

Text-based Analysis: a brief introduction

by Duane Truex, Assistant Professor of CIS
College of Business Administration
Georgia State University

I. Forms of Textual Analysis: from traditional to emergent language analysis

One view of textual analysis allows that the text may stand alone and that the words or idiomatic expressions themselves speak its meaning. In general Content Analysis assumes this position. However other forms of textual analysis include the wider context in which the text was constructed including, but not limited to, understanding of the author(s) and their environment and of the sources of power and domination in that environment. Still other forms allow for the unfolding awareness and learning of the text's reader/interpreter to be considered in the process of uncovering the meaning(s) of the text. This panel presentation focuses on three forms of textual analysis: Content Analysis, Conversation Analysis and Discourse Analysis.

In general Content and Conversation Analysis search for structures or patterned regularities within the text. In the case of *Content Analysis* the examination of these regularities is thought to allow the researcher to draw inferences from the text. Formal definitions of Content Analysis vary, but the general assumption is that intention and meaning are discoverable in the frequency with which words, phrases, idioms or ideas occur in a text and the meaning can be captured in a set of predefined content variables. Meanings are assumed to be inherent in the word or idiom. Those meanings are defined in a limited set of definitions in a dictionary or concordance of meanings and are, therefore, relatively fixed. Classes of meanings are assignable to a predefined 'content variable' and it is the frequency counts of word/idiom meanings assigned to the content variables that are studied and analyzed statistically to look for patterns of meanings. A formal definition of a content variable is provided by Krippendorff when he says, "... a variable that partitions a set of all recording units into mutually exclusive classes." (p.88) The recording unit is: "the specific segment of content that is characterized by placing it in a given category." (p. 58) Furthermore, in traditional content analysis the unit of analysis is generally limited to a rather small unit of text. Clearly it is a structural tool, typically used to search for patterns and structures in the data at the word or phrase level.

Conversation Analysis evolved as a means to study verbally conducted interactions the meanings of which are shaped by a variety of contextual factors. Conversation Analysis, in contrast to content analysis, does not presume the existence of fixed meanings in words and idioms. It presumes that meanings are embedded in layers of contexts, negotiated interpretations and lifeworld knowledge. It is grounded on the assumptions of philosophical hermeneutics that meanings are understood through repetitive readings and interpretations of a text. Meanings are built up through iterative, continuing conversations in which actors have built up layers of shared interpretations. If a researcher (or any third party) wishes to recover the meanings communicated through the many layers of interaction, s/he needs to reverse this process. This requires the researcher to immerse himself in the language of the organization and in the work context and work in which that language was created.

Discourse Analysis builds on the elementary ideas of both content and conversation analysis. It retains the notion of a content variable based on Krippendorff's definition, specifically the idea that a content variable partitions the set of all recording units into mutually exclusive classes. However, Discourse Analysis differs from Content Analysis in the way it defines the content variables: 1) the content variables are not predetermined and fixed; they evolve during the interpretation of the collected records and are adjusted during the analysis; 2) the unit of analysis varies in length from single utterances to whole conversations possibly spanning more than one meeting. In Discourse Analysis content variables are taken as a starting point. Discourse analysis differs from conversation analysis in the assumption that patterns of meaning may be traceable, can be assigned to an open set of content variables and thus serve as exemplars in further analysis from different organizational settings. Discourse analysis helps the researcher by encouraging him or her to define an emergent set of content variables. But like Conversation Analysis multiple readings and analysis of the records allow the cross-checking of interpretations through iterative "hermeneutic circles", thus improving analytic rigor and inter-coder reliability.

Methods which support formal text-based analysis of organizational events provide means by which we may empirically filter and measure theoretical constructs which have not been amenable to formal testing. As such these approaches provide a important development in examining the efficacy and truth claims made by broad theories of society now being applied to social organization such as critical social theory and structuration theory. Such methods also provide means of examination and comparisons of frames of references such as language games, ideal speech acts, and organizational emergence. However, these text-based methods are themselves imbued with epistemological and ontological assumptions which constrain their general applicability to certain conditions and thus require examination. Among those issues one must consider, 1) what actually qualifies as a 'text', and; 2) what forms of filtering of these texts are fair, and meaningful.

II. The Organization as a Text

Methods which support formal text-based analysis of organizational events provide means by which we may empirically filter and measure theoretical constructs which have not been amenable to formal testing. As such these approaches provide a important development in examining the efficacy and truth claims made by broad theories of society now being applied to social organization such as critical social theory and structuration theory. Such methods also provide means of examination and comparisons of frames of references such as language games, ideal speech acts, and organizational emergence. However, these text-based methods are themselves imbued with epistemological and ontological assumptions which constrain their general applicability to certain conditions and thus require examination. Among those issues one must consider, 1) what actually qualifies as a 'text', and; 2) what forms of filtering of these texts are fair, and meaningful.

The notion of a text may also be extended to mean any shared social behavior which may be interpreted to have multiple levels of meaning. The concept of a 'text' may be commonly understood to be a written, and possibly formal, document expressing some aspect from a shared social realm such as a directive from a superior, a memorandum of

the results of a shared decision, meeting minutes, operating procedures or a report of on-going research and so on. Common characteristics of such texts are: a clear notion of authorship and document purpose; a sense that the text may be viewed by readers other than the author and a prose which was constructed according to some conventions of style, form and grammar. Discourse, on the other hand, refers to the free-flowing exchange of meanings in face-to-face communication. It may take sudden shifts, leaving ideas incomplete to be rejoined or to disappear in the process of creation. It has a life of its own which is given shape as others join in the enactment of the discourse. A discourse may be inconclusive, it is generally informal and semi-structured, it may have multiple levels of meaning and never follows the rules of conventional written grammatical structures. It is a shared medium. It is meaning creation in process.

Much of what constitutes organizational life resembles a discourse. Organizational life takes place in language. It is the process of meaning creation and of meaning sharing. Meaning generation is work in process. Thus, as a continuous process of meaning creation and enactment the interactions which constitute organizational life may be seen as a form of organizational 'text' which may also be subject to forms of textual analysis.

III. Choosing an appropriate means of textual (a.k.a meaning) analysis

If one believes in the relative clarity, simplicity and stability of a given text then methods which are oriented towards extracting commonly understood meanings of words, idioms and events may be sufficient. However one may often find that 1) conversational textual records are too rich to be analyzed by traditional content analysis, and; 2) each organizational text is ideographic, individual and therefore difficult to compare with other organizational texts. In the literature it is quite common to compare various textual records. But such comparisons are not supported by any systematic guidelines of record collection and interpretation. Rather they are left to an informal discussion which solely draws on the ingenuity and experience of the author who is working backward from the data. There are few systematic approaches which allow one to compare results of one project with another apart from the creative and integrative abilities of the action researcher.

=====

Klein, Heinz and Truex, Duane P. "Discourse Analysis: An approach to the Investigation of Organizational Emergence", in *Signs of Work Semiosis and Information Processing in Organizations*, Holmqvist et al editors, Walter de Gruyter: Berlin, 1996, pp. 228-268.

Krippendorff, K. (1980) *Content Analysis: an introduction to its methodology*, Beverly Hill, CA: Sage Publications.

=====