

# 1://A first look at electronic discourse

## *On defining electronic discourse*

Electronic discourse is one form of interactive electronic communication. In this study, we reserve the term for the two-directional texts in which one person using a keyboard writes language that appears on the sender's monitor and is transmitted to the monitor of a recipient, who responds by keyboard. The recipient may actually be a single individual or a group, large or small, of receivers. Like any other way that humans use language for interaction and communicative purposes, electronic discourse is multifaceted and complex. Since the textual artifact resulting from electronic discourse is written language, both language in general and written forms of language are included in its study. Inasmuch as electronic discourse involves interaction among people, the text implies something about the variety of social interactions among its composers.

Electronic discourse is not a surrogate for language, such as whistle or drum systems, but a different context for its use. It is the interaction of that context with language that is interesting. While people have invented or evolved a series of conventions within different types of electronic discourse, it is probably premature to claim that it is a new genre within the repertoire of a language's performance possibilities. Instead, in this study, we begin with a basic question: Is there anything that might differentiate the ways people use language in electronic discourse from those in, for example, an exchange of signals by flags, a series of postcards, letters to newspapers, or successive sections of an epistolary novel? Before deciding whether to classify electronic discourse in any

particular language as a genre or a register of that language, we must begin with preliminary description and analysis of what electronic discourse seems to be or do, and what people choose to do with it. To do this, we draw on the field of discourse analysis: first, because the various approaches within discourse analysis are suitable for the task of describing and analyzing text, and second, because discourse analysis is, in itself, multidisciplinary. The kinds of questions we ask of the text and of ourselves are being asked by scholars in several disciplines.

The term *electronic discourse* focuses on how individuals use language to exchange ideas rather than on the medium or channel by which they transfer and deliver their messages. Using this term, as opposed to the term *computer-mediated communication*, emphasized our focus on language above the sentence—on language as utterances (Schiffrin 1994), whether written or spoken. Writing is often seen as space-bound, static, and permanent, whereas speaking is viewed as time-bound, dynamic, transient. Electronic communication, written on keyboards and read on computer screens, has many characteristics of both speaking and writing. Like telephone conversations, it is transmitted by a technology that replaces face-to-face communication, in the case of the telephone conversation with voices speaking and in the case of electronic discourse with images on a screen. Like letters, electronic discourse is supported by a delivery system that replaces face-to-face communication with writing that stands in place of voices. As a consequence, electronic discourse is writing that very often reads as if it were being spoken—that is, as if the sender were writing talking.

### *Writing that reads like conversation*

Electronic conference discourse exchanged by university students participating in mainframe conferences as part of a course is multiparty interaction through extemporaneous, rapidly written keyboard composition. It reads like and to a certain extent acts like conversation. However, the discourse cannot be analyzed only or primarily by the methods used for the analysis of conversation because—aside from being unspoken—electronic conference discourse is asynchronous. It thus has a different kind of immediacy of

feedback or response. That is, interactivity is delayed: the time between the creation of text between sender and responder may range anywhere from several seconds to several weeks, or even longer, depending on the length of time that the conference is available to its participants.

Electronic discourse also differs from face-to-face communication in turn taking. Turn taking is constrained for electronic conference discourse, both by time and by the computer software, which also delineates boundaries between utterances and archives each written utterance as received. Hence the interruptions and overlaps so characteristic of conversation are not possible. Electronic discourse can alter or rearrange the sequential ordering of conversation's adjacency pairs, which speakers and analysts use to track the sequence of conversational interactions or discern topical or thematic shifts. For example, Lee is reading and responding to Carter's note received on a Tuesday at 3 P.M., and pauses before saving the message within the program. On the same day and at the same time, Biff composes a message in response both to Cane's from the preceding Friday and to Carter's. It is completed on Tuesday at 3:01 so that it is saved by the software before Lee's. The messages, then, are distributed as coming in the order of Cane, Carter, Biff, Lee. Thus, their chronological distribution will not represent the actual timing of the utterance exchanges. In electronic conference discourse, interactivity draws on two time frames: that of the sender and that of the responder. Since conferences typically involve multiple people, the two time frames may intersect but are not necessarily immediate in their interaction.

Electronic conference discourse is like conversation in that it presents a number of performance features generally characteristic of in process or in situ communicative events and behaviors, such as repetition, direct address, disfluencies, and markers of personal involvement. These features include syntactic and lexical items on which Biber (1988) performs multivariate analysis in order to derive dimensions characterizing genres of written and spoken text.

The features may also be graphic. Wilkins (1991) notes the use in electronic mail discussion lists of all-capital letters, the creation of emoticons, the use of punctuation to signal humor or irony or a sense of intimacy. In her collection of multiparty conversation on a conferencing network for an electronic communications utility over

a three-months period, Wilkins (1991) observed that what kept the conversation flowing was not references by name or number or established conversational sequences. Instead “the conversational topic was maintained through lexical repetition, synonyms and shared cultural knowledge” (63).

### *Speaking and writing: Biber's dimensions*

Lexical repetition and variation is probably the most immediately noticeable feature in electronic mail and in electronic conference discourse. Biber's series of analyses of variation across spoken and written forms of languages has shown that “linguistic variation in any language is too complex to be analyzed in terms of any single dimension” (1988: 22). His goal was not to set up or to confirm an absolute distinction between spoken and written language. Rather, it was to “specify the multidimensional relations among the many different types of speech and writing in English” (25). He identified four notions that are useful in discussing both speech and writing—integration, fragmentation, involvement, and detachment (43). He used these notions to identify six dimensions of variation (expressed as two opposing poles on a continuum), which characterize genres in both speaking and writing. These dimensions collapse co-occurrences of features that serve to typify or characterize different genres of discourse as being *involved* in nature as opposed to *informational*; *narrative* as opposed to *nonnarrative*; *explicit* as opposed to *situation-dependent*; offering *overt expression of persuasion*; *abstract* as opposed to *nonabstract* in terms of information; and presenting *on-line informational elaboration*. The application of Biber's dimensions can be extended to the study of electronic discourse.

According to Biber, integration (“the way in which a large amount of information is packed into relatively few words”) is present in writing, but not generally in speech, which “cannot be highly integrated because it is produced and comprehended on-line” (43). Here, *on-line* refers to the constraints of time and immediacy accompanying speech. Careful word choice, in other words, is expected in written discourse, but does not characterize most kinds of speech situations. Electronic messaging in real time, or very brief synchronous interactive electronic communication, is more like

informal speech situations. Because of its *integration*, electronic conference writing, though extemporaneous, is more like written discourse.

*Fragmentation*, shown by such features as clauses connected by “and” as opposed to subordinating conjunctions, characterizes text produced under severe time constraints, such as a typical speech interaction. *Involvement* is suggested by linguistic features of interaction that refer directly to the recipient of the text. It contrasts with *detachment*, often marked by agentless passives and nominalizations, which characterizes situations that are not two-way interactions. Like other forms of speech and writing, electronic discourse may be analyzed for these functions.

Is it sufficient to consider electronic discourse only as writing? It is both possible and instructive to analyze a corpus of electronic conference discourse as notes—longer than a comment, shorter than an essay—that all deal with one or more aspects of a topic and which cluster into several very general themes and probably some fairly distinct narrative patterns. In that sense, we could reconstruct and study the text of the interaction as if it were a thematically organized sonnet sequence, perhaps, or a series of commentaries glossing a series of specific texts, as with midrash.

Problems arise, however, with the order in which different entries in electronic discourse can be read and analyzed. The electronic conference sustains an event of language contact. Multiple texts, and through them, their writers, are in contact in a variety of ways. The entries in electronic conference discourse can be arranged in chronological order for the whole conference, but that order jumbles topical or thematic threads of discussion and omits the role of the individual text or writer. Entries can be arranged by the order of entries keyed to one specific text or topic, but that order jumbles the chronology of the whole, though it can present an array by each topic. Entries can also be arranged in terms of the writings by each writer in order to track connections across topic or time. One order facilitates looking at the group through synchronic arrangement of contiguous texts. Another order facilitates looking at change in individual writings over time. Reading the artifact of an electronic conference—the text remaining after the interactive performance of its writers and readers—is not a straightforward task. In this study, we examine features presented by the group, by the topic, and by

the individual, turning the kaleidoscope of discourse in order to look at the role played by the individual within the group.

Biber's analyses of a wide range of spoken and written varieties of English have demonstrated that "there is no single, absolute difference between speech and writing in English; rather there are several dimensions of variation, and particular types of speech and writing are more or less similar with respect to each dimension" (1988: 199). Therefore, we cannot look only at speech or only at writing in order to characterize electronic discourse, because the two share many characteristics. Electronic discourse, an interesting and important example of language use and text, is different from the conventional sense of both spoken and written language.

### *Multidisciplinary perspectives*

Scholars in several disciplines, such as sociology, anthropology, rhetoric, psychology, composition theory, and folklore, currently ask similar questions and often present complementary perspectives about writing, social situations, and conversation, drawing on insights from earlier work by linguists and anthropologists. Analysis of electronic discourse is keyed to how one thinks about everyday conversation or narrative, or about the ways that special kinds of writing and special features of ordinary speaking are interrelated.

Scholars from several disciplines have begun to apply the findings of linguists who examine contemporary spoken and written language and their relationships to each other (see, for example, Chafe 1986; Chafe and Danielewicz 1987; Tannen 1990; Biber 1988). Schiffrin (1994) and Johnstone (1996) have called for multidisciplinary perspectives on the study of discourse. Scholars in allied disciplines endorse the interdisciplinary examination of communicative discourse in general and electronic communication specifically (Baym 1995; Collot and Belmore 1993; Eldred and Hawisher 1995; Ferrara, Brunner, and Whittemore 1991). From the perspective of communications studies, Rice (1982, 1987), Galegher, Kraut, and Egidio (1990), Dunlop and Kling (1991), and Foulger (1990) emphasize different aspects of collaborative work, usually within organizations or businesses, in order to see if and how electronic discourse supports and even improves the work, the product, or

different kinds of interactions within a group or team. Hiltz (1984), Sproull and Kiesler (1991), and McGrath (1990) have assessed electronic interactive communication in organizations and businesses (see Finholt and Sproull 1990). Theories underlying the process of collaboration in writing tasks have been discussed in the fields of composition and rhetoric, particularly in terms of notions of social construction; that research is reviewed from a variety of perspectives by Lanham (1992), Tuman (1992a), Faigley (1986), Bolter (1991), Duin and Hansen (1994), and Selfe and Meyer (1991). Studies of teaching applications involving interactive electronic communication, whether within a single classroom or spread out across a county, a country, or several continents, include Harasim (1990), Hawisher and LeBlanc (1992), Harrison and Stephen (1996b), Baldwin (1996), and Kaye (1992). Studies of setting (Foulger 1990), gender (Herring 1993), and other aspects of social organization discerned in or effected by electronic discourse (Eldred and Hawisher 1995) have begun to influence research, particularly as scholars note the interdisciplinary nature of such efforts.

In order to determine where and how electronic language fit Biber's dimensions, Collot and Belmore analyzed a corpus of electronic messages that were organized into nine subject areas—chit-chat, current events, science, science fiction, finance, film and music, photo and cooking, medical, and sports (1992:45). They found the features characterizing electronic language to be most like those in the dimensions for two genres Biber had studied: public interviews and personal and professional letters. While they identified situational features affecting language use, Collot and Belmore were not able to determine the extent to which the situational features affected the "overall linguistic configuration of Electronic Language" or the "relationship of the participants to the text" (Collot and Belmore 1993: 53). In her study of a Usenet newsgroup that discussed daytime soap operas, Baym concluded that the "complex and dynamic process" involved in the development of patterns for identity, norms, and communication point to the need for more "naturalistic, ethnographic, and microanalytic research" in order to "refine our understanding of both influences and outcomes" (1995: 161).

Studies like these point to the need to look at electronic discourse from the microlevel as well as the macrolevel. Whether

people are using computer networks to interact and communicate in “real time” (synchronous) as on the telephone or in delayed time (asynchronous) as with letters, whether their computers are side by side, in the next room, in the same town, or a continent away, they are using computers for the purpose of communicating with each other. One focus on these various uses can be the channel being used—that is, the way the networks, the computers, the hardware and software, and the connections support the act of communication. Another focus can be the context in which the communication takes place. This study focuses on the latter in order to examine how people adapt their approaches to and their ways of communicating in different contexts. Electronic discourse is discourse that takes place in those contexts. What the individual participants do with language in order to discuss issues and create a community is the basis for understanding the various levels of electronic discourse. How the “linguistic individual” (Johnstone 1996) relates to the text and adapts to the electronic context provides the key to understanding the linguistic and social features of electronic discourse. The fabric of electronic discourse is language; the weavers of that fabric are the individual participants.

### *Selected approaches to discourse analysis*

Just as there is no longest sentence, there is also no longest stretch of discourse, which is generally considered to be language beyond the sentence. There is no single theoretical approach to analyzing discourse. As Deborah Tannen notes (1989: 6–7), discourse analysis may seem sprawling, even “heterogeneous,” because “it does not grow out of a single discipline . . . it is by nature interdisciplinary.” In *The Linguistic Individual*, Barbara Johnstone (1996: 22) reminds us that discourse analysis can also be defined, following Coulthard, as “situated speech or, following Brown and Yule, as “language in use.” Johnstone frames her study of the individual’s consistency and idiosyncrasy, undertaken “to show how paying attention to individual voice helps in understanding language” (4) by calling for “the kind of language study I find most compelling, the work A.L. Becker . . . calls ‘modern philology.’ As was the traditional philol-



ogy that gave rise to modern linguistics, modern philology is centered on particular texts" (4). Johnstone calls on scholars to look, then, not only at speakers seen as a group or at an idealized system, but also particular utterances and at approaches that "locate language and dialect in the individual's creative choices for how to talk and understand" (13). Her approach to discourse analysis has the goal of understanding language through understanding the text. Speaker-centered linguistics, a linguistics "once again willing" to draw on philology as "the close reading of texts considered to have historical or literary value" (180) will find that "one's text or texts, rather than one's theory, tends to be the source of discipline" (24).

Looking at text, and specifically at written text, is one of the concerns of discourse analysis. According to Tannen, discourse analysis is a term that "describes the object of the study" and which developed "in order to make legitimate types of analysis of types of language that do not fit into the established subfields of linguistics, more narrowly focused, which had come to be regarded by many as synonymous with the name of the discipline, and to encompass work in other disciplines that also study language"(1989:6-8). In her survey of the different approaches to the analysis of discourse, Schiffrin (1994) identifies six key methodologies, each drawn from different areas of linguistics. Each of these methodologies views language in its social context from a slightly different perspective. Speech act analysis "focuses upon knowledge of underlying conditions for production and interpretation of acts through words" (6). Pragmatics analyzes "speaker meaning at the level of utterances" (9). Interactional sociolinguistics examines "how language is situated in particular circumstances of social life and on how it adds (or reflects) different types of meaning . . . and structure . . . to those circumstances" (7). Ethnography of communication analyzes "the structures and functions of communicating that organize the use of language in speech situations, events, and acts" (185). Conversation analysis looks at how "members of a society produce a sense of social order" (232; see 9-10). While interactional sociolinguistics, ethnography of communication, and conversational analysis assume a language-as-interaction model, variation analysis assumes a language-as-code model (385; see 405) to examine the distribution and variation of forms across text types (331).

In this study, we draw on insights and perspectives from interactional and variationist sociolinguistics, the ethnography of communication, and historical linguistics, in whose bosom philology still nestles: “in Old Irish, or Hittite, or Vedic Sanskrit, or Indo-European studies, everyone does both linguistics and philology on a daily basis, and it’s no big deal” (Watkins 1990: 22). As Watkins reminded the Linguistic Society of America in his presidential address of 1989, both philology and pragmatics are “the study of the meaning of language forms as these depend on the linkage of signs to the context in which they occur” and that “good comparatists like a Saussure, a Wackernagel or a Delbrück moved freely and effortlessly between diachrony and synchrony” (1989:785). The examination of electronic discourse involves the analyst with both planes as well as with both older and newer approaches to discourse.

Suzanne Fleischmann’s (1990) comments about texts in Old French are, we think, particularly appropriate to the study of electronic discourse with its features of both oral and written language:

As a linguistically oriented philologist, I am convinced that many of the disconcerting properties of medieval vernacular texts—their extraordinary parataxis, mystery particles, conspicuous anaphora and repetitions, “proleptic” topicalizations, and jarring alternations of tenses, to cite but a few—can find more satisfying explanations if we first of all acknowledge the extent to which our texts structure information the way a spoken language does, and then proceed to the linguistic literature that explores the pragmatic underpinning of parallel phenomena in naturally occurring discourse. (Fleischmann 1990: 23)

As linguists interested in synchronic and diachronic aspects of language as they occur in text and discourse, and thus in philology, we adapt to the study of electronic discourse Ochs’s (1990: 289) definition of discourse as “a set of norms, preferences and expectations relating language to context, which speaker-hearers draw on and modify in producing and making sense out of language in context.” Our primary goal throughout the study is to reconstruct the electronic conference text and to identify and describe features important to its structure. “Structure, or regularity, comes out of discourse and is shaped by discourse as much as it shapes discourse in an ongoing process” (Hopper 1987: 142).

Schiffrin's (1994: 39) principles for discourse analysis assume that discourse arises "not as a collection of decontextualized units of language structure but as a collection of inherently contextualized units of language use." Discourse analysis is empirical—that is, it is sequential, distributional, and predictive. It is more than the sequence of its linguistic units, in that "forms and meanings work together with social and cultural meanings, and interpretive frameworks." It assumes that discourse is interactive, sequentially situated, and is "guided by relationships among speaker intentions . . . conventionalized strategies for making intentions recognizable . . . the meanings and functions of linguistic forms within their emerging contexts . . . the sequential context of other utterances . . . the properties of the discourse mode. . . the social context . . . [and] a cultural framework of beliefs and actions" (316). Our examination of electronic discourse is sequential and distributional. For example, we look at sequences of utterances as units and as sequences within texts, at how individuals shift styles keyed to their own intentions, and at how they develop conventions and appropriate strategies from each other. We describe the computer conference as the cultural framework for actions by student writers.

In 1989 we devised an electronic conference and repeated it for the next four years. The texts included in the conference focused on one topic—the original newspaper reports from the early days of the 1960 Sit-Ins. The teacher-directive, which served as the initial prompt introducing the Sit-Ins conference, asked the students to examine certain aspects of language use in the newspaper stories. Language use was the first "topic" that students encountered, though seldom the first they addressed. Instead, they were more likely to respond to issues keyed to both the text and the subtext issues in the newspaper stories. As reports of racial confrontation during the civil rights movement in the U.S.A., these stories presented historical and affective, even inflammatory, subtexts about, for example, race, confrontation, civil liberties, and violence. Each student's writing had its own hierarchy of issues, depending on how it addressed the teacher-directive, the general topic of the conference texts as a whole ("the Sit-Ins"), a particular newspaper story, ideas and issues in writings by other students, and the student's own comments in a previous writing. As soon as the conference discussion began, students became emotionally as well as intellectually involved. Although the students

wrote about the verbs and modifiers used by the writers of the newspaper stories, they also analyzed the meaning or significance of those forms in terms of the news writers' stance and reflections of attitudes conveyed through those forms. In terms of modern philology, they engaged in "the study of the meaning of language forms as these depend on the linkage of signs to the context in which they occur" (Watkins 1989: 785).

When we introduced the Sit-Ins mainframe computer conferences to our students in 1989, we saw the conferences as a reserve shelf in a library that never closed, a chalkboard that was never erased, and a fieldwork record that included everybody's entries. As students began to write in the conferences, we recognized that they had engaged themselves in the difficult task of establishing norms for their temporary electronic community of writers as they chose ways to write about a socially sensitive subject. We chose to look at those aspects of students' uses of language that shift within specific contexts as the students wrote about practices and attitudes conveyed through language recorded before most of them were born.

As Stubbs (1996: 152–53) recommends in his principles for text analysis, our analysis of the corpus of the students' writing is comparative. We look at the full corpus, the individual in differently situated conferences, the individual topics, similar topics across all conferences, the first entries in all topics across a single conference, the narrative schema of all entries in a single topic, and all writings by single students across topics. We analyze specific features that sometimes characterize the corpus, sometimes a cohort of writers, sometimes a type or style of text, and most frequently, the individual writer. We archived and examined a corpus of language—collected from computer-naïve users who wrote extemporaneously at the computer keyboard and interfaced with other texts—so that this corpus could be compared with features from other corpora of text, both within the electronic domain (such as electronic mail lists, netmail, and messages) and without.

From interactional sociolinguistics, we drew on methods for examining the multiparty and interactive nature of our text. From variationist sociolinguistics, we drew from methods of looking at the distribution of repetition and variation of features for individual and groups of writers. From the ethnographic and historical paradigms, we drew on the emphases of comparison across contexts, of natural description, of reconstruction of text, in order to examine

communicative competence and performance. We began with the word, since that is where electronic conference discourse begins.

Electronic discourse is patterned, structured in multiple layers that can have, in the minimally hypertextual discourse of the electronic conference, multiple links. Throughout this study, we look at the ways the student writers legitimate their claims, opinions, ideas, insights, or responses by appeals either to the newspaper text, to some body of cultural knowledge about either the time of the Sit-Ins or the current scene that they take as given or shared, or to their own memories and life experiences. We note that their shorter writings generally fall into either a mode that is “guarded” or one that is “self-disclosing.” If their writings are 150 words or longer, the individual writing will generally present both modes, or text types. We chose these modes or text types rather than Labov’s terms, “casual” and “careful,” because either (or both) of the text types can be presented in either casual or careful style. We show that the student writers’ alternations between types, and the boundaries for those text types, are signaled by the shifts in type-token ratio that index density, repetition, and patterned variation at the lexical level.

We also analyze direct and indirect address because they are an important part of the interactive nature of this kind of writing. How students present themselves at the most general level through direct and indirect address involves a form of negotiation of identity as presented in, by, and through interactive electronic text. We look above the sentence level at successive levels of language and text, in order to delineate how the writers develop conventions for themselves and enculturate each other. The index for such conventions and enculturation is the use of repetition and emulation—repetition at the level of words or lexical collocation, and emulation for larger patterns, including syntactic features, narrative schemata, and communicative orientation. Both repetition and emulation are analyzed by tagging and sorting the features in our corpus. We hand tagged each feature or pattern described and used the computer to sort and generate a concordance of the tagged features.

### *Description of the corpus*

Our corpus of electronic discourse consists of a set of three main-frame conferences from the first three consecutive semesters (1990–

1991) that we incorporated the conference into linguistics courses at both universities' campuses. The social situation varied each time. The first semester's conference (Stand Alone) had no interaction between campuses; the second conference (Exchange) was completed on each campus and then exchanged with the other, to elicit additional replies; the third (Transparent) linked students from both campuses simultaneously. The newspaper stories and our only directional prompt that appeared as the first item remained the same for each conference. Table 1.1 profiles the writers, entries, and words for the total corpus of 116,929 words. Based on 350 words of typescript for an 8 1/2 x 11 inch page, the corpus contains approximately 334 pages of conventional text. Table 1.1 gives a profile of the conferences.

Excluding dual writers, students who worked in dyads but used only one user name to write and send their comments, the proportion of female to male writers was almost 3 to 1, with females represented by 197 writers, with 89,474 words, and males represented by 71 writers, with 27,455 words. Females, however, wrote slightly more; although they represented 73.5 percent of the writers, they wrote 76.5 percent of the words. Additional correlations with gender are shown in Table 1.2, which organizes the corpus by lexical characteristics. "Token" represents the total of all words, "type" means the total of unique words, and "chunk" means the number of fifty-word segments.

Since the corpus is an artifact of written text, we approached it first at the lexical level, with two ways of organizing the text for analysis. We indexed the total corpus with a software program that enabled us to create a concordance of a word or grammatical tag

Table 1.1. Profile of the Electronic Conferences

	<i>Entries</i>	<i>Writers</i>	<i>Words</i>
Stand Alone UNCC fall 1990	163	30	20,174
Stand Alone UNCG fall 1990	79	23	12,049
Exchange UNCC spring 1991	230	74	25,048
Exchange UNCG spring 1991	281	73	31,785
Transparent fall 1991	226	68	27,873
Totals	979	268	116,929

*Note:* Writers from both campuses participated in the Exchange conferences.

Table 1.2. Lexical Characteristics

	<i>Exchange UNCC</i>	<i>Exchange UNCG</i>	<i>Stand Alone UNCC</i>	<i>Stand Alone UNCG</i>	<i>Transparent</i>	<i>Totals</i>
No. of Females	51	53	24	21	48	197
Tokens	16,999	31,785	16,207	11,130	20,783	48,120
Max	376	495	505	570	634	
Min	0	0	0	4	0	
Mean	108	113	117	150	127	
Types	5,667	12,063	6,532	5,836	9,838	22,206
Max	324	387	389	468	599	
Min	0	0	0	0	0	
Mean	36	43	47	79	60	
Chunks (est.)*	340.1	635.9	324.2	222.7	416.3	963.2
Max	7.5	9.9	10.1	11.4	12.7	
Min	0.0	0.0	0.0	0.1	0.0	
Mean	2.2	2.3	2.3	3.0	2.5	
No. of Males	23	20	6	2	20	71
Tokens	8,049	7,430	3,967	919	7,090	11,976
Max	450	347	446	346	303	
Min	0	0	8	8	0	
Mean	110	118	159	184	114	
Types	3281	2959	2775	622	2,224	5,621
Max	325	272	371	264	230	
Min	0	0	0	0	0	
Mean	45	47	111	124	36	
Chunks (est.)*	161.2	148.4	79.4	18.4	142.1	239.9
Max	9.0	6.9	8.9	6.9	6.1	
Min	0.0	0.0	0.2	0.2	0.0	
Mean	2.2	2.4	3.2	3.7	2.3	

\*The number of estimated chunks = tokens/50

(done by hand for selected portions of the corpus). We also divided every writing into fifty-word segments of text in order to run type-token ratios on every writing presenting three or more fifty-word chunks.

Type-token ratio indicates the lexical diversity within a text by dividing the number of different words, or types, by the total number of words, or tokens, in segments of text. It has been used for studies of written and spoken language. For example, Carpenter (1990) based his study of depositions, oral testimony, and cross-examinations during a trial on earlier studies of written text as well

as oral speech. He noted that for spoken text, the segments subjected to type-token ratio (TTR) analysis need to be at least twenty-five words in length: "Segments any shorter . . . are not fruitful for TTR analyses because the statistical probability is that native speakers of English do not utter statements of 10 to 15 words in length without repeating one of those words, and variations between such smaller segments do not lend themselves to meaningful interpretations" (Carpenter 1990, fn. 19: 16–17).

For our TTR analysis of electronic discourse, we used segments of fifty words. This segment length is practical and useful in analyzing electronic discourse because the analyst can study features used by individual writers as well as features used by groups of writers. This segment length is also roughly equivalent to a minute of uninterrupted "speech" and half a screen of single-space writing in the "reply" function of VAXNotes, the mainframe software program used for our conferences. Electronic discourse—conferences, e-mail, discussion lists, or forums—consists of the statements and perspectives of individual writers. Analyzing all conference or discussion list writing as a unified whole would be equivalent to analyzing a conversation, which by definition must include at least two speakers, as if only one speaker participated. For example, the sixty-eight writers in the Transparent conference discussion would be presented as a generalized, single individual. This approach eliminates the distinguishing feature of electronic discourse—the individual writer.

The fifty-word segment length allowed writers to present elaboration, setting a baseline for our comparisons and allowing us to examine variation and repetition across individuals with some degree of precision. TTR measures the variation for an individual speaker or writer rather than for a group, in that it indicates the individual's variation from his or her mean. A segment with a TTR higher than the mean for that writer's statements would indicate more diversity and less repetition. Changes of TTR within successive segments of discourse signal that the level of lexical diversity has changed, suggesting some sort of shift in style.

Throughout our discussion, we call these segment "chunks," for the same reason as Barbara Johnstone in her *Stories, Community and Place*; we want "to avoid prejudging the issue of what these larger units are" (Johnstone 1990: 41). We found that these fifty-word chunks had a nice fit with the average length of segments of



narrative when we reanalyzed the writings using a framework adapted from Labov (1972): orientation, narrative(s), and coda. Table 1.3 uses different type fonts and markings to show this fit.

Comparing features of electronic discourse has led us to hypothesize something about an otherwise unobservable pair of phenomena, which we call by the metaphoric term “chiming.” In interactive electronic discourse, the writer is a reader, a writer, and a thinking communicator. Writers chime into text that they have read, calling on their repertoires of styles and competencies in both speaking and writing. As they write responses, they (probably unconsciously) present features into which their presumed readers may chime. In effect, they “charm” a reader into reading the whole of their text, using repetition to signal direct assent or alignment with a position—an opinion or a stance—and emulation to present more indirect agreement.

### *Using the concordance: An example*

Interaction among a number of factors underlie shifts of style among segments of a text; TTR signaled that shifts had taken place, and the concordancer provided in the WordCruncher software program allowed us both to look in more detail at specific aspects that might underlie a particular shift and to examine larger patterns of usage for lexical or grammatical features. Our analysis of the use of the intensifier “a lot” can illustrate how we used the corpus. Both “a lot”, written as two words, and “alot”, written as one word, occur in our data. We consider the two spellings in free variation and thus as the same collocation. The choice of a particular intensifier, such as “a lot” will be maintained throughout entries by a specific person; its appearance is usually in segments that depart from the writer’s mean TTR, in either direction, by one or more standard deviations. “A lot” occurs in three well-attested usages in our data: (1) as an intensifier, as in “a lot more carefully” (see Quirk, Greenbaum, Leech and Svartvik 1972: 295); (2) as a partitive, as in “a lot of hype” (ibid.); and (3) as a noun phrase, as in “gave them a lot” (ibid., cf. section #5.77).

Twenty of the sixty-eight writers in the Transparent Conference presented twenty-eight uses of *a lot/alot*, with females more frequent

Table 1.3. Narrachunks: The Overlap of Narrative and TTR Segments

This figure illustrates the correlation of onset/orientation, narrative and coda with TTR-segments

**Transparent Conference, Entry 2.17**

'Shandie' uses her first name in the title (coda) as part of the "Titling Game" that took place in Topic 2 of the Transparent Conference. Her orientation indirectly refutes Writing 212 with an empathetic reframing of the 1960s newspaper writer's position. Her frame, by setting up her location of self as being born after the 60s (Writing 209 had located self in that time period), mitigates her claims to 'given knowledge' about that time period while drawing on her own personal experience, and allows her to project and hedge how she might or might not have acted at that time (Writings 213 and 216 had engaged in projection). She evaluates and sets up a new orientation, reflecting on the newspaper article, to frame her final segment. This segment looks at the wording in the article itself, beginning with "orderly" (a word from the newspaper story studied in Writing 202) ending with an evaluation of the style as being in conflict with the time period of the 60s: her final evaluation echoes the time set forth in the initial orientation, and she writes the Title as her Coda to the whole writing.

**Coding:**

// marks end of line      \* marks TTR-chunk separation      # marks beginning, new sentence

**Narrative Schema:**

Coda 3: Title for whole writing [last thing written]

Orientation to first and all successive narratives until the SHIFT

Narrative 1

Narrative 2

Coda 1, evaluation from Narratives 1 and 2 / Orientation to third narrative

*TTR SHIFT: FROM MEDIUM TO HIGHEST TTR*

Narrative 3      Highest TTR

*TTR SHIFT: FROM HIGHEST TO LOWEST TTR*

Coda 2      Evaluation of Narrative 3, lowest TTR

Coda-3	1	<- shandie's thoughts >-
Orientation	2	*i thought that the article was sympathetic for the time in which it was //
	3	written. #i don't think that they gave enough information about the //
	4	facts as they were but then again, i don't think it would be possible //
	5	seeing as how everyone views things differently
		#the issues of civil //
Narr-1	6	*rights are interesting to me yet i really don't understand why the //
	7	black population was treated this way to begin with. #i know this is //
	8	probably because i did not grow up in this span of time and i have //
	9	always gone to desegregated schools.
		#it makes me feel strange to think//
Narr-2	10	*about not having gone to school with all different races.#i think it //
	11	has helped me get a better outlook on different cultures and to become //
	12	a more well-rounded person. #i would like to think that i would not //
(Eval)	13	have treated the afro-americans this way if i had been alive back then //
Coda-1	14	*but i cannot honestly say that because i don't know what it was like//
Orientation	15	#the article itself surprises me by being sympathetic at a time when no //
	16	one of the white population wanted to have anything to do with civil //
<b>HI TTR</b>	17	rights. #it really surprised me to read the word "orderly" describing //
<i>Narr3</i>	18	*the blacks [sic] students as they were leaving the store. #the wording used//
	19	makes the white people look worse mannered than the black students. # i //
<b>LO TTR</b>	20	think the writer of this article deserves great credit for his/her //
	21	choice of words. #the style chosen is in direct conflict with the actual //
Coda-2	22	*tumultuous times of the 60's civil rights movement.//

in terms of total tokens, but males more frequent relative to size of cohort (nine of twenty males; eleven of forty-eight females). Table 1.4 lists writers by gender and a number assigned to mask student identity; the number following the colon is the number of times that student used *a lot* in a particular way.

Notice that only five of these uses are “true” intensives; the only writers who use the nominalized formation are also those who write more than 150 words. If a writer uses *a lot* more than once, the writer (as with F2, F6, M6, F8) will tend to continue using it in the same way. In other words, the partitive is more common than either the intensifier or the nominalization from the partitive, but the use of any of these is driven by idiolectal preference.

One of the most interesting aspects of our study has been the notion of idiolect and the individual’s discourse signature in electronic discourse. Writing in the electronic universe, people adapt conventions of oral and written discourse to their own, individual communicative ends, as when they draw on modality or develop interactive strategies keyed to the formulaic properties of rhetorical questions or the dramatic potential of “asides.” Electronic discourse presents us with texts in contact, and through those texts, their writers. We have sought evidence for change, some trace of being and human interaction in a domain where footprints are not in moon dust but in ether.

Table 1.4. Use of *a lot* by type and gender

	Type 1 <i>Intensive</i>	Type 2 <i>Partitive</i>	Type 3 <i>Nominal</i>
TTR-writings (150+ words)	F1: 1	F2: 4 M1: 1 M3: 1 F3: 1	F2: 1 M2: 2 M4: 1 F4: 1
non-TTR writings (fewer than 150 words)	F5: 1 M5: 1 F7: 1 M7: 1	F6: 2 M6: 2 F8: 2 M8: 1	M9: 1 F9: 1 F10: 1 F11: 1