Abstract

CMC has long been recognized as blurring the distinction between written and oral modes of communication. Analysis of genre has traditionally focused on either the structure of the text or an ethnographic description. This study suggests a new approach for investigating organizational genre by combining ethnographic methodology and corpus analysis techniques to examine a previously unexplored genre: the technical trouble ticket. This yields a richer analysis and illuminates not only what happens in trouble ticket discourse, but also how and why. Particular attention is paid to the relative orality of this homely CMC genre.

1. Introduction

To consider as potential genres such homely discourse as the letter of recommendation, the user manual, the progress report, the ransom note, the lecture, and the white paper … is not to trivialize the study of genres; it is to take seriously the … situations in which we find ourselves. ([10], p. 155)

While there has been much investigation into communicative genres in the fields of Rhetoric [10], Organizational Behavior [13], Computer Mediated Communication or CMC [7] and Applied Linguistics [12], an extensive literature review revealed no published research on trouble ticket discourse. Perhaps the closest relation is Ackerman and Halverson’s [1] work on organizational memory and a telephone hotline call.

The call center environment in which trouble tickets are written is one of the fastest growing segments in today’s workplace. Datamonitor reports that in 1998 three percent of the entire US workforce was employed in call centers [3]. Although growth was flat in the last two years, double digit expansion is now expected through 2006 [6].

In general CMC discourse does not follow the traditional patterns of either spoken or written discourse. One of the key lines of study in CMC research has been investigation of this blurring between spoken and written modes of communication [4], [7]. This study of trouble tickets genre sheds light on this question with three key observations:

1) Trouble ticket authors, under pressure to write quickly and professionally, tend to favor speed over form.

2) Like many forms of CMC, trouble tickets exhibit an economy of style which results in fragmentary texts filled with acronyms, abbreviations and symbols.

3) Trouble tickets strongly exhibit many characteristics of oral communication.

This paper suggests and illustrates a bimodal approach which fuses the rich qualitative nature of ethnographic investigation with the precise quantitative measurement of corpus analysis, enabling illumination of questions of what, how, and why.

Directions for future study include the evolution of the trouble ticket genre over time, the affect of gender on communicative style, readers’ reactions to trouble tickets, and trouble tickets as organizational memory.

2. Ethnographic Analysis

This study focuses on written discourse in trouble tickets generated in a technical call center (also called a Network Operations Center or NOC) in a suburban area of the northeastern United States. The ethnographic portion of this study is based on eighteen months the author spent working in the NOC (first as a senior technician, then as a shift supervisor). The NOC has operated as a Managed Service Provider since the late 1980s, currently boasting more than 150 customers, and 30,000 managed sites world-wide. Its primary function is to manage Wide Area Network communication lines and equipment on behalf of its customers. Technicians are equipped
with a graphical display of networks and remote access to network equipment.

When a problem occurs (e.g. communication to one or more locations is lost) the technician works to resolve the issue while simultaneously keeping the customer up to date on the situation. Troubleshooting involves interfacing with telecommunication carriers (telco), field support representatives and customers as well performing hardware/software diagnostics and repair. Trouble tickets are used to track progress and record all actions and interactions.

3.1 Trouble tickets

Johnson [9] describes trouble tickets in RFC 1297. Tickets are digital, stored in a trouble ticket database. This database is fronted by Clarify®, a customer relationship management software package. Trouble tickets vary in length from 10 to 10,000 words. Tickets are typically opened and updated by automated tools and/or NOC personnel.

Trouble tickets have a standard format, beginning with identifying information (customer name, site affected, status, severity, primary technician, customer contact and an initial assessment of the problem). This is followed by a series of individual entries, usually made by different people at different times. Each entry is marked with a header that includes a time/date stamp and uniquely identifies the author. All headers and some entries are generated by the system in response to information from remote devices (See Figure I).

3.1.1 Purpose

The purpose of a trouble ticket is, by definition, to track an unusual situation or problem. NOC policy, related to every potential NOC employee during the interview and repeatedly thereafter, requires technicians to log into the ticket everything related to the issue (e.g. phone calls and trouble-shooting steps).

Tickets are the NOC equivalent of a hospital chart [9]; they keep all interested parties on the same page. Closed tickets are used to generate monthly per-formance reports, to review billing issues and to perform post-hoc evaluations when customers complain about the service they have received. Closed tickets are also systematically queried each time a new ticket is opened to ascertain if a given site is chronically problematic. While management recognizes the potential benefits of turning trouble tickets into a knowledge warehouse for troubleshooting, this has not yet become a reality.

3.1.2 Audience

There are several audiences for NOC trouble tickets, the customer being the primary focus of both technicians and management. Technicians are frequently admonished to keep their ticket discourse professional by avoiding non-standard grammar and abbreviations. While all customers have read/write access to their tickets via a Web portal, use of the portal varies widely from those who monitor tickets...
in real time to those never access it. Management, another important audience in the mind of technicians, typically reads tickets only when an issue is escalated. Even the ticket is may only be reviewed briefly, as the primary technician typically explains the issue. The primary users of trouble tickets are the technicians themselves.

3.2 Study Environment

All NOC technicians work together in a single room, with most of the workspaces around the perimeter. Senior technicians sit in the middle of the room and offer support wherever it is needed. At full capacity, the NOC can seat 35 technicians. The x’s in Figure ii show the seating arrangement. Each workspace is approximately three feet wide and contains four monitors, stacked two high. A keyboard, two telephones, and hardcopy network binders (each describing the network of a single customer) fill each workstation, leaving little space for technical manuals or personal items. There are no barriers between adjacent workplaces. On off-shifts, when there are fewer technicians, a single person may be responsible for the networks displayed in multiple workspaces.

During peak hours (12PM – 4PM), there are typically 30 or more people working in the NOC. Space is tight, the noise level is high, and the pace is hectic. It is not unusual for a technician to simultaneously type in a ticket, balance two telephone conversations (one with a handset and the other with a headset), and scribble a note to pass on to a technician on the other side of the room. Ambitious technicians may eat lunch at the same time. Although the number of technicians and the accompanying noise and space discomforts decrease on other shifts, the pace and stress level remains high as fewer people manage the same number of networks. Approximately 15,000 tickets are created in the NOC every month.

3.3 Genre Repertoire

While this study focuses on trouble tickets alone, there is a rich genre repertoire at work in the NOC. As mentioned above, there are network books for each customer containing both technical and customer service-related information (e.g. including main contacts, hours of operation, business functions of the network, network diagrams, equipment serial numbers, circuit identification numbers, protocols used). All this information is also available in softcopy. Technical manuals are also available in both hard and softcopy format. Current and historical software images, configuration and memory files and software are also accessible. Additionally, Web links to various vendors’ technical support portals and various training materials are commonly shared between technicians. Handwritten notes, shift turnover forms, and emails are a few other genres that are employed in the day-to-day life of a NOC technician.

3.4 Authors

Although customers have the potential to make entries in trouble tickets, this happens infrequently (there was only one customer generated entry in the sample described in section 4). NOC technicians are the primary authors of trouble tickets. Approximately 70 technicians are spread over five shifts. Although call centers have high turnover ratios, at the time of the study, more than half of the technicians had been with the NOC for a year or more and over ten percent had worked there more than five years. Technicians ranged in age from 21 to 60. Most have high school education and on-the-job experience. Many take technical certification or degree courses in the evenings, but there are no specific certifications required to obtain a job in the NOC or for promotion to senior technician. Management places the most emphasis on hiring people with customer service skills and an aptitude for technology, on the tenet that soft skills are more difficult to learn than technical knowledge.

3.5 Training

Training in the NOC is primarily on-the-job. New employees are given one to two weeks to sit with an experienced technician and learn the job. Management expects new hires to listen and observe for one day and then to pick up the phone and begin interfacing with customers (with their mentor silently listening in) by the end of their second day. No formal training is provided beyond the mandatory Human Resources review of benefits and company policies. There are no manuals describing appropriate ticket
entries and no prototypical samples are provided to new employees.

3.6 Problems

Tickets are often fragmented and difficult to process for customers, management and even experienced technicians. They are full of acronyms, abbreviations and symbols, as well as non-standard punctuation and grammar. They typically contain only bare details (two hours of troubleshooting may be condensed into 25 words) and are frequently repetitious with much documentation that a telco was called or a customer updated. The cryptic nature of tickets leads to problems including customer complaints, time lost troubleshooting, and field technicians arriving on site with the wrong equipment.

3.7 Policy Enforcement

Key policies for trouble ticket writing are (a) Keep tickets professional and (b) Log every action into a ticket. Neither of these policies is enforced on a regular basis. Trouble ticket discourse is not a part of performance reviews, nor the subject of (infrequent) team meetings. Technicians are reprimanded for deviation from these policies only when a ticket is escalated to management and the non-standard or incomplete nature of a ticket causes a problem.

While neither policy is regularly monitored, the second is taken more seriously than the first. Technicians who write poorly may still be promoted if their technical and customer skills are good. In fact, senior technicians often delegate ticket updates to junior technicians. In contrast, technicians who fail to log their actions into tickets may be terminated.

Another reason technicians log their actions in tickets is for self-protection. Though very abbreviated accounts are recorded, technicians want to be able to prove what actions they take. Although the vast majority of tickets go unread by management, no one can know in advance which issues will be escalated.

One indication of technicians’ desire to protect themselves is found in the increase in word counts per entry as tickets increase in number of entries. The likelihood that a ticket will draw management’s attention increases with the number of entries: more entries are indicative of more problems that had to be addressed. In the ticket sample for this study (see section 4) tickets with fewer than ten entries averaged 8.3 words per entry, while tickets with ten or more entries averaged 19.8 words per entry.

3.8 Summary

In summary, the trouble tickets examined in this study originate in a large and well-established NOC. Trouble tickets act like hospital charts, keeping all interested parties on the same page as problematic or unusual situations are investigated. While customers are the nominal target audience for trouble tickets, the technicians who are the primary authors appear to focus more directly on NOC management as they write and technicians are the principle users.

The physical environment of the NOC is crowded and loud. Technicians work under pressure, performing multiple tasks simultaneously and using many different resources with only a minimum of training. This results in their composing fragmented text that is often hard to understand. While NOC management has created policies to enhance ticket readability, these policies are not regularly enforced.

4. Trouble Ticket Features

This section and the next describe analyses of trouble tickets based on a random sample of 30 tickets selected from all tickets created during the month of February 2003. Since the purpose of this study is primarily to examine user-entered text, tickets that contained only computer generated text were discarded and replaced by another randomly selected ticket. Nine tickets were discarded. The sample was analyzed for significant features in overall ticket construction and was then compared on the grammatical level with four other genres. Five software programs were used in the analysis: a Part-of-Speech tagger (CLAWS®), Concordancer®, MS Word® and Excel®. With regard to ticket construction, the blend of automated and user-entered text is examined first, then the use of punctuation markers is investigated.

4.1 Automated vs. User-Entered Text

Nearly half of the trouble ticket text in the sample is automated. Automated entries are defined here as entries that are placed into a ticket by the system, independent of any user action. Any text which is purposefully entered by a user is considered a user-entry. This includes three forms of machine generated text: (a) information captured in equipment logs and copied into tickets, (b) form headings and (c) boilerplate text. There are a total of 392 machine generated tokens in the sample and an overall total of 3,340 user-generated tokens. The total count of automated tokens in the sample is 2,443. Tokens are defined here as a lexical item such as a word, number or symbol. Interestingly, though the environment and
subject matter is quite technical, only 173, or five percent, of the total tokens are words that appear in Newton’s Telecom Dictionary [11].

Automated entries may appear in two ways: as headers at the beginning of user text and as system generated entries. Systemic entries are currently used only for a subset of NOC customers. For customers with this service, the system enters status changes whenever connectivity is lost or restored. While the automated text of neither header nor standalone entry is a focal part of this study in itself, it is a feature of the text and its presence is part of the overall experience in using this genre.

Of the thirty tickets selected for this study, only 20 had the potential for systemic entries. Of these, the mean scores were 1.9 for systemic entries and 3.3 for user entries. The remaining ten tickets (opened for networks without the capacity for systemic entries) shared 118 entries, but the distribution was skewed by a single ticket with 67 user entries. There were two tickets in this category with only two entries. The median number of entries for the tickets without systemic entries was four (see Table A).

### 4.2 Punctuation

Punctuation in trouble ticket discourse does not follow traditional proscribed rules and is therefore of interest to this study.

Since sentence boundaries are often ambiguous in trouble tickets, this study focuses on t-unit final punctuation for this portion of the analysis. A t-unit is the main clause of a sentence plus all subordinate clauses and nonclausal structures attached to or embedded in it [8]. Example 1 highlights the nonstandard use of punctuation in tickets with only one main clause but three ellipses. All examples maintain their original format, grammar and punctuation. To ensure privacy, pseudonyms are used for all personal and company names.

**Example 1:**

Naomi .. aa michaels ..
requesting status ...
Example 3:
Tim says the backup was working even though we found that the modem on site does not have the RESTORAL Config option...?... Tim does not believe this unit was swapped out...

The most obvious motivation for omissions and ellipses is the desire of trouble ticket writers to minimize their typing efforts. Herring [7] points out that only a relatively small percentage of the non-standard forms in CMC are caused by a lack of attention or knowledge of the standard form. Instead, she suggests most are deliberate choices made by users in order to minimize the typing required, to simulate features of spoken language, or to express themselves creatively.

While there is little room in trouble tickets for creative expression, strategies to minimize effort are evident throughout trouble tickets. For example, the omission of punctuation is more common than the ellipsis mark, even though the ellipsis enhances the legibility of the text by breaking the text into smaller units. The non-standard use of punctuation is one indication that time pressure outweighs the stated policy that technicians should write professionally.

4.3 Summary

General characteristics of the trouble tickets in this study were discussed in this section. Nearly half of the total text in the sample was created by a computer system rather than an individual. In the user-generated text, the most interesting non-grammatical feature is non-standard punctuation which is quite common in spite of NOC policy to keep tickets professional.

5. Grammatical Analysis

The main purpose of a trouble ticket, as discussed in section 3.1.1, is to provide information on a problem. Biber [2] examines differences between genres of spoken and written communication on six dimensions, each of which consisted of several grammatical features. His foundational dimension, “Involved versus Informational Production”, marks high informational content as well as “the extent to which (texts) are oral or literate in terms of their production characteristics and primary communicative purposes” (p. 108). Collot [4] follows Biber with a study of Bulletin Board System (BBS) discourse using the methodology. Chafe and Danielewicz [3], employ a different methodology to find several grammatical features which indicate orality or the “rapidity and evanescence of speaking as opposed to the deliberateness and editability of writing” (p. 105). Five features appear in both Biber and Chafe and Danielewicz’ studies: type/token ratios, and the frequency of nouns, contractions, prepositional phrases and hedges. These five features are analyzed in trouble ticket discourse using corpus analysis methodology.

Technicians who write in trouble tickets are under pressure to write quickly, but are also often instructed to make their tickets both informational and professional. These two opposing pressures will be considered as the analysis proceeds. Quantitative results from trouble ticket discourse will be normalized to count per 1,000 words and compared with three of the genres described in Biber’s [2] study (academic papers, personal letters, and face to face conversations), and with Collot’s [4] data on BBS discourse. Due to differences in methodology, Chafe and Danielewicz’s [3] results are not directly comparable to Biber and Collot’s and are not included here. Additionally, only user-generated text is included in the grammatical analysis.

Results on each of the five features are presented side by side in numerical format in Table C. Graphical displays of each dimension are provided as each dimension is discussed. For the sake of continuity the top of each scale indicates Informational content and the bottom indicates Involved production. High instances of contractions and lexical hedges indicate involvement in production, so the high values for these features are found at the bottom of the scale rather than at the top as they are for the other three features.

<table>
<thead>
<tr>
<th>Table C: Summary of Grammatical Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type/Token</td>
</tr>
<tr>
<td>Conversations</td>
</tr>
<tr>
<td>Letters</td>
</tr>
<tr>
<td>Academic Papers</td>
</tr>
<tr>
<td>BBS Conferences</td>
</tr>
<tr>
<td>Trouble Tickets</td>
</tr>
</tbody>
</table>
5.1 Type/Token Ratio

Chafe and Danielewicz [3] suggest that speakers don’t have time to “sift through all the (lexical) choices they might make, and may typically settle on the first words that occur to them.” The result, they continue, “is that the vocabulary of spoken language is more limited in variety” (p. 88). Biber [2] expands on this saying that a high type/token ratio marks high density of information and also “very precise lexical choice resulting in an exact presentation of informational content” (p. 104). The trouble ticket text in this sample has a very low type/token ratio, even lower than conversation. This is in stark contrast to the BBS discourse which has the highest ratio of the genres contrasted here. Lexical items like circuit (and the abbreviation ckt) have a high frequency in trouble tickets, but the low ratio also indicates that the informational content is low and may imply that technicians are willing to sacrifice density of information in order to enter their text rapidly (see Figure III).

5.2 Nouns

According to Biber [2], “nouns are the primary bearers of referential meaning in a text, and a high frequency of nouns thus indicates great density of information” (p. 104). With relatively few nouns, trouble tickets again appear at the involved end of the spectrum (see Figure IV). The low frequency of nouns may be partially due the fragmentary nature of trouble ticket discourse. Most meaning units (many cannot be called sentences in the standard sense) are short and exhibit deletion of one or more forms. Many of these are subject deletions:

Example 4:

[The circuit] cleared while testing

Though often the subject position is filled (or would be filled were it not deleted) by a pronoun:

Example 5:
[I] updated Phil....

Perhaps the most interesting use of nouns in the sample is a proper noun simultaneously acting as an object of one clause and the subject of another:

Example 6:
Talked with George is happy now

5.3 Prepositions

Chafe and Danielewicz [3] find that, in spite of the relative ease in producing them, prepositional phrases occurred nearly twice as often in academic writing as in conversation. They suggest that prepositional phrases are used as a strategy to increase the size of language units. Biber correlates prepositions with informational production as they “serve to integrate high amounts of information into a text” (p. 104). Although the frequency of prepositions in trouble tickets is lower than the other genres in this study, again indicating involved production (see figure V), they do serve to integrate information, often in the most economical fashion available:

Example 7:
I spoke with Joe/Site @ 555-999-9999... Telco LL being displayed on the modem....CSU
in a loop back... Telco checking line...

Each of the four prepositional phrases in example 7 contain critical information in a very concise manner. Of these four phrases, only one (8b) could have felicitously been deleted, and its deletion would have required additional punctuation (infelicitous forms are preceded by an asterisk):

Example 8:

a. * I spoke [ ] Joe/site
b. I spoke with Joe/Site[:]
   555-999-9999
c. * Telco LL being displayed
   [ ] the modem
d. * CSU [ ] a loopback

5.4 Contractions

Contractions (and other reductions) result “in a more generalized, uncertain content” ([2], p. 106) and are indicative of involved production. Given the loose usage of punctuation in trouble tickets and the time pressure under which trouble tickets are written, one might expect to see frequent examples of contractions. This is not the case here. There were only six uses of the contraction marker found in the entire sample and five of these were used to contract words into non-standard abbreviations.

Example 9:

telnet'd to the router...

One possible explanation could be the pressure to keep trouble tickets professional but given the frequency of non-standard punctuation and grammar it seems quite unusual that this text shortening technique would not be used.

It may be that pronoun, subject and verb deletion are easier for technicians to process (cf. [7]) so that when reduction is possible, total deletion of a form (or multiple forms) is preferred over contraction:

Example 10:
[1] Will monitor

Example 11:
ckt [is] up

Of the categories examined, trouble tickets are most closely aligned with BBS discourse on this feature. The relative informational tendency of BBS discourse across all five features, however, contrasts with trouble tickets’ appearance on both ends of the scale. This may indicate that there are different underlying motivations for avoiding contractions in BBS and trouble ticket discourse.

5.5 Lexical Hedges

For Biber [2] a hedge is an indication of involved production. A hedge is used when a language producer is not happy with the exact words they have chosen, prefacing them with “sort of” or “kind of”. While Chafe and Danielewicz (cf. [7]) found these common in both conversations and lectures, they found only one example in letter writing and none in academic prose. They suggested that: “the constraints inherent in the speaking and writing processes are dominate here and are not overridden by different uses of the two” (p. 89).

Collot [4], however, found this type of hedge common in her corpus of written electronic language. In contrast, no occurrences of lexical hedges appeared in the sample collected for this study, placing tickets on the extreme informational end of this feature’s spectrum (see figure VII). It may be that technicians feel it is important to find the correct word, or at least may prefer to appear confident by avoiding the use of a lexical hedge. Another possible explanation goes back to the ease of production for the writer. Lexical hedges are only be used when communicators take time to evaluate their word choice. This may not be the case in the fast paced NOC environment.

5.6 Summary

Trouble ticket discourse was examined on five dimensions which have been used in two methodologies to rank the relative informational content (literacy) and involved production (orality) of texts. On three of the dimensions (Type/Token Ratio, Nouns and Prepositions), trouble tickets appear on the ex-
treme involved end of the scale. On the remaining dimensions (Contractions and Lexical Hedges), they appear on the extreme informational end of the spectrum. Discussion of likely causes for the style of use on each dimension was informed by observations made in the ethnographic analysis.

No consistent relationship between the two types of CMC (BBS and trouble ticket discourse) was observed in this study.

6. Conclusions

This study has introduced a new, bimodal method for analyzing organizational genre. By combining the rich qualitative character of the ethnographic approach with the precise quantitative nature of corpus linguistics, this study has addressed the questions of what, how, and why with regard to a previously unexamined CMC genre: the technical trouble ticket.

While this study is based on a small sample of trouble tickets, the homogeneity of trouble ticket discourse supports its representativeness. This study indicates the following:

• Faced with conflicting pressures both to write quickly and to make their discourse informational and professional, technicians typically choose speed. As might be expected, information density becomes more important with the increased likelihood that an issue will be escalated to management.
• While parallels exist between this discourse and other forms of CMC, particularly with regard to the fragmentary nature of the text, ticket text differs significantly from BBS discourse.
• Trouble tickets weigh heavily toward involved production, or orality, on three of five features examined. On the remaining features, they appear dramatically to the informational end of the scale. Though possible explanations have been offered for this discrepancy, the need for further study is indicated.

Other directions for future study include the evolution of the trouble ticket genre over time, the effect of gender on communicative style, readers’ reactions to trouble tickets, and trouble tickets as organizational memory and boundary objects.

7. Acknowledgements

Many individuals have contributed to the evolution of thought and writing in this study. My sincere thanks to Herb Stahlke, Kris Fleckinstien, Jensen Zhao, Jim Nyce, Kevin Crowston, Bryn Marlow and two anonymous reviewers for their comments.

8. References

9. Glossary

The following is a list of acronyms, technical terms, and typographical errors that are found in the examples.

*AA Micheals*
  *Pseudonym for a customer*

*All 0's*
  Possible test result on a circuit. Indicates that there is no response from the communication device on site.

*Backup*
  A solution to restore communication, often using a regular telephone line.

*Ckt*
  Circuit. The telecommunication line that connects one site to another.

*Config*
  Configuration. Settings in the software that control how a piece of hardware responds to various situations.

*CSU*
  Channel Service Unit. Converts the digital information from an enterprise level signal to a telecommunication carrier signal (and vise versa).

*Hard-down*
  Total loss of communication to a site.

*LEC*
  Local Exchange Carrier.

*LL*
  Link Light. Indicates whether or not the circuit is active.

*Loopback*
  A test mode on the CSU that takes the signal from telco and sends it back to the sender (similar to an Internet ping).

*NI*
  Network Interface, also called a smartjack. The jack where a customer may plug his telecommunication line into the carrier’s network.

*Restoral Config Option*
  Software in a device that enables a back-up solution to be configured.

*RFO*
  Reason For Outage

*Router*
  A device which directs packets of information between two points on a network.

*Telco*
  A telecommunications carrier.

*Telco A, B, C*
  Pseudonyms for telecommunication carriers. As carriers may provide long distance, local services, or both, no distinction is made between companies that are acting as LECs and other carriers.

*Telnet*
  A protocol which governs the remote connection to a communication device. In example 9, is noun is used as a verb in the past tense.

*tom (in example 2)*
  typographical error for “to”

*TOK*
  Tests Okay.