

Content Analysis

**Kathleen Carley
Associate Professor of Sociology
Department of Social and Decision Sciences
Carnegie Mellon University**

30 July 1990

Edinburgh, UK: Pergamon Press.
)

in R. E. Asher et al (eds) The Encyclopedia of Language and Linguistics.
Edinburgh, UK: Pergamon Press, Vol. 2 pp. 725-730.

.

.

.

.

.

.

.

Content Analysis

Introduction

In the early days of the General Inquirer (Stone, 1968a; Stone, 1968b) content analysis centered around word counts; i.e., did a certain set of words occur in the text and what was the frequency with which they occurred. This tradition is continuing (Sullivan, 1973; Markoff, 1974; Robinson Sheehan, 1983; Namenwirth and Weber, 1987; Weber, 1984a; Neuman, 1989, see for examples). In addition, a plethora of new tools and approaches are emerging that facilitate analyzing content in ways well beyond word counts. The spread has become so great that, currently, any procedure for analysis of a text, regardless of the origin of the text (book chapters, interviews, essays, discussions, and so on), that goes beyond syntactic analysis to semantics, is only minimally concerned with the conversational protocols, and admits empirical analysis can claim to fall under the general heading of content analysis. Despite this broadening there are a common set of methodological concerns and only a few basic strategies. This paper describes these basic strategies.

Content analysis as it was envisioned in the 1960's focuses on the frequency with which words or concepts occur in texts or across texts and the development of scales of meaning against which the content of texts can be compared (North, 1963; Gerbner, 1969; Holsti, 1969; Gottschalk and Gleser, 1969; Gottschalk Winget, 1969; Janowitz, 1969; Stone, 1968a; Stone, 1968b; Berelson, 1952). Originally, such analysis required extensive data preparation, access to large mainframe computers and access to special-purpose programs. This work met with limited success largely due to factors such as limited access to software and the necessary processing power, time-consuming data preparation, difficulty of relating such data to other data and lack of a strong theoretical underpinning (Namenwirth and Weber, 1987; Roberts, 1987; Neuman, 1989). Today, doing content analysis is technically simpler (Rosengren, 1981; Weber, 1984b; Weber, 1985; Ogilvie, Stone and Kelly, 1982). Further, the increasing interest in the cognitive foundations of social behavior is leading to a revival of interest in the analysis of texts. Such work is spread across a variety of social science and humanistic fields, among them organizations (Eden, Jones and Sims, 1979; Namenwirth, 1981), sociology (Carley, 1986a), rhetoric and composition (Gere, 1987; Gere and Stevens, 1985; Langer and Applebee, 1986; Michaels, 1987; Palmquist, 1990), sociolinguistics (Edwards and Mercer, 1986; Stubbs, 1983a), political science (Axelrod, 1976; Neuman, 1989). One consequence of the current interest is that content analysis now goes beyond frequency counts to more qualitative procedures that capture aspects of the text other than word frequency or occurrence, yet admit empirical analyses. And, a variety of new tools for doing both traditional and non-traditional content analysis are appearing spurred on by recent advances in spreadsheets and artificial intelligence coupled with the existence of powerful personal computers.

These new tools have been applied to a wide variety of texts, including interview data gathered through free- and cued-recall (Chiesi, Spilich and Voss, 1979; Reitman and Rueter, 1980), free association (Moshe et al., 1986), open-ended interviewing (Carley, 1986a; Finch et al., 1987; Palmquist, 1990), as well as work-sheets (Gussarsky and Gorodetsky, 1988), pattern notes (Jonassen, 1987), textbooks (Palmquist, 1990), concept circles (Palmquist, 1990; Carley and Palmquist, 1990), stories (Rumelhart, 1978; Lehnert, 1987; Cicourel and Carley, 1990), biographies (vanMeter and Mounier, 1989), newspaper clippings (Cullingford, 1981; Wilensky, 1981) and press releases (Fan, 1988). These tools have been applied in a wide variety of disciplines, including political science (Janowitz, 1969; Fan, 1988; Neuman, 1989), sociology (Carley, 1986a; Roberts, 1987), artificial intelligence (Lehnert, 1977; Lehnert, 1987), communication (Danowski, 1980; Danowski, 1988; Alexander and Danowski, 1990), cultural studies (Namenwirth and Weber, 1987), cognitive science (Means and Voss, 1985; Jonassen, 1987; Reitman and Rueter, 1980; McKeithen et al., 1981) and rhetoric (Palmquist, 1990). Each type of text and discipline present their own special difficulties and needs. Yet, across all these differences, there is a common set of methodological concerns and basic strategies.

Current strategies can be roughly divided into three categories: (1) conceptual analysis, (2) procedural analysis and (3) relational analysis. Conceptual analysis centers on extracting what concepts are explicitly or implicitly present in the text. Procedural analysis centers on extracting the procedures that the author of the text uses to perform some task. Relational analysis centers on extracting the mental model implicit in the text. Each of these three strategies will be discussed in turn. In addition, a special section on extracting affectual content and the utilization of affect in analyzing texts is provided.

There are a wide variety of textual analysis techniques that, although related to the works discussed herein, will not be directly addressed here. These techniques include, but are not limited to: conversational analysis (Sacks, 1972), discourse analysis (Meyer, 1975; Stubbs, 1983b; Polanyi, 1985; Sozer, 1985), computational hermeneutics (Alker, 1985; Mallery, 1985; Mallery, 1986), data-base techniques for doing ethnographic and qualitative studies (Conrad and Reinharz, 1984; Sproull and Sproull, 1982; 1985, 1989; Seidel and Clark, 1984; Brent, 1984; Gerson, 1984; Dennis, 1984; Tesch, 1989; Drass, 1989) and natural language approaches (Dyer, 1983; Reddy, 1973; Fennel and Lesser, 1977).

1. Conceptual Analysis

Content analysis has traditionally been thought of as determining what words or concepts are present in a text or set of texts. In this paper, I refer to this approach as conceptual analysis. A concept is a single idea regardless of whether it is represented by a single word or a phrase. Two types of analyses fall under this rubric: (1) extraction of explicit concepts and (2) extraction of implicit concepts. Explicit concepts are words or phrases that actually occur in the text. Thus, the phrase *There is a U.S. peacekeeping force in Libya* explicitly contains the concept *peacekeeping force*. Implicit concepts are words or phrases that occur in the text only by implication. The issue here is what level of implication is allowed. For example, the concept *peacekeeping force* can be thought to imply the concept *troop* in the sense that it is an alternate phrase for the same basic idea. In contrast, a less tightly coupled implication of the foregoing sentence might be the concept *U.S. places military sanctions*.

Explicit concept analysis locates what words or phrases are explicitly in the text, or the frequency with which they occur. This type of analysis is easy to automate. Consequently, numerous programs have been developed -- ranging from basic UNIX utilities like *grep* to the General Inquirer (Stone, 1968a; Stone, 1968b) to assorted smaller or special purpose programs developed by individual researchers such as that by Garson (1985) or Carley (1990) -- that differ primarily in the length of text that they can analyze. Much of the meaning within the text, however, may be lost by relying only on explicit concepts (Ogilvie, Stone and Kelly, 1982; Woodrum, 1984).

In contrast, implicit concept analysis makes explicit implicit concepts and then locates the frequency of concept occurrence. This procedure facilitates text comparison, admits a richer definition of meaning, but is much more difficult to automate. In order to understand most texts, however, the researcher will need to explicate implicit information (Collins, 1977; Schank, 1977; Schank, 1981; Bruce, 1978; Merleau, 1964; Bar-Hillel, 1960). As noted by Merleau-Ponty (1964: 29) "The totality of meaning is never fully rendered: there is an immense mass of implications, even in the most explicit of languages". However, it is methodologically more difficult to extract implicit than explicit concepts. In order to extract implicit concepts impressionistic judgments must be made (Holsti, 1969; Krippendorf, 1980) which causes human errors to make more mistakes (typically errors of omission) when extracting implicit than explicit concepts (Carley, 1988). In addition, there are few software procedures for extracting implicit concepts and, of those that exist, each takes a very different tack thus inhibiting scientific cumulation in this area. The difficulty in extracting implicit concepts follows from the fact that it is a more difficult recognition task and from the fact that such extraction often requires subjective judgments on the part of the coder as to what is really meant by the text's author.

Two basic methodological approaches are currently used -- dictionaries and translation rules. In the dictionary approach, a dictionary or thesaurus is constructed that contains key concepts (or categories) and their associated synonyms and "equivalent" phrases or words (Fan, 1988; Carley, 1984). Example dictionaries are the Harvard IV (Kelly and Stone, 1975; Dunphy, 1974) and the Lasswell Value Dictionary (Namenwirth and Weber, 1987, ch. 2). Following the compilation of a dictionary the investigator must apply a mapping procedure (involving either a special purpose pre-processing program or human coders) that maps from the explicit concepts in the text to the concepts (or categories) in the dictionary. Difficulties in constructing such dictionaries are discussed in detail by (Namenwirth and Weber, 1987, ch. 8). And the reliability of using humans and computers when the dictionary approach is taken is analyzed by Grey, Kaplan and Lasswell (1965) and Saris-Gallhofer, Saris and Morton (1978).

When the dictionary approach is used, all phrases or alternate concepts must occur exactly as listed in the dictionary in order for the concept to be coded as present. Thus, if the same words occur but not in the proper order, the concept will not be coded as present. Nor can the dictionary approach distinguish between homographs. In contrast, when a rule-based approach is used words need not occur in a particular order and homographs can be distinguished. For example, Fan (1988) uses a rule-based approach in which the desired words need merely be physically proximal (such as within the same paragraph) in order for the concept to be coded as present. As another example, Carley (1988a) employs an expert system which uses the shared social knowledge of the subjects to make explicit those concepts that are implicit in the text. In this case, a concept is coded as present if it is judged by the coder to be present in the text. In addition, should any of the concepts coded as present by the coder evoke in other members of the society a different concept, this different concept is coded as present by implication in the text. Frequently, researchers employ both dictionaries and rule sets (Fan, 1988, for example). In fact, the most recent versions of the Harvard (Dunphy, 1974) and Lasswell (Namenwirth and Weber, 1987, ch. 2) dictionaries contain rules, such as those for distinguishing between common homographs.

Creating dictionaries and rule sets is a time consuming process. Thus, a general-purpose or universal set of dictionaries and rules that could be used to translate a wide variety of texts would seem to be the ideal from a pragmatic perspective. Both the Harvard and Lasswell dictionaries are viewed as such general-purpose dictionaries consisting of many, according to Namenwirth and Weber (1987, p. 211), "commonsense" categories of meaning chosen to "reflect the wide range of human experience and understanding encoded in language". It should be noted that a similar argument was forwarded by Roget about the original categories in his thesaurus.

There is, however, an alternate camp that argues against the use of, and even against the possibility of constructing, such a general-purpose dictionary. The basic arguments are: that particular research questions and micro-level analyses require finer grained and special purpose dictionaries; what is commonsense changes over time; and a true universal dictionary is impossible to construct. The basic idea is that extracting implicit concepts requires the researcher to have extensive knowledge of the socio-cultural environment that generated the texts being coded. Or, in other words, content analysis is similar to translation from one language to another (Andren G., 1981). In 1960, Bar-Hillel argued that computers could not do such translation as it depends on social knowledge:

"What such a suggestion amounts to, if taken seriously, is the requirement that a translation machine should not only be supplied with a dictionary but also a universal encyclopedia. This is surely utterly chimerical and hardly deserves any further discussion."

Given that universal dictionaries and rule sets are an impossibility, at least in the near future, the process of extracting implicit concepts can be facilitated by general purpose software that enables the researcher to easily create dictionaries and rule sets and process texts. Steps in this direction are provided by (Fan, 1988; Carley, 1988; Iwanska, 1989).

The main strength of the conceptual analysis strategy is that it can be totally automated and so applied to vast numbers of texts thus making possible cross-subjects and cross-time comparisons at an empirical level. Nevertheless, there is a fundamental problem with simply conducting a conceptual analysis. Simply stated, the presence of concepts may not be sufficient to denote meaning. People can use the same words or concepts with the same frequency and have very different meanings. Consequently, this methodology does not lend itself to exploring the nature of shared knowledge within the social group. When a conceptual analysis strategy is used, attempts to locate meaning or shifts in meaning are made by applying factor and/or cluster analysis (Namenwirth and Weber, 1987; Gallhofer and Saris, 1988a), analysis of variance (Potter, 1982) or other statistical techniques to the frequencies with which concepts occur in a set of texts. It is assumed, in such cases, that meaning can be measured by the covariance of the concepts across or within texts. Such a procedure seems to work reasonably well at macro levels where the interest is in broad cultural concepts (Gallhofer and Saris, 1988b; Namenwirth and Weber, 1987) but it appears to be less effective at the micro level where it is necessary to examine more psychological or cognitive phenomena (Ogilvie, 1966; Namenwirth and Weber, 1987, p. 195). At the micro level it is necessary to examine how people inter-relate concepts in order that information on the semantic and syntactic structure not be lost (Carley, 1986a; Carley and Palmquist, 1990; Roberts, 1987).

2. Procedural Analysis

Procedural analysis focuses on what procedures or actions are present in the text. Procedural analysis, like conversational analysis (Sacks, 1972), discourse analysis (Stubbs, 1983b) and computational hermeneutics (Alker, 1985; Mallery, 1985; Mallery, 1986), treats the content of text as action. Arguments like Austin's (1962), that utterances are actions, are often made. Procedural analysis provides information about the structure of a given task and the repertoire of actions that an individual can draw upon when engaging in a task. Unlike conceptual analysis, it allows the researcher to focus on process. Procedural analysis, thus focuses the researchers attention on the domain, action sequences, and decision sequences exhibited by the author or actor in the text.

In addition to taking an action perspective, procedural analysis extracts from the text action sequences and goals. The transcript is treated as plot and sequences of action and re-action. Consequently, unlike conceptual analysis, the specific order of the sentences in the text matters to the analysis. There are two distinct procedural approaches -- decision-based and plot-based. Decision-based procedural analysis typically uses texts of a single actor engaged in a task such as chess or mathematics and focuses on what it is the actor is thinking and doing in the performance of that task (Newell, 1972; Ericsson and Simon, 1984; Simon, 1979; VanLehn and Brown, 1980). Decision-based procedural analysis is often referred to as protocol analysis (Ericsson and Simon, 1984). The goal is to locate the explicit and implicit "rules" that the speaker/author uses to perform a task such as chess (Newell, 1972; Kilpatrick, 1968; Goor, 1974; Goor and Sommerfeld, 1975; Ericsson and Simon, 1984). The researcher typically focuses on how is it that people perform a particular task and what errors they are likely to make. In many cases, rules extracted from such protocols are used as the basis for, or contrasted with, rules employed by artificial intelligence programs for doing the same tasks. In contrast, plot-based procedural analysis typically uses articles, books, or stories and focuses on the story or plot (Abelson, 1976; Abelson, 1969; Schank, 1977; Cullingford, 1981; Wilensky, 1981; Rumelhart, 1976a; Rumelhart, 1978; Lehnert and Ringle, 1982; Lehnert, 1981; Heise, 1987).

The main difficulty with the decision-based approach to procedural analysis is automation sufficiently robust to work across a variety of tasks. Recently, automatic procedures such as Cirrus (VanLehn and Garlick, 1987; Kowalski and VanLehn, 1988) and ACM (Langley and Ohlsson, 1984) have emerged providing hope that larger number of texts can be analyzed quickly and economically. In contrast, plot-based procedural analysis lends

itself to automation. That is, it is possible to develop elementary syntactic units such as basic action sets (Schank, 1977; Abell, 1984; Abell, 1988; Abell, 1989) or plot units (Lehnert, 1981) that make possible the automatic coding of texts (Cullingford, 1981; Lehnert, 1975; Wilensky, 1981; Dyer, 1983). Difficult issues in this area include handling large numbers of actors, emotions, social knowledge, and extremely novel plots. Work by Dyer (1983) and Lehnert and Vine (1987) provides an exciting approach to the coding of narrative that takes into account the underlying affective meaning. However, even those plot-based procedures that are automated or semi-automated are still at the stage of being used on extremely small numbers of texts. The software is generally used to exhibit the ways in which representation scheme makes possible the interpretation and coding of a single text. Thus, although the methodologies are generalizable in principle, they have not been shown to be so in fact. In addition, research in this area is hampered by the fact that different research questions seem to require the creation of special purpose action or plot sets. This is a problem that is analogous to the creation of general-purpose dictionaries and rule sets when conducting conceptual analysis. Nonetheless, this methodology may potentially prove quite valuable to social scientists interested in analyzing texts that contain multiple actors -- such as debates, group meeting transcripts, and so forth -- or when the concern is with action sequences.

3. Relational Analysis

Relational analysis goes beyond conceptual analysis in that it focuses both on what concepts are present in the text and on the relations between those concepts. Researchers engaging in relational analysis typically treat concepts as ideational kernels, single ideas totally bereft of meaning except as they are connected to other concepts (Carley, 1986b). All concepts in isolation are thought to be meaningless (Carley, 1986a; Carley, 1986b). Thus, concepts are nothing more than symbols whose meaning is dependent on their relationship to other symbols (Osgood, 1963; Gollob, 1968; Heise, 1969; Heise, 1970; Minsky, 1975; Carley, 1986a; Carley, 1986b; Carley, 1990b; Roberts, 1987). A large number of relational analysis techniques exist: affect extraction, proximity analysis, and assorted cognitive mapping techniques. All of these techniques are concerned with extracting a conceptual network (a set of concepts and the relationships between them), but differ in what they count as a relationship. In addition, the cognitive mapping techniques have, to a large extent, been made possible by recent advances in artificial intelligence, cognitive psychology and network analysis. Such advances provide a theoretical foundation for such a methodology and suggest a set of techniques that can be employed to represent and analyze cognitive maps or mental models.

Affect extraction (Gollob, 1968; Heise, 1969; Heise, 1970) provides an emotional evaluation of the text based on the relationship between concepts that are explicitly present in the text. For Gollob (1968) and Heise (1969, 1970) the meaning of a concept is numerically determined as an additive function of the combined affective value of the related concepts. This approach differs from the other relational approaches in that concepts in isolation can have an affective evaluation. Thus, isolated concepts are not totally bereft of meaning. Affective evaluations of isolated concepts are, however, difficult to determine and may vary across time and populations.

Proximity analysis (Danowski, 1980; Danowski, 1982; Danowski, Barnett and Friedland, 1986; Danowski, 1988; Hildum, 1963; Salem, 1988; vanMeter and Mounier, 1989) is concerned with the co-occurrence of explicit concepts in the text. Typically the text is treated as a single string of words. Then a window is defined that determines how many contiguous words will be examined to locate relationships. This window is moved along the string and the co-occurrence of words within that window are counted. A concept-by-concept matrix is created which simply contains the number of times the row-word preceded the column-word within the text. This approach has been used with either a predefined list or dictionary of concepts or on all concepts within the text. This procedure is totally automated, and permits the rapid analysis of large numbers of texts. Difficulties with this approach

stem from the fact that it is concerned with only explicit concepts and treats meaning as proximal co-occurrence.

Cognitive mapping involves extracting from text and then representing, the "mental models" or "cognitive structures" that individuals at one time had in their memory. The idea of mental models has appeared under various guises in the work of Mead (Mead, 1962; Mead, 1964), symbolic interactionists (Blumer, 1969; Stryker, 1980), Goffman (Goffman, 1974; Goffman, 1963), social constructivists (Latour and Woolgar, 1979; Knorr-Cetina, 1981), philosophers (Giere, 1988), cognitive and social psychologists (Fiske and Taylor, 1984), researchers in artificial intelligence (Minsky, 1975; Schank, 1977; Johnson-Laird, 1983; Sowa, 1984), and many others. Cognitive mapping techniques are based on three theoretical assumptions: (1) mental models are internal representations; (2) language is the key to understanding mental models; and (3) mental models can be represented as networks. Within a cognitive map, meaning is located within the organization of, and hence relationships between, concepts (Axelrod, 1976; Shavelson, 1972; Sowa, 1984; Carley, 1986a; Garnham, 1987). Mental models are dynamic structures that are constructed as individuals make inferences (Johnson-Laird, 1983) and gather information (Carley, 1986a; Carley, 1990b; Carley, 1990c). They contain both specific knowledge about particular items, such as *Cassandra is two years old*, and general (or social) knowledge, such as *A mother is older than her children*. The text contains a snapshot of an individual's mental model at a particular point in time and so can be thought of as a sample of the information in the individual's memory (Carley, 1986a; Carley, 1990b).

A variety of schemes for representing mental models as some type of cognitive map have been proposed, all of which share a basic network orientation: conceptual structures (Sowa, 1984; Johnson-Laird, 1983; Garnham, 1987), hierarchical dependency maps (Means and Voss, 1985), conceptual distance maps or clusters (Jonassen, 1987; Reitman and Rueter, 1980; McKeithen et al., 1981), schemes (Anderson J., 1973; Bobrow, 1976), schemata (Rumelhart, 1976b; Tesser, 1977; Tversky, 1980), structured frames (Minsky, 1975; Charniak, 1972), dynamic frames (Goffman, 1974), transition networks (Collins, 1975; Clark, 1977; Wyer, 1979; Bobrow, 1969), semantic nets (Simmons, 1973; Schank, 1973), scripts (Schank, 1977; Abelson, 1976; Abelson, 1969), plot units (Lehnert, 1981; Lehnert, 1987), semantic planning nets (VanLehn and Brown, 1980; Leinhardt, 1985), decision and belief networks (Axelrod, 1976; Eden, Jones and Sims, 1979), and cognitive maps (Carley, 1986a; Carley, 1990b; Carley, 1986b; Carley and Palmquist, 1990).

These approaches tend to characterize mental models as semantic structures in which verbal statements are translated into visual structures composed of concepts and the relationships between them. All of these approaches have techniques for representing the text's content; but, very few have associated procedures, or software, for actually coding, comparing, contrasting, combining, and empirically analyzing the resultant cognitive maps. In addition, the procedural analysis techniques previously discussed are cognitive mapping techniques that focus on action and the presented order of actions in order to inter-relate concepts. By contrast, the three general types of cognitive mapping techniques that will now be discussed – linguistic content analysis, decision maps and mental model extraction – are typically less concerned about action and the order in which statements occur in the text. In addition, all three techniques admit empirical analysis of the cognitive map.

Linguistic content analysis (Lindkvist, 1981; Hutchins, 1982; Gottschalk, Hausmann and Brown, 1975; Salton and Smith, 1988; Salton, 1989; Roberts, 1987) is concerned with the empirical analysis of syntactic information. Linguistic content analysis techniques typically take the text and analyzes it one clause at a time, converting clauses into the structure or numeric equivalent of their clause-type. For example, Gottschalk et al. (1975) have an automated procedure that evaluates each clause in a text and determines its numerical score on each of several emotional/psychological scales (see also Gottschalk and Gleser, 1969). As another example, relying on the recent work in linguistics such as the work conducted by Searle (Searle, 1975; Searle, 1979; Searle and Vanderveken, 1985), Roberts (1987) proposes a

strategy, based on the grammatical parsing of texts into clauses and parts of speech, that can be used to convert a text into a matrix representation. This latter approach is not automated and is extremely costly as it requires extensive training of coders.

Decision maps represent the inter-relationship between ideas, beliefs, attitudes, and information that the author of the text considers when making a decision on a particular topic. Links between concepts generally represent logical, inferential, causal, sequential relationships (Axelrod, 1976; Axelrod, 1972; Heise, 1987; Neuman, 1986) or mathematical relationships (Eden, Jones and Sims, 1979). Gallhofer and Saris (1988) utilize a decision tree approach in which all concepts are forced into a hierarchical structure. In contrast, other approaches use a more generalized network structure (Axelrod, 1976; Axelrod, 1972; Bonham, Shapiro and Nozicka, 1976; Eden, Jones and Sims, 1979; Heise, 1987; Carley, 1984). These approaches differ in the methodology for evaluating the conceptual network to produce a decision. Axelrod (1976) and Heise (1987) do a predominantly symbolic evaluation using the logical and sequential links between concepts; whereas, Bonham, Shapiro, and Nozicka (1976), Eden, Jones and Sims (1979) and Carley (1984) use various numerical evaluation schemes. Many of these procedures are automated (Heise, 1987; Bonham, Shapiro and Nozicka, 1976; Eden, Jones and Sims, 1979; Carley, 1984). In addition, Eden, Jones and Sims discuss network-based techniques for evaluating cognitive maps that are generally useful regardless of whether the map of interest is a decision map.

Mental model extraction is a semi-automated procedure for extracting, comparing, combining, and numerically and graphically analyzing cognitive maps (Carley, 1986a; Carley, 1990c; Carley, 1990a; Carley and Palmquist, 1990). When this approach is used the researcher converts a text into a map of concepts and relations that can be analyzed at two levels -- concepts and statements (where a statement is two concepts and the relation between them (Minsky, 1975; Carley, 1986a; Carley, 1986b)). Each concept is an index to a larger network of meaning (Cicourel, 1974; Carley, 1986b). The procedure is highly general and can be used to extract many different types of maps where the concepts are either explicit or implicit within the text and where relations between concepts can be of many types. When the approach is taken the researcher engages in a four step process: (1) identify concepts, (2) define relationship types, (3) codes specific texts using these concepts and relationships, and (4) graphically display or numerically analyze the resultant maps.

There are several unique features to this approach. Many of these features derive from the fact that this approach is based on a cognitively motivated theory of the nature of knowledge, knowledge acquisition, and knowledge dissemination that provides direction for locating implicit concepts. One such feature is an expert system that utilizes social knowledge to make explicit implicit information and thereby increase the reliability of the coded text (Carley, 1988). Another feature is software for combining and comparing mental models derived from multiple sources. This makes it possible to locate maps that represent socially shared cognitions (Carley, 1986b). A third feature is software that, given a set of cognitive maps, extracts information from each of them and puts it into a format that can be analyzed by standard statistical packages (Carley, 1990a).

The diverse techniques and analyses employed in cognitive mapping suggest the breadth of interest in this methodology. A common limitation is the lack of complete automation when dealing with implicit concepts. As a result, these techniques tend to be fairly time consuming. More fully automated procedures, however, are appearing (Danowski, 1988; Carley, 1990a) which should ease the coding burden. A goal behind cognitive mapping techniques is to enable the researcher to construct and compare representations of mental models in a rigorous fashion without losing the richness of detail present in more qualitative approaches. By specifying the nature and structure of relationships in the mental models being represented, the approach emphasizes meaning and allows the researcher to determine, through an examination of the ways in which people relate concepts together, whether people mean the same thing by the words that they use. Thus, such techniques should make it possible for the researcher to represent the dynamic

properties of the model and distinguish general from specific knowledge. In addition, such techniques should admit aggregation and comparison of the cognitive maps across time and individuals. While all of the techniques permit such comparisons in principle, only Eden, Jones and Sims (1979) and Carley (1990) provide procedures that actually facilitate the empirical comparisons of maps.

4. Emotional Analysis

Many texts have an affective or emotional content, which if ignored may reduce the validity of the analysis but which if attended to may increase the researcher's ability to interpret the text. There are, however, a number of difficulties in attending to the affective content of the text many of which stem from the fact that a large number of "emotions" are not semantically distinct. Consequently, researchers are attempting to classify emotions (Mees, 1985; Johnson-Laird, 1986; Johnson-Laird, 1988), create definitions for emotions (Wierzbicka, 1972; Wierzbicka, 1973), and create affective lexicons (Ortony, 1987). For example, Ortony, Clore, and Foss (1987) used componential analysis to develop a taxonomy based on 500 emotion related words. Given this taxonomy, they propose that the best examples of terms denoting emotions are those that refer to internal mental conditions rather than external or physical conditions, are clear cases of states of being rather than frames of mind, and have a predominant referential focus that is affective rather than behavioral or cognitive. Examples would be "happy", "broken hearted", and "contented". These attempts are leading to the development of semantically-based cognitive theories of emotion (Johnson-Laird, 1986; Ortony, *ress*).

Emotions create particular problems in the analysis of texts. That is, in texts the emotional states of characters *John loves Mary* and the emotional reactions of characters *John was broken-hearted when Mary eloped with Tony* appear frequently and so procedural analyses that do not take emotions into account are suspect. Moreover, affect is necessary for text comprehension for a variety of reasons: emotions reveal underlying goals (but are not tied to a specific goal), emotions signal the occurrence of expectation failures (Joe is worried about x), emotions indicate the status of interpersonal relations, emotions influence what thematic structures become instantiated in episodic memory (e.g., is this a romance or a mystery) (Dyer, 1983). Consequently, in many plot-based schemes (Rumelhart, 1975; Mandler, 1977; Stein, 1979) affect is included along with goals, cognitions, and internal states and events as one of the components into which text statements are parsed. When this approach is taken, the researcher creates a categorization scheme for each of the components, e.g., there are 2 goals -- becoming un-hungry and achieving success, and there are 5 emotions -- love, hate, fear, anger, oblivion. These categories are then used to map out the underlying meaning of the story. Each proposed scheme differs in how components are characterized. Given a particular scheme the researcher might develop a computer program to code, and hence analyze, the text. An example is BORIS (Dyer, 1983), a computer program that based upon the conceptual dependency approach developed by Schank and Abelson (Schank, 1977). In BORIS the emotional states of story characters are represented as knowledge structures referred to as AFFECT. These structures are used to follow the emotional states of story characters throughout the story.

The difficulty with the categorization approach is that it forces the researcher to predefine a general-purpose set of universal emotions. In contrast, Lehnert (1981, 1987) has developed a scheme for coding text in terms of plot units that relies on the underlying affective state of the story characters. Lehnert takes a dimensional, rather than categorical, approach to emotions and thus simply codes the story character's emotional state as positive (one that pleases), negative (one that displeases), or neutral. The specific type of emotion present in the character, e.g., love or hate, is not noted. The emotional state then determines the meaning and nature of the plot units present in the text.

For researchers interested in types of texts other than stories, a dimensional

approach to emotions is the most common. For example, Gottschalk (1969) scores the content of texts in terms of a series of emotional/psychological scales, such as anxiety, hostility and social alienation, thus creating a multi-dimensional view of the emotional content of the text. Gollob (1968) and Heise (1969, 1970) evaluate affective content by numerically determining the meaning of a concept in terms of the combined affective value of the related concepts. And, still others use the Bale's categories (1950) in order to disambiguate socio-emotional from task content (Rice and Love, 1987).

5. Discussion

For over a quarter of a century computer-based tools for analyzing content have been available and have exhibited a wide range of applicability. Nevertheless, today, researchers in this area are faced with a confusing panoply of tools that often fail to generalize across projects, often suffer from lack of complete automation, are rarely portable across machines, are rarely theoretically grounded, and are generally poorly documented. Applying these techniques, as is true for most ethnographic approaches, is generally an extremely time consuming process unless the researcher focuses exclusively on explicit concepts. Nevertheless, regardless of the types of concepts examined, use of automated and semi-automated procedures reduces coding time, enables the analysis of larger numbers of texts and facilitates statistical analysis. Conceptual analysis procedures are easily automated and continue to be used by researchers in a wide variety of areas. Relational analysis procedures are currently less commonly used but may evolve as the methodology of choice as like conceptual analysis many of the procedures involved in cognitive mapping can be automated and relational analysis techniques tend to subsume conceptual analysis techniques in that they also can provide information concerning the frequency with which concepts occur in texts. Unlike conceptual analysis techniques, however, relational approaches can also be used to explore the relationships between those concepts thus allowing for a more thorough exploration of the meaning of a given text.

It should be noted that research in this area is hampered not only by the lack of methodological tools, but also by the lack of a clearly defined theoretical foundation. The types of theoretical issues that need to be addressed include: the relationship between mental models and language, the relationship between words and meaning, the role of emotions in text, the role of syntax in textual analysis, and the nature of social knowledge or shared meaning. Although, recent work has increased our theoretical understanding in these areas and provided a new set of methodological tools consistent with these theoretical advances there is still a long way to go.

Areas that will be particularly important to develop in order to extend the usefulness of content analysis techniques will be the integration of text and video (Palmer, 1989), utilization of syntax to augment semantic analysis (Roberts, 1987), automated relational analysis (Danowski, 1988; Carley, 1988), and utilization of affectual and social content when coding meaning (Lehnert, 1987; Dyer, 1983; Carley, 1988; Cicourel and Carley, 1990). Further, the entire procedure needs further automation, particularly when implicit concepts are employed and when the researcher is interested in the area of cognitive mapping. As new tools emerge, it is important that they be tied to graphical output procedures and standard statistical packages so that the data can be analyzed visually as well as empirically. Only under such conditions will content analysis techniques admit rich as well as quantitative text descriptions.

REFERENCES

- Garson D. (1985). Content analysis machine-readable data file V 1.0.
- Qualis Research Associates. (1989). *The Ethnograph*.
- Abell P. (1984). Comparative Narratives, Some Rules for the Study of Action. *Journal for the Theory of Social Behavior*, 14, 309-332.
- Abell P. (1988). The Structuration of Action: Inference and Comparative Narratives. In Fielding N.G. (Eds.), *Actions and Structures*. London, England: Sage.
- Abell P. (1989). The Theory and Method of Comparative Narratives. *Connections*, 12, 51-79.
- Abelson, R.P. and Reich, C.M. (1969). Implicated Molecules; A method for extracting meaning from input sentences. In Walker, D.E. and L.M. Norton (Eds.), *Proceedings of the International Joint Conference on Artificial Intelligence*. : .
- Abelson, R.P. (1976). Script Processing in Attitude Formation and Decision-Making. In Carroll, J.S. and J.W. Payne (Eds.), *Cognition and Social Behavior*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Alexander M. and J. Danowski. (1990). Analysis of an ancient network: Personal communication and the study of social structure in a past society. *Social Networks*, , .
- Alker H., G. Duffy and J. Mallery. (1985). From Quantity to Quality: A New Research Program in Sequential Prisoner's Dilemmas.
- Anderson J. and G. Bower. (1973). *Human Associative Memory*. Washington DC: Winston-Wiley.
- Andren G. (1981). Reliability and Content Analysis. In Rosengren K.E. (Eds.), *Advances in Content Analysis*. Beverly Hills, CA: Sage.
- Urmson, J. (Eds.). Austin, J. (1962). *How to Do Things with Words*. London, ENGLAND: Oxford University Press.
- Axelrod, R.M. (1972). *Framework for a General Theory of Cognition and Choice*. Berkeley, CA: University of California at Berkeley.
- Axelrod, R.M. (1976). *Structure of Decision*. Princeton, NJ: Princeton University Press.
- Bales R.F. (1950). *Interaction Process Analysis: A Method for the Study of Small Groups*. Reading, MA: Addison-Wesley.
- Bar-Hillel, Y. (1960). The Present Status of Automatic Translation of Languages". In Alt, F.L. (Eds.), *Advances in Computers 1*. New York, NY: Academic Press.

- Berelson B. (1952). *Content analysis in communication research*. New York, NY: Hafner.
- Blumer, H. (1969). *Symbolic Interactionism*. Englewood Cliffs, NJ: Prentice Hall.
- Bobrow, D. G. and B. Fraser. (1969). An Augmented state transition network analysis procedure. In Bobrow, D. G. and B. Fraser (Eds.), *Proceedings of the First International Joint Conference on Artificial Intelligence*. .
- Bobrow D. G. and D. A. Norman. (1976). Some principles of memory schemata. In Bobrow, D. G. and A. Collins (Eds.), *Representation and Understanding: Studies in Cognitive Science*. New York, NY: Academic Press.
- Bonham G.M., Shapiro M.J. and G.J. Nozicka. (1976). A Cognitive Process Model of Foreign Policy Decision Making. *Simulation and Games*, 7, .
- Brent E. (1984). Qualitative Computing: Approaches and Issues. *Qualitative Sociology*, 7, 34-60.
- Bruce B. and D. Newman. (1978). *Interacting Plans* (Tech. Rep.). Champaign, IL., Center for the Study of Reading.
- Carley K.M. (1984). *Constructing Consensus*. Doctoral dissertation, Harvard,
- Carley K.M. (1986). An Approach for Relating Social Structure to Cognitive Structure. *Journal of Mathematical Sociology*, 12(2), 137-189.
- Carley K.M. (1986). Knowledge Acquisition as a Social Phenomenon. *Instructional Science*, 14, 381-438.
- Carley K.M. (1988). Formalizing the Social Expert's Knowledge. *Sociological Methods and Research*, 17, 165-232.
- Carley K.M. (August 1990). Computer Analysis of Qualitative Data - Copy of Overheads.
- Carley K.M. (1990). The Social Construction of Knowledge. In Carley K.M., Lawler R. and G. Drescher (Eds.), *Experiments in Epistemology*. Norwood, NJ: Ablex.
- Carley K.M. (1990). Knowledge, Interaction, and Language. In Carley K.M., Lawler R. and G. Drescher (Eds.), *Experiments in Epistemology*. Norwood, NJ: Ablex.
- Carley K.M. and Palmquist M. (1990). Extracting, Representing and Analyzing Mental Models. *Working paper*, , .
- Charniak E. (1972). *Toward a Model of Children's Story Comprehension*. Doctoral dissertation, Massachusetts Institute of Technology,
- Chiesi H.L., Spilich G.J., and Voss J.F. (1979). Acquisition of domain-related information in

- relation to high and low domain knowledge. *Journal of Verbal Learning and Verbal Behavior*, 18, 257-273.
- Cicourel, A.V. (1974). *Cognitive Sociology*. New York, NY: The Free Press, Macmillan Publishing Co.
- Cicourel A. and K.M. Carley. (1990). The Coder of Narrative as Expert: The Ecological Validity of Coding Practices. to be presented at the American Anthropological Association Meetings, , .
- Clark, H.H. and Clark E.V. (1977). *Psychology and language: An introduction to psycholinguistics*. New York, NY: Harcourt Brace Jovanovich.
- Collins, B.E. and Loftus, E.P. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, 82, 407-428.
- Collins, A., J.S. Brown, and K.M. Larkin. (1977). *Inference in Text Understanding* (Tech. Rep. Report #3684). Bolt Berenek and Newman,
- Conrad P. and Reinhartz S. (1984). Computers and qualitative data. *Qualitative Sociology*, 7,
- Cullingford, R. (1981). SAM. In Schank R.C. and C.K. Riesbeck (Eds.), *Inside Computer Understanding*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Danowski J. A. (1980). Computer-mediated communication: a network-based content analysis using a CBBS conference. In Burgoon M (Ed.), *Communication Yearbook 6*. Beverly Hills, CA: SAGE.
- Danowski J. (1982). A Network-based Content Analysis Methodology for Computer-Mediated Communication: An Illustration with a Computer Bulletin Board. In (Eds.), *Communication Yearbook*. : Sage.
- Danowski J. (1988). Organizational Infographics and Automated Auditing: Using Computers to Unobtrusively Gather and Analyze Communication. In Goldhaber G. and G. Barnett (Eds.), *Handbook of Organizational Communication*. Norwood, NJ: Ablex.
- Danowski J., Barnett G. and M. Friedland. (1986). Interorganizational Network Position, Media Coverage, and Publics' Images. In McLaughlin M. (Eds.), *Communication Yearbook*. : Sage.
- Dennis D.L. (1984). 'Word Crunching': An Annotated Bibliography on Computers and Qualitative Data Analysis. *Qualitative Sociology*, 7, 148-156.

- Drass K.A. (1989). Text Analysis and Text-Analysis Software: A Comparison of Assumptions. In Blank G., McCartney J.L. and Brent E. (Eds.), *New Technology in Sociology*. New Brunswick, NJ: Transaction Publishers.
- Dunphy D.C., Bullard C.G. and E.E.M. Crossing. (1974). Validation of the General Inquirer Harvard IV Dictionary.
- Dyer, M. (1983). The Role of Affect in Narratives. *Cognitive Science*, 7, 211-242.
- Eden C., Jones S. and Sims D. (1979). *Thinking in Organizations*. London, England: Macmillan Press.
- Edwards D. and Mercer N. (1986). Context and continuity: Classroom discourse and the development of shared knowledge. In K. Durkin (Eds.), *Language development in the school years*. Cambridge, MA: Brookline Books.
- Ericsson K.A. and H.A. Simon. (1984). *Protocol Analysis: Verbal Reports as Data*. Cambridge, MA: MIT Press.
- Fan D. (1988). *Predictions of public opinion from the mass media: computer content analysis and mathematical modeling*. New York, NY: Greenwood Press.
- Fennel R.D. and V.R. Lesser. (1977). Parallelism in AI Problem-Solving: A Case Study of HEARSAY II. *IEEE Transactions on Computers*, . .
- Finch L., Landry J., Monarchi D., and Tegarden D. (1987). A knowledge acquisition methodology using cognitive mapping and information display boards. *Proceedings of the Twentieth Annual Hawaii International Conference on Systems Sciences*. Hawaii, .
- Fiske S. and S. Taylor. (1984). *Social Cognition*. Reading MA: Addison-Wesley.
- Splichal S. and Ferligoj A. (1988). Ideology in International Propaganda. In Saris W.E. and Gallhofer I.N. (Eds.), *Sociometric Research VI: Data Collection and Scaling*. New York, NY: St. Martin's Press.
- Gallhofer I.N. and Saris W.E. (1988). A Coding Procedure for Empirical Research of Political Decision Making. In Saris W.E. and Gallhofer I.N. (Eds.), *Sociometric Research VI: Data Collection and Scaling*. New York, NY: St. Martin's Press.
- Garnham A. (1987). *Mental models as representations of discourse and text*. Chichester, West Sussex England: Halsted Press .
- Gerbner G. (Ed). (1969). *The Analysis of communication content; developments in scientific theories and computer techniques*. New York, NY: Wiley.

- Gere A.R. (1987). *Writing groups: History, theory, and implications*. Carbondale, IL: Southern Illinois University Press.
- Gere A.R. and Stevens R.S. (1985). The language of writing groups: How oral response shapes revision. In S.W. Freedman (Eds.), *The acquisition of written language: Response and revision..* Norwood, NJ: Ablex.
- Gerson E.M. (1984). Qualitative Research and the Computer. *Qualitative Sociology*, 7, 61-74.
- Giere R.N. (1988). *Explaining Science: A Cognitive Approach*. Chicago, IL: University of Chicago Press.
- Goffman E. (1963). *Behavior in Public Places: Notes on the Social Organization of Gatherings*. Glencoe, IL: Free Press.
- Goffman E. (1974). *Frame Analysis: An Essay on the Organization of Experience*. New York, NY: Harper and Row.
- Gollob H. (1968). Impression Formation and Word Combination in Sentences. *Journal of Personality and Social Psychology*, 10, 341-353.
- Goor A. (1974). *Problem Solving Process of Creative and Non-Creative Students*. Doctoral dissertation, University of North Carolina,
- Goor A. and R.E. Sommerfeld. (1975). A Comparison of Problem-Solving Processes of Creative Students and Non-Creative Students. *Journal of Educational Psychology*, 67, 495-505.
- Gottschalk L.A. Gleser G.C. (1969). *The measurement of psychological states through the content analysis of verbal behavior*. Berkeley, CA: University of California Press.
- Gottschalk L.A. and Winget C.N. (1969). *Manual of instructions for using the Gottschalk-Gleser content analysis scales*. Berkeley, CA: University of California Press.
- Gottschalk L.A., Hausmann C. and J.S. Brown. (1975). A Computerized Scoring System for Use with Content Analysis Scales. *Comprehensive Psychiatry*, 16, 77-90.
- Grey A., Kaplan D. and H.D. Lasswell. (1965). Recording and Context Units -- Four Ways of Coding Editorial Content. In Laswell H.D., Leites N. and Associates (Eds.), *Language of Politics*. Cambridge, MA: MIT Press.
- Gussarsky E., and Gorodetsky M. (1988). On the chemical equilibrium concept: Constrained word associations and conception. *Journal of Research in Science Teaching*, 25, 319-333.

- Heise D. (1969). Affectual Dynamics in Simple Sentences. *Journal of Personality and Social Psychology*, 11, 204-213.
- Heise D. (1970). Potency Dynamics in Simple Sentences. *Journal of Personality and Social Psychology*, 16, 48-54.
- Heise, D. (1987). *Ethno Documentation* (Tech. Rep.). Indiana University,
- Hildum, D.C. (1963). Semantic Analysis of Texts by Computer. *Language*, 39, 649-653.
- Holsti O.R. (1969). *Content Analysis for the Social Sciences and Humanities*. Reading, MA: Addison-Wesley.
- Hutchins W.J. (1982). The Evolution of Machine Translation Systems. In Lawson V. (Eds.), *Practical Experience of Machine Translation: Proceedings of a Conference, London Nov. 1981.* : North-Holland.
- Iwanska L. (1989). *Automated processing of narratives written by bilingual students: the BILING program* (Tech. Rep.). Urbana-Champaign, IL: University of Illinois at Urbana-Champaign Dept. of Computer Science Report UIUCDCS-R-89-1508,
- Janowitz M. (1969). Content Analysis and the Study of the 'Symbolic Environment'. In Rogow A.A. (Eds.), *Politics, Personality, and Social Science in the Twentieth Century*. Chicago, IL: University of Chicago Press.
- Johnson-Laird P. (1983). *Mental models: Toward a cognitive science of language, inference, and consciousness*. Cambridge, MA: Harvard University Press.
- Johnson-Laird, P.N. and K. Oatley. (1986). The meaning of emotions: A cognitive theory and a semantic analysis. Unpublished Manuscript.
- Johnson-Laird P.N. (1988). *The language of emotions: an analysis of a semantic field* (Tech. Rep.). Princeton, NJ: E&S Tech report PRC 33,
- Jonassen D.H. (1987). Assessing cognitive structure: Verifying a method using pattern notes. *Journal of Research and Development in Education*, 20, 1-14.
- Kelly E.F. and Stone P.J. (1975). *Computer Recognition of English Word Senses*. Amsterdam: North Holland: .
- Kilpatrick J. (1968). *Analyzing the Solution of Word Problems in Mathematics: An Exploratory Study*. Doctoral dissertation, Stanford University,
- Knorr-Cretina K. (1981). *The Manufacture of Knowledge*. Oxford England: Pergamom.
- Kowalski B. and K. VanLehn. (1988). Cirrus: Inducing Subject Models from Protocol Data.

- Patel V. (Eds.), *Proceedings of the Tenth Annual Conference of the Cognitive Science Society*. Hillsdale, NJ, Erlbaum.
- Krippendorff K. (1980). *Content Analysis: An Introduction to its Methodology*. Beverly Hills, CA: Sage.
- Langer J.A. and Applebee A.N. (1986). Reading and writing instruction: Toward a theory of teaching and learning. In E.Z. Rothkopf (Eds.), *Review of Research in Education*, Vol. 13. Washington, DC: AERA.
- Langley P. and S. Ohlsson. (1984). Automated Cognitive Modeling. (Eds.), *Proceedings of the Fourth National Conference on Artificial Intelligence*. Austin, TX, Morgan-Kaufman.
- Latour B. and S. Woolgar. (1979). *Laboratory Life*. Beverly Hills, CA: Sage.
- Lehnert, W.G. (1977). Human and computational question answering. *Cognitive Science* 1, ,
- Lehnert W. (1981). Plot Units and Narrative Summarization. *Cognitive Science*, 5, 293-331.
- Lehnert, W. and E. Vine. (1987). The Role of Affect in Narrative Structure. *Cognition and Emotion*, 1, 299-322.
- Lehnert W. and M. Ringle (Eds.). (1982). *Strategies for Natural Language Processing*. Hillsdale, NJ: Lawrence Erlbaum.
- Lehnert, W.G. (1975). What makes SAM run? In Schank, R.C. and Nash-Webber, B.L. (Ed.), *Theoretical Issues in Natural Language Processing*. Cambridge, MA: .
- Leinhardt G. (in press). Situated knowledge and expertise in teaching. In Calderhead J. (Eds.), *Teacher's professional learning*. Lewes, Essex, England: Palmer Press.
- Lindkvist K. (1981). Approaches to Text Analysis. In Rosengren K.E. (Eds.), *Advances in Content Analysis*. Beverly Hills, CA: Sage.
- Mallery, J. (1985). Universality and Individuality: The Interaction of Noun Phrase Determiners in Copular Clauses. *23rd Annual Meeting of the Association for Computational Linguistics: Proceedings of the conference*. Chicago, IL.
- Mallery, J. and G. Duffy. (1986). *A Computational Model of Semantic Perception* (Tech. Rep.). MIT Artificial Intelligence Laboratory,
- Mandler, J. and N. Johnson. (1977). Remembrance of Things Parsed: Story Structure and Recall. *Cognitive Psychology*, 9, 111-151.

- Markoff J., Shapiro G. and S. Weitman. (1974). Toward the Integration of Content Analysis and General Methodology. In Heise, D. (Eds.), *Sociological Methodology*. San Francisco, CA: Jossey-Bass.
- McKeithen K.B., Reitman J.S., Rueter H.H., and Hirtle S.C. (1981). Knowledge organization and skill differences in computer programmers. *Cognitive Psychology*, 13, 307-325.
- Mead G.H. (1934,1962). *Mind, Self, and Society*. Chicago, IL: University of Chicago Press.
- Reck A. (Eds.). Mead G. (1964). *The Genesis of the Self and Social Control*. New York, NY: Bobbs-Merrill.
- Means M.L., and Voss J.F. (1985). Star Wars: A developmental study of expert and novice knowledge structures. *Journal of Memory and Language*, 24, 746-757.
- Mees U. (1985). What do we mean when we speak with feelings? On the psychological texture of words denoting emotions. *Sprache und Kognition*, 1, 2-20.
- Merleau-Ponty, M. (1964). *Consciousness and the Acquisition of Language*. Evanston, IL: Northwestern University Press.
- Meyer B.J.F. (1975). *The organization of prose and its effects on memory*. Amsterdam: North-Holland Pub. CO. .
- Michaels S. (1987). Text and context: A new approach to the study of classroom writing. *Discourse Processes*, 10, 321-346.
- Minsky, M.A. (1975). A framework for representing knowledge. In Winston, P. (Ed.), *The Pshchology of Computer Vision*. New York, NY: McGraw-Hill.
- Moshe N.B., McKeachie W.J., Lin Y.G., and Tucker D.G. (1986). Inferring students' cognitive structures and their development using the "ordered tree technique.". *Journal of Educational Psychology*, 78, 130-140.
- Namenwirth J.Z., Miller R.L. and R.P. Weber. (1981). Organizations have Opinions: Towards a Redefinition of Publics. *Public Opinion Quarterly*, 45, 463-476.
- Namenwirth A. and R. Weber. (1987). *Dynamics of Culture*. Boston, MA: Allen & Unwin.
- Neuman W.R. (1986). *The Paradox of Mass Politics*. Cambridge, MA: Harvard University Press.
- Neuman W.R. (1989). Parallel Content Analysis: Old Paradigms and New Proposals. *Public Communication and Behavior*, 2, 205-289.

- Newell A. and H.A. Simon. (1972). *Human Problem Solving*. Englewood Cliffs, NJ: Prentice-Hall.
- North R.C. (1963). *Content analysis; a handbook with applications for the study of international crisis*. Evanston, IL: Northwestern University Press.
- Ogilvie D.M. (1966). Procedures for Improving the Interpretation of Tag Scores: The Case of Windle. In Stone P., Dunphy D.D., Smith M.S. and D.M. Ogilvie (Eds.), *The General Inquirer: A Computer Approach to Content Analysis*. Cambridge, MA: MIT Press.
- Ogilvie D.M., Stone P.J. and E.F. Kelly. (1982). Computer-Aided Content Analysis. In Smith R.B. and P.K. Manning (Eds.), *New York, NY. A Handbook of Social Science Methods*: Ballinger.
- Ortony A., Clore G.L. and M.A. Foss. (1987). The Referential Structure of the Affective Lexicon. *Cognitive Science*, 11, 341-364.
- Ortony A., Clore G.L., and A. Collins. (In Press). *Principia pathemtica: The cognitive structure of emotions*. New York, NY: Cambridge University Press.
- Osgood, C. (1963). On Understanding and Creating Sentences. *American Psychologist*, 18, 735-751.
- Palmer R.B. (1989). *The Cinematic text: methods and approaches*. New York, NY: AMS Press.
- Palmquist M.E. (1990). *The lexicon of the classroom: Language and learning in writing classrooms*. Doctoral dissertation, Carnegie Mellon University,
- Polanyi L. (1985). *Telling the American story: a structural and cultural analysis of conversational storytelling*. Norwood N.J.: Ablex Publishers.
- Potter R.G. (1982). Reader Responses and Character Syntax. In Bailey R.W. (Eds.), *Computing in the Humanities*. : North-Holland.
- Reddy, D.R., L.D. Erman, R.D. Fennell, and R.B. Neely. (1973). The HEARSAY speech understanding system: An example of the recognition process. *Proceedings of the International Joint Conference on Artificial Intelligence*, 3, 185-193.
- Reitman J.S., and Rueter H.R. (1980). Organization revealed by recall orders and confirmed by pauses. *Cognitive Psychology*, 12, 554-581.
- Rice R. and G. Love. (1987). Electronic Emotion: Socioemotional Content in a Computer-Mediated Communication Network. *Communication Research*, 14, 85-108.

- Roberts C.W. (1987). *Other Than Counting Words: A Linguistic Approach to Content Analysis*. *Social Forces*, 68, .
- Robinson M. and M.A. Sheehan. (1983). *Over the Wire and on TV*. New York, NY: Russell Sage Foundation.
- Rosengren K.E. (1981). *Advances in Content Analysis*. Beverly Hills, CA: Sage.
- Rumelhart D. (1975). Notes on a Schema for Stories. In Bobrow, D. and A. Collins (Eds.), *Representation and Understanding*. New York, NY: Academic Press.
- Rumelhart, D.E. (1976). A Problem Solving Schema for Stories. In LaBerge, D. and S.J. Samuels (Eds.), *Basic Processes in Reading*. Hillsdale, NJ: Lawrence Erlbaum Assoc.
- Rumelhart, D.E. and A. Ortny. (1976). The representation of knowledge in memory. In Anderson, Spiro and Montague (Eds.), *Schooling and the Acquisition of Knowledge*. Lawrence Erlbaum Associates.
- Rumelhart, D.E. (1978). Understanding and Summarizing Brief Stories. In LaBerge, D. and S.J. Samuels (Eds.), *Basic Processes in Reading*. Hillsdale, NJ: Lawrence Erlbaum Assoc.
- Sacks H. (1972). On the analysability of stories by children. In Gumpres J. and D. Hymes (Eds.), *Directions in Socio-linguistics*. New York, NY: Holt, Rinehart & Winston.
- Salem A. (1988). *Le Lexicloud, Fontenay-Saint Cloud*. : Ecole Normale Supérieure.
- Salton G. (1989). *Automatic text processing: the transformation, analysis, and retrieval of information by computer*. Reading, MA: Addison Wesley.
- Salton G. and Smith M. (1988). *On the application of syntactic methodologies in automatic text analysis* (Tech. Rep.). Ithaca, NY: Cornell University. Dept. of Computer Science Tech Report 88-954,
- Saris-Gallhofer I.N., Saris W.E., and E.L. Morton. (1978). A Validation Study of Holsti's Content Analysis Procedure. *Quality and Quantity*, 12, 131-145.
- Schank R. and K. Colby. (1973). *Computer Models of Thought and Language*. San Francisco, CA: W. H. Freeman.
- Schank, R. and R. Abelson. (1977). *Scripts Plans and Goals and Understanding*. New York, NY: Wiley.
- Schank R.C. and C.K. Riesbeck (Eds.). (1981). *Inside Computer Understanding*. Hillsdale, NJ: Lawrence Erlbaum Associates.

- Searle J.R. (1975). Indirect Speech Acts. In Cole P. and Morgan J.L. (Eds.), *Speech Acts*. : Academic Press.
- Searle, J.R. (1979). *Expression and Meaning*. New York, NY: Cambridge University Press.
- Searle J.R. and Vanderveken D. (1985). *Foundations of Illocutionary Logic*. Cambridge, England: Cambridge University Press.
- Seidel J.V. and J.A. Clark. (1984). THE ETHNOGRAPH: A Computer Program for the Analysis of Qualitative Data. *Qualitative Sociology*, 7, 110-126.
- Shavelson R.J. (1972). Some aspects of the correspondence between content structure and cognitive structure in physics instruction. *Journal of Educational Psychology*, 63, 225-234.
- Simmons, R. (1973). Semantic networks: Their computational use for understanding english sentences. In Schank, R. and Colby K. (Eds.), *Computer Models of Thought and Language*. San Francisco, CA: W. H. Freeman, Publisher.
- Simon H. (1979). *Models of Thought*. New Haven, CN: Yale University Press.
- Sowa, J.F. (1984). *Conceptual Structures*. Reading, MA: Addison-Wesley.
- Sozer E. (1985). *Text connexity, text coherence: aspects, methods, results*. Hamburg, Germany: Buske.
- Sproull L.S. and Sproull R.F. (1982). Managing and analyzing behavioral records: Explorations in non-numeric data analysis. *Human Organization*, 41, 283-290.
- Stein N. and C. Glenn. (1979). An Analysis of Story Comprehension in Elementary School Children. In Freedle, R. (Eds.), *New Directions in Discourse Processing*. Norwood, NJ: Ablex.
- Stone, P.J. and Cambridge Computer Associates. (1968). *User's Manual for the General Inquirer*. Cambridge, MA: MIT Press.
- Stone, P.J., D.C. Dunphy, M.S. Smith, and D.M. Ogilvie. (1968). *The General Inquirer: A Computer Approach to Content Analysis*. Cambridge, MA: MIT Press.
- Stryker S. (1980). *Symbolic Interactionism*. Menlo Park, CA: Benjamin Cummings.
- Stubbs M. (1983). *Language, schools and classrooms (2nd ed.)*. London: Methuen.
- (1983). *Discourse Analysis: The Sociolinguistic Analysis of Natural Language*. Chicago, IL: University of Chicago Press.
- Sullivan M.P. (1973). *Perceptions of symbols in foreign policy: data from the Vietnam case* (Tech. Rep.). Ann Arbor MI: Inter-University Consortium for Political Research,

- Tesch R. (1989). Computer Software and Qualitative Analysis: A Reassessment. In Blank G., McCartney J.L. and Brent E. (Eds.), *New Technology in Sociology*. New Brunswick, NJ: Transaction Publishers.
- Tesser A. (1977). Toward a theory of self-generated attitude change. In Berkowitz, L. (Ed.), *Advances in Experimental Social Psychology*(10 ed.). New York, NY: Academic Press.
- Tversky, A. and D. Kahneman. (1980). Causal Schemas in Judgments Under Uncertainty. In Fishbein, M. (Eds.), *Progrss in Social Psychology*. Hillsdales, NJ: Erl-Baum.
- VanLehn K. and Brown J.S. (1980). Planning Nets: A representation of formalizing analogies and semantic models of procedural skills. In Snow R.E., Frederico P.A. and Montague W.E. (Eds.), *Aptitude learning and instruction: Volume 2, Cognitive process analysis and problem solving*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- VanLehn K. and S. Garlick. (1987). Cirrus: An Automated Protocol Analysis Tool. Langley P. (Eds.), *Proceedings of the Fourth International Workshop on Machine Learning*. Los Altos, CA, Morgan-Kaufman.
- vanMeter K.M. and Mounier L. (1989). East Meets West: Official Biographies of Members of the Central Committee of the Communist Party of the Soviet Union Between 1981 and 1987, Analyzed with Western Social Network Analysis Methods. *Connections*, 12, 32-38.
- Weber R.P. (1984). Content Analytic Cultural Indicators. In Melischek, G., Rosengren, K. and J. Stappers (Eds.), *Cultural Indicators: An International Symposium*. Vienna: Austrian Academy of Science.
- Weber R. (1984). Computer-Aided Content Analysis: A short primer. *Qualitative Sociology*, 7, 126-147.
- Weber, R. (1985). *Basic Content Analysis*. Beverly Hills, CA: Sage.
- Wierzbicka A. (1972). Emotions. In Wierzbicka, A. (Eds.), *Semantic Primitives*. Frankfurt, Germany: Atheneum Verlag.
- Wierzbicka A. (1973). The semantic structure of words for emotions. In Jakobson, R., Van Schooneveld, C.H., and D.S. Worth (Eds.), *Slavic Poetics: Essays in honor of Kiril Taranovsky*. The Hague: Mouton.
- Wilensky, R. (1981). PAM. In Schank R.C. and C.K. Riesbeck (Eds.), *Inside Computer Understanding*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Woodrum E. (1984). 'Mainstreaming' Content Analysis in Social Science: Methodological Advantages, Obstacles, and Solutions. *Social Science Research*, 13, 1-19.

Wyer R.S. Jr., and D.E. Carlston. (1979). *Social Cognition, Inference, and Attribution*. Hillsdale, NJ: Lawrence Erlbaum Associates.