

January - 2001

## ***Addressing some Common Problems in Transcript Analysis***

**Patrick J. Fahy**

Athabasca University – Canada's Open University

### **Background**

Computer conferencing is one of the more useful parts of computer-mediated communications (CMC), and is virtually ubiquitous in distance education. The temptation to analyze the resulting interaction has resulted in only partial success, however (Henri, 1992; Kanuka and Anderson, 1998; Rourke, Anderson, Garrison and Archer, 1999; Fahy, Crawford, Ally, Cookson, Keller and Prosser, 2000). Some suggest the problem is made more complex by failings of both technique and, more seriously, theory capable of guiding transcript analysis research (Gunawardena, Lowe and Anderson, 1997).

We have previously described development and pilot-testing of an instrument and a process for transcript analysis, call the the TAT (Transcript Analysis Tool), based on a model originally developed by Zhu (1996). We found that the instrument and coding procedures used provided acceptable – sometimes excellent – levels of interrater reliability (varying from 70 percent to 94 percent in pilot applications, depending upon user training and practice with the instrument), and that results of pilots indicated the TAT discriminated well among the various types of statements found in online conferences (Fahy, et al., 2000).

That work clarified some of the conceptual and theoretical problems with transcript analysis. Initial development concluded with the realization that these issues needed resolution, or at least serious exploration, as our testing of the TAT proceeded.

This paper is a brief commentary on some of the issues which we and others have encountered in transcript analysis, with a description of steps we are taking, or exploring, to address them. It is intended to contribute to what Garrison (2000) referred to recently here as a theoretical framework, "a broad paradigmatic set of assumptions that provides the elements of a theory, but without the detail and completeness (nuances) of a comprehensive theory (p. 3)." Included here are two previously reported problem areas: how the obvious orthographic and syntactic features of transcripts might serve as useful indicators of underlying interaction patterns, and the progression of topics in conferences as revealed by structural indicators.

### ***Common Problems***

Discriminant capability and reliability among users have been major problems in previous transcript analysis work. Discriminant capability means a coding instrument readily and

unambiguously permits placing of conference content into discrete and useful categories. It has not proven to be easy to achieve, as demonstrated by the prevalence of admitted coding problems in transcript research. Gunawardena et al. (1997), for example, found problems using Henri's (1992) model to distinguish between cognitive and metacognitive activities in conferences. They concluded that a large number of units could have been coded as either (p. 404). As a result of this experience, Gunawardena et al. developed their own analytic tool, but concluded that it was a poor discriminator: over 90 percent of transcript postings fell into a single category (p. 425). In another study, Kanuka and Anderson (1998) reported two problems attributable to weak discriminant capability of their instrument: an "overwhelming" number of messages were coded into one category, and messages could often be coded into more than one category (p. 65). Zhu (1996) also had acknowledged that her classification system permitted postings to fit into several categories (p. 837).

Reliability is directly affected by lack of discriminant capability: if categories are not clear, discrepancies in coding will occur. In fact, reliability is often either low or not simply mentioned at all in published reports of transcript analysis research, and to improve reliability researchers often resort to convenient but inefficient and expensive strategies such as collaborative coding (Kanuka and Anderson, 1997; Rourke et al., 1999). Such strategies may meet the need for consensus in an specific research context, but they do not argue for the reliability of the coding instrument.

Problems with discriminant capability may be attributed to two causes: complexity of the instrument (both too many categories or codes, and lack of mutual exclusiveness among them), and use of an inappropriate unit of analysis (anything other than the sentence).

Complexity is directly related to the number of codes available. Some coding tools simply contain too many categories, forcing users to make many excessively fine discriminations. Gunawardena et al.'s (1997) model included over twenty categories grouped into five "phases"; Cookson and Chang (1995) employed four main groups of criteria, with each further subdivided into four more categories; Higgins' (1998) model used as many as twenty; Rourke's (1999) model has twelve indicators, in three groups; and Zhu (1996) used 8 categories. Obviously, with more categories there is more likelihood of ambiguity, definitions of differences among categories must be made unambiguously clear, and there is more need for training and practice for potential users.

Some researchers have rejected the sentence as the unit of analysis, and then faced uncertainty about what to code. Henri (1992), in her early and influential work, argued for "units of meaning . . . rather than messages proper" (p. 126). She argued that "CMC messages harbour more than one unit of meaning," and that each analytic purpose could and should "define its own relevant unit of meaning" (p. 134.) Objecting to the "mechanistic" in Henri's approach, Gunawardena et al. (1997) rejected Henri's methods and attempted to code whole messages in a single category or "phase," with the poor results already mentioned. Rourke et al. (1999) judged sentences and paragraphs were "artificial and arbitrary" (p. 60), and instead used a 12-point, 3-group analytic system (which in fact does appear actually to have been applied to the coding of sentences).

While focus is on the meaning of the interaction of the conference, the unit of analysis must be something obvious and constant within transcripts. In our work with the TAT we have concluded that this is the sentence (or, in the case of highly elaborated sentences, independent clauses which, punctuated differently, could be sentences). Sentences are, after all, what conference participants produce to convey their ideas, and are what transcripts consist of.

We find support for our position on the importance of the sentence for analysis of transcripts in two concepts from linguistic analysis of electronic communications: the macrosegment (Herring, 1996) and the discourse topic (Witte, 1983). Macrosegments are trans-sentence components of texts which consist of both notional and surface components. Notional coherence in texts is achieved by the writer through choices made about words, sentences and paragraphs, including the structure of the writing conveyed by orthographic and syntactic features. Notional coherence is contained in, but transcends, the merely orthographic, syntactic and structural features of the transcript. Put another way, the macrosegments containing a conference's ideas and themes are not bounded by the limits of the text, including sentences and paragraphs, though they are constructed from them.

While notional meaning transcends textual structures, those structures should not be ignored. Skillful writers provide structural and syntactic clues to help readers accurately get the point. Structural elements of text help form and convey the notional relationships of the argument, and good readers are therefore alert to them.

Obviously, the transaction between a writer and a reader is complex, involving both parties. The concept of the discourse topic recognizes the mutual participation of both parties in textual communication. The discourse topic is "what the writer intends to communicate" (Witte, 1983) – what the writer hopes the reader will understand. What makes discourse topics problematic – in fact, what makes seeking "meaning units" of any kind a perilous, even impossible, task – is their inherent subjectivity: the "meaning" of any "meaning unit" depends largely on what the reader brings to it. Since a writer's meaning does not reside in the simple textual elements alone, discourse topics must consequently be inferred or constructed by the reader; meaning reflects the interaction of the reader's knowledge and experience with the text of the message. Thus, regardless of what writers intend, what readers understand is based on the interaction between the message and the readers' experiences, knowledge, and capability for understanding the topic (Witte, 1983).

An element of threaded transcripts potentially adds to our understanding of the progression of discourse topic as it emerges and evolves in a conference: the threading structure itself. Using the date- and time-stamps of postings, the progression of the discussion can be reconstructed in detail. Our ongoing research is exploring how these structural features might be used, perhaps in combination with the TAT analysis of message content, to illuminate important internal patterns of interaction in the transcript.

### ***Transcript analysis research with the TAT***

The theoretical assumptions provided by the framework above have resulted in the following strategic decisions in our work in progress with CMC conference transcripts:

- The sentence is the unit of analysis
- The TAT is the method of analysis
- Interaction is the criterion for judging conference success
- Topical progression (types and patterns) is the focus of analysis

The TAT classifies transcript content (usually sentences) into one of four categories: 1) questioning, 2) statements, 3) reflections, and 4) interpersonal coaching and scaffolding.

- **Questioning:** includes vertical questions (there is a "correct" answer somewhere), and horizontal questions (there is no one right answer; all input welcome.)
- **Statements:** contain no self-revelation, transmits information impersonally, and usually do not especially invite dialogue.
- **Reflections:** the speaker displays trust by revealing usually guarded material (values, beliefs, doubts, reasoning processes, experiences; both what he or she thinks, and why).
- **Coaching and scaffolding:** intended to encourage, support, model, provide hints and help, and generally support others in difficulties, new or unfamiliar experiences, crises, or moments of doubt, insecurity or high emotion.

Using the chronological information provided in the transcripts, we focus in our analysis on the types and patterns of progression of the topics under discussion. Several features of each posting are coded (these are steadily evolving and changing):

- **Level:** initiating or following another posting. If following, we note how far into the progression of the discussion or topic the posting occurs.
- **Progression type:** parallel or sequential. Parallel posts are those made to the same initiating posting or topic; sequential postings provide extension or depth to the discussion or the topic. For example:

#### **Parallel progression**

Comment A

Comment B

Comment C

Etc.

#### **Sequential progression**

Comment A

Comment B

Comment C

- **Response rate:** high or low. ("Unusual" exchanges or single postings are particularly of interest. High response generators, resulting in unusually large numbers of subsequent

parallel or sequential postings, are called "response triggers." Postings which receive no responses are also studied.)

- **Interaction patterns:** how individual members of the conference perform and interact. (This analysis exposes the leaders and followers, those who initiate or "trigger" interaction, those who follow, and those who have strong response-type preferences. Patterns are being investigated in relation to variables such as gender, personal communication patterns, and other features of observed conferencing behaviour.

While this work is ongoing, some promising, intriguing or confirming results have already been noted in a pilot application of this approach with a transcript of 71 postings, consisting of 8283 words, created by 14 graduate student participants (seven women, six men). Importantly, the TAT succeeded in discriminating among the types of possible responses, detecting the following distribution of types: questions, 7 percent; statements, 45 percent; reflections, 28 percent; scaffolding and coaching, 21 percent.

Compared with results reported by Herring (1996) in her analysis of electronic communications, we found the following in our pilot analysis:

- Women were more assertive in the test sample, and men less so: women made 53 percent of the initiating posts, and accounted for 55 percent of the words posted, while equal numbers of men and women never made initiating postings.
- Men, however, were more overtly assertive: All the postings expressing outright disagreement (6) were made by men.
- Finally, there was little support for a key finding of Herring's work: students in the sample made more statements than reflections. Herring's students were more interested in exchanging views than information.

## Conclusion

We expect to report shortly on an analysis of a much larger amount of transcript material. We are also assessing other instruments and techniques; Rourke, Anderson, Garrison and Archer's (1999) tool for assessing social presence, for example, has some promising features we intend to explore further.

## References

- Cookson, P. S. and Chang, Y. (1995). The multidimensional audioconferencing classification system (MACS). *The American Journal of Distance Education* 9(3), 18 – 36.
- Fahy, P. J., Crawford, G., Ally, M., Cookson, P., Keller, V. and Prosser, F. (2000). The development and testing of a tool for analysis of computer-mediated conferencing transcripts. *The Alberta Journal of Educational Research* 66(1), 85 – 88.
- Garrison, Randy. (2000). Theoretical Challenges for Distance Education in the Twenty-First Century: A shift from structural to transactional issues. *International Review of Research in Open And Distance Learning* 1(1), Retrieved September 29, 2000 from: <http://www.irrodl.org/index.php/irrodl/article/view/2/333>

- Gunawardena, C., Lowe, C., and Anderson, T. (1997). Analysis of a global online debate and the development of an interaction analysis model for examining social construction of knowledge in computer conferencing. *Journal of Educational Computing Research* 17(4), 395 – 429.
- Henri, F. (1992). Computer conferencing and content analysis. In A. Kaye (Ed.) *Collaborative Learning Through Computer Conferencing: The Najaden papers* (p. 117-136). Berlin: Springer-Verlag.
- Herring, S. (1996). Two variants of an electronic message schema. In Susan C. Herring (Ed.) *Computer-Mediated Communication*. Philadelphia, PA.: John Benjamins, 81 – 106.
- Higgins, R. (1998). *Computer-Mediated Cooperative Learning: Synchronous and asynchronous communication between students learning nursing diagnosis*. Retrieved November 5, 2000, from: <http://www.cybercorp.net/rhiggins/thesis/higtoc.html>
- Kanuka, H., and Anderson, T. (1998). Online social interchange, discord, and knowledge construction. *Journal of Distance Education* 13(1), Spring, 57 – 74.
- Rourke, L., Anderson, T., Garrison, R., and Archer, W. (1999). Assessing social presence in asynchronous text-based computer conferencing. *Journal of Distance Education* 14(2), 50 – 71.
- Witte, S. (1983). Topic structure and revision: an exploratory study. *College Composition and Communications* 34(3), 313 – 341.
- Zhu, E. (1996). Meaning negotiation, knowledge construction, and mentoring in a distance learning course. In Proceedings of Selected Research and Development Presentations at the 1996 National Convention of the Association for Educational Communications and Technology at Indianapolis, IN. ERIC Doc Number: ED397849.

