 1964). In principle, one could ask the same questions of a large corpus of texts as one asks of a smaller one; the patterning of linguistic elements across time or context that can be seen in a large corpus of materials could, in principle, be seen by a human or a team of humans. But analyses of large corpora are practical only with the help of computers. Thus most corpus analysis, in this specialized sense, is assisted by computers, which can find repeated words and – if the words in the corpus are appropriately labeled, or "tagged" – structures many times faster than humans can.

The output of a computer-aided corpus analysis varies according to the computer's instructions for searching and for displaying its findings. “Concordance” software finds all the occurrences of a word and displays them with their surrounding text. Other programs produce lists of words arranged by the frequency of their occurrence (the most frequent are sometimes identified as "key words") or find repeated collocations (words that are used together). “Factor analysis” sorts word-classes into groups based on whether they co-vary. (A very influential study using this technique is Biber 1988.) These results are commonly (although not always) submitted to statistical tests to discover which are robust and which are probably accidental.

Peter Cramer’s chapter describes a computer-aided corpus analysis. Cramer shares with all the other authors in this volume the assumption that linguistic
choice at all levels is rhetorical. Unlike the others, Cramer explores not a set of choices, but one choice: the print news media's choice of a word to use as the cover term for what went on with respect to a museum exhibit in New York that included a painting that some people, among them the city's mayor, found objectionable. Focusing on a single instance of word choice, Cramer was able to track this choice in 273 texts, totaling 204,203 words. Cramer started by using a computer search tool to locate all the articles in three New York newspapers that talked about these events. He then read all of them "manually" (or, perhaps better, humanly) locating words commonly used to categorize the events, words such as battle, outcry, flap, fuss, and controversy. Then he used another electronic tool – concordance software – to locate and pull out every instance of each of these words in all the newspaper articles that constituted the corpus.

Cramer then focused on the most common of these terms (controversy was at the top of the list), exploring how they worked in the sentences they were in. Because, "in contrast to this sort of CDA study design, [Cramer's] study examined a single feature over many instances, and across large corpus of texts," he developed a way of coding each instance using a four-item list of possibilities as to what kind of event each term could be used to represent.

Cramer's methods are in some ways the inverse of Gilpin's. Gilpin examines a very small number of interactions in great detail; getting at them through her own detailed, minute-by-minute observations and through interviews with the people involved and subsequent conversations with them. Her aim is to see how the kind of proto-political discourse she calls heuristic engagement is shaped by constraints and affordances of all kinds; to answer this question, she looks at a single instance under multiple lenses. Cramer examines a very large number of texts, but focuses on a single thing about them, the word that is chosen for one set of events. Cramer looks at his data through one lens, asking how the meaning of a word is shaped by co-text, by the surrounding words. He uses his findings to illustrate the contingency and discursive constructedness of an event category, "controversy," that rhetoricians have commonly taken for granted.

References


