Studying Complex Discursive Systems
Centering Resonance Analysis of Communication

STEVEN R. CORMAN
Arizona State University
TIMOTHY KUHN
University of Colorado, Boulder
ROBERT D. MCPHEE
Arizona State University
KEVIN J. DOOLEY
Arizona State University

Scholars increasingly theorize about the power of communication to organize and structure social collectives. However, two factors threaten to impede research on these theories: limitations in the scope and range of existing methods for studying complex systems of communication and the large volume of communication produced by even small collectives. Centering resonance analysis (CRA) is a new text analysis method that has broad scope and range and can be applied to large quantities of written text and transcribed conversation. It identifies discursively important words and represents these as a network, then uses structural properties of the network to index word importance. CRA networks can be directly visualized and can be scored for resonance with other networks to support a number of spatial analysis methods. Following a critique of existing methodologies, this paper describes the theoretical basis and operational details of CRA, describes its advantages relative to other techniques, demonstrates its face validity and representational validity, and demonstrates its utility in modeling organizational knowledge. The conclusion argues for its applicability in several organizational research contexts before describing its potential for use in a broader range of applications, including media content analysis, conversation analysis, computer simulations, and models of communication systems.

Steven R. Corman (Ph.D., University of Illinois at Urbana-Champaign, 1988) is an associate professor in the Hugh Downs School of Human Communication at Arizona State University. *Timothy Kuhn* (Ph.D., Arizona State University, 2000) is an assistant professor at the University of Colorado at Boulder. *Robert D. McPhee* (Ph.D., Michigan State University, 1978) is a professor in the Hugh Downs School of Human Communication, and *Kevin J. Dooley* (Ph.D., University of Illinois at Urbana-Champaign, 1987) is a professor in management and industrial engineering, both at Arizona State University. This research was supported by grants from the College of Public Programs and the Office of the Vice-Provost for Research at Arizona State University. The authors would like to thank Ulrik Brandes for his valuable advice related to developing CRA software. Address correspondence to the first author at Hugh Downs School of Human Communication, Box 871205, Arizona State University, Tempe, AZ, 85287-1205, email: steve.corman@asu.edu.

Human Communication Research, Vol. 28 No. 2, April 2002 157-206

Although communication scholars know that the exchange and interpretation of messages is a key to human social organizing, recent theoretical developments have left little doubt that the relationship between communication and social structure is profound. Our theories emphasize the collective as a whole more than ever before, taking account of how it constitutes and is constituted by social discourse. Ellis (1999) calls discourse the "empirics of social organization and structure" (p. 69) and argues that we can only understand collective level constructions (like class and ethnicity) by studying their production in micro-practices of communication. Collective discourse is an important issue in critical studies of communication, Foucault's idea of discursive formations being one influential example. In a recent review of organizational communication theory and research, Taylor, Flanagin, Cheney, and Seibold (2001) call for elaboration of structurational approaches, attention to overlapping discursive fields, study of groups as intermediate structures, and research on new organizational forms; each of these reflects an increased focus on organization-level communication.[1] Even in the traditionally microanalytic area of interpersonal communication and personal relationships, there is a growing concern with the influence of collective-level phenomena (Adams & Allen, 1998; Miller, Cody, & McLaughlin, 1994).

As our theories become more sophisticated, they increasingly treat social collectives (if sometimes implicitly) as *complex systems*. Complexity refers to the assumption that the whole consists of more variables, processes, subsystems, and activities than we can comprehend at one time, and that many of these elements are subject to influences we can neither predict nor control (Daft, 1995; Thompson, 1967). Collectives are sets of interdependent elements that comprise a whole, itself interdependent with a broader environment and adaptable to changing circumstances. Although differing on particulars, scholars from a range of perspectives express growing hospitality toward the idea that social organization is an emergent property of a (complexly) structured system of communicating individuals (Anderson, 1999; Contractor & Grant, 1996; Corman & Scott, 1994; DeSanctis & Monge, 1999; Dooley, 1997; Eisenberg, 1990; Ford & Ford, 1995; McPhee, 1988; McPhee & Corman, 1995; McPhee & Zaug, 2000; Taylor & Van Every, 1999; Tsoukas, 1996; Weick & Roberts, 1993).

These developments point to an empirical problem on the horizon that we hope to address, at least partially, in this paper: We are ill equipped to study manifest communication behavior beyond the most modest scales of social aggregation. The more complex a collective, the more ambiguous its operation and the less knowable it is (Perrow, 1967). This principle may be behind the apparent inverse relationship between level of aggregation and attention to interaction in communication research. At the interpersonal level, we find the greatest focus on communicative interaction, but even here the difficulty of the methods (Bakeman & Gottman, 1986; Schiffrin, 1994) makes conversation and interaction analysis less common than perhaps it should be in the communication discipline. Moving to the small group level, the detailed study of interaction is more difficult and rare still, and at the multigroup (i.e. organizational or cultural) level it is virtually unheard of. As our emerging theories and models of communication grow in scope to embrace complex collective phenomena, we risk making them unworkable as guides for empirical research. Put simply, we worry that the existing body of communication research methods is incapable of handling the complexity being theorized in the discipline.

In this article, we propose a general analytical framework called centering resonance analysis (CRA), a flexible means of representing the content of large sets of messages, and assisting in their analysis. We present CRA as a text analysis method, suited to studying formalized communication like reports, letters, memos, emails, and other written texts. However, as we illustrate below, CRA is also applicable to transcribed conversations. Developments in voice recognition technology promise the ability in the near future to "textualize" live conversation, making it amenable to CRA methods as well. In establishing the need for CRA, we ground our arguments in the concerns of organizational communication because this is where the need to study large volumes of communication is especially acute. However, this does not mean CRA is only useful in research on formal organizations. As we explain in the conclusion, CRA has many applications in the broader discipline to problems in rhetoric, mass communication, microlevel analysis of dyadic and group interaction, and simulation of communication systems.

In the next section we argue that existing research methods, including ethnographies, conversation analysis, questionnaires, and computational models, are inadequate for the task of testing claims about complex organizational communication systems. Following that, we describe CRA, its theoretical and operational definitions, and its relationships to existing methods. Then we illustrate its validity in three applications, analysis of conversations in a meeting, comparison to human readings of interview texts, and uncovering structure in an interdisciplinary research network. In the conclusion we discuss the likely advantages of CRA for studying complex organizational communication systems, and explain its potential for application in other areas of communication research.

STUDYING COMPLEX ORGANIZATIONAL DISCOURSE

As an illustration of the difficult task of understanding complex organizational communication phenomena with existing methods, let us consider Browning and Beyer's (1998) grounded theory analysis of the development of shared standards in SEMATECH, a consortium of suppliers and manufacturers in the U.S. semiconductor industry. Drawing on structuration theory (Giddens, 1979, 1984), the authors aim to show how communication related to several events led companies in the industry to construct coordinated meanings and build new cooperative structures for developing shared standards. Their underlying claim that communication processes led to the emergence of structures and the adoption of standards is at the heart of the organization/communication relationship.

Browning and Beyer (1998) use a critical incident technique, in which they first analyze the outcomes of four years of interviews, observations, and archival material. They then conduct follow-up interviews, member checks, and document searches to provide support for their claims. They document the ways industry problems are discovered, interpreted, and responded to by the SEMATECH team, and the resulting changes in "consciousness, action, frameworks, meanings, constraints, power and authority, obligations, symbols, and outcomes that interacted reflexively over time to produce new structures" (p. 221). We believe the processes they report and document are both interesting and convincing.

At the same time, the scope of understanding provided by the Browning and Beyer study is limited by characteristics of their method. Interpretive-ethnographic methods seek insights expressed in specific accounts and links between accounts, rather than deeper or more subtle overall patterns in micropractices throughout the organization. Browning and Beyer often build on the developments experienced, reflexively noticed, and articulated directly by their informants. These naturally exhibit selection and demand effects, which tend to filter out critical details about communication processes and how they work. This helps us understand the collective perception of a past event, an important phenomenon in itself and one perfectly suited to study by ethnographic methods. However, something additional is needed if we are to understand how such collective perceptions are built over time.

We illustrate this claim with Browning and Beyer's (1998) observations on "discovering the costs of non-standardization." Their major finding in this regard was that SEMATECH participants came to a realization: Secrecy practices adopted by participating companies were preventing standardization that would benefit the entire industry. Although, this was readily apparent in the research data,

> [o]ther outcomes were more subtle. The most general was that the successful collaboration that emerged for choosing equipment provided an important precedent for later cooperation. For example, during the discussions that occurred, assignees discovered that it was surprisingly easy to "switch hats" and change from their parent company's viewpoint to that of SEMATECH as the

> decision seemed to require. They learned they could base their decisions on what was best for the shared mission rather than on what was best for the parent company. (p. 226)

Interestingly, the authors go on to note that this "sense of shared missions was not developed equally by all assignees in the early days of SEMATECH" (p. 241, note 8). So, although we can accept the idea that initial collaboration must have functioned as a precedent for later cooperation, we lack understanding of how this process worked. We know there was resistance, at least in the early going. How is it that the collaborative spirit overcame this resistance by member companies to become a resource for use in later interaction? How and why did the assignees suddenly find it easy to "switch hats"?

We cannot discern the answers from Browning and Beyer's data because "individuals seldom remember, and organizations rarely record, how behaviors and interpretations stabilize over the course of the structuring process. As an interaction order solidifies, one's analytic focus [necessarily] shifts back to the institutional realm" (Barley, 1986, p. 83). In other words, critical incident analysis treats the collaboration phenomenon as a *fait accomplis,* postinterpreted through institutional structures of the organization. This is an important enough phenomenon to understand, but we have little idea how it happened.

A more process-oriented complex systems approach to the SEMATECH case would require a wider ranging study, both in scope and time. The collective structures and developing standards emerged in parallel and serial fashions over a long period. They interpenetrated other structures at multiple levels of analysis, were influenced by numerous variables and events, and were subject to environmental exigencies that impinged on members' actions. The reports of these incidents in texts and member accounts interact with a complex system of personal cognitive biases, serial transmission effects, political processes, organizational structure, and the like.

Without access to the actual organization-wide communication behavior as it happened, it is difficult to see how researchers could ever really understand these organization-wide processes. Worse, access to actual organization-wide communication behavior is a problem in itself. Detecting and describing complex patterns spread out over a vast field of discourse may well be too difficult a task for informants, or for human analysts of accounts and residual texts. Indeed, we believe it is beyond any analytic approach that does not fundamentally summarize, transform, and re-present ongoing organizational discourse.

To be clear, our mission here is not to criticize or fault the Browning and Beyer study, which makes valuable contributions to our understanding of the communication processes at SEMATECH. Rather, our point is that even this carefully conducted, multiyear study—much more extensive than is the norm—leaves unexplained significant elements of the underlying organization-wide communication/organization relationship. We also note that this is not a problem unique to Browning and Beyer's study. Several esteemed studies exhibit the same methodological drawbacks. For instance, Barley's (1986) well-known research contrasting technological change processes in two hospitals likewise lacks attention to potentially important communication processes. Barley seeks to explain the impact of new scanning technology on hospital employees' scripts for interaction involving the machines. Assuming the role of participant-observer, he recorded 91 scanning sessions and participants' interpretations of these events over the course of a year, finding dramatically different outcomes for each hospital's interaction order. In this way, Barley was able to link institution and action during a structuring process that considered interactions between occupants of two organizational roles during CT scans. The more profound claim, that members' communication and interpretation reproduced and transformed structure, however, requires further evidence. In particular, an assessment of members' non-CT interactions is needed, as is an assessment of the influence of other institutional and organizational discourse, such as memoranda, newsletters, and hospital strategy statements, on the emergent interaction order.

Another illustration of an exemplary study (celebrated as such in Frost & Stablein, 1992) is Gersick's (1988) study of punctuated equilibrium in work teams. Rather than examining organization-wide or even department-wide structuring, this study shows how communication moves small, task-oriented groups through stages of development as they work on their tasks. Gersick studied detailed transcripts of eight groups' meetings; she also condensed meanings of sequences of utterances and conducted interviews with members of four additional groups. In this substantial data set, Gersick searched for milestones in groups' decision-making interaction, and found that the groups were characterized by a long period of inertia created by interaction in their first meeting, punctuated by brief periods of revolutionary change at the midpoint of their "lives." These revolutionary periods hold the potential to alter group structures and future interaction.

Despite the detailed and rigorous nature of her approach, Gersick's (1988) claims about the transitions do not capture the complexity of communication's structuring quality. In examining the critical events associated solely with decision-making interaction, other "activity threads" (Poole, 1983; Poole & Roth, 1989), such as socio-emotional interaction, are ignored. Moreover, as groups interact over time, they appropriate structures that may accumulate and lead to an "avalanche" (Bak & Chen, 1991; see also Brunt, 2000). This may interact with existing communication patterns and time pressures in the production of revolutionary change.

An approach that is able to support the sort of claims made by Browning and Beyer (1998), Barley (1986), and Gersick (1988) is needed if the emerging theories about the equivalence of organization and communication are to be compelling. Even these exemplary studies are forced to bracket-out significant portions of complex communication processes in generating their insights. If methods used in these studies are wanting, the pertinent question will be what methods could be used to gain rich, multilevel access to organization-level communication processes in all their complexity.

The simple answer is that no methods in our existing disciplinary arsenal are up to the job. Let us consider a few likely candidates. Ethnography using participant observation is one way to study complex organizational systems. It is clearly up to the task of understanding complexity, but is limited in scope to the local context of the ethnographer. At best, a skilled ethnographer is no less limited by the "horizon of observability" (Friedkin, 1983) than the most competent organization member: Both are necessarily only partly aware of operations of the collective system (Weick & Roberts, 1993; Sandelands & Stablein, 1987). Studying collective-level communication process as they develop would require an impracticably high number of ethnographers.

The situation is much the same for conversation analysis (Boden, 1994; Fairhurst & Chandler, 1989; Putnam & Fairhurst, 2001). It is perhaps even better suited than ethnography to understanding the specific communicative practices that enact the environment, instantiate activities, and reproduce perceived relationships. However, the intense micro-interpretation it entails (Schiffrin, 1994) makes it even more limited in range than ethnography. In particular, it cannot cope with the problem of system-level simultaneity (McPhee, Corman, & Dooley, 1999). The impact on the system of a conversation in the here and now is dependent on parallel conversations at other times and places, many unknown to the conversants and analysts.

Questionnaires can potentially be much broader in range than ethnography or conversation analysis, but suffer their own limitations. As a practical matter, one could never administer questionnaires fast enough to study message behavior.[2] Questionnaires therefore yield data about communication as it has been perceived by the respondent, often referred to as "self-report" data. Granted, self-reported perceptions are valuable data for many purposes (Howard, 1994; Richards, 1985). They are probably even necessary to help understand the behavior of organization members. However, it is a serious mistake to treat perceptions of communication as a substitute for observation of message behavior. The actions of individual communicators have both unacknowledged conditions and unintended consequences (Giddens, 1984) and, *ipso facto,* these cannot be accounted for in their self-reports. Moreover, people perceive organizational communication behaviors in ways that are systematically biased by their local context and organizational conditions (Corman & Bradford, 1993), so self-reports are not isomorphic with communication behavior at the level of the system.[3] This is a critical point because explanations of complex communication systems require accurate and detailed data about sequences of behaviors (Sandelands & Stablein, 1987; Poole, Van de Ven, Dooley, & Holmes, 2000), the sort of interlocked interaction Weick (1979) shows to be fundamental to organizing.

One interesting possibility for studying complex systems claims is the application of computational models. These approaches deal with organizational complexity by modeling its consequences in computer simulations (Carley & Prietula, 1994; Contractor & Grant, 1996; Corman, 1996; Drazin & Rao, 1996; Hyatt, Contractor, & Jones, 1997; Latane, 1996). Simulations have no difficulty handling large volumes of information, yet they are not as deft at observing large volumes of communication. For one thing, contemporary usage of simulation is geared toward hypothesis generation, not descriptive analysis or hypothesis testing (Dooley, in press; Hyatt et al., 1997). Simulations are used to ground narrow hypotheses (relative to the complexity of the system) about how the system should behave empirically if the assumptions behind the simulation are true. Simulations therefore create demand for follow-up observations, and actually increase our need for methods that can analyze large volumes of actual discourse. We also note that, like other broad-range methods, simulations invariably gloss important discursive details by treating communication as an unproblematic transfer of information between simulated agents. As we argue in the conclusion, CRA offers one possibility for incorporating more sophisticated models of communication in computer simulations of complex organizations.

The problem, in a nutshell, is that we have some methods that are broad in understanding but restricted in range, and other methods that are restricted in understanding but broad in range. To study complex systems of organizational communication, we need both. To have range, the method must be applicable across different scales of aggregation and contexts, and must be able to operate in these simultaneously. To provide deep understanding of the organization/communication relationship, we must have access to the words people speak or write. At the same time, we cannot reduce communication to message transmission. We need empirical evidence of the communicative coordination and control of activity to be able to claim that communication and organization are mutually constitutive. It would seem that the only way to simultaneously meet these needs is to listen in on all the communication in an organization for a considerable period of time, no matter when or where it occurs.

"BE CAREFUL WHAT YOU WISH FOR"

Unfortunately, indications are that organizations of even modest size produce scary quantities of messages. As an illustration of this, let us extrapolate from a high-resolution interaction study of a single organization member. In Gronn's (1983) study, "the talk of a school principal and everyone with whom he spoke over two days was received by an unobtrusive radio microphone attached to the principal's lapel" (p. 3). The talk was recorded and transcribed. Gronn reported that the recordings "yielded in excess of 300 pages of [trans]scripts" (p. 3).

In honor of what must have been a challenging study at the time it was conducted, let us define one *Gronn* as the volume of utterances produced by one organization member in a 5-day workweek, expressed as the number of transcript pages those messages would occupy. To provide an initial estimate of the value of the *Gronn,* let us assume that the principal's utterances accounted for half of the page count reported in the study, the other half being occupied by utterances of his conversation partners. This suggests that the principal produced 75-pages worth of utterances per workday, pegging the value of the *Gronn* at 375.

The *Gronn,* therefore, helps us get an idea of what an investigation of the communication/organization relationship might entail. Suppose we secured the cooperation of a small, 50-person organization, fitted its members with wireless microphones, and taped and transcribed everything for a week. If they were all like the principal, they would generate 50 *Gronn* or 18,750 pages of transcripts during this 1-week period. That is enough to fill 37.5 reams of paper, a stack over six feet tall. Realistically, all organization members are not going to be as talkative as a school principal. Yet even if we have estimated the value of the *Gronn* at 200% of its true value, our small organization would still generate such a volume of utterances that no one could realistically expect to read the whole transcript, much less analyze it in detail.

Documents also form an important corpus of discourse that must be analyzed in order to understand complex organizational phenomena. A single large corporation may have several hundred thousand to over a million web pages on its corporate intranet. Various databases and data warehouses store and archive millions of documents that may be relevant for some future reference; each individual in the company may receive and send hundreds of pieces of electronic mail and other memos daily. These correspondences are not trivial. Not only is important information and opinion carried in written documents, but reading and responding to such memos can encompass as much as half a manager's available work time.

These illustrations show why we singled out organizational commu-

TABLE 1
Summary of Text Analysis Approaches

Approach	*Goal of analysis*	*Representative works*
Inference	Draw conclusions about what is not given in the text	Connolly et al. (1997) Dong & Agogino (1997) Eizirik et al. (1993) Federici & Pirelli (1997) Hahn et al.'s MEDSYNDIKATE (1999) Jacobs & Rau (1993) Leacock et al. (1998) Mitkov (1997) Perez-Carballo & Strzalkowski (2000)
Positioning	Position texts in a field of other texts	Bartell et al. (1992) CATPAC (Terra Research and Computing, 1994) Doerfel & Barnett (1999) Fraenkel & Klein (1999) LSA (Landauer et al., 1998) Lund & Burgess (1996) Nomoto & Nitta (1997) Stephen (2000) Treadwell & Harrison (1994)
Representation	Extract or distill an efficient representation of the content	Alterman & Bookman (1990) Carley (1997b) Danowski (1982) Galal et al. (1999) General content analysis approaches Humphrey (1999) Keyword frequency approaches Rice & Danowski (1993) WORDij (Danowski, 1992) TACT (Siemens, 1993)

nication research as a leading-edge case of the need for high-volume methods. Theories are quickly orienting toward larger scale, complex systems, but the empirical data that the theories demand exceeds the capacities of our methods. Thus it does us little good to wish for organizational (in the sense of organization-wide) communication data unless we have some means of coping with this empirical phenomenon. CRA is a method capable of analyzing a 50-*Gronn* set of transcribed conversations. However, there are other existing methods that could be applied to such data too. Before describing CRA in detail, we review existing automated text analysis approaches in the next section.

APPROACHES TO AUTOMATED TEXT ANALYSIS

A time-tested method of dealing with data floods is computer processing, which will be the approach advocated here. Specifically, we believe that a computerized analysis can serve as a substitute (though not necessarily a replacement) for a complete reading of a text. This goal has for some time been a priority in the well-established field of text analysis (Diefenbach, 2001; Popping, 2000; Roberts, 1997). Text analysis is a crowded field, and this article is not the place for a detailed review. However, to position CRA with respect to existing efforts, we distinguish three general approaches based on inference, positioning, and representation. Table 1 lists representative works from each.

Inference

A first approach to textual analysis seeks to assign meanings to linguistic inputs, but those meanings are usually at a level of abstraction above the word content directly given in the text. For instance, if a text includes words such as branches, leaves, and roots, the method might conclude that the passage is likely about trees. To arrive at such a conclusion, inference approaches apply rules, learned patterns, or ontologies to a text in ways that allow a computer program to distinguish important from unimportant material, and thereby to infer meanings. A foundational assumption in this approach is that the meanings of words are discernable based on probabilistic causal relationships among other words in a text; this assumption underlies the common use of Bayesian networks (Eizirik, Barbosa, & Mendes, 1993) and leads this approach to be commonly linked with artificial intelligence (Dong & Agogino, 1997).

A primary methodological distinction in inference-based approaches is between syntax-driven lexical analysis and semantic grammars (Jurafsky & Martin, 2000). In syntax-driven lexical analysis, a textual input is parsed into its constituent linguistic units to arrive at a structural representation of the text, which is then passed through a semantic analyzer (software program) to arrive at a meaning representation. This semantic analyzer plays a crucial role in determining meaning, but the analyzer must be "trained" to look for particular word co-occurrences or grammatical forms.

The second inference approach, semantic grammars, develops rules based on entities and relations in a particular domain from which the text is drawn with the goal of resolving word ambiguities in particular contexts. Therefore, the rules for making sense of ambiguous words are specific to the domain and are often quite effective in automatically identifying the meanings of metaphors, pronouns, and the like. As in the case of the semantic analyzers mentioned above, these rules must be "trained" into the systems, usually through the use of a domain-specific dictionary or by running a program through a corpus of text native to that domain to gain a familiarity with word use.

An example of the inference approach is provided by Leacock, Miller, and Chodorow (1998), who focus on identifying the senses of words. Their Topical/Local Classifier (TLC) is designed to statistically infer the senses of verbs, nouns, and adjectives using words in the textual neighborhood. The TLC uses a Bayesian approach to find the word meaning that is most probable, given a set of cues contained in a researcher-selected "window" of *N* words around a polysemous word. It then tags the text by separating grammatical units and reducing word variants to a base form. These tags and the meanings of the cue words were trained into the TLC through human coding and the WordNet lexical database (Miller, 1995), which includes a dictionary of words arranged conceptually that can "teach" the TLC about synonyms, antonyms, hyponyms (superordinate-subordinate relationships), and the like. This training enables the classifier to recognize topics and meanings of words when analyzing subsequent texts.

Inferential approaches such as the TLC are most useful when the goal involves making sense out of ambiguous words or finding related texts in a given domain. Increases in computing power allow these approaches to be trained and used relatively inexpensively, meaning that the accuracy of identifications can improve over time. Shortcomings of such approaches, however, relate to the need to train a domain-specific grammar to the program such that the rules it learns are probably not transportable to other contexts. In addition, relatively few approaches here employ a linguistic or semiotic theory to arrive at semantic representations, instead assuming that lexical co-occurrence provides the information necessary to make sense of word meanings.

Positioning

The second automated text analysis approach produces a characterization of a text in relation to a collection of other texts in a set or corpus; this approach has become increasingly common in knowledge-management technologies for organizations (Adams, 2001). Here, the key element is the creation of a semantic space: "a space, often with a large number of dimensions, in which words or concepts are represented by points; the position of each point along each axis is somehow related to the meaning of the word" (Lund & Burgess, 1996, p. 203). This space can be constructed either by human raters' definitions of words or by computerized analyses of lexical co-occurrence, producing a vector for the focal unit that is placed within the semantic space (Salton & Buckley, 1988). For instance, CATPAC (Terra Research and Computing, 1994) is a positioning approach that scans a text to find the most frequently used words, then creates a word adjacency matrix indicating the frequency with which each pair of words co-occurs in the text using a window that slides across the text, similar to that described in the inference approach. This matrix can then be analyzed using clustering techniques and multidimensional scaling to facilitate comprehension of the relationships (Woelfel, 1993). Positioning methods assume that (a) authors of texts in the domain structure word co-occurrence in ways similar to other authors in the domain such that relationships among words are similar across a body of text, (b) that windows sliding across texts capture consistent patterns of co-occurrence, and (c) that the semantic space into which word or text vectors are inserted is substantively important. In line with the example mentioned above, a text analyzed through a positioning approach would create a semantic space in which leaves, branches, and roots would be found close together, but would be distanced from hooves, tails, and manes.

A related positioning approach in the communication discipline is Galileo, created by Joseph Woelfel in association with several collaborators (see Woelfel & Fink, 1980). This theory-method complex uses a variety of techniques, including relatively casual content analysis (similar to our description of representational approaches below), to determine a set of important concepts relevant to some topic. This is followed by questionnaires to induce subjects to report psychological distance between the concepts. The distance scores are then averaged to generate data for an MDS representation of collective conceptual space. The creators' goal was to find relatively simple laws of motion that described changes in this psychological space when messages and other influence processes led concepts in the space to move closer to, or farther from one another.[4]

A prominent positioning approach in the text analysis field is latent semantic analysis (LSA), a theory and method designed by Landauer and colleauges (Foltz, Kintsch, & Landauer, 1998; Landauer, Foltz, & Laham, 1998). LSA uses a large corpus of machine-readable language to construct its semantic space, a matrix in which each row represents a unique word, and each column is a text passage or other context; the cell entries are the frequencies with which the row elements appear in the column elements. This matrix is analyzed by singular value decomposition (a type of factor analysis) so that the meanings of words can be represented as vectors in the resulting space. In contrast to the inference approaches discussed above, LSA "uses no humanly constructed dictionaries, knowledge bases, semantic networks, grammars, syntactic parsers, morphologies, or the like, and takes as its input only raw text" (Landauer et al., 1998, p. 263). However, LSA requires that the semantic space be trained with a large corpus of written text (Kintsch, 2001). The purpose of LSA differs from inference approaches in that it does not attempt to distinguish alternative senses of a word, but rather constructs and modifies vectors and semantic spaces with subsequent texts, producing overall semantic spaces of about 300 dimensions, on average (Kintsch, 2001). LSA, then, can represent both word and document meanings through vector techniques in multidimensional space based on semantic relatedness, and new entities can be represented in the same space.

Positioning approaches are based on the (often implicit) theoretical assumption that since humans learn the meanings of words through reading and hearing them used in particular combinations, word co-occurrence is a suitable basis for representing meanings (Kintsch, 2001; Landauer et al., 1998; Spence & Owens, 1990). Positioning approaches are particularly useful when researchers use textual sources to characterize the symbols used in a given social system, such as the International Communication Association (Doerfel & Barnett, 1999), or the gender, feminist, and women's studies community (Stephen, 2000). It allows placement of words or documents in semantic space, facilitating understanding of individual texts, comparisons between texts, and a consideration of the larger corpus. When, however, the methods require training to construct the semantic space (as in LSA), or the use of terms varies across social or temporal divisions, the positioning of vectors is open to questioning. Therefore, the validity of the placement of vectors in semantic space depends on the quality of the construction of the semantic space itself.

Representation

A final type of text analysis method attempts to produce characterizations of a text by extracting or distilling meaningful content from its words, without reference to a training set, corpus, semantic network, or ontology. In other words, these representations are meaningful by themselves. Returning to the example employed above, a representational approach would identify leaves, branches, and roots as particularly prominent words important to the overall meaning of the text in question, and would produce a characterization of that text based on the information provided by these and other words. The selection of important words is based on a set of criteria that describe the functions of words in texts, such as the claim in keyword indexing that the importance of a word is inversely related to the frequency of its occurrence (Baeza-Yates & Ribiero-Neto, 1999). In contrast to the approaches described above, representational approaches would not seek to make inferences about words' common referent, or to understand word meanings in a conceptual space, though they may be incorporated in other analyses that involve inference or positioning.

Representational approaches tend to be similar to content analysis procedures that use manifest content (as opposed to latent content) to arrive at descriptions of texts. Although such content analysis procedures often accept a wide range of symbolic material in addition to texts, their emphasis on objective, quantitative measurement of features of texts often detracts from their ability to produce stand-alone statements of texts' meanings, particularly through attempts to identify units of analysis and to generate coding categories.

In communication studies, an example of a representational approach that avoids the problems of manifest content analysis is provided by Danowski and colleagues (Danowski, 1982, 1988, 1993; Rice & Danowski, 1993), who have developed a sophisticated approach to the representation of texts. Their method uses an automated content analysis that creates networks of words based on their co-occurrence in a researcher-selected "window" of *N* words. This window slides across a text, incrementing the frame one word each step, to locate clusters of words that frequently co-occur, as well as words that link clusters together. From this set of co-occurring words, a word network (as the representation of the text) is constructed. The content and structure of this network can then be linked to a variety of relevant information, such as organizational market share or financial performance (Danowski, 1993), or can be used to manage organizational discussions through the introduction of optimal messages (Danowski, 1982, 1988).

This technique has an important drawback: The use of a window (also used in several other methods described above) created and sized without reference to the intent of the speaker or writer or a theory of linguistic behavior. We question the validity of the window as a unitizing scheme for analyzing text. Representational methods (and unitizing schemes) should mirror as much as possible the way authors or speakers represent their thoughts in text. It is hard to imagine a corresponding mental process where authors create connections among words in sliding windows of four or five words. The demands of coherent communication, as we argue below, require a higher-level focus, at least considering the composition of whole utterances. Even granting the basic validity of the windowing scheme, the size of the windows is a worrisome issue. If the windows are too small, they miss important connections intended by the authors. However, if the windows are too large, they are liable to span structural features in utterances, generating substantial "noise" in the form of connections among words with little discursive importance (Carley, 1997a). Window size may also not just be a matter of determining a single correct value. Different texts could require different window sizes, or the optimal size might vary across different portions of the same text. With no theoretical basis for sizing windows, it is difficult to see how such differences could be identified and accounted for.

These shortcomings notwithstanding, the work by Danowski and colleagues should be recognized as an important innovation in text analysis. The method produces a network of interconnected words that is used in discerning the referential meanings of the text. Appropriately so, as concepts networks are a well-recognized and respected format for accounting for meaning in the social sciences (Axelrod, 1976; Schank & Abelson, 1977). Unlike some inference and positioning approaches that also employ network metaphors, Danowski's networks are intrinsically meaningful and do not depend on a context of rules, training sets, or other documents. Once constructed, the networks can be analyzed according to their global and local characteristics, including qualities of the relationships among words and the differing conceptions of the same word among networks (Carley, 1997b; Carley & Kaufer, 1993). Compared to other representational approaches, word networks are very rich data structures that preserve significantly more information about a text than keywords or word frequency statistics.

Representational approaches tend to be weak, however, in that they are rarely based on a linguistic theory about text production or interpretation, opening their methodological choices, such as the "window" sliding across the text, and conclusions to question. Further, some representational techniques, such as TACT (Siemens, 1993) require a researcher to manually tag the text before analysis, marking up particular elements for attention in subsequent analysis. Although this aids in the retrieval of specific passages in a text, it is both labor-intensive and introduces the possibility of error in the process, because the results of the program depend on the accuracy and comprehensiveness of the tagging. Wretched would be the researcher (and his or her graduate assistants) who set out to pre-tag a 50-*Gronn* dataset.

In sum, the approaches to automated text analysis discussed here involve a variety of assumptions and methodological choices. We conclude that attempts to deal with data floods must use computing power wisely to respond to the limitations of these methods. In particular, we see a need for a representational method that meets three criteria. First, it should be based on a network representation of associated words to take advantage of the rich and complex data structure offered by that form. Second, it should represent intentional, discursive acts of competent authors and speakers. Units of analysis and rules for linking words should be theoretically grounded in discursive acts. Third, it should be versatile and transportable across contexts. Making text representations independent of dictionaries, corpora, or collections of other texts would provide researchers maximum flexibility in analyzing and comparing different groups, people, time periods, and contexts.

CENTERING RESONANCE ANALYSIS

CRA is a new kind of network text analysis that meets the criteria just outlined. Here we give a brief overview of our argument for the method before explaining it in more detail. CRA uses linguistic analysis to identify important words in utterances and to link these into a network. Important words are those making up noun phrases, which are potential centers in the utterance. Accumulating these words and their links over a set of utterances making up a text (or segment of conversation) yields a network that represents the aggregate of intentional acts by the author or speaker to deploy words and connect them to other words. Some words in this network are especially influential due to their location in the structure, tying together many other words and helping organize the whole. Thus, by analyzing the CRA network structure, we can index the structural importance of words without reference to other texts, corpora, rule sets, training data, and so forth.

CRA is grounded in a theory of communicative coherence, specifically centering theory (Grosz, Weinstein, & Joshi, 1995; Walker, Joshi, & Prince, 1998). Competent authors or speakers generate utterances that are locally coherent by focusing their statements on conversational "centers" (McKoon & Ratcliffe, 1998; Lecoeuche, Robertson, Barry, & Mellish, 2000). Centers are words and noun phrases constituting the subjects and objects of utterances, and are generally entities such as objects, events, or persons (Gordon, Grosz, & Gilliom, 1993). In a written text, for example, each sentence except the first has a backward-looking center that refers to a preferred forward-looking center expressed in the previous utterance. The author/speaker also establishes an ordered set of forward-looking centers to which the next utterance can coherently refer (Gordon & Hendrick, 1998; Grosz, et al. 1995). A given utterance is made locally coherent by connecting the backward-looking center in a predictable way to previous forward-looking centers.

Under the assumptions of centering theory, then, communicators speak or write coherently by creating utterances that deploy a stream of words comprising centers (more specifically, noun phrases) in a strategic way, creating a semantic structure of words.[5] Coherence, in turn, is a fundamental criterion for understandable, relevant communication (Jackson, Jacobs, & Rossi, 1987; Kellerman & Sleight, 1989; Sperber & Wilson, 1995). This notion provides the basis for an efficient automatic coding system for the content of communication, grounded in centering theory, replacing the arbitrary windows of the network text analysis approaches described above.

A stock issue in coding is unitizing, breaking a stream of communication into codable units. In CRA, we unitize communication in terms of words contained in the noun phrases that make up utterances. Utterances are sentences or the conversational equivalent thereof (Auld & White, 1956), and they represent finite groups of centers constructed by communicators to fit into a coherent stream of other utterances. Noun phrases identify the centers, and the words making them up are the codable (linkable) units.

Let us emphasize that in CRA we are not doing an analysis of the centering process itself. That would require looking at connections between utterances. In CRA, we assume that centering or a process like it is operating, and we are concerned with the deployment of a stream of centers within utterances. CRA's attention to center-related words within utterances is justified because our intent is to represent the essential content of messages and their impact on the coordination and control of activity, not to examine the ability of interactants to achieve coherence, though it is likely that the former goal subsumes the latter. We assume coherent communication by competent writers and speakers, and attempt to extract some of its associated structure. To explain the method in greater depth, we next describe the four steps involved in generating a CRA representation of a text.

Selection

CRA categorizes texts in terms of a pattern of connections among words that are crucial to the centering process. Compiling the words and their connections across all utterances in a text yields a CRA network representing the text. The procedure is in the spirit of earlier network text analysis methods, but represents a more restricted form of linking that takes account of the discursive structure of the utterance. It begins with selection. Rather than linking all words that fall within an arbitrarily sized window of text, CRA parses an utterance into its component noun phrases. A noun phrase is a noun plus zero or more additional nouns and/or adjectives, which serves as the subject or object of a sentence. Determiners (the, an, a, etc.), which can also be parts of noun phrases, are dropped in CRA analyses. Thus, a noun phrase is represented by one or more words, and a sentence can consist of one or more noun phrases. Because the centering process operates largely through noun phrases, this step acts as a filter that turns sentences into sequenced sets of words contained in noun phrases.

Before moving to the next step of the CRA method, we address a few important issues. First, CRA intentionally excludes the other main component of utterances, verb phrases. In the linguistic model underlying CRA, verb phrases would be the "action" components linking different noun phrases in an utterance. As such, they are really a different kind of information, explaining the contexts of action that link the centers. Given our concern to represent the manifest content of texts, rather than provide inferences about the significance of particular utterances in ongoing interaction, the exclusion of verb phrases is logical. Noun phrases, according to linguistic semanticists (e.g., Frawley, 1992), are the only elements that can be unambiguously classified as entities in discourse.[6] Nouns denote conceptual categories that provide more salient discourse information than verbs and generally control the use and expression of verb phrases (Hopper & Thompson, 1984, 1985; Langacker, 1987). Moreover, nouns are less likely than verbs to be temporally situated, and thus more likely to be portrayed as entities (i.e., concepts) in discourse (Givón, 1984). In short, the parsing of texts into networks of noun phrases, and the concomitant exclusion of verbs, aligns with both our guiding model of discourse coherence and our desire to represent the manifest content of texts.[7]

Second, is the question of whether to include pronouns in our analysis. Although a significant amount of research in linguistics is devoted to disambiguating the referents of pronouns in discourse (Cloitre & Bever, 1988; Connolly, Burger, & Day, 1997; Gerken & Bever, 1986), our approach makes the inclusion or exclusion of pronouns contingent on the purpose of the investigation and the quantity of texts involved. A similar logic guided Rice and Danowski's (1993) study of individuals' open-ended comments regarding voice-mail. They excluded pronouns due to a concern with semantic structures across types of users, rather than those characterizing individual respondents. In similar cases, CRA can safely forego disambiguation of pronouns, dropping them from the analysis. In most spoken and written texts, proper nouns or referents are introduced before pronouns (Lecouche et al., 2000) and topic shifts are introduced by specific nouns (Passonneu, 1998), meaning that little textual information is lost by dropping pronouns (as backward-looking centers) that appear later. In other cases, however, an analyst may have reasons to consider the identification of actors pronominalized (by words such as she, he, it, I, we, or they) relevant to the analysis, and make appropriate substitutions of disambiguated nouns for the pronouns that represent them.[8]

Third, is the question of whether we should use stemming to convert words to more basic root forms before analyzing them. In general, we are wary of this technique. It is easy to think of cases where stemming might obscure important shades of meaning. The statements "the negotiators connected on the issues" and "there was a disconnect between the negotiators on the issues" would stem to the same set of objects, despite quite opposite meanings. Indeed, the effectiveness of stemming in general is an open question (Frakes & Bazea-Yates, 1992), and more sophisticated forms of stemming depend on sources of training data that are not always available or techniques that are not practical (Bazea-Yates & Ribeiro-Neto, 1999, p. 168). Therefore, we adopt only minimal affix stemming, going from plural to singular forms by removing "s" or "es" suffixes.

Linking

The second step, linking, converts the word sequences into networks of relationships among words. According to the rationale above, the author or speaker of a text being analyzed with CRA intentionally groups the words into noun phrases and strings these phrases together (using verbs, pronouns, determiners, etc.) to form an utterance. CRA linking rules attempt to embody those choices. To begin with, all words comprising the centers in the utterance are linked sequentially. In the majority of cases, where noun phrases contain one or two words, the sequential connections capture all the linkage intended by the author because no higher-order connections are possible without crossing the boundaries of the noun phrases. However, there are cases where three or more words are contained in a single noun phrase. In that case, the sequential links do not exhaust the connectedness possible in the set created by the author. Hence, we link all possible pairs of words within the noun phrases. For example, the phrase "complex discursive system" would generate the links: complex-discursive, discursive-system, and complex-system.[9]

Accumulating these links over a set of utterances comprising a text (or a paper, a collection of papers, a transcribed speaking turn or set of turns, and so on) yields a symmetric, valued, undirected network whose nodes represent the center-related words. In a CRA network, the link values represent the number of times the words were linked in the text according to the rules above. This network, when indexed as we describe in the next section, becomes a fundamental representation of the text and forms the basis for all applications of CRA.

Indexing

The third step in CRA is indexing. Here, the network of word associations is analyzed to determine the relative influence of each node. This is a key step in differentiating the words, so it deserves some discussion here. Network metaphors are always based on some abstract notion of flow. In the case of CRA networks, we would say there is a flow of meaning. To the extent that a CRA network is structured, some words are more influential than others in channeling flows of meaning. They are literally more meaning-full than other words in the network. Thus, identifying the structural influence of the words allows one to measure this property. We operationalize this idea of influence as the centrality of a given word in the CRA network. Although a variety of measures could be used, centering theory points us most clearly toward betweenness centrality.[10] To our knowledge, the concept was first formalized by Anthonisse (1971), who described it as the rush in a graph: "The rush in an element is the total flow through the element, resulting from a flow between each pair of vertices" (p.1). Freeman (1979) contrasted betweenness centrality with other classic measures in a way that is instructive. Consider a minimal network of four peripheral nodes that are all connected to a single node in the middle (but not to each other). There are at least three senses in which the node in the middle is central. It is connected to a lot of nodes, relative to the others, which is the notion of degree centrality. It is also very directly connected to all of the other nodes, whereas the peripheral nodes are at least two steps away from each other. This reflects the notion of closeness centrality, usually measured as the average number of steps required to reach other nodes in the network from a focal node. The middle node is also central in the sense that any kind of resources flowing in the network (meaning, in the case of CRA networks) must flow through it. This is the idea of "rush" or betweenness centrality described above. Each of these measures can be computed for the network as a whole, as well as for the individual nodes (Wasserman & Faust, 1994).

Of the three kinds, betweenness centrality is the most appropriate for estimating the influence of words in CRA. Degree centrality, the most often applied measure in earlier network text analysis efforts (e.g., Carley & Kaufer, 1993), takes only the local connections of each node into account. Closeness centrality is better in that it considers the entire network structure. However, it cannot be computed for disconnected graphs, which, in CRA, are not only possible but likely for low-coherence texts. More important, closeness undervalues the influence of nodes lying on paths connecting disparate parts of the network because nodes in the center of large, densely connected clusters will have higher closeness, on the average. From the standpoint of maintaining coherence in a structure of words, this "tying-together" function is crucial. Betweenness centrality therefore best represents the extent to which a particular centering word (represented by a network node) mediates chains of association in the CRA network. It tells us more than any other measure about how a given node channels the "rush" of meaning through a network of centering words. Therefore (adapting notation in Wasserman & Faust, 1994), the influence I of a word i in text T is operationalized as:

$$I_i^T = \frac{\sum_{j<k} g_{jk}(i)/g_{jk}}{[(N-1)(N-2)/2]} \tag{1}$$

where g_{jk} is the number of shortest paths connecting the jth and kth words, $g_{jk}(i)$ is the number of those paths containing word i, and N is the number of words in the network.

Resonance is a latent property of the structure of a CRA network. Although resonance is a property of a single network, it is only realized in the presence of an external signal (i.e., another network), just as a physical material only resonates when brought into contact with an external vibrating wave. To the extent that other texts or utterances deploy words in the same way as a given network, they may be said to resonate with it. To understand how we make operational the resonance of one text with another, assume that texts A and B have been represented as CRA networks. The two texts may be of similar nature, or one may be considered a query and the other a text potentially relevant to the query. There are two ways of measuring resonance, one less specific and based on the words common in the two documents, the other more specific and based on word pairs common in the two documents.

Word resonance is calculated directly from the influence scores of the words in the two texts. Let the (unique) words (after parsing into phrases) for text A be represented by $[w_1^A, w_2^A, \ldots w_{N(A)}^A]$ with corresponding influence scores of $[I_1^A, I_2^A, \ldots I_{N(A)}^A]$, where N(A) is the number of (unique) words in text A. Similarly, text B has words $[w_1^B, w_2^B, \ldots w_{N(B)}^B]$ with influence scores $[I_1^B, I_2^B, \ldots I_{N(B)}^B]$. In general $N(A) \neq N(B)$. Let the indicator function α_{ij}^{AB} be equal to 1 if w_i^A and w_j^B are the same words, and be equal to zero if w_i^A and w_j^B are not the same words. The word resonance between texts A and B, WR_{AB}, is defined by:

$$WR_{AB} = \sum_{i=1}^{N(A)} \sum_{j=1}^{N(B)} I_i^A \bullet I_j^B \bullet \alpha_{ij}^{AB} \tag{2}$$

The more two texts frequently use the same words in influential positions, the more word resonance they have. The more word resonance they have, the more the communicators used the same words, and the more those words were prominent in structuring the text's coherence. Word resonance is a more general measure of the mutual relevance of two texts, and has applications in the modeling of large corpora.

This measure is unstandardized in the sense that resonance will increase naturally as the two texts become longer in length and contain more common words. There are cases, however, where a standardized measure is more appropriate. For example, in positioning documents relative to one another (as described below), one does not necessarily want to overemphasize differences in document length, number of words, and so on. In these cases, the appropriate standardized measure of resonance is given by:

$$WR'_{AB} = WR_{AB} \,/\, \sqrt{\sum_{i=1}^{N(A)} (I_i^A)^2 \bullet \sum_{j=1}^{N(B)} (I_j^B)^2} \tag{3}$$

which is structurally equivalent to the manner in which the covariance between two random variables is standardized to a measure of correlation.

Pair resonance is estimated using co-occurring word pairs, as opposed to co-occurring words. Let the frequency weighted pair influence of words i and j in text T be given by:

$$P_{ij}^{T} = I_{i}^{T} \bullet I_{j}^{T} \bullet F_{ij}^{T} \tag{4}$$

where I_i^T is the influence of w_i^T, I_j^T is the influence of w_j^T, and F_{ij}^T is the number of times that w_i^T and w_j^T co-occur (their corresponding nodes are connected directly by an edge) in text T. If text T has N (unique) terms, then there will be (N • (N-1) / 2) pairs, but many of them will have a value of $F_{ij}^T = 0$ as they will not represent connected terms. Let the indicator function β_{ijkl}^{AB} be equal to 1 (a) if the two word sets (w_i^A, w_j^A) and (w_k^B, w_l^B) are equivalent (regardless of the manner in which the set elements are ordered), and (b) F_{ij}^A and F_{kl}^B both are equal to one (the sets represent connected nodes); otherwise the indicator is zero. In other words, the indicator function β_{ijkl}^{AB} is 1 when the corresponding pairs of co-occurring words co-occur in both texts. The pair resonance PR_{AB} is defined by:

$$PR_{AB} = \sum_{i=1}^{N(A)-1} \left(\sum_{j=i+1}^{N(A)} \left(\sum_{k=1}^{N(B)-1} \left[\sum_{l=k+1}^{N(B)} P_{ij}^{A} \bullet P_{kl}^{B} \bullet \beta_{ijkl}^{AB} \right] \right) \right) \tag{5}$$

The more pair resonance two texts have, the more their authors assembled words in the same ways, in order to make their communication coherent. Pair resonance is a more sensitive measure of the mutual relevance of two texts than word resonance, because it takes account not only of the words and their position in the network, but also how they were assembled in the utterances. As such, pair resonance has applications in high-accuracy information retrieval tasks.

For the same reasons discussed previously, it may be desirable to form a standardized measure of pair resonance:

$$PR'_{AB} = PR_{AB} \;/\; \sqrt{\left(\sum_{i=1}^{N(A)-1} \sum_{j=i+1}^{N(A)} \left(P_{ij}^{A} \right)^{2} \right)} \bullet \sqrt{\left(\sum_{k=1}^{N(B)-1} \sum_{l=k+1}^{N(B)} \left(P_{kl}^{B} \right)^{2} \right)} \tag{6}$$

In summary, CRA is a representational technique that describes the extent to which words are prominent in creating a structural pattern of coherence in a text. The description provided to this point shows that CRA possesses distinct advantages over other text analysis approaches. First, because CRA networks are independent of text corpora and training sets, they are highly transportable. The influence values for words are calculated only once for a given text, and CRA networks can be computed for single texts, parts texts, or sensible aggregation of texts. Second, because it does not depend on training or rules sets, CRA accommo-

TABLE 2
Arguments by Advocates and Opponents of Change

*Arguments by advocates**	*Arguments by opponents***
(James, 82) In all our previous discussions, and Topteam had before we came up here to Palomino, we thought, and many individuals raised the point, that we should be trying, difficult as it is for all of us, to be objective about what we say in these reports. We should think of the organization not in terms of the incumbency of any one position, but as to how the organization itself should be best structured from the point of view. And we are dodging the issue because we are saying this. It may possibly point the finger to any one of us, and that is too delicate an area for us to discuss.	(Arnold, 88) It seems to me that while there may be some benefit from the exercise I really see that we will go through the exercise for several hours, maybe several days, and Topteam will have a very heated discussion. It is inconceivable to me that it could be resolved without a heated discussion. Then the whole thing can be a complete waste of time because of the relationships between the chairman and the President and how they see the job.
(Jack, 85) Topteam ought to have enough nerve, or gumption, to look at the company and say to our present President in writing, this. Look, Mr. President, you have to be organized because we have certain weaknesses, and if we do not do this, the company is not going to move. Topteam was willing to do it, but when it comes to individuals, you want to check it out, because you may have to give and you may have to take losses.	(Sam, 89) When a man is made the chief executive officer, and I'm just using the President by way of example, then he is going to determine the kind of a structure he is going to operate in effectively in order to achieve the desired goals. You're saying to him, Topteam will make you the President but this is the way you're going to have to operate.
(Jack, 92) But Topteam has a responsibility as a group to put this kind of recommendation on the board the same way as Topteam did other recommendations, rather than leave it again for a haphazard putting together without the resources to put together or call another meeting for that purpose.	(Arnold, 94) Traditionally the President of the United States or the Prime Minister under the parliamentary system alone chooses his own cabinet and for the most part the choice of cabinet depends on the skills of that particular individual. I think it was obvious that Kennedy chose a very weak secretary of state because he himself wanted to do the secretary of state's job. I think that for a President to come unto the job without this choice being made by him puts him at a very serious disadvantage.
	(Irving, 110) I think that the groupings that are made are really the prerogative of the chairman and whoever he nominates to be the President. And those groupings must be made on two bases, and I do not know if we can go much deeper with it over here. Number one is what is a natural grouping business wise, and number two is the competence of the people available, in the judgment of the chairman and his President. That will obviously have to determine to some degree the groupings fundamentally based upon the natural groupings that are available to us. But I think that beyond that, you have to take people into account and into consideration and Topteam should leave here ready to say that whatever these people deem to be in the best interest of the corporation, this is what Topteam will have to go along with. I do not think we can go beyond that point.

* These statements are edited slightly from the transcript. They include only the statements made after the suggestion to let the next President decide on how his or her own structure was made.
** These statements are edited slightly from the transcript.

dates emergence of new terms or shifts in relationships among existing words and concepts, as should be expected in knowledge development and other forms of innovation. Third, relative to other representational techniques, CRA is structurally sensitive in that it accounts for all likely chains of association among the words that make texts and conversations coherent. This makes the technique more sensitive to complex associations in the text than statistical methods based on word frequency or local co-occurrence. Fourth, CRA is based on a theory of communicative coherence that avoids the imposition of an arbitrary "window" sliding over text to locate word co-occurrence.

Application

The final step in CRA is application, wherein the indexed CRA network or a set of these are used for some analysis task. CRA networks are useful for a wide variety of tasks. One is visualization of the CRA network for text understanding purposes. It is possible to "read" a CRA network and get a good, though necessarily compressed, sense of the content of the original text. In the next section, we illustrate how that can be done in one application, and then empirically test a key presumption behind the visualization. After that, we turn to a third application, spatial modeling of resonance scores, which shows how CRA networks can be used to analyze the intellectual organization of a set of scholars. Due to limitations of space and medium, other applications of CRA networks, such as information retrieval and thematic analysis of collections, are not demonstrated here.

CRA APPLICATIONS

Application 1: Analyzing Group Interaction

To demonstrate the face validity of CRA network visualizations, we conducted an analysis of a transcribed videotape of organizational decision making. "After Mr. Sam" (Hammond & Pearson, 1974) is a documentary film compiled from a long discussion, at a resort called Palomino, by managers of Steinberg Limited, a Canadian retail chain. The meeting took place in the early 1970s, just before the founder and CEO, Sam Steinberg, appointed a successor and retired.[11] In an early segment of the discussion, some managers argued that structural changes were needed for the company before Mr. Steinberg's successor was appointed; others argued against such structural changes. In Table 2, important statements supporting the need for change (by a group we have called "the advocates") and opposing it (by "the opponents") are cited in slightly edited form. Figures 1 and 2 display the CRA networks of those sets of statements (one for the text in each column). Figure 3 shows the CRA network derived from the whole relevant section of the transcript (acts 77 through 110, pp. 6–9 of the transcript), including the arguments collected in the samples for the two groups.

In producing the CRA networks, we followed the general procedures described above. We noted above that pronoun disambiguation should be decided on a case-by-case basis. In this analysis, we did pronoun disambiguation, substituting the word *topteam* for the pronoun "we," in cases where the speaker clearly was using "we" to refer to the top management committee as an empowered group. The disambiguation was appropriate in this case because the discussion was focused on an issue of whether the topteam group itself, or the new president would make key personnel decisions.

In the advocates'network shown in Figure 1, the most influential and most frequently appearing concept is *topteam* (words appearing in the CRA network are shown here in italics). It is linked to other concepts, like *individual, nerve,* and *recommendation,* the three next-most influential words. As the graph shows, these words are in turn linked to other influential words including *point*, *gumption*, *company,* and *haphazard.*[12] A look at Table 2 shows why *topteam* is influential: It is linked to a number of other words in various sentences, and these other words get some of their influence by being linked directly to *topteam.* The pattern in Figure 1 clearly represents the focus of the advocates' arguments that *topteam* has the duty (established earlier in meetings of this group) to make change *recommendations,* but is lacking the *gumption* to do this, and is approaching the task in a *haphazard* way.

Figure 2 represents the arguments by the opponents who do not want to make any definite change recommendation. They emphasize the *prerogative* of the next *president,* in concert with future board *chairman* Sam Steinberg, to decide on changes himself. They believe he needs to have freedom of *choice* to do so, much as the U.S. President can make *cabinet* appointments. The word *topteam* is also influential in the opponents' network, but achieves that influence through its place in arguments that *topteam* is in danger of usurping presidential choice and should be aware of that danger.

Figure 3 exhibits a graph based on a larger text sample, about three pages of transcript from which the advocates' and opponents' arguments were sampled. The network is striking in showing the top words for each group (*topteam* and *president*) as central, yet distinct foci of somewhat different, yet connected subnetworks. In the upper portion *topteam* is the anchor, and is connected with other high influence words, *individual, recommendation,* and *people.* In the lower portion, we find *president* connected

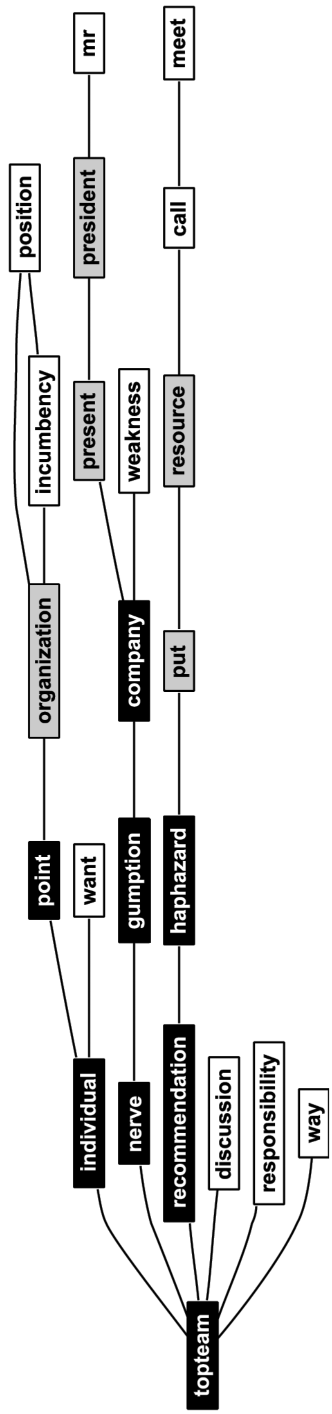

Figure 1. CRA Network of the Advocates
NOTE: Highest influence words are black, next highest are gray, and next highest are white. Many words lower in influence are not shown.

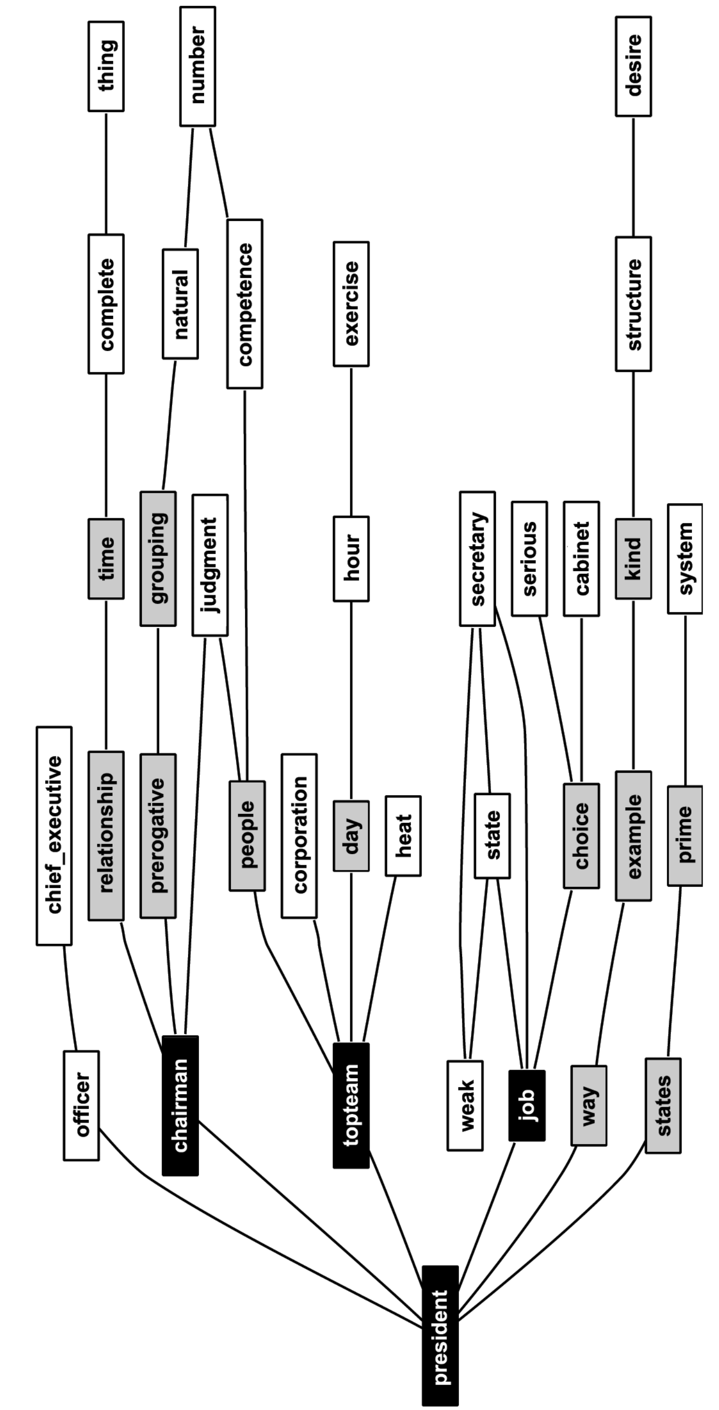

Figure 2. CRA Network of the Opponents
NOTE: Highest influence words are black, next highest are gray, and next highest are white. Many words lower in influence are not shown.

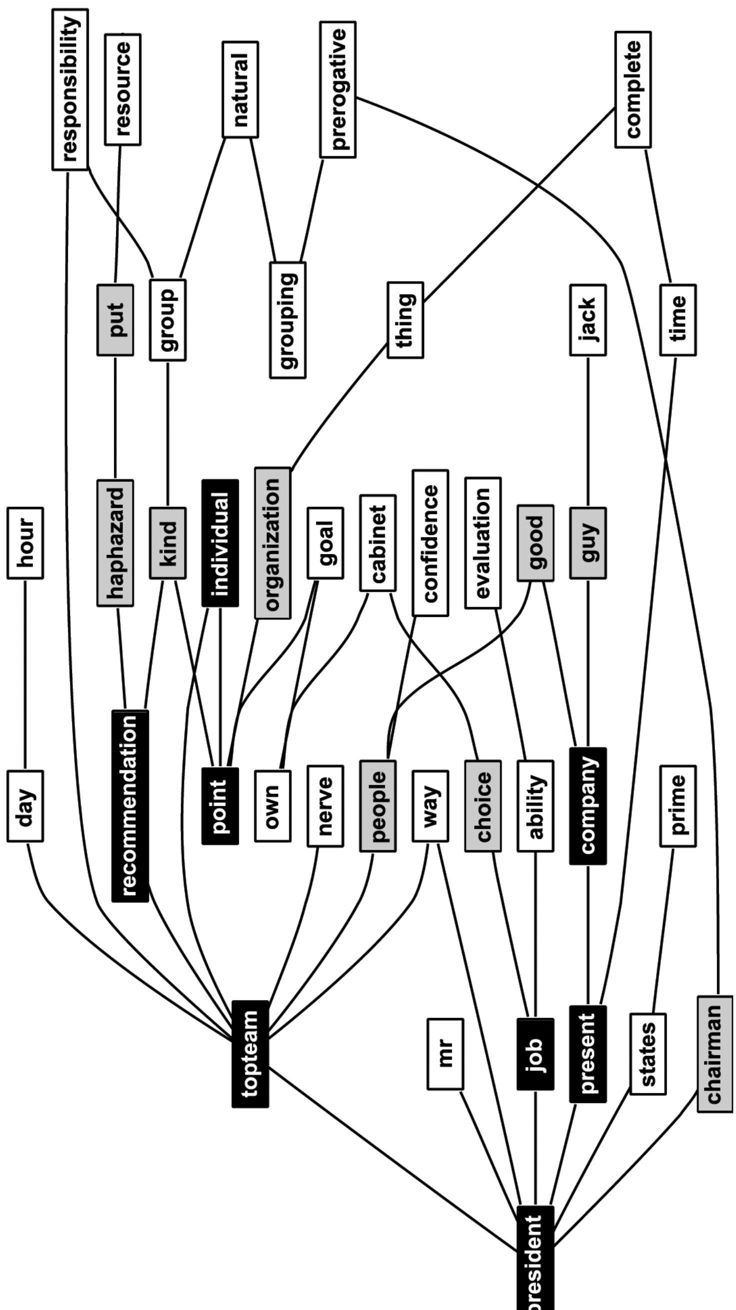

Figure 3. CRA Network of the Larger Transcript Segment
NOTE: Highest influence words are black, next highest are gray, and next highest are white. Many words lower in influence are not shown.

to *job*, *present*(—*company*), *chairman*, and *choice*. Interestingly, it is not the case that all words on the bottom were spoken only by opponents and vice versa; for example the chain, *president—present—company,* can by found in the advocates' network in Figure 1. The word *recommendation* is rather influential because both groups used it in ways that linked it to various other influential concepts, as they advocated quite different forms of recommendations.

Even though words in the two parts of the network were not exclusively spoken by one side or another, the top and bottom parts of the graph align pretty well with the arguments of the two sides. The advocates sought to link *topteam* with the *responsibility* to make a *recommendation* and not allow some *haphazard* process to govern changes. The opponents focused on the *president*'s prerogative of *choice* regarding who will do key *jobs*. We conclude, then, that the overall graph conveys the discursive division between the *recommendation* role of *topteam* on one hand, and the *choice*-making *job* of the *president* (and the *chairman*) on the other. Of course, this composite example is based on very limited samples of texts. The best use of CRA will depend on much larger bodies of text, in which important concept linkages will be more frequently repeated.

Application 2: Collective Perception of Texts

We hope that readers will be convinced by the foregoing example of the face validity of CRA visualization. However, there is a question of whether CRA "maps" a set of words in the way human readers would. This is an issue of representational validity (Folger, Hewes, & Poole, 1984; Poole & Folger, 1981): Does CRA represent texts in a way that is related to the way humans represent them?[13] There is no unambiguous standard against which to assess representational validity, so our approach was to compare the CRA representation of a text with a similar form of representation produced by human readers.

Toward this end, we presented 63 undergraduate students with one of two excerpts of employee responses to a set of identical interview questions. These interviews were conducted as part of a larger study of planned change within one division of a municipal government in the southwestern U.S. (Kuhn, 2000). In this test, there were two conditions corresponding to the two interview texts: 21 participants read a text produced by employee A, and 42 read a text produced by employee B.[14] The interview texts were drawn from verbatim transcripts of the interviews with the employees, but were prepared so that the interviewer's questions were removed; this step was necessary to avoid a response bias by the readers, who might be influenced by the words employed by the interviewer rather than the interviewee.

Several steps were taken to enhance the quality of participants' responses. To aid in their comprehension of their task, we gave each participant an unrelated and smaller example of how an author connects words in a text.[15] To assist in participants' understanding of the context of the interviews they would be reading, we gave verbal and written explanations of the organization and its planned change. Finally, to increase motivation, each participant was promised and given course credit upon completion of the task.

For the research task, participants were asked to read the interview text. Next, they were presented with an empty lower-half matrix labeled with the 10 most frequently occurring words in the interview. The words were arranged so that all possible combinations of the words were represented by the cells of the lower-half matrix. Participants were instructed to "Place a check mark in the boxes that indicate how (interviewee name) connects words together in his or her thoughts about re-engineering." In effect, then, each participant produced an adjacency matrix that could be compared to the adjacency of the same words as estimated by CRA.

To make this comparison, we first generated a CRA network for the interview text, then extracted a subnetwork whose nodes corresponded to the words given to the participants. Next, we generated a lower-half matrix of frequency-weighted pair influence values (See equation 4) for all possible pairs of the 10 words. If words were not linked in the text, the value of the cell was zero. The more influential the words were and the more they co-occurred, the larger the values for those words in the matrix. We reasoned that if the centering and discourse assumptions behind CRA were true, then participants should be unlikely to perceive links between pairs of words with zero cells in the matrix. The larger the value in a cell, the more likely participants should be to perceive a link between the words.

To compare the matrices produced by the participants with the matrices produced by the CRA technique, we employed the Quadratic Assignment Procedure (QAP). This is a nonparametric technique that can assess the degree of similarity between two matrices of the same size, even if the observations are nonindependent (Baker & Hubert, 1981; Hubert & Schultz, 1976; Krackhardt, 1987, 1988). The technique permutes the rows and columns of one matrix (using UCINET, 2,500 permutations) to generate a distribution of possible correlations, against which the observed correlation with the original matrix is compared. A test of significance comparable to that for Pearson's Product-Moment correlation is derived.

Using QAP, each individual's matrix of connected words was compared with the appropriate CRA matrix as a criterion. Then, for each condition, the average of correlations among participants' matrices and the CRA-derived matrix was computed. For readers of the text produced by employee A, the average relationship was weak ($r = .25$, $p > .05$, $N = 21$).

Likewise, for readers of employee B's text, there was little association ($r = .22$, $p > .05$, $N = 42$). These results indicate that the individual naïve readers only duplicate the pattern of association among words generated by CRA weakly if at all.

At first glance, then, there seemed to be little reason to conclude that CRA's representation matches the interpretation of participants. Yet, as students of social networks have found, individuals' perceptions of communicative phenomena are often inaccurate at the level of the individual, but can provide valuable knowledge when considered in the aggregate (Krachardt, 1987; Krackhardt & Kilduff, 1990). Following this logic, we considered the entire set of responses in each condition, and an interesting result emerged. To accomplish this analysis, we first pooled all participants' responses for each condition, aggregating their matrices and then averaging the result in each cell of the ten-by-ten matrix. This resulted in a figure between 0 and 1.0 in each cell; the entries represented the percentage of readers in each condition who believed that the words were connected by the employee in the interview text. These aggregate matrices were then compared with the corresponding CRA matrices, as described above.

In both cases, the resulting correlation was significant, in stark contrast to the nonsignificant results from the averaged correlations presented previously. The relationship between readers' aggregated responses and the CRA representation was strong both in condition A ($r = .570$, $p = .006$, $N = 21$) and condition B ($r = .509$, $p = .016$, $N = 42$). These results provide evidence of a meaningful association between the matrices generated by the aggregated responses of participants, on the one hand, and the CRA representation, on the other. We interpret this as evidence in favor of the representational validity of CRA. We also regard the size of the correlations as impressive, given the wide range of variables not controlled in this simple study. For example, reading comprehension skills vary widely in student populations such as the one used in this test. Taking factors like this into account would likely reveal an even bigger association between the CRA analysis and the aggregate participant data.

The results suggest that CRA captures an emergent property of a text that is not reducible to a single individual's reading. Instead, CRA's representation may be similar to how readers in the aggregate attribute coherence to discourse. This finding mirrors Folger et al.'s (1984) notion that the validation criterion should be an intersubjective representation that "transcends any individual instantiation" (p. 149). It is also consistent with arguments and findings reported by Woelfel and Fink (1980). They report several studies showing that conceptual relations measured as interconcept distances achieve high and increasing reliability and stability, while retaining valuable information and structure, when averaged over increasing numbers of respondents. They argue that averaged relations represent a collective cognitive structure, with the averaging process eliminating noise in the measures (but also real individual and subgroup differences).

CRA networks, then, do seem to represent texts in a way that is related to how humans represent texts. Though it is a matter for further formal testing, we believe the representational validity extends to other CRA-based analyses of multiple texts and volumes larger than a few transcript pages. In the next section, we show how CRA can be used to analyze the written products of a larger heterogeneous group of scholars and organize them into practically useful structures.

Application 3: Positioning of Authors

Resonance, as described above, is a measure of the mutual relevance of two texts based on their CRA networks. The more they resonate, the more their CRA networks are similar. Therefore, computing scores for all pairs of texts in some set yields a similarity matrix. Given a set of objects and similarities between them, a number of useful spatial modeling techniques can be applied to help organize the objects, highlighting important similarities and differences between them. Applying this idea to texts, we can characterize the conceptual structure of the sources from which the texts were drawn. This approach to the analysis of texts has already received a good deal of attention in the communication discipline. As mentioned above, analysts have examined the conceptual structure of the International Communication Association and this journal (Barnett & Danowski, 1992; Doerfel & Barnett, 1999; Stephen, 1999), conceptions of participation among European managers (Stohl, 1993), users' interpretations of a voice-mail system (Rice & Danowski, 1993), image and groups in a college culture (Treadwell & Harrison, 1994), and gender/feminist/women's studies research in the field (Stephen, 2000). Spatial modeling, then, is a recognized procedure in communication studies for analyzing the relationship among texts.[16]

To illustrate the application of CRA in spatial modeling, we briefly describe its use for clustering university faculty in an interdisciplinary research area. It is as true for university faculty as for business people: "Effective knowledge seekers almost always need to cross corporate boundaries and ignore representative structures to get what they need. . . . Expertise not reflected in a position's job description doesn't show up on the organizational chart" (Davenport & Prusak, 1998, p. 73). Although traditionally most academic research occurs within the boundaries of departments, colleges, and disciplines, the majority of "tough problems" that are important today are interdisciplinary. This requires the university and its faculty to be able to identify experts and possible cohorts across departmental and collegiate boundaries, especially as a means of exploiting shifting funding priorities.

Environmental health is an example of such an interdisciplinary area. Its issues require expertise from both the physical and social sciences, and cut across numerous disciplinary boundaries. The Office of the Vice Provost for Research (OVPR) at a large university provided us with information about some 55 faculty in the university who OVPR believed might have interest and expertise in environmental health. The faculty represented over 15 different departments and were distributed over five different colleges. It was the belief of OVPR that there was little existing interaction among these individuals outside their departments, and that most were probably not aware of each other or their corresponding interests.

OVPR gave us a database consisting of personal statements of research interests, grant proposal summaries, and abstracts of published works of the researchers. From this database, we decided to use article titles, article abstracts, personal statements, and abstracts of funding proposals as evidence of each individual's manifest knowledge. In so doing, we chose to include all works listed on the database, regardless of the date of publication or the authorship position of the faculty member. We recognize that these texts do not exhaust any researcher's knowledge of a content area; however, the combination of these readily available public documents provides a legitimate source to help us understand the professional expertise that any individual researcher is likely to claim. In the 55-person database provided by OVPR, there was too little information on 7 researchers for their records to be meaningfully analyzed. In some cases, article titles but not abstracts were available; in other cases, there were simply too few entries. Thus, our final set consisted of 48 researchers. Using the technique described above, we computed standardized word resonance scores for each faculty member scored against all the others.

A matrix of similarities among the 48 researchers was submitted to hierarchical cluster analysis using Ward's method for agglomeration. Results of this analysis are shown in Figure 4. The tightest clusters are shown in the shallow brackets toward the left. These are gradually combined into higher-order clusters as the linkage distance (similarity) criterion is relaxed, until that criterion is relaxed so far that we are left with one large cluster. In interpreting these plots, one looks for clean breaks that yield a manageable number of distinct clusters. Figure 4 clearly shows two distinct groups that are not merged until the very end. The top main cluster contains two clear subclusters that remain separate until linkage distance is relaxed to over 20. In the bottom main cluster, four such subclusters are discernable at roughly the same linkage distance.

To aid in visualization of these clusters, we applied multidimensional

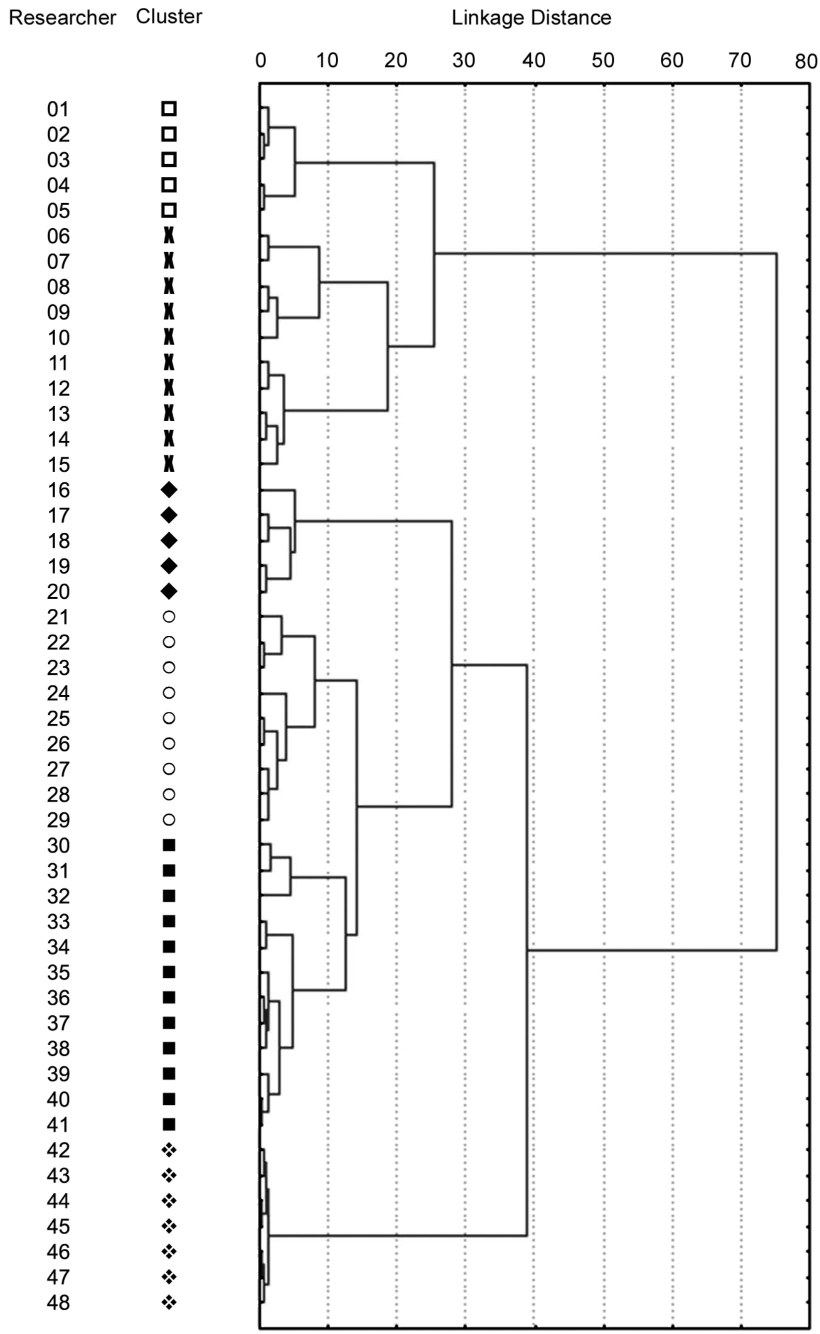

Figure 4. Hierarchical Cluster Analysis of CRA Dissimilarities for OVPR Data
NOTE: Distances are based on word resonance. Special characters indicate the cluster groupings and correspond to those used in Figure 5.

scaling (MDS) to the similarity data. MDS is a descriptive technique that, like cluster analysis, starts with a set of distances between objects. It uses an iterative procedure to produce a smaller dimensional space that optimally represents the original distances. In this application, we used nonmetric multidimensional scaling because it has a tendency to produce more interpretable spaces. We used the program KYST2a (Kruskal & Wish, 1978) to scale the distances. A 2-dimensional representation produced an adequate fit to the data (stress = 0.17).

MDS produces a set of points in n-dimensional space as its output. The cluster analysis described above reveals groupings of the scaled objects. Thus, we present a hybrid in Figure 5, with the clusters superimposed on a scatter plot of the MDS results. It is apparent from Figure 5 that the results of the MDS and clustering procedure closely agree. The main clusters are distinct in the space and the subclusters are spatially distinct within these. There is, however, some disagreement among techniques. The right main cluster "bulges" into the left main cluster somewhat along the horizontal axis. Within the right cluster, two members of the subcluster represented by small circles are separated from the bulk of their group. Still, the overall agreement between the two techniques is good, and we judge this to be a "clean" clustering.

To interpret the clusters, we looked at the curricula vitae and CRA maps of researchers, looking for similarities among members of the clusters and subclusters. The clearest-cut distinction is between two main clusters, separated on the horizontal axis. We interpret this axis in hard-soft science terms: Researchers in the left cluster are physical science oriented, and studying or measuring small-scale chemical and biological processes. Those toward the right are concerned with larger scale phenomena affecting humans and populations of humans.

Within the microscience cluster on the left, we find two subclusters. Researchers in one group clearly study cell-level processes, especially DNA and genetic processes, and related disorders. We label this cluster cellular generics. In the second subcluster, the researchers are concerned with measurement and sensing issues, as applied to physical, biological, and biochemical systems. We label this cluster measurement of physical and biological systems.

Within the cluster on the right, there are four subclusters. In one, we find researchers who study basic biological processes. They predominately do experimental studies using humans and animal models. We label this subcluster experimental biology and biochemistry. Another group includes researchers interested in diseases, stressors, pathogens, and patterns in such dysfunctions. This cluster we label diseases and disorders. Next we have a large group that studies psychology, sociology, and communication in individuals, families, and larger social groups. Social sciences seems

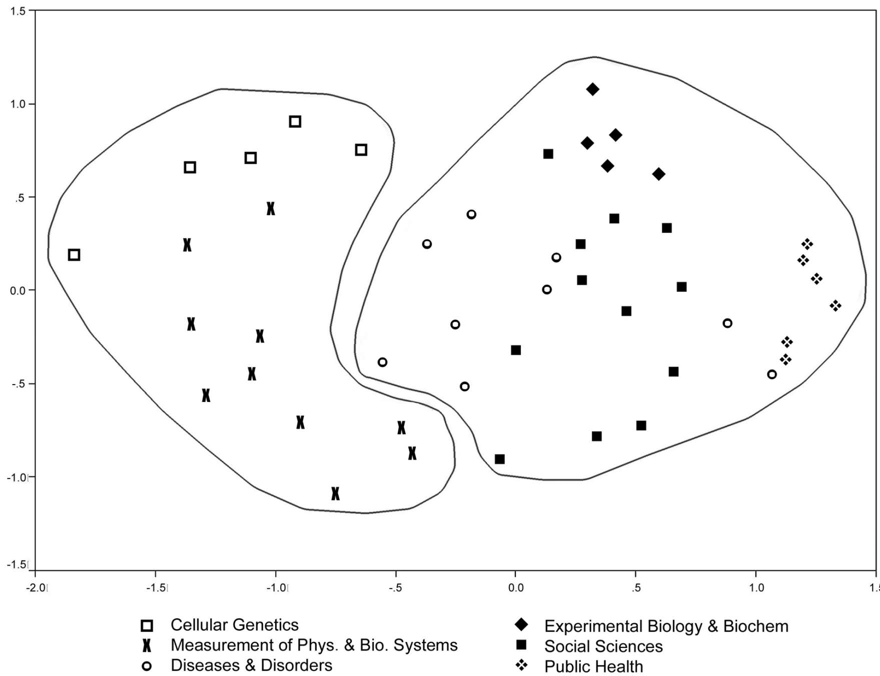

Figure 5. Hierarchical Clusters Superimposed on a Two-Dimensional Scaling of OVPR Data
NOTE: Distances are based on word resonance.

to be the appropriate label for this subcluster. Finally, there is a group of faculty whose work clearly centers on public health.

Maps like these have heuristic value in that they provide a sense of how a group of authors is organized with respect to a common field of knowledge or activity domain. The maps have practical applications as well, and OVPR could use the map in Figure 5 for practical purposes. What is the best interdisciplinary team to pursue a grant in environmental health? A "breadth" strategy would dictate including members from each cluster, to best tap diverse knowledge resources of the organization. A "depth" strategy might dictate focusing grant-getting efforts in teams from particular clusters. A more general application would be to fill structural holes in the organizational network (Burt, 1992) by ensuring that members of clusters know and have the opportunity to interact with their cluster-mates. Alternatively, cluster maps could be used to identify structural holes that could be filled by hiring employees with the requisite attributes or skills.

APPLICATIONS IN COMMUNICATION RESEARCH

Organizational Communication

We began this essay by arguing that research methods currently available are not adequate for studying complex discursive systems. We advanced CRA as a method with the potential to fill some of the gap, particularly in endeavors involving large quantities of textual data. We want to emphasize here that CRA is neither a panacea nor an all-purpose method. It does, however, have the potential to provide new insights into traditional problems and to make significant contributions to the understanding of complex systems of interaction. To illustrate this, we turned to particular examples from the three studies critiqued in the introduction, explaining how we think CRA-based methods could be applied to good effect.

We began with Browning and Beyer's (1998) observations on "discovering the costs of non-standardization." What light could CRA cast on this phenomenon? We speculate that with sufficient access to live interaction, we could have used CRA to track the build-up of a sense of shared mission in discourse. For example, in the early going, we would expect the centering words related to the member companies to be influential and not particularly interrelated. Conversely, we would expect the word SEMATECH to have low influence in structuring organizational discourse. As participants come to the realization that they are all reinventing the same wheels, words related to particular companies would become more interrelated with one another and with words representing jointly used technologies. We could track changes in linkage and influence of words, starting in personal and group conversations, then spreading through the organization, perhaps as a contagion (Carley, 1991; Contractor & Grant, 1996), bandwagon (Drazin & Rao, 1996), or some other diffusion process.

As new discursive structures spread, we would expect to see the words associated with particular companies decrease in influence, in favor of words like *collaboration, cooperation,* and *SEMATECH*. The finding that participants were "switching hats" also suggests the emergence of a new identity resource that participants can use in manifest acts of identification (Scott, Corman, & Cheney, 1998) with SEMATECH. Accordingly, we would have observed changes in pronominalization, especially a shift in the use of "we" references from the individual companies to the consortium.

A second example of CRA's potential for organizational communication is in the study of structural change discussed by Barley (1986). An investigation utilizing CRA could have expanded the units of analysis to include discourse of other types, such as hospital memoranda and records of interactions unrelated to the CT scanners. In conjunction with Barley's notes taken during conversations, tape recordings, and postexamination interviews, CRA would be able to aid in a fine-grained assessment of the changes in interaction patterns due to its ability to extract and link centering words across a wide range of texts. For instance, much of Barley's argument revolves around actors' changing evaluations of radiologists' *experience* and *qualifications;* with CRA, one could examine the web of terms generally associated with words such as these. Claims about changes in the interaction order would be significantly bolstered with evidence that shows modifications in members' discursive structures over time; this is just the sort of support that CRA can provide.

Likewise, studies of the communicative construction of group development, such as Gersick's (1988) study of time-based transitions in work groups, could be aided by CRA. The immense set of textual data Gersick collected (primarily transcribed team meetings) was made manageable in her analysis by searching for milestones and condensing sequences of utterances. Such moves would not be necessary using CRA, since the method is designed to handle large quantities of text and thereby avoid the data reduction and bracketing seen in Gersick's study. Substantively, CRA could trace how words related to time and the group's sense of urgency become manifest in collective conversation, including those discussions not explicitly related to the group's overall task. Data of this sort could demonstrate how group interaction structures accumulate over time in ways not always evident to either observers or participants (Seyfarth, 2000), and could also show how these subtle interaction patterns can lead to discontinuous and complex changes in group dynamics, providing greater insight into the ways communication constitutes both decision making and group structures (Poole & Hirokawa, 1996). If Gersick's application of the punctuated equilibrium model characterizes other teams, changes based on the recognition of time pressure should manifest themselves in collective conversation; in other words, a method capable of analyzing all conversation and surfacing complex discursive structures is required to round out the theory. As explained throughout this paper, CRA exhibits this capability.

Broader Applications of CRA

Though we have framed this paper in terms of applying CRA to the study of organizational texts, we believe it potentially has wide applicability in human communication research. To begin with, the areas of mass communication and rhetoric have longstanding interests in the content analysis of texts. They have traditionally taken a more qualitative approach, but quantitative content analysis has proponents in both mass communication (Riffe, Lacy, & Fico, 1998) and rhetoric (Hart, 1985, 2001). There is a steadily increasing volume of rhetoric and mass communication research seeking to analyze large volumes of text (especially on the

Internet; see McMillan, 2000), so quantitative techniques that can be applied by computer will look increasingly attractive.

CRA offers a good option here because it can produce meaningful abstractions of news stories or rhetorical acts, representing their main concepts and their interrelationships. These can be compared to one another, and analyzed for change over time. For example, there is growing concern that media are covering cancer research in a way that sensationalizes new treatments, raises false expectations on the part of patients, and interferes with the orderly progress of clinical trials. One activist describes this coverage as the "pornography of cancer" (Groopman, 2000). CRA could be used to track cancer coverage in print media, and in broadcast media where transcripts are available. Simply having a broadband (national) picture of how the media write/talk about cancer would be interesting enough. One could detect the extent to which new research findings were portrayed in the media as, say, life-saving breakthroughs or magic bullets, and gain insight into how the portrayals propagate through the media. More sophisticated analyses are possible, too. CRA could be used to identify clusters in stories or sources of stories. Using a dynamic approach, it would be possible to assess the impact of new announcements on the conceptual structure of the media stories. It may also be feasible to infer the networks that connect different sources and media outlets by looking at commonalities in their CRA networks.

Moving to the other end of the scale, we see a number of potential applications for CRA in the analysis of conversation. We have already demonstrated in the "After Mr. Sam" example above that it can be profitably applied to aggregates of utterances in group interaction. We further envision a dynamic form of CRA that could be applied to interaction in groups or dyads. Because the unit of observation in CRA is the utterance, each utterance (or series of utterances comprising a turn) can be thought of as a fragment of a complete CRA network (representing, for example, a portion of conversation). Accumulating these fragments over a series of utterances (to account for conversational memory; Stafford, Burggraf, & Sharkey, 1987) yields a CRA network. The network changes when new utterances occur and those nodes and links are added, whereas the oldest utterances have their nodes and links decay and fall out of the network. The dynamic network could be visualized and interpreted qualitatively, to understand how important words in the conversation and their interrelations change over time. Time series generated from the dynamic network would support a number of additional qualitative analysis approaches.[17]

The idea that utterances are fragments of CRA networks might also find application in the simulation of communication systems. As noted in the critique section above, existing simulations of communication systems (for example, Corman, in press; Contractor, Danner, Palazzolo, Serb,

& She, 2001) assume that communication is an unproblematic exchange of information among agents. This assumption has been roundly criticized (e.g., Axley, 1984), yet simulation researchers have little choice but to make it in the absence of a suitable modeling scheme for communicative exchanges.

We can imagine a model (Dooley, Corman, McPhee, & Kuhn, in press) in which simulation agents have CRA networks as properties and exchange fragments of these networks as they communicate. An agent would incorporate a received fragment (after suitable filtering) into its existing network. This could connect nodes that had not been connected before, leading to a more structured representation that we might associate with clarity or focus of understanding. However, a new fragment might also short-circuit the network in a way that decreases the influence of existing nodes, making reorganization more likely. Undoubtedly, many of the fragments, lacking any existing structure to "hook up" with, would have very little impact.

We feel obligated to add that there are less benign uses for CRA, too. It does not take much imagination to turn the dynamic analysis and modeling tools just described into Big Brother technologies. Even on a less dramatic scale, employers could use CRA analyses like the one in Figure 5 to identify and eliminate "peripheral" personnel. Narratives about technology tend to bifurcate into these extreme positions, dwelling on existential harm to individuals on the one hand, and utopian benefits to the collective on the other (Alder, 1998). In reality, most technologies have the potential for both, and technologies like CRA need not be oppressive if they are implemented ethically (Trethewey & Corman, 2001).

CONCLUSION

We began by arguing that trends in communication theory are creating observational demands that existing methods cannot satisfy. As one solution to this problem we proposed CRA, which represents texts or transcribed conversations as networks of centering words. These are components of utterances (noun phrases) that authors or speakers deploy in such a way as to make their utterances coherent. A CRA network can be derived for any text and abstractly represents its main concepts, their influence, and their interrelationships. These networks form the basis for a number of visualization and spatial-modeling techniques that eventually support a wide array of applications. Three of these were illustrated in this paper as a way of establishing the face and representational validity of CRA: analysis of factionalism in a group decision-making episode, identification of an emergent "group reading" of a text and its substantial correlation to a CRA network, and the clustering of members of an interdisciplinary research network. The examples suggest that CRA has the ability to resolve some key shortcomings of existing methods in organizational communication, and to be useful in other areas of human communication research as well.

NOTES

1. This is a marked shift from the 80s and early 90s, when the lion's share of organizational communication research was focused on interpersonal relations in organizational contexts (Allen, Gotcher, & Seibert, 1993).

2. Some techniques, especially those used in marketing research, attempt this. However, they are applicable only in highly controlled, laboratory conditions.

3. Specifically with regard to reticulation theory, the questionnaire itself would be an instance of observable communication in the context of an enacted activity. The response it provokes would be, by assumptions of the theory, reflective of phenomena in the structural domain.

4. The Galileo approach produced a number of conceptual insights and made several interesting empirical findings, but its goals and approach are quite different from those of CRA . First, the identification of a concept set is a relatively unimportant preliminary task for Galileo research, but is a central task—indeed, a major purpose of—CRA work. Second, the Galileo research mainly used questionnaires to assess distances that were avowedly latent and (collectively) psychological. In contrast, CRA is based on relationships that summarize manifest patterns in text or conversation, so CRA's words are more tangible and avowedly discursive in essence. Finally, the Galileo researchers were committed to finding a general space in which concepts relate and move (Craig, 1977), whereas CRA depends on the assumption that concepts are connected primarily by a network of specific relations, using spatial representations mainly as simplified summaries or visualizations to aid in comprehension.

5. Although our present concern is limited to centering in English-language texts, the theory has been extended to other languages that utilize similar noun verb patterns, such as Italian (Di Eugenio, 1998), Turkish, (Turan, 1998), and Japanese (Iida, 1998).

6. As Frawley (1992) notes, "Nouns are not always persons, places, or things, but persons, places, or things *always turn out to be nouns*" (p. 63).

7. That said, verbs do give important information about the contexts of action in which the noun phrases were deployed. It may be possible to generate "layered" CRA networks where words are linked under the influence of verbs. Clearly, such a scheme could not allow different layers for every known verb. Some theory that could reduce verbs to a small number of categories would have to be identified.

8. Researchers interested in this type of information may substitute a proper noun for each instance of a pronoun as the sort of "semantic normalization" often employed in natural language processing (Lewis & Jones, 1996). This technique can also be useful in less structured discourse formats such as casual conversations or interviews, in which speakers may assume shared knowledge or pronominalize elements of an interviewer's questions in their responses.

9. We note that, in some cases, this practice could lead to stray links, not intended by the author. For example, someone referring to a "white paper abstract" probably does not intend to evoke the pair: white-abstract. We are less concerned about this in the case of CRA than in traditional network text analysis because: (a) CRA relies on a theoretical rationale for unitizing noun phrases, so it is more reasonable to assume that all their elements are meaningfully related; (b) in CRA, no higher-order linking is possible across noun phrases,

so those present zero possibilities for stray links; and (c) our informal observations suggest that cases like "white paper abstract" are relatively rare, and that any stray links they might generate are unlikely to be structurally consequential in the CRA network.

10. A reviewer raised the question of whether, given the flow metaphor we invoked, a measure like flow betweenness (Freeman, Borgatti, & White, 1991) might be more appropriate. Whereas betweenness centrality measures the extent to which a given node lies on the shortest paths through the network, flow betweenness measures the extent to which a node lies on all paths through the network, and, in this sense, is similar to the information centrality described in Wasserman & Faust (1994, p. 192). The rationale for this measure is credited to Stephenson & Zelen (1989) who argued that, in certain networks, such as human communication networks, it might be possible for information to take a more indirect path than the shortest one. We know of no basis for making this assumption in CRA networks. Indeed, we think it most likely that, if anything, communicators strive for efficiency and on this criterion the more parsimonious betweenness measure wins. It is also worth noting that before recent algorithmic developments (Brandes, 2001), computation of even the simpler betweenness measure was not practical for networks of more than a few hundred nodes, and that the flow measure is even more computationally intensive. We view this issue, and the one of the value of other centrality measures like degree and closeness, to be a matter for further research.

11. The videotape was the focus of a special preconference of the International Communication Association, "Interaction and Organizing: A Dialogue Between Organizational Communication and Language and Social Interaction Scholars," May, 2001, Washington, DC, organized by Francois Cooren, SUNY-Albany. The analysis presented here is based on results presented at that preconference (McPhee, Iverson, & Corman, 2001). Analysis was conducted on a transcript prepared by Professor Cooren and associates, and is available from him. Page and act numbers are based on that transcript.

12. Table 2 shows that these chains of concepts correspond to specific chains of nouns in specific sentences. For instance, the chain "topteam," "recommendation," "haphazard," "put," and "resource" is found in the sentence segment "same way as Topteam did other recommendations, rather than leave it again for a haphazard putting together without the resources to put together or call another meeting for that purpose." Single sentences such as this are clearly represented in the graph because the discourse sample is so small; for that reason also, the links in the graph are very low in influence.

13. We note that the value of CRA does not hinge on a positive answer to this question. It could be that the representations are useful for other reasons.

14. The uneven distribution of participants was an effect of sampling students in university courses. Although an equivalent number of transcripts were presented to students (approximately 60 of each), varying levels of instructor support and student willingness to contribute resulted in these unequal sample sizes.

15. The sample text was drawn from an online article about dietary guidelines. The text was reproduced in its entirety, and participants were instructed to look for nouns and adjectives that were most critical in giving the text meaning or important in the structure of the text; they were then told to examine how these terms were connected within sentences. This was followed by a five-by-five matrix of influential terms as identified by CRA, with links between the terms marked within relevant cells. Important to acknowledge here is that the purpose of the test was to validate CRA through a comparison of readers' and CRA's positioning of associations between terms, rather than to test readers' ability to assign influence to words.

16. We have not yet conducted benchmark tests to evaluate the performance of CRA relative to these other techniques. However, there are theoretical grounds for believing that resonance scores are less noisy measures of association. CRA focuses on words that are used to make text coherent, so they are representative of intentional acts by the communicator. Then it employs an explicit measure of the meaningfulness of elements in texts (influence, or betweenness centrality). It therefore distinguishes between cases when a word is

merely used, and cases where it is used to tie together other important concepts in the text. In other words, we claim that resonance scores provide a more accurate measure of the conceptual association between texts than existing methods.

17. Describing these is beyond the scope of this paper. In general, dynamic CRA would yield series of influence values for words, which could be analyzed with time- or frequency-domain techniques to assess the impact of conversational events, cycles of influence, and so forth. Also, any number of whole-network measures, like density or graph centralization, could be applied to the individual frames and related to conversational events (Poole et al., 2000).

REFERENCES

Adams, K. A. (2001). Word wranglers: Automatic classification tools transform enterprise documents from "bags of words" into knowledge resources. *Intelligent KM, 1*. Retrieved February 21, 2002, from http://www.intelligentkm.com/feature/010101/feat1.shtml

Adams, R. G., & Allen, G. (Eds.). (1998). *Putting friendship in context.* Cambridge, UK: Cambridge University Press.

Alder, G. S. (1998). Ethical issues in electronic performance monitoring: A consideration of deontological and teleological perspectives, *Journal of Business Ethics, 17*, 729–743.

Allen, M. W., Gotcher, J. M., & Seibert, J. H. (1993). A decade of organizational communication research: Journal articles 1980–1991. In S. A. Deetz (Ed.), *Communication yearbook 16* (pp. 252–330). Newbury Park, CA: Sage.

Alterman, R., & Bookman, L. (1990). Some computational experiments in summarization. *Discourse Processes, 13,* 143–174.

Anderson, P. (1999). Complexity theory and organization science. *Organization Science, 10,* 216–232.

Anthonisse, J. M. (1971, October). *The rush in a directed graph.* Amsterdam: Stichting Mathematisch Centrum.

Auld, F., & White, A. M. (1956). Rules for dividing interviews into sentences. *Journal of Psychology, 42*, 273–281.

Axelrod, R. (Ed.). (1976). *Structure of decision: The cognitive maps of political elites.* Princeton, NJ: Princeton University Press.

Axley, S. R. (1984). Managerial and organizational communication in terms of the conduit metaphor. *Academy of Management Review, 9,* 428–437.

Baeza-Yates, R., & Ribiero-Neto, B. (1999). *Modern information retrieval.* New York: Addison-Wesley.

Bak, P., & Chen, K. (1991, January). Self-organized criticality. *Scientific American, 264,* 46–53.

Bakeman, R., & Gottman, J. M. (1986). *Observing social interaction: An introduction to sequential analysis.* Cambridge, UK: Cambridge University Press.

Baker, F. B., & Hubert, L. J. (1981). The analysis of social interaction data. *Sociological Methods and Research, 9,* 339–361.

Barley, S. R. (1986). Technology as an occasion for structuring: Evidence from observations of CT scanners and the social order of radiology departments. *Administrative Science Quarterly, 31*, 78–108.

Barnett, G. A., & Danowski, J. A. (1992). The structure of communication: A network analysis of the International Communication Association. *Human Communication Research, 19,* 264–285.

Bartell, B. T., Cottrell, G. W., & Belew, R. K. (1992, June). *Latent semantic indexing is an optimal case of multidimensional scaling.* Paper presented at the 15th annual SIGIR, Denmark.

Boden, D. (1994). *The business of talk: Organizations in action.* Cambridge, UK: Polity Press.

Botan, C. (1996). Communication work and electronic surveillance: A model for predicting panoptic effects. *Communication Monographs, 63,* 293–313.

Brandes, U. (2001). A faster algorithm for betweenness centrality. *Journal of Mathematical Sociology, 25(2),* 163–177.

Browning, L. D., & Beyer, J. M. (1998). The structuring of shared voluntary standards in the U.S. semiconductor industry: Communicating to reach agreement. *Communication Monographs, 65,* 220–243.
Brunt, G. G. (2000). Understanding self-organized criticality as a statistical process. *Complexity, 5,* 26–33.
Burt, R. S. (1992). *Structural holes: The social structure of competition.* Cambridge, MA: Harvard University Press.
Carley, K. M. (1991). A theory of group stability. *American Sociological Review, 56,* 331–354.
Carley, K. M. (1997a). Extracting team mental models through textual analysis. *Journal of Organizational Behavior, 18,* 533–558.
Carley, K. M. (1997b). Network text analysis: The network position of concepts. In C. W. Roberts (Ed.), *Text analysis for the social sciences: Methods for drawing statistical inferences from texts and transcripts* (pp. 79–100). Mahwah, NJ: Erlbaum.
Carley, K. M., & Kaufer, D. M. (1993). Semantic connectivity: An approach for analyzing symbols in semantic networks. *Communication Theory, 3,* 183–213.
Carley, K. M., & Prietula, M. J. (1994). Computational organization theory: An introduction. In K. M. Carley & M. J. Prietula (Eds.), *Computational organizational theory* (pp. xi–xvii). Hillsdale, NJ: Erlbaum.
Cloitre, M., & Bever, T. G. (1988). Linguistic anaphors, levels of representation, and discourse. *Language and Cognitive Processes, 3,* 293–322.
Connolly, D., Burger, J. D., & Day, D. S. (1997). A machine learning approach to anaphoric reference. In D. B. Jones & H. L. Somers (Eds.), *New methods in language processing* (pp. 133–144). London: UCL Press.
Contractor, N. S., Danner, N., Palazzolo, E., Serb, D. A., & She, Y. (2001, May). *Co-evolution of knowledge networks as public goods and transactive memory systems for theoretical integration and extension.* Paper presented at the annual conference of the International Communication Association, Washington, DC.
Contractor, N. S., & Grant, S. (1996). The emergence of shared interpretations in organizations. In J. Watt & A. VanLear (Eds.), *Dynamic patterns in communication processes* (pp. 215–230). Thousand Oaks, CA: Sage.
Corman, S. R. (1996). Cellular automata as models of unintended consequences of organizational communication. In J. Watt & A. VanLear (Eds.), *Dynamic patterns in communication processes* (pp. 191–212). Thousand Oaks, CA: Sage.
Corman, S. R. (in press). Network dynamics in self-organizing communication and activity systems. In G. A. Barnett & R. Houston (Eds.), *Progress in communication sciences: Self-organizing systems* (Vol. 17). Westport, CT: Ablex.
Corman, S. R., & Bradford, L. B. (1993). Situational effects on the accuracy of self-reported organizational communication behavior. *Communication Research, 20,* 822–840.
Corman, S. R., & Scott, C. R. (1994). Perceived communication relationships, activity foci, and observable communication in collectives. *Communication Theory, 4,* 171–190.
Craig, R. T. (1977). Limiting the scope of the spatial model of message effects. *Human Communication Research, 3,* 309–325.
Daft, R. L. (1995). *Organization theory and design.* (5th ed.). St. Paul, MN: West.
Danowski, J. (1982). A network-based content analysis methodology for computer-mediated communication: An illustration with a computer bulletin board. *Communication yearbook 6,* 904–925.
Danowski, J. A. (1988). Organizational infographics and automated auditing: Using computers to unobtrusively gather as well as analyze communication. In G. Goldhaber & G. Barnett (Eds.), *Handbook of organizational communication* (pp. 385–433). Norwood, NJ: Ablex.
Danowski, J. A. (1992, November). *WORDij: A word-pair approach to information retrieval.* Paper presented at the Text Retrieval Conference (TREC-1), Gaithersburg, MD.
Danowski, J. A. (1993). Network analysis of message content. In W. D. Richards & G. A. Barnett (Eds.), *Progress in communication sciences* (Vol. 12, pp. 197–221). Norwood, NJ: Ablex.

Davenport, T. H., & Prusak, L. (1998). *Working knowledge: How organizations manage what they know.* Boston: Harvard Business School Press.

Desanctis, G., & Monge, P. (1999). Introduction to the special issue: Communication processes for virtual organizations. *Organization Science, 10,* 693–703.

Diefenbach, D. L. (2001). Historical foundations of computer-assisted content analysis. In M. D. West (Ed.), *Progress in communication sciences: Theory, method, and practice in computer content analysis* (Vol. 16, pp. 13–41). Westport, CT: Ablex.

Di Eugenio, B. (1998). Centering in Italian. In M. Walker, A. Joshi, & E. Prince (Eds.), *Centering theory in discourse* (pp. 115–138). Oxford, UK: Clarendon Press.

Doerfel, M. L., & Barnett, G. A. (1999). A semantic network analysis of the International Communication Association. *Human Communication Research, 25,* 589–603.

Dong, A., & Agogino, A. M. (1997). Text analysis for constructing design representations. *Artificial Intelligence in Engineering, 11,* 65–75.

Dooley, K. J. (1997). A complex adaptive systems model of organization change. *Nonlinear Dynamics, Psychology, & Life Sciences, 1,* 69–97.

Dooley, K. J. (in press). Simulation research methods. In J. Baum (Ed.), *Companion to organizations.* Oxford, UK: Blackwell.

Dooley, K., Corman, S., McPhee, R., & Kuhn, T. (in press). Modeling high-resolution broadband discourse in complex adaptive systems. *Nonlinear Dynamics, Psychology, & Life Sciences.*

Drazin, R., & Rao, H. (1996). The adoption of competing technologies as an organizing process. *Technology Studies, 3,* 1–29.

Eisenberg, E. M. (1990). Jamming: Transcendence through organizing. *Communication Research, 17,* 139–164.

Eizirik, L. M. R., Barbosa, V. C., & Mendes, S. B. T. (1993). A Bayesian-network approach to lexical disambiguation. *Cognitive Science, 17,* 257–283.

Ellis, D. G. (1999). *Crafting society: Ethnicity, class, and communicaiton theory.* Mahwah, NJ: Erlbaum.

Fairhurst, G. T., & Chandler, T. A. (1989). Social structure in leader-member interaction. *Communication Monographs, 56,* 215–235.

Federici, S., & Pirrelli, V. (1997). Analogy, computation, and linguistic theory. In D. B. Jones & H. L. Somers (Eds.), *New methods in language processing* (pp. 16–34). London: UCL Press.

Folger, J. P., Hewes, D. E., & Poole, M. S. (1984). Coding social interaction. In B. Dervin & M. Voight (Eds.), *Progress in communication sciences* (Vol. 4, pp. 115–161). Norwood, NJ: Ablex.

Foltz, P. W., Kintsch, W., & Landauer, T. K. (1998). The measurement of textual coherence with latent semantic analysis. *Discourse Processes, 25,* 285–307.

Ford, J. D., & Ford, L. W. (1995). The role of conversations in producing intentional change in organizations. *Academy of Management Review, 20,* 541–570.

Foucault, M. (1979). *Discipline and punish.* New York: Vintage Books.

Fraenkel, A. S., & Klein, S. T. (1999). Information retrieval from annotated texts. *Journal of the American Society for Information Science, 50,* 845–854.

Frakes, W. B., & Bazea-Yates, R. (1992). *Information retrieval: Data structures and algorithms.* Englewood Cliffs, NJ: Prentice Hall.

Frawley, W. (1992). *Linguistic semantics.* Hillsdale, NJ: Erlbaum.

Freeman, L. C. (1979). Centrality in social networks: Conceptual clarification. *Social Networks, 1,* 215–239.

Freeman, L. C. (1992). Filling in the blanks: A theory of cognitive categories and the structure of social affiliation. *Social Psychology Quarterly, 55,* 118–127.

Freeman, L. C. (2000). Visualizing social networks. *Journal of Social Structure, 1.* Retrieved March 18, 2002, from http://www.heinz.cmu.edu/project/INSNA/joss/vsn.html

Freeman, L. C., Borgatti, S. P., & White, D. R. (1991). Centrality in valued graphs: A measure of betweenness based on network flow. *Social Networks, 13,* 141–154.

Friedkin, N. A. (1983). Horizons of observability and limits of informal control in organizations. *Social Forces, 62,* 54–77.

Frost, P. J., & Stablein, R. E. (Eds.). (1992). *Doing exemplary research.* Newbury Park, CA: Sage.

Galal, G. M., Cook, D. J., & Holder, L. B. (1999). Exploiting parallelism in a structural scientific discovery system to improve scalability. *Journal of the American Society for Information Science, 50,* 65–73.

Gerken, L., & Bever, T. (1986). Linguistic intuitions are the result of interactions between perceptual processes and linguistic universals. *Cognitive Science, 10,* 457–476.

Gersick, C. G. J. (1988). Time and transition in work teams: Toward a new model of group development. *Academy of Management Journal, 31,* 9–41.

Giddens, A. (1979). *Central problems in social theory.* Berkeley: University of California Press.

Giddens, A. (1984). *The constitution of society: Outline of the theory of structuration.* Berkeley: University of California Press.

Giddens, A. (1990). *The consequences of modernity.* Stanford, CA: Stanford University Press.

Givón, T. (1984). *Syntax: A functional-typological introduction, 1.* Amsterdam: Benjamins.

Gordon, P. C., Grosz, B. J., & Gilliom, L. A. (1993). Pronouns, names, and the centering of attention in discourse. *Cognitive Science, 17,* 311–347.

Gordon, P. C., & Hendrick, R. (1998). The representation and processing of coherence in discourse. *Cognitive Science, 22,* 389–424.

Gronn, P. C. (1983). Talk as the work: The accomplishment of school administration. *Administrative Science Quarterly, 28,* 1–21.

Groopman (2000, June 4). The thirty years war: Have we been fighting cancer the wrong way? *New Yorker,* pp. 52–63.

Grosz, B. J., Weinstein, S., & Joshi, A. K. (1995). Centering: A framework for modeling the local coherence of a discourse. *Computational Linguistics, 21,* 203–225.

Hahn, U., Romacker, M., & Schulz, S. (1999). Discourse structures in medical reports—watch out! The generation of referentially coherent and valid text knowledge bases in the MEDSYNDIKATE system. *International Journal of Medical Informatics, 53,* 1–28.

Hammond, A., & Pearson, G. (Producers), (1974). *After Mr. Sam* [Motion picture]. (Available from the National Film Board of Canada, 350, Fifth Avenue, Suite 4820, New York, NY 10118)

Hart, R. P. (1985) Systematic analysis of political discourse: The development of DICTION. In K. Sanders (Ed.), *Political communication yearbook* (pp. 97–134). Carbondale: Southern Illinois University Press.

Hart, R. P. (2001). Redeveloping diction: Theoretical considerations. In M. West (Ed.), *Progress in communication sciences: Theory, methods, and practice in computer content analysis* (Vol. 16, pp. 43–60). New York: Ablex.

Hopper, P., & Thompson, S. (1984). The discourse basis for lexical categories in universal grammar. *Language, 60,* 703–752.

Hopper, P., & Thompson, S. (1985). The iconicity of the universal categories of "noun" and "verb." In J. Haiman (Ed.), *Iconicity in syntax* (pp. 151–183). Amsterdam: Benjamins.

Howard, G. S. (1994). Why do people say nasty things about self-reports? *Journal of Organizational Behavior, 15,* 399–404.

Hubert, L., & Schultz, J. (1976). Quadratic assignment as a general data analysis strategy. *British Journal of Statistical Psychology, 29,* 190–241.

Humphrey, S. M. (1999). Automatic indexing of documents from journal descriptors: A preliminary investigation. *Journal of the American Society for Information Science, 50,* 661–674.

Hyatt, A., Contractor, N., & Jones, P. (1997). Computational organizational network modeling: Strategies and an example. *Computational and Mathematical Organization Theory, 2,* 285–300.

Iida, M. (1998). Discourse coherence and shifting centers in Japanese texts. In M. Walker, A. Joshi, & E. Prince (Eds.), *Centering theory in discourse* (pp. 161–182). Oxford, UK: Clarendon Press.

Jackson, S., Jacobs, S., & Rossi, A. M. (1987). Conversational relevance: Three experiments on pragmatic connectedness in conversation. In Margaret L. McLaughlin (Ed.), *Communication yearbook 10* (pp. 323–347). Newbury Park, CA: Sage.

Jacobs, P. S., & Rau, L. F. (1993). Innovations in text interpretation. *Artificial Intelligence, 63,* 143–191.

Jurafsky, D., & Martin, J. H. (2000). *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition.* Upper Saddle River, NJ: Prentice Hall.

Kellerman, K., & Sleight, C. (1989). Coherence: A meaningful adhesive for discourse. In J. A. Anderson (Ed.), *Communication yearbook 12* (pp. 95–129). Newbury Park, CA: Sage.

Kintsch, W. (2001). Predication. *Cognitive Science, 25,* 173–202.

Kohonen, T. (1997). *Self-organizing maps* (2nd ed.). New York: Springer Verlag.

Krackhardt, D. (1987). QAP partialing as a test of spuriousness. *Social Networks, 9,* 171–186.

Krackhardt, D. (1988). Predicting with networks: Nonparametric multiple regression analysis of dyadic data. *Social Networks, 10,* 359–381.

Krackhardt, D., & Kilduff, M. (1990). Friendship patterns and culture: The control of organizational diversity. *American Anthropologist, 92,* 142–154.

Kruskal, J. B., & Wish, M. (1978). *Multidimensional scaling.* Beverly Hills, CA: Sage.

Kuhn, T. R. (2000). *The complex process of planned organizational change: Developing a model of knowledge, activity, and communication networks.* Unpublished doctoral dissertation, Arizona State University, Tempe.

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes, 25,* 259–284.

Langacker, R. (1987). Nouns and verbs. *Language, 63,* 53–94.

Latane, B. (1996). Dynamic social impact: The creation of culture by communication. *Journal of Communication, 46*(4), 13–23.

Leacock, C., Chodorow, M., & Miller, G. A. (1998). Using corpus statistics and WordNet relations for sense identification. *Computational Linguistics, 24,* 147–165.

Lecoeuche, R., Robertson, D., Barry, C., & Mellish, C. (2000). Evaluating focus theories for dialogue management. *International Journal of Human-Computer Studies, 52,* 23–76.

Lewis, D. D., & Jones, K. S. (1996). Natural language processing for information retrieval. *Communications of the ACM, 39,* 92–101.

Lund, K., & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instruments, & Computers, 28,* 203–208.

McKoon, G., & Ratcliff, R. (1998). Memory based language processing: Psycholinguistic research in the 1990s. *Annual Review of Psychology, 49,* 25–42.

McMillan, S. J. (2000). The microscope and the moving target: The challenge of applying content analysis to the World Wide Web. *Journalism & Mass Communication Quarterly, 77,* 80–98.

McPhee, R. (1988). Vertical communication chains: Toward an integrated approach. *Management Communication Quarterly, 1,* 455–493.

McPhee, R. D. (1995). Cognitive perspectives on communication: Interpretive and critical responses. In D. E. Hewes (Ed.), *The cognitive bases of interpersonal communication* (pp. 225–246). Hillsdale, NJ: Erlbaum.

McPhee, R. D. (2000, June). *Do surveillance developments lead to increasing proprietary control of ideas and communication in organizations?* Paper presented at the annual conference of the International Communication Association, Acapulco, Mexico.

McPhee, R. D., & Corman, S. R. (1995). An activity-based theory of communication networks in organizations, applied to the case of the local church. *Communication Monographs, 62,* 132–151.

McPhee, R. D., & Corman, S. R. (2001, May). *"We ought to have . . . gumption . . . :" An analysis of an excerpt from the videotape "After Mr. Sam."* Paper presented at a preconference of the annual conference of the International Communication Association, Washington, DC.

McPhee, R. D., Corman, S. R., & Dooley, K. J. (1999, May). *Characteristic processes and discursive methods in the study of organizational knowledge.* Paper presented at the annual conference of the International Communication Association, San Francisco, CA.

Miller, G. A. (1995). WordNet: A lexical database for English. *Communications of the ACM, 38*(11), 39–41.

Miller, L. L., Cody, M. J., & McLaughlin, M. L. (1994). Situations and goals as fundamental constructs in interpersonal communication research. In M. L. Knapp & G. R. Miller (Eds.),

Handbook of interpersonal communication. (2nd ed., pp 162–198). Thousand Oaks, CA: Sage.
Mitkov, R. (1997). A new approach to center tracking. In D. B. Jones & H. L. Somers (Eds.), *New methods in language processing* (pp. 282–287). London: UCL Press.
Nomoto, T., & Nitta, Y. (1997). Structuring raw discourse. In D. B. Jones & H. L. Somers (Eds.), *New methods in language processing* (pp. 288–298). London: UCL Press.
Passonneau, R. J. (1998). Interaction of discourse structure with explicitness of discourse anaphoric noun phrases. In M. A. Walker, A. K. Joshi, & E. F. Prince (Eds.), *Centering theory in discourse* (pp. 327–358). Oxford, UK: Clarendon.
Perez-Carballo, J., & Strzalkowski, T. (2000). Natural language information retrieval: Progress report. *Information Processing & Management, 36,* 155–178.
Perrow, C. (1967). A framework for the comparative analysis of organizations. *American Sociological Review, 26,* 854–866.
Poole, M. S. (1983). Decision development in small groups III: A multiple-sequence theory of decision development. *Communication Monographs, 50,* 321–341.
Poole, M. S., & Folger, J. P. (1981). A method for establishing the representational validity of interaction coding schemes: Do we see what they see? *Human Communication Research, 8,* 26–42.
Poole, M. S., & Hirokawa, R. Y. (1996). Introduction: Communication and group decision making. In R. Y. Hirokawa & M. S. Poole (Eds.), *Communication and group decision making* (2nd ed., pp. 3–18). Thousand Oaks, CA: Sage.
Poole, M. S., & Roth, J. (1989). Decision development in small groups V: Test of a contingency model. *Human Communication Research, 15,* 549–589.
Poole, M. S., Van de Ven, A. H., Dooley, K., & Holmes, M. E. (2000). *Organizational change and innovation processes: Theory and methods for research.* New York: Oxford University Press.
Popping, R. (2000). *Computer-assisted text analysis.* Thousand Oaks, CA: Sage.
Putnam, L. L., & Fairhurst, G. T. (2001). Discourse analysis in organizations: Issues and concerns. In F. M. Jablin & L. L. Putnam (Eds.), *New handbook of organizational communication* (pp. 78–137). Thousand Oaks, CA: Sage.
Rice, R. E., & Danowski, J. A. (1993). Is it really just a fancy answering machine? Comparing semantic networks of different types of voice mail users. *Journal of Business Communication, 30,* 369–397.
Richards, W. A. (1985). Data, models, and assumptions in network analysis. In R. D. McPhee & P. K. Tompkins (Eds.), *Organizational communication: Traditional themes and new directions.* Beverly Hills, CA: Sage.
Riffe, D., Lacy, S., & Fico, F. G. (1998). *Analyzing media messages: Using quantitative content analysis in research.* Mahwah, N.J.: Erlbaum.
Roberts, C. W. (Ed.). (1997). *Text analysis for the social sciences: Methods for drawing statistical inferences from texts and transcripts.* Mahwah, NJ: Erlbaum.
Salton, G., & Buckley, C. (1988). Term weighting approaches in automatic text retrieval. *Information Processing and Management, 24,* 513–523.
Sandelands, L. E., & Stablein, R. E. (1987). The concept of organization mind. In S. Bacharach & N. DiTomaso (Eds.), *Research in the sociology of organizations* (Vol. 5, pp. 135–162). Greenwich, CT: JAI Press.
Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals, and understandings: An inquiry into human knowledge structures.* Hillsdale, NJ: Erlbaum.
Schiffrin, D. (1994). *Approaches to discourse.* Oxford, UK: Blackwell.
Scott, C. R., Corman, S. R., & Cheney, G. (1998). Development of a situated-action theory of identification in the organization. *Communication Theory, 8,* 298–336.
Seyfarth, B. (2000). Structuration theory in small group communication: A review and agenda for future research. In M. Roloff (Ed.), *Communication yearbook 23* (pp. 341–379). Thousand Oaks, CA: Sage.
Shapiro, G., & Markoff, J. (1997). A matter of definition. In C. Roberts (Ed.), *Text analysis for the social sciences: Methods for drawing statistical inferences from texts and transcripts* (pp. 9–31). Mahwah, NJ: Erlbaum.
Siemens, R. (1993, October). *Practical content analysis techniques for text retrieval in large, untagged*

text-bases. Paper presented at the annual ACM conference on systems documentation, Kitchener, Canada.
Spence, D. P., & Owens, K. C. (1990). Lexical co-occurrence and association strength. *Journal of Psycholinguistic Research, 19,* 317–330.
Sperber, D., & Wilson, D. (1995). *Relevance: Communication and cognition.* Cambridge, MA: Blackwell.
Stafford, L., Burggraf, C. S., & Sharkey, W. F. (1987). Conversational memory: The effects of time, recall, mode and memory expectancies on remembrances of natural conversations. *Human Communication Research, 14,* 203–229.
Stephen, T. (1999). Computer-assisted concept analysis of HCR's first 25 years. *Human Communication Research, 25,* 498–513.
Stephen, T. (2000). Concept analysis of gender, feminist, and women's studies research in the communication literature. *Communication Monographs, 67,* 193–214.
Stephenson, K., & Zelen, M. (1989). Rethinking centrality: Methods and applications. *Social Networks, 11,* 1–37.
Stohl, C. (1993). European managers' interpretations of participation: A semantic network analysis. *Human Communication Research, 20,* 97–117.
Taylor, J. R., Flanagin, A. J., Cheney, G., & Seibold, D. R. (2001). Organizational communication research: Key moments, central concerns, and future challenges. In W. Gudykunst (Ed.), *Communication yearbook 24* (pp. 99–137). Thousand Oaks, CA: Sage.
Taylor, J. R., & Van Every, E. J. (1999). *The emergent organization: Communication as its site and surface.* Mahwah, NJ: Erlbaum.
Terra Research & Computing. (1994). *The Galileo computer program.* [Computer Software] Birmingham, MI: Terra.
Thompson, J. D. (1967). *Organizations in action.* New York: McGraw-Hill.
Treadwell, D. T., & Harrison, T. (1994). Conceptualizing and assessing organizational image: Model images, commitment, and communication. *Communication Monographs, 61,* 63–85.
Trethewey, A., & Corman, S. R. (2001). Anticipating K-commerce: E-commerce, knowledge management, and organizational communication. *Management Communication Quarterly, 14*, 619–628.
Tsoukas, H. (1996). The firm as a distributed knowledge system: A constructionist approach. *Strategic Management Journal, 17,* 11–25.
Turan, U. D. (1998). Ranking forward-looking centers in Turkish: Universal and language-specific properties. In M. Walker, A. Joshi, & E. Prince (Eds.), *Centering theory in discourse* (pp. 139–160). Oxford, UK: Clarendon.
United Nations. (2001). *World population prospects: The 2000 revision highlights.* New York: Population Division, Department of Economic and Social Affairs, United Nations.
von Bertalanffy, L. (1968). *General system theory: Foundations, development, applications.* New York: George Brazilier.
Walker, M. A., Joshi, A. K., & Prince, E. F. (Eds.). (1998). *Centering theory in discourse.* New York: Oxford.
Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications.* New York: Cambridge University Press.
Weick, K. E. (1979). *The social psychology of organizing* (2nd ed.). New York: McGraw-Hill.
Weick, K. E., & Roberts, K. H. (1993). Collective mind in organization: Heedful interrelating on flight decks. *Administrative Science Quarterly, 38,* 357–381.
Woelfel, J. (1993). Cognitive processes and communication networks: A general theory. In W. D. Richards & G. A. Barnett (Eds.), *Progress in communication sciences: Vol. 12* (pp. 21–42). Norwood, NJ: Ablex.
Woelfel, J., & Fink, E. (1980). *The measurement of communication processes: Galileo theory and method.* New York: Academic Press.