

# Written Communication

<http://wcx.sagepub.com>

---

## **Interactive Written Discourse as an Emergent Register**

KATHLEEN FERRARA, HANS BRUNNER and GREG WHITTEMORE

*Written Communication* 1991; 8; 8

DOI: 10.1177/0741088391008001002

The online version of this article can be found at:

<http://wcx.sagepub.com/cgi/content/abstract/8/1/8>

---

Published by:



<http://www.sagepublications.com>

**Additional services and information for *Written Communication* can be found at:**

**Email Alerts:** <http://wcx.sagepub.com/cgi/alerts>

**Subscriptions:** <http://wcx.sagepub.com/subscriptions>

**Reprints:** <http://www.sagepub.com/journalsReprints.nav>

**Permissions:** <http://www.sagepub.com/journalsPermissions.nav>

**Citations** <http://wcx.sagepub.com/cgi/content/refs/8/1/8>

*Text transmitted electronically through computer-mediated communication networks is an increasingly available yet little documented form of written communication. This article examines the syntactic and stylistic features of an emergent phenomenon called Interactive Written Discourse (IWD) and finds that the concept of "register," a language variety according to use, helps account for the syntactic reductions and omissions that characterize this historical juxtaposition of text format with real-time and interactive pressures. Similarities with another written register showing surface brevity, the note taking register, are explored. The study is an empirical examination of written communication from a single discourse community, on a single topic, with a single recipient, involving 23 experienced computer users making travel plans with the same travel advisor by exchanging messages through linked computers. The study shows rates of omissions of subject pronouns, copulas, and articles and suggests that IWD is a hybrid, showing features of both spoken and written language. In tracing variable use of conventions such as sentence initial lower case and parentheses, the study shows that norms are gradually emerging. This form of written communication demands study because, as capabilities expand, norms associated with this medium of communication may come to influence or even replace those of more traditional writing styles.*

# *Interactive Written Discourse as an Emergent Register*

**KATHLEEN FERRARA**

*Texas A&M University*

**HANS BRUNNER**

*U.S. West Advanced Technologies*

**GREG WHITTEMORE**

*Electronic Data Systems Research, Inc.*

In human language, processes of conventionalization are always operating, as social groups develop shared norms of the varieties of lan-

---

Authors' Note: The data collection and some of the scoring and analysis were performed at Microelectronics and Computer Technology Corporation, Austin, TX, where each of the authors was affiliated with its Human Interface Laboratory.

WRITTEN COMMUNICATION, Vol. 8 No. 1, January 1991 8-34

© 1991 Sage Publications, Inc.

guage appropriate for different occasions and as individuals acquire and modify these norms.

—C. A. Ferguson (1983, p. 170)

Traditional views of writing as a noninvolved, solitary activity lacking a copresent audience are rapidly exploding as new technologies make available new modes of composing and communicating. Such technology may be supporting an already existing natural collaborative capacity of human beings. The past quarter of a century has witnessed the emergence of the computer as a vibrant force, first in government agencies, then corporate and university contexts, and increasingly in homes and small business offices. Even more important, the technological advances allowing linkage of computer users through communication networks (e.g., BITNET, ARPANET, INTERNET, and TELENET) are initiating far-reaching changes in the composing process as well as in conventions of style. The prospect for interactive creating, revising, and editing of documents by participants who are separated physically in time and space and may have no previous acquaintance with one another makes a systematic appraisal of what Kiesler, Siegel, and McGuire (1984) called Computer-mediated Communication (CmC) an appealing subject of inquiry for teachers and researchers of written communication.

As Faigley (1985) argued, writing is a social act that takes place in a structure of authority and changes constantly as society changes (p. 236). Because writing is an act not easily separated from its functions in a particular discourse community (Faigley, 1985, p. 246), it is incumbent on those who seek to understand or teach the use of written language to keep up with such changes and to expand research beyond the familiar territory of composition research to encompass the wide range of nonacademic written communication. Currently, very little is known about the characteristics or conventions of what we term Interactive Written Discourse (IWD), the written language occurring in simultaneous terminal-to-terminal typed dialogues.<sup>1</sup> Nonetheless, written communication through keyboards and computers is frequent in the workplace.

In general, "people do not receive either formal or informal instruction in an etiquette of electronic communication," according to Kiesler et al. (1984, p. 1125). We can ask, do computer users constitute a "discourse community" (Bizzell, 1982)? Are there indeed shared norms? If so, how are these conventions developed and acquired,

what are they like, and are they differentially distributed throughout the community?

The growing awareness among rhetoricians that context is a powerful shaping influence has a parallel in linguistic studies in the notion of register, defined as a variety of language according to use, as contrasted with dialect, a variety of language according to user (Halliday, MacIntosh, & Strevens, 1964, p. 87). We will draw on the concept of register to help explain the development of this new form of written communication.

The goal of this article is to provide insight into the probable nature of the emerging language variety that we term Interactive Written Discourse (IWD) based on a corpus of computer-linked written communication from 23 subjects. The findings support three claims about IWD: first, that it is a naturally occurring register, perhaps a reduced register; second, that it is a hybrid language variety, displaying characteristics of both oral and written language; and third, that norms of its use are in the process of becoming conventionalized.

We contend that this new development represents an amalgam forged out of existing genres with similar characteristics. We argue that IWD is a hybrid register that resembles both speech and writing, yet is neither. Previous research on the similarities and differences between speech and writing that regarded the two types as dichotomized (Biber, 1986; Goody & Watt, 1968) or opposite ends of a continuum (Tannen, 1982) did not examine the amalgam IWD. The fact that manuals for computer etiquette or many other registers are neither readily available nor commonplace suggests that norms for registers are acquired during use from other users.

The corpus we have chosen for empirical examination is the typed interactive dialogues of 23 computer professionals or their spouses with a single travel advisor, linked through terminals in separate rooms. This group is especially valuable for study because their livelihood depends on computer use. Their presence, practices, and prior experience will certainly make themselves felt in the discourse community; they are likely to form a subculture whose norms have the potential to influence a wide range of other computer users in government, schools, homes, board rooms, and the marketplace.

Computer-mediated human-to-human communication is changing not only the way we think about the possibilities for communication but most probably the ways we use written language and compose in it (cf. Murray, this issue) as well as the ways we interact through

language, as Wilkins (this issue) shows. Understanding these developments will require the tools and approaches of many disciplines.

### PREVIOUS RESEARCH ON REGISTER

One intriguing aspect of the study is the question of how discourse communities establish conventions of use. Do these norms arise *de novo*, or are they fabricated from threads of other types of discourse? Rubin (1984) maintained that "writers are guided by sociolinguistic norms that associate stylistic variables with contexts by force of convention" (p. 226). Martin's (1983) study of the development of register observed that "registers are developed not by acquiring structures but by adjusting, appropriate to the context, the probabilities attached to features in already developed systems" (p. 36).

The term "register" was suggested by Reid (1956) to refer to a variety of a language appropriate for a particular situation. Halliday et al. (1964) further developed the concept that situation shapes language. Ferguson (1982) stipulated five working assumptions regarding the theoretical construct "register":

1. Register variation is universal. All speech communities show variation in structures (phonological, morphological, lexical, syntactic, or discursive) according to use.
2. Registers exist, although the establishment of discrete, uniquely characteristic registers is problematic.
3. Register systems differ cross-linguistically and change diachronically. We can discover typologies and principles of change.
4. A given register is variable in the degree of its distinctiveness. Structural and situational boundaries are often blurred. This variation in degree of distinctiveness from other registers or an unmarked "neutral" variety may itself be used as a marker of adjustment and eventually become conventionalized.
5. Competence in register variation is acquired as part of language development. (pp. 57-58)

Critics of the concept of register variation have pointed to the lack of discreteness of different registers as a problem. How is bureaucratic language different from legalese, for example. Don't registers such as Baby Talk and Foreigner Talk both have features such as exaggerated pitch or tempo in common? It is true that analysts have found the

boundaries to be blurred between various registers, and their features often overlap. We suggest that this lack of discreteness, more than being a problem, is actually a solution. The overlapping edges, the similarities between registers, provide a clue to the puzzle of linguistic competence and to the question of how registers, which are not taught by conscious instruction, come to be acquired in a natural fashion.

When confronted with a new situation, people draw on previous knowledge of partially similar activities to form an amalgam. Similarly, it appears that competent users of a language have an extended language repertoire, and when new situations arise, they create new appropriate language varieties out of existing language varieties: They form hybrids. For example, a person newly arrived in psychotherapy draws on knowledge of interviews, medical discourse, and conversation to fashion communicative behavior appropriate to the new situation. These elements are displayed in column 1 of the following table. In a similar way, IWD has the appearance of being forged out of elements of postcardese, headlines (studied by Straumann, 1935), and telegraphese, as indicated in column 2.

*Elements of Therapeutic Discourse*

Part interview  
Part medical discourse  
Part conversation

*Elements of IWD*

Part postcardese  
Part telegraphese  
Part headlines

Although the concept of "register" is widely accepted by both British and American sociolinguists, some linguists have the misconception that registers are characterized solely by vocabulary differences and that such lexical specialization is hardly worth notice. On the contrary, work by Ferguson on Baby Talk (1977), Foreigner Talk (1981), Sports Announcer Talk (1983), and special language registers (1985), Janda (1985) on Note-taking, Finegan (1982) on Testament Language seen in wills, Charrow (1982) on Bureaucratic Language, and others show systematic variation on every level of language: morphological, phonological, syntactic, and discoursal as well as lexical. Such work led Hudson (1980) to suggest that "the amount of variation due to register differences (if it could somehow be quantified) may be quite comparable with that due to differences in dialect" (p. 49).

A comparison of the register examined here shows that Interactive Written Discourse differs from other registers such as Baby Talk and Foreigner Talk in that they occur in the oral mode but IWD occurs in

a typed medium. In this respect, it is similar to the Note-taking register that Janda (1985) described as a written variety of English displaying surface brevity. The Note-taking register and IWD are similar in that both have real-time processing constraints. They both appear to try to save time and to minimize effort, but IWD has greater demands for clarity or comprehensibility because the audience is other than the self.

IWD differs from the Note-taking register in that it is interactive, and the potential exists for immediate feedback from the message receiver. In contrast, note-taking is not interactive, and the audience is usually oneself. These interactive and immediate features make IWD different from any previously studied variety of written communication.

#### PREVIOUS RESEARCH ON COMPUTER-MEDIATED COMMUNICATION

Outside of "dialogue journals" (Staton, cited in Rubin, 1984, p. 218), a writing assignment that strategically includes interactive turn-taking, or the phenomenon of classroom note-passing by school children, very little attention has been paid to the types of writing that involve interaction, rapid feedback, or copresent audiences.<sup>2</sup> Many researchers previously assumed that context of use in written communication was "eventual, not concurrent with the production of discourse, as it is with spoken language" (Nystrand, 1987, p. 205). While in the past it may almost always have been true that "writers, unlike speakers, do not produce language in the company of a language receiver," (Nystrand, 1987, p. 198), studies of computer-mediated discourse such as IWD show that this assumption no longer holds. Such conceptions are changing, as Kiesler et al. (1984), Murray (1985, 1988), Maynor (in press), and Wilkins (this issue) demonstrated in their investigations of the various subgenres of CmC. Murray (1988) made a careful distinction between subtypes of CmC, such as e-mail, e-messages, b-board, lists, and forums (fora). Perhaps the least directly interactional subtype is e-mail. Although it can be instantly transmitted, it is frequently stored for later attention. In fact, feedback is never guaranteed. Its increasing popularity for short messages lies in the very fact that, like traditional ("snail-") mail, it does not require attention in real time but can be stored. E-mail can be sent to one recipient or to a mass mailing list, even on a prearranged schedule.



In contrast, e-messages (the subject of this study) are typically not stored. Users are usually logged on simultaneously in separate locations. The interchanges scroll in real time at various rates across the screen (according to baud rate). Murray's (1988) study showed professional computer users' preference for this variety, which is the least planned and most heavily interactive type of CmC. Technological advances now allow full-screen interactive e-messages to be sent rather than the one line (120 characters) which was transmitted when the user hit the "enter" key in Murray's study.

Bulletin boards or b-boards serve as public electronic announcements, often having wide and unconstrained circulation. Lists facilitate special-interest discussion groups while forums (fora) support interactive chatting. Based on wide distribution lists, they do not require simultaneous log-on but allow users to contribute commentary sequentially over time. Kiesler et al. (1984) examined both e-mail and fora, while Maynor (in press) examined e-style, the type of language present in BITNET exchanges (an e-mail network that allows information exchange spanning time and space). Wilkins (this issue) provides an analysis of language behavior in a public conferencing network on a commercially available electronic communication utility by a relatively homogeneous group of discussants. Wilkins's study is especially valuable in that it attempts to control for "topic." As Rubin (1984, p. 272) and others have observed, composition researchers have recognized for some time the importance of controlling the topic of writing in empirical studies.

While the various types of CmC appear to differ, they have several important characteristics in common. Each requires written communication using a keyboard. Typing is by nature a slower process than speaking but, for many people, faster than writing longhand. The faster rates have undoubted effects on the composing process and the ability of writers to revise or edit. Each type is rendered in text only, that is, visually, without auditory, tactile, or other sensory information. All five types of CmC lack the full range of paralinguistic cues, providing few if any nonverbal or social clues because interactants are not visually or auditorially present. In IWD in particular and CmC in general, acquaintance with interlocutors and facts of identity are nonessential matters with profound implications for the future. All such types of CmC disregard distance as a barrier. Written communication is possible even with those physically separated in time and space.



## THE STUDY

### Data Collection: The Corpus

The present study examines the general features, including syntactic and stylistic dimensions, of one type of CmC: extended dyadic written communication through simultaneously linked computer terminals. This type of online terminal-to-terminal communication is increasingly common in commercial, government, and academic institutions as well as in small and large companies where employees send real-time messages to persons in separate buildings and receive rapid responses through simultaneously linked networks. We designate this language variety as Interactive Written Discourse.

Three factors make the present study unique. Sharing the assumption that audience is a shaping factor in the production of discourse (Bell, 1984), the conviction that topic should be held constant in textual research (Rubin, 1984), and the belief that professional computer users constitute an influential discourse community (Kiesler et al., 1984, p. 1125) with the potential for establishing conventions of use, we sought to constrain the data collection to a single topic, a single recipient, and a single discourse community. The study selected the travel information domain and the making of travel plans as the central topic for conversation. This topic had the benefit of being common enough to not require extensive prestudy training. Travel planning—the booking of flights, hotel reservations, and ground travel—is a familiar activity. Just as in the oral travel dialogues studied by Coupland (1980), subjects asked questions and tried to book flights, hotels, and rental cars. They also asked about prices, suitable times for flights, the relative quaintness of lodgings, and conversion rates of international currency. The corpus is part of a large ongoing study concerned with the dynamics of information retrieval based on natural language queries from users (see Brunner, Whittemore, Ferrara, & Hsu, 1989, for additional details).

### Subjects

The audience for all of the discourse from the 23 subjects was the same person, Tom C., a reference librarian and online data base specialist at Microelectronics and Computer Technology Corporation (MCC), a large computer research consortium in Austin, Texas, where

the data were collected. Because one goal of the large ongoing study was to examine natural language queries and answers, someone who was not a professional travel agent but had strong skills in electronic data access was selected.

The 23 subjects for the study were all MCC employees or their spouses who volunteered to participate in the study. Their occupations varied from technical to administrative. All had extensive prior experience with computers and keyboard text entry, and all had very sophisticated exposure to advanced computers and advanced computer concepts. Thus, in terms of knowledge of and experience with using computers, they constituted a relatively homogeneous discourse community. Participants ranged in age from 26 to 56 years.

### Experimental Procedure

The study, like those of Guindon and Schuldborg (1986), Guindon and Sladky (1985), Guindon, Sladky, Brunner, and Conner (1986), and Hill (1987), was based on a Wizard of Oz paradigm, a technique in which a confederate of the researchers, the wizard, simulates a target data base system (or virtual information server) in a written dialogue with a subject (Belkin 1986; Bobrow et al., 1977). The written protocols are then stored for use as a corpus for analysis.

Tom, the wizard for this study, performed five functions for the user:

1. Interpreted the user's request for travel information
2. Selected the appropriate information source for each request
3. Performed the necessary retrieval
4. Selected the subset of the returned data appropriate for the user request
5. Presented the information requested by the user

His work station was connected via TELENET to a Sabre dial-in travel services data base (used by American Airlines), which provided information on international reservations at hotels, as well as schedule information and fares for flights and auto rentals. The wizard also had copies of the *Hotel and Travel Index* (1986) and *OAG European Travel Planner* (1986), which contained information on smaller, lesser known hotels and airfields in book form.

Because the 23 subjects were all computer research firm employees or spouses, no attempt was made to convince them that they were actually dealing with a next-generation advanced computer system. Instead, they were told, ambiguously, that they would be soliciting travel information from a "human-assisted computer system."

Prior to participating, each subject was mailed a brochure for the 10th World Computer Congress to be held in Dublin, Ireland, with instructions to think about an itinerary for a 2-week vacation to be taken anywhere either before or after the Congress. In thinking about the trip, subjects were asked to plan as they normally would for travel, using the kinds of budgetary or other constraints that they would regularly adhere to. Most chose side trips to European destinations such as England, Scotland, Spain, or France (see Brunner et al., 1989, for further details of the experimental conditions that applied).

For the study, the subjects and the wizard were seated at terminals in separate rooms. Each of the individual sessions lasted about 2 hours, with the last 15 minutes reserved for completing a satisfaction questionnaire and undergoing debriefing. Subjects, who were told to use the travel information to develop a specific itinerary and make bookings, were videotaped. During the individual sessions, an assistant prompted them to "think out loud" while they were ruminating.

## Apparatus

The CRT display for the subjects was divided into two 12-line areas; outgoing messages were automatically typed into the upper area, and incoming messages were displayed in the lower half of the screen. Such display formatting allowed each individual to see information from the last response while simultaneously composing the current query or response. Subjects pressed the "transmit" button to send a message.

Text transmission between the subject and the wizard was controlled by a VAX 750 running text exchange software. The subject was seated at a DEC VT 220 terminal. The wizard was seated at a Sun 2/50 work station with a display surface subdivided into two windows, one for written dialogue with the subject and one for access to the Sabre dial-in travel services data base.

## FINDINGS

### Dialogue Structure and General Characteristics

Largely because of the slowness of the Sabre system, 2 hours of written interactive dialogue between the wizard and each subject yielded an average of 79.5 sentences per travel dialogue. On average, these sentences were spread across 61.9 messages. A message consisted of a turn before the "enter" button was pressed. Messages were brief, averaging approximately 1.28 sentences apiece. An average of 30.9 exchanges between the wizard and each subject occurred. An exchange consisted of a two-pair part, like query-reply or statement-response. Only the written communication of the 23 subjects is examined here (see Brunner et al., 1989, for discussion of the wizard's language use). Subjects' sentences contained, on average, 8.95 words, somewhat shorter than the 12.8 words per sentence found for standard fiction by Finegan (1982).

These findings are informative as to the quantity of discourse and tend to indicate a general preference for surface brevity. Before discussing the constituent nature of Interactive Written Discourse or describing the conventions which typify this hybrid register, we will provide some background to the concept of register and the relationship of IWD to other registers.

### Structural Similarities to Other Registers

The central idea behind the concept of "register" is that the structural properties of any register follow from the circumstances of use. A characteristic of registers such as Baby Talk, Foreigner Talk, and Note-taking is that they frequently omit copulas, articles, pronouns, and the auxiliary *do*, a fact that led Ferguson (1982) to claim that they are "simplified registers."<sup>3</sup> However, Trudgill (personal communication) suggested that "reduced register" is a more appropriate label because no consensus exists on what constitutes simplicity in language. Nonetheless, the present study shows that IWD shares four characteristics with Note-taking (Janda, 1985), the other written register that is produced under real-time constraints and appears to involve reduction of some sort:

1. Omission of (unstressed) pronouns, especially subject pronouns

2. Omission of articles, both the definite and indefinite
3. Omission of finite forms of the copula
4. Shortening of words through use of abbreviations and symbols

Like several registers, including Sports Announcer Talk (Ferguson, 1983), IWD displays a type of prosiopesis or omission of expected sentence initial material. For example, the expected subject is absent in the sentence "Want time." One of the two most prominent characteristics of IWD is omission of subject pronouns, as in the following examples where absence of an expected feature is noted by the null sign Ø and all are shown as typed, with no corrections for standardized capitalization, punctuation, or spelling:

1. Ø need hotel reservations for 2 in London, august 24 to august 27, moderate price range.
2. Ø requesting reservations for the 10th World Computer Congress beginning 8/30/86 through 9/6/86, Dublin, Ireland, basic package, category B, Gresham Hotel, double room occupancy, \$380.
3. Ø want a reservation at a hotel in the center of London.
4. Ø need reservation for 1 at Trinity College, Dublin for Ifip conference

A few subjects omitted both subject pronouns and the accompanying auxiliary (e.g., finite forms of *have* or *to be*) in the same sentence:

1. ØØ returned from break
2. ØØ taking Ø 5 minute break

In the present corpus, one quarter of the subjects' displayed subject pronoun deletion; such deletion was found in 26% (6) of the dialogues. In interactive situations, pronominal information is recoverable from context, and linguists have found that some languages, such as Japanese and Spanish, have what Chomsky and others called the pro-drop parameter, in which subject pronouns are not obligatory. Users of IWD also find subject pronouns dispensable,<sup>4</sup> and the omitted pronouns are exactly those that pro-drop languages omit.

A second and even more frequent characteristic of Interactive Written Discourse is omission of articles, both definite and indefinite. Keyboard dialogues frequently showed deletion of *a*, *an*, and *the*. One out of 3 subjects omitted these determiners. The corpus shows that 35% (8) of the participants favored this apparent economy of effort. Because some languages (e.g., Latin) have no indefinite article, and

other languages (e.g., Chinese and Korean) lack articles altogether, we know that articles are not necessary for a human language, and it is therefore not surprising that they disappear in IWD since definiteness or indefiniteness can be inferred from context. The examples below from the corpus involve deletion of articles:

1. Did you receive Ø message asking about Ø location of Ø Westbury hotel?
2. verify if Ø \$429 Aer Lingus fare is Ø round trip fare
3. can you tell me if Ø hertz compact car is available and Ø price?
4. I would like to stay near Trinity College where Ø conference is being held.

Copula deletion, omission of finite forms of the verb *to be*, also occurred, as in the following first two examples. Furthermore, such omission often co-occurred with omission of existential *there* or an article:

1. What flights ØØ from London to Dublin on August 30 after 12 noon?
2. how much plane change time ØØ at JFK?
3. What ØØ price

The copula was missing in the dialogues of 27% (5) of the subjects. Not surprisingly, linguists have found that copula forms of *to be* are also frequently dispensed with by languages.

Contractions, informal spellings (e.g., *nite* for *night*), and clippings (e.g., *info* for *information*) as well as the use of symbols for words (/ for *per*; *pm* for *afternoon*, and *&* for *and*) also appear to involve reduction or simplification since they require fewer keystrokes. The following are some examples from the corpus:

1. I'd like some *info* on your charming, small hotel.
2. I would like to reserve a room in a hotel in Salzburg for the *nite* of Sept. 8 single occupancy, cost should be around \$75/*nite*.
3. provide departure information for *pm* of 9/20/86
4. I would like the 800 #, please.

Such written shortcuts were present in the data for 30% (7) of the subjects. The frequent use of dashes, signifying a type of ellipsis, also characterized the data.

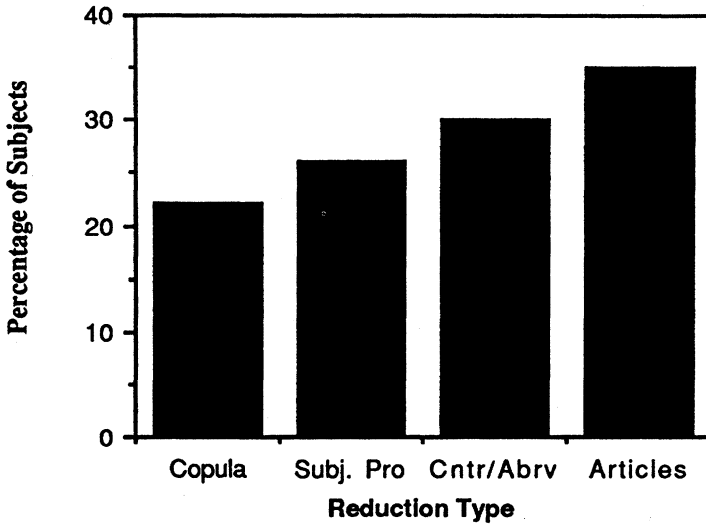


Figure 1: Frequency of Type of Reduction in IWD

A profile of the rates of reduction is given in Figure 1.

These reductions and omissions lead us to consider IWD, like Baby Talk, Foreigner Talk, and Note-taking, as a "reduced register." At issue is the question of why these particular characteristics occur if situation determines or influences the type of language. When we combine our knowledge of the nature of language universals and what is considered dispensable with what we know about the pragmatics of the context of use for computer-mediated communication, we can easily account for the omissions and shortcuts detailed earlier.

An important issue for future research to explore is whether such reductions stem from the belief that the discourse recipient (like a baby or a foreigner) is somehow linguistically less competent, or whether (as in note-taking) real-time production constraints compel users to omit items for the sake of speed. The answer may well depend on further refining the distinction, as yet little understood, between simplification and reduction.



## Similarity of IWD to Spoken and Written Language

Over the past decade or two, one question that many students of language have been concerned with is the relationship between spoken and written language. The literature contains a variety of positions: dichotomy, bipolar continuum, interacting continua, and—the most sophisticated—Biber's (1986, 1988) statistical analysis of text type involving multidimensional scaling.<sup>5</sup> As an emergent language variety, IWD has not been included in previous comparative studies of text types and thus offers a unique opportunity to examine again the relationship between written and spoken language. We can ask, "Is it more like spoken or written language or is it like neither?" "Is it an intersection?" While this type of communication occurs in a typed medium and could be expected to resemble other written language by being elaborated and expanded, it is clearly also interactive (defined as involving give and take, the sending and receiving of information, and occurring in real time). According to Biber (1986), the linguistic correlates of interactive language use are (a) use of first- and second-person pronouns (*I/You*), (b) presence of WH questions (e.g., What flights . . . , Which hotels . . . , When . . . , and (c) presence of yes/no questions (e.g., Do you have any hotels in the palace area for under \$60/nite?).

The highly interactive nature of the present dialogues can be demonstrated by observing that the words *you* and *I* were among the 10 most frequently occurring vocabulary items in the 1,478 unique words and 18,769 terms present in the corpus. These interactive pronouns ranked fifth and sixth in the corpus, constituting together 5% of the lexicon. Similarly, the query-based format of the dialogues yielded a high number of both WH and yes/no questions.

Although it is commonly thought that "involvement with one's audience . . . is lacking in any kind of writing" (Chafe & Danielewicz, 1987, p. 110), the language of terminal-to-terminal exchanges such as IWD shows features of heavy involvement (e.g., deictics, adverbs of time, and direct questions) traditionally associated with oral language and face-to-face interaction. As Nystrand (1987) observed, one begins to suspect that characteristics of speech are due to its immediacy, evanescence, and interactive nature rather than to the fact that it is produced in the oral mode. As even written communication, now that computer technology has advanced, can duplicate immediate response, fading over time, and back and forth messages, the old distinction between written and oral language becomes less viable.

An important finding of this study is that computer-mediated human-to-human written discourse, an emerging hybrid variety of language, displays features of both oral and written language.

In addition to the presence of first- and third-person pronouns (Blankenship, 1974; Chafe, 1982; Chafe & Danielewicz, 1987) and direct questions (Marckworth & Baker, 1974), the use of general (informal) emphatics (e.g., *just*, and *real*; Chafe & Danielewicz, 1987) is an oral characteristic of IWD, as seen in the following:

1. Please give me the name of a large hotel, which is not likely to be *real* strict in its cancellation policy. Sorry, give me 2 or 3 names with nightly rates for a single room.
2. How about *just* the cheapest hotel?

Observe the question format, the use of informal discourse particles, such as *okay*, *sure*, *sorry*, and *now*, and colloquial usage, such as *Let's* and *How about*, in the following extracts:

Subject: Which hotel is closest to the conference location?

Wizard: Where is the conference?

S: *Sorry*, the conference is located on the campus of Trinity College.

S: *Let's* take the Aer Lingus flight at 8:55A.

W: The reservation will be made for you.

S: *Now*, about the hotel near Prestwick. Do you have any information.

W: Do you have a particular hotel in mind?

S: No-I asked you for recommendations.

W: There is only one hotel in Prestwick.

S: What can you tell me about the only hotel in Prestwick?

W: There is the Carlton at \$27-44 for a double.

S: *How about* a nice hotel in Glasgow instead?

At the same time that informal spellings and informal usage abound in the data, occasional examples of very formal language also occur:

Is that fare available on the day I *wish* to travel?

I *wish* to spend a vacation in Portugal and Spain. I want to fly to Lisbon on Saturday, rent a car and drive into Spain and then leave for Austin. *Let us* start with Lisbon. What flights leave Dublin Saturday morning for Lisbon (Sat, Sept. 6)?

Careful examination also shows that IWD exhibits features of written language in that it is elaborated and expanded. It shows frequent use of relative clauses, adverbial clauses, and subordination. For instance, in the following example, the relative clause is a lengthy 32 words:

Is there another flight *that would allow us to leave at a later time from Istanbul on Sept. 20 and stop overnight at some nice place in Europe, continuing our journey to Austin on Sept. 21?*

Cataphora, or forward reference, is considered in standard grammars to be characteristic of written language (cf. Martin, 1983, p. 12; Quirk, Greenbaum, Leech, & Svartvik, 1980). Cataphoric reference was quite common in the corpus. Phrases such as "the *following* criteria," "the *following* two airlines," or "the *below mentioned* connections," which required linking information in subsequent (rather than prior) discourse, were used by about one fifth (22%) of the computer professionals or their spouses. The following are examples from the corpus that resemble written language with respect to their use of cataphoric references:

book me on *the below mentioned* flight in an aisle seat in the non smoking section.

I want you to find me the cheapest airfare for *the following* trip: Leave Austin on the 28th of August, stop in Washington for a couple of days, fly to Dublin on the 30th, and return to Austin from Copenhagen about 3 weeks later.

IWD presents a challenge to Biber's (1986) characterization of text types, which posited that one of three dimensions that distinguish texts is the dimension Interactive versus Edited, because IWD is both interactive (involving sending and receiving of messages in real time) *and* edited, in that before sending a message, users can peruse their typed input and emend it. Ample textual evidence from this corpus reveals this editing, both in the use of parentheses to add or change material and in self-correcting phrases, such as "Rather" or "Sorry," as in the following examples:

Subject: I also need hotel reservations from 8/30/86 (*guaranteed late arrival*) to 9/6/86.

S: Give me a listing of hotels in London that meet the following criteria about \$50-60 per night, per room (*for two people*), one double bed is ok in the Imperial College or Trafalgar Square area vacancies on Sept. 5th, 6th, 7th.

S: I would like to fly back to Austin from Munich the following morning, Sept. 10.

(57 seconds elapse)

S: *Rather*, Sept. 10 for the hotel room in Munich and flight home on Sept. 11.

S: What will I be flying on for theses flights?

(1 minute 17 seconds later, before intervening message is received)

S: *Rather* what airlines will I be flying for these flights?

Thus IWD may be the first type of language use to be studied that is both edited *and* interactive.

In the last interchange above, the subject apparently noticed not only the absurdity of his question (which could have been answered sarcastically, "An airplane, dummy") but also his typing errors. Typing mistakes occurred in 2.6% of the items in the corpus (488 misspellings out of 18,769 terms). However, as will be shown, users showed a preference for retyping the misspellings they noticed rather than deleting the error and erasing any trace of their mistake. This willingness of users to permit others to see their mistakes gives their communicative partners a seldom equaled glimpse into communicators' composing misfires; in most other types of written communication, first drafts do not see light. This "first draft quality" appears to be an accepted convention of IWD.

The directly preceding extracts provide overt linguistic evidence of the editing process, but typed computer data are always editable before the sender presses "enter." Paralinguistic evidence from videotapes suggests that most people briefly scan or reread their message before pressing "return" or as they press "enter" or "transmit." The following example demonstrates an online correction of a typing error and also shows the use of parentheses (the written equivalent of lowered pitch and loudness) to represent asides.

Subject: What is the cosat of the next larger room (or suite)? (The word is "cost.")

Parentheses were used by 22% (5) of the participants. Those who availed themselves of parentheses did so repeatedly. This finding is consistent with Wilkins's (this issue) observation that 11% of primarily novice computer users took advantage of this strategy to compensate for the paralinguistic cues of spoken language. Maynor (in press) also noted parentheses, dashes, and trailing dots as written equivalents of paralinguistic clues.

The findings of this study also support Murray's (1988, p. 364) claim that there was no evidence of flaming, the use of inflammatory language such as swearing or insults, which Kiesler et al. (1984) identified. No inflammatory language occurred, and, as Murray (1988, p. 364) found, some language was actually tempered. For example, in the present corpus, participants used phrases such as "if you like," and the word "please" was among the 20 most frequently occurring words in the corpus, ranking 19th.

We can conclude that IWD defies characterization as belonging on either end of a hypothetical oral/written continuum, and that, unlike any genre studied by Biber (1986, 1988), it is both interactive and edited, both interactive and written. It is a language variety that never existed before. IWD demonstrates both oral and written characteristics.

### CONVENTIONS IN TRANSITION

We turn now to questions of the acquisition of the norms of IWD and the question of diachronic change. As Murray (1988) reported, conventions of use for computer-mediated communication are still being formed. However, it is inaccurate to presume, as contended by Kiesler et al. (1984, p. 1126), "the absence of norms governing the social interaction" while norms are in the process of being established. Language is dynamic, ever changing, and no doubt the norms of the present will yield to others as technology advances. Nowhere is there a clearer indication of the norms of IWD in transition than in practices of capitalization. Participants in Interactive Written Discourse are unsure whether to capitalize or use lower case for sentence beginnings and proper names. As Murray (1988, p. 353) noted, formerly systems converted all characters into upper case so that even messages composed in lower case were received in upper case. Similarly, messages were previously restricted in size, sometimes to one line. With technological advances, these situations have changed, and, as in the

current corpus, both capital and lower case letters are readily transmittable. Subjects in this study had access to both lower and upper case, and the wizard in this study communicated with them using standard written practice. Yet examination of the practices of all 23 subjects shows a normative influence that we will describe. Capitalization was examined separately in both sentence-initial material (first word of a sentence) and in the first letter of proper names. There was abundant opportunity to examine proper names since 22% of the nouns in the written travel dialogues were proper names (e.g., of hotels, cities, and airlines).

Thirty percent of the subjects were consistent in using standard capitalization for both sentence-initial words and proper names. However, the profile shows a difference between these two types of capitalization for the remaining 70%. There appears to be a residual norm of using lower case for sentence-initial material. About two thirds of all subjects (15) used lower case for sentence beginnings at least some of the time. Of all subjects, 39% (9) consistently used lower case for sentence beginnings, while 26% (6) were ambivalent, switching back and forth in an apparently inconsistent manner. Only 35% (8) used standard capitalization for sentence beginnings. The high incidence of lower case at the start of a sentence may well indicate a convention that will prevail. Such a norm may be favored because one or two less keystrokes (shift or "caps lock") saved may give the essence of rapid transmission of messages. Registers of all types evolve over time, often with historical remnants intact (cf. the seemingly redundant lexical couplets, such as "aid and abet," "give and bequeath," and "cease and desist," in legal and bureaucratic registers, Charrow, 1982).

Nonetheless, the data showed more conventional practices where proper names were involved. Of the 23 subjects, 7 (30%) were inconsistent, switching back and forth, and 3 (13%) consistently used lower case for the initial letters of proper names. Over half, 57% (13), consistently maintained standard capitalization for the initial letters of proper names, indicating for some reason, less inclination to abandon capitals for names than in the first letter of a sentence. The differences between the two types of capitalization are shown in Figure 2.

It would be misleading not to mention the vast individual differences found among subjects. This variation provides additional evidence that norms are still being formulated. Three subjects used terser commands than did the other subjects, apparently modeling their discourse on a type of telegraphese. For example, the mean sentence

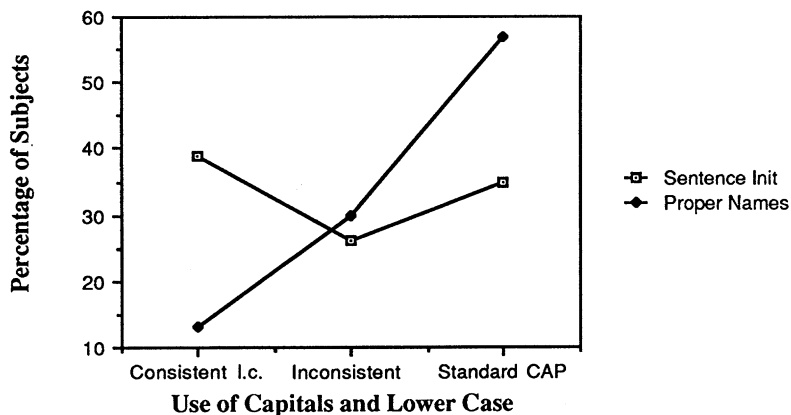


Figure 2: Norms of Capitalization in Transition

length of 1 subject (S4) was 2.35 words per sentence. This terse style contrasts sharply with that of other subjects, such as S5, who used deeply subordinated, elaborate syntactic structures and had a mean sentence length of 10.3 words. An example of this individual difference can be seen in the following extracts in which the two subjects, S4 and S5, are conducting essentially the same business: checking which flights are available and asking about their prices:

Subject 4 (Terse)

S: what flights ØØ from London to Dublin on August 30 after 12 noon?

W: There is a Dan-Air flight that leaves Gatwick at 130p and arrives in Dublin at 240p.

S: price?

W: Round trip London to Dublin on Dan-Air is \$127.00

S: other flights?

W: The only other direct flight from Gatwick leaves at 800p and arrives in Dublin at 915p. It is an Aer Lingus Flight.

Subject 5 (Elaborated)

S: i need information on plane travel to Dublin Ireland on 8/30/86. I will be traveling alone and will leave from Austin, Tx. What flights are available in an economy class leaving in the morning?



W: There is a TWA flight that leaves Austin at 1120a arrives at JFK at 500p, which connects to an Aer Lingus flight that leaves JFK at 645p and arrives in Dublin at 725a.

S: Are there other flights available?

W: There is a United Flight that leaves Austin at 805A arrives in Chicago at 1029a which connects to a Northwest flight that leaves at 240p and arrives in Dublin at 845a.

S: What are the costs for these two routes?

This individual variation reinforces our notion of Interactive Written Discourse as an emergent language variety. Ferguson (1983) observed that some registers are "learned differentially through the speech community . . . not explicitly taught, but picked up in an unplanned, 'natural' manner" (p. 169). Clearly, as more people use a register, social groups develop shared norms for the variety of language appropriate for different occasions. Rubin (1984) stated that "writers are guided by sociolinguistic norms that associate stylistic variables with contexts by force of convention" (p. 226). We can argue that the force of convention contributes to the retention of sentence initial lower case as an indicator of register. This convention, now no longer technologically required, was introduced by early users of computers, such as the professionals in this study.

In response to constraints on time, memory, and general effort, those who engage in IWD often use syntactically reduced forms, abbreviations, shorthand symbols, and terse phrasing, possibly modeling such features of IWD on those of other registers with severe constraints on space, such as telegraphese, headlineese, or the postcard register. In discussing the ontogenesis of register, Martin (1983, p. 36) argued that registers grow out of adaptations of previously developed systems. If such is the case, as we believe, then IWD can be seen as a hybrid register whose structural properties follow from circumstances of use.

## CONCLUSION

In human language, processes of conventionalization are always operating as social groups develop shared norms of the varieties of language appropriate for different occasions and as individuals ac-

quire and modify these norms (Ferguson, 1983, p. 170). We suggest that study of Interactive Written Discourse offers an opportunity to examine what Labov called language change in progress.

The study of a new register such as IWD furthers our understanding of systematic language variation across social and situational dimensions and supports the claims of Ferguson (1982) that registers do exist, are subject to diachronic change, and that competence in register variation is a normal part of language development.

Interactive Written Discourse is a novel writing style that appears to be an emergent register, a language type shaped by situation and context. Empirical examination of a corpus of written dialogues between 23 experienced computer users and the same travel advisor linked through real-time computer networks shows that IWD contains some syntactic features of reduction, such as omission of articles, subject pronouns, and copula, and that conventions of usage (e.g., use of sentence-initial lower case and parentheses) are in transition. Like all registers, the edges of IWD overlap those of other registers, such as Baby Talk and Sports Announcer Talk. The similarities between another written register favoring surface brevity, the Note-taking register discussed by Janda (1985), were noted. However, the interactive nature of this novel form, with the copresence of an unseen audience, the opportunity for immediate feedback, and the evanescence over time, make this type of written communication unique. This variety characteristically displays features of both written and spoken language, suggesting that it is a hybrid. We further claim that it arises not *de novo* but as an amalgam of components taken from other language varieties.

Interactive Written Discourse is a historically unique juxtaposition of text format with real-time interaction pressures. This form of written communication is relatively undocumented but undoubtedly functionally tailored to context of use. There is ample reason to believe that this type of written communication will proliferate in the future, perhaps even influencing or outdating the stylistic conventions of traditional writing styles. What is certain is that the technological advances of computer networks that have permitted IWD also afford the opportunity for simultaneously creating, editing, and revising documents for people separated by large distances. The pedagogical implication for work with the mobility impaired or those who have

difficulty communicating orally (e.g., the deaf and stutterers) as they seek to become better writers is enormous. For these reasons, a study of computer-mediated communication in general and Interactive Written Discourse in particular is of interest to researchers, writers, and teachers alike. The focus of this study has been on language as code, but we acknowledge that to gain a full understanding of the changes that computer-mediated human-to-human communication may bring to both the way we think about the possibilities for communication and the ways we use written language and compose in it, the perspectives and analytical approaches of many disciplines will be required.

## NOTES

1. Ultimately researchers will want to distinguish carefully between typed computer dialogues *through* a computer, that is, between human beings, and typed computer message *to* a computer system, that is, between a human being and a computer system. This distinction will become more important as technology advances to the stage that human-machine interface in extended contexts becomes a reality. It is becoming increasingly more common to talk to machines, and the next leap is for machines to talk back. Workers in artificial intelligence have actually devised systems that can sustain conversations in a limited domain, but these exchanges are not extended discourse.

2. B. Johnstone has brought to my attention the related and no doubt widespread phenomenon of interactive note writing by professional colleagues at conferences. The copresent audience and potential for immediate response in written form make this type of language use suitable for inclusion under the label IWD. To my knowledge, no empirical studies of this text type exist.

3. Not all registers are "reduced registers"; some are elaborated, involving extensive syntactic subordination. Bureaucratic language (Charrow, 1982), testament language found in wills (Finegan, 1982), and a learned lecture style, which Neal (1990) called "academic register," are examples of elaborated registers.

4. Compare the practice of deleting the subject pronoun in postcardese: Ø Having a good time. Ø Wish you were here. Love, Sandy.

5. Biber's (1986, 1988) comprehensive factor analysis examined the distribution of 41 linguistic features in 545 text samples from 16 major text types. These included press reportage, mystery and detective fiction, *belles lettres*, and academic prose as well as conversations, broadcasts, and public speeches. His 1986 study identified three dimensions, tentatively labeled as Interactive versus Edited Text, Abstract versus Situated Context, and Reported versus Immediate Style. These labels were emended somewhat in Biber's 1988 work.

## REFERENCES

- Belkin, N. J. (1986). Designing an intelligent information system interface. In *Intelligent design aids in process environments*. New York: Springer Verlag.
- Bell, A. (1984). Language style as audience design. *Language in Society*, 13, 145-204.
- Biber, D. (1986). Spoken and written textual dimensions in English. *Language*, 62, 384-414.
- Biber, D. (1988). *Variation across speech and writing*. Cambridge: Cambridge University Press.
- Bizzell, P. (1982). Cognition, convention, and certainty: What we need to know about writing. *Pre/Text*, 3, 213-243.
- Blankenship, J. (1974). The influence of mode, submode, and speaker predilection on style. *Speech Monographs*, 41, 85-118.
- Bobrow, D. G., Kaplan, R. M., Kay, M., Norman, D. A., Thompson, H., & Winograd, T. (1977). GUS: A frame-driven dialog system. *Artificial Intelligence*, 8, 155-173.
- Brunner, H., Whittemore, G., Ferrara, K., & Hsu, J. (1989). *An assessment of written/interactive dialogue for information retrieval applications* (MCC Tech. Rep. No. ACT-HI-245-89). Austin, TX: Microelectronics and Computer Technology Corporation.
- Chafe, W. (1982). Integration in involvement in speaking, writing, and oral literature. In D. Tannen (Ed.), *Spoken and written language: Exploring orality and literacy* (pp. 35-53). Norwood, NJ: Ablex.
- Chafe, W., & Danielewicz, J. (1987). Properties of spoken and written language. In R. Horowitz & S. J. Samuels (Eds.), *Comprehending oral and written language* (pp. 83-113). New York: Academic Press.
- Charrow, V. P. (1982). Language in the bureaucracy. In R. J. DiPietro (Ed.), *Linguistics and the professions* (pp. 173-188). Norwood, NJ: Ablex.
- Coupland, N. (1980). Style shifting in a Cardiff work setting. *Language in Society*, 9, 1-12.
- Faigley, L. (1985). Nonacademic writing: The social perspective. In L. Odell & D. Goswami (Eds.), *Writing in nonacademic settings* (pp. 231-248). New York: Guilford.
- Ferguson, C. A. (1977). Baby talk as a simplified register. In C. E. Snow & C. A. Ferguson (Eds.), *Talking to children: Language input and acquisition* (pp. 209-235). Cambridge: Cambridge University Press.
- Ferguson, C. A. (1981). "Foreigner talk" as the name of a simplified register. *International Journal of the Sociology of Language*, 28, 9-18.
- Ferguson, C. A. (1982). Simplified registers and linguistic theory. In L. K. Obler & L. Menn (Eds.), *Exceptional language and linguistics* (pp. 49-66). New York: Academic Press.
- Ferguson, C. A. (1983). Sports announcer talk: Syntactic aspects of register variation. *Language in Society*, 12, 153-172.
- Ferguson, C. A. (1985). Special language registers. *Discourse Processes*, 8, 391-393.
- Finegan, E. (1982). Form and function in testament language. In R. J. DiPietro (Ed.), *Linguistics and the professions* (pp. 113-120). Norwood, NJ: Ablex.
- Goody, J. & Watt, I. (1968). The consequences of literacy. In J. Goody (Ed.) *Literacy in traditional societies* (pp. 53-84). Cambridge: Cambridge University Press.
- Guindon, R. & Schulderberg, K. (1986). *Grammatical and ungrammatical structures in user-advisor dialogues: Implications for the design of natural language interfaces* (MCC Tech. Rep. No. HI-86-P). Austin, TX: Microelectronics and Computer Technology Corporation.

- Guindon, R. & Sladky, P. (1985). *Analysis of user/expert dialogues: Task networks and subdialogue markers* (MCC Tech. Rep. No. HI-084-85). Austin, TX: Microelectronics and Computer Technology Corporation.
- Guindon, R., Sladky, P., Brunner, H., & Conner, J. (1986). The structure of user-advisor dialogues: Is there method in their madness? *Proceeding of the 24th Meeting of the Association for Computational Linguistics*.
- Halliday, M.A.K., McIntosh, A., & Stevens, P. (1964). *The linguistic sciences and language teaching*. London: Longman.
- Hill, W. (1987). *Two hidden operator studies of advice giving for a direct representation interface* (MCC Tech. Rep. No. HI-169-87-P). Austin, TX: Microelectronics and Computer Technology Corporation.
- Hotel and travel index*. (1986) New York: News America Publishing, Inc.
- Hudson, R. A. (1980). *Sociolinguistics*. Cambridge: Cambridge University Press.
- Janda, R. (1985). Note-taking English as a simplified register. *Discourse Processes*, 8, 437-454.
- Kiesler, S., Siegel, J., & McGuire, T. W. (1984). Social psychological aspects of computer-mediated communication. *American Psychologist*, 39, 1123-1134.
- Marckworth, M. L., & Baker, W. J. (1974). A discriminant function analysis of covariation of a number of syntactic devices in five prose genres. *American Journal of Computational Linguistics*, Microfiche 11.
- Martin, J. R., (1983). The development of register. In J. Fine & R. O. Freedle (Eds.), *Developmental issues in discourse* (pp. 1-40,) Vol. 10 in the series *Advances in Discourse Processes*, R. O. Freedle, Ed.). Norwood, NJ: Ablex.
- Maynor, N. (in press). The language of electronic mail: Written speech? In M. Montgomery & G. Little (Eds.), *Centennial usage studies* (Publications of the American Dialect Society Series). University: University of Alabama Press.
- Murray, D. E. (1985). Composition as conversation: The computer terminal as medium of communication. In L. Odell & D. Goswami (Eds.), *Writing in nonacademic settings* (pp. 203-228). New York: Guilford.
- Murray, D. E. (1988). The context of oral and written language: A framework for mode and medium switching. *Language in Society*, 17, 351-373.
- Neal, M. (1990). *Academic register*. Unpublished manuscript, Department of English, Texas A&M University.
- Nystrand, M. (1987). The role of context in written communication. In R. Horowitz & S. J. Samuels (Eds.), *Comprehending oral and written language* (pp. 197-214). New York: Academic Press.
- OAG European travel planner*. (1986). Oak Brook, IL: Dun & Bradstreet, Inc.
- Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1980). *A grammar of contemporary English*. Harlow: Longman.
- Reid, T.B.W. (1956). Linguistics, structuralism and philology. *Archivum Linguisticum*, 8, 28-37.
- Rubin, D. L. (1984). The influence of communicative context on stylistic variations in writing. In A. Pellegrini & T. Yawkey (Eds.), *The development of oral and written language in social contexts* (pp. 213-232, Vol. 13 in the series *Advances in Discourse Processes*, R. O. Freedle, Ed.). Norwood, NJ: Ablex.
- Straumann, H. (1935). *Newspaper headlines: A study in linguistic method*. London: Allen & Unwin.

Tannen, D. (1982). The oral/literate continuum in discourse. In D. Tannen (Ed.), *Spoken and written discourse: Exploring orality and literacy* (pp. 1-16). Norwood, NJ: Ablex.

*Kathleen Ferrara is Assistant Professor of Linguistics in the Department of English at Texas A&M University. Her research is in the areas of discourse analysis and sociolinguistics. She holds a Ph.D. in linguistics from the University of Texas at Austin.*

*Hans Brunner is Research Supervisor at U.S. West Advanced Technologies in Englewood, CO. A psychologist, his work is in the area of natural language processing. He received a Ph.D. from the University of Colorado.*

*Greg Whittemore is a Senior Researcher at Electronic Data Systems in Albuquerque, NM, specializing in ambiguity in natural language. He received a doctorate in linguistics from the University of Texas at Austin.*