Features of English in CMC and Their Implications for Language Learning

SU-HSUN TSAI

Thesis submitted for the degree of Doctor of Philosophy (Ph.D.)
of the University of London

English
Institute of Education
University of London

March 2001

ACKNOWLEDGEMENT

Upon the completion of my Ph.D. study, I would first give my sincere thanks to my supervisor, Dr. Tony Burgess. His enlightening support and guidance have sustained my confidence and courage to carry out the whole thing under very heavy mental and material pressure. I should also thank Dr. Jennifer Jenkins and Mr. Martin Rich, the examiners of the thesis, and Ms. Anita Pincas, the examiner for the upgrading exam, for their valuable comments and suggestions, which are indispensable to the completion and maturity of the thesis.

I feel obliged to many other people who have helped my research in one way or another. Though I can not list all the names, but they should include at least Prof. Douglas Biber, Prof. Chung-tien Chou, Prof. Der-hsing Fan, Dr. Jane Hurry, Mr. Ming-chi Lan, Dr. Brian Lienard, Prof. Brian MacWhinney, Dr. Oliver Mason, and Mr. Chun-liang Wu. I learned a lot from several electronic discussion lists, TESLCA-L, NETEACH-L, well, Corpora, BNC, Linguist-L, info-childes, info-chibolt, and netskills. They are my favourite academic partners. I would thank all the participants in those lists.

I have shared my tears and laughter, complaints, and happiness with those friends who have been working in Room 836 with me. I won't forget Gertrude Shotte, Clare Bentall, Claudia Flores-Moreno, Grace Wang and Frank Wang. I look forward to the good news of their completing the studies soon. My appreciation is also due to John Egan, Lenny Bristow, Ade Dedeke, and Joe Anyemedu. As security guards at the Institute of Education, University of London, they ensured that I was safely protected even when I worked in the computer room till midnight.

My lovely friends should include the BBC Radio 4. Her rich programmes educated me far beyond my research. Every night she talked to me nicely before I fell into sleep after a long working day.

Last but not the least, my thanks and love go to my family. My dear children, Junior and Sarah, have been so brave and independent while their mother was working hard far away from home. My mother and my elder brother have done their best to provide me with mental and financial support. My deceased father was the one to inspire me with the desire to seek academic growth. This Ph.D. thesis is dedicated to him.

ABSTRACT

The similarities and differences between written and spoken forms of language have been a focus of interest of many scholars. There is agreement that instead of being a dichotomy or one single continuum, the differences between spoken and written forms can be measured along several dimensions. The coming into existence of computer-mediated communication (CMC) has made the line of distinction even less obvious. It is technically a writing (key-pressing) behaviour but may be used to carry out spontaneous communication.

This study is intended to investigate the special linguistic features of CMC versus non-CMC texts. The study adopts a corpus linguistic approach to analyse a host of 67 linguistic features in synchronous and asynchronous CMC genres and finds interesting differences in the use of these features when used in different temporalities of CMC contexts. A comparison of these features in CMC genres with those in non-CMC texts also reveals some special characteristics of language developed through the use of CMC.

The study suggests that, within the general development of CMC, there are emerging genres reflecting particular contexts. As CMC may soon become a major means of communication, and corpus linguistics is an innovative linguistic approach, awareness of CMC is likely to be of increasing importance for language learning. Some pedagogical suggestions are proposed from the experience and findings that have been gained.

Table of Contents

Chapter 3: Observation and Description of CMC Contexts, language Variation and Language Teaching

Chapter 4: Methods and Procedures of Investigation

Chapter 5: Findings and Discussion

Chapter 6: Pedagogical Implications

Chapter 7: Conclusion

List of Tables

List of Figures

List of Appendices

Chapter 1: Introduction

1.1 Background and Motivation for the Study

The motivation for the present study was largely stimulated by my interest in the many varieties of the English language, especially those that have developed as a result of the new technology of computer-mediated communication (CMC). With the continuous development of society and technologies, the contexts for communication are constantly changing, and new varieties of language use keep emerging. As a language teacher, I have a strong interest in exploring the characteristics of the English language as it is used in different contexts, and believe it is a major responsibility of mine to help students to learn how to best understand and use the language adequately in different circumstances. In the past decade, CMC has developed rapidly, and this led me to speculate that new varieties of the English language have emerged. Besides, as the use of the computer proliferates exponentially in the world, it seems certain that CMC will develop into a major form of communication in the new century. Therefore, the study of the language varieties associated with CMC is necessary and important.

Among the many studies on language varieties, a very large number concentrate on the variation between speech and writing. The similarities and differences between written and spoken forms of language have been the focus of many scholars. Some scholars have viewed orality and literacy as a dichotomy. The distinction, as has been argued, may be that writing is marked for its permanence (Olson, 1996), that writing is rational, while oral is emotional (Goody, 1986), and that written language is analytical and sequential, while oral language is less orderly (Ong, 1982). In these studies, certain social and cognitive consequences are believed to be derived from the development of literacy, including the ability to decontextualise knowledge, to create autonomous discourse and to think analytically.

However, other studies consider orality and literacy to be on a continuum. That is, one or two dimensions are established, along which the difference between spoken and written forms is measured. These dimensions cover word choice, lexical density, readability, clausal complexity, message structure complexity, etc. (Botta, 1993; Halliday, 1989; Kress, 1994; Ljung, 1991; Smeltzer, 1992; Stubbs, 1986; Ure, 1971). The researchers who have taken this approach believe that language has a repertoire of many

dimensions that cannot be, nor should be, categorised dichotomously, with spoken and written forms its only nature.

In addition to the many possible dimensions along which language use may be depicted as being more oral or written, different genres and media also have a direct impact on its characteristics (Ferrara, Brunner & Whittemore, 1991; Maynor, 1994). Radio broadcasting (Oksenholt, 1977; Scanlan, 1980) and tape recording (Frommer & Weitz, 1981; Long & Povey, 1982; Wherritt, 1979), for instance, have broken the barriers of space and time that are usually associated with the spoken mode of language; the fax machine has brought about certain changes in the style of written messages (Basch, 1995; Coyle & Spitzer, 1992; Gerrey, Brabyn, & Crandall, 1990; Strasser, 1995). The development of a new medium, it seems, then, inevitably results in some new styles of language use (McLuhan & Nevitt, 1974).

The evolution of computer-mediated communication (CMC), now in full development at the turn of the new century, has brought a new medium to human interaction (Adkins & Brashers, 1995; Baron, 1984; Hardy, Hodgson, & McConnell, 1994). CMC is defined as any message sent via a networked-computer (Rosenbaum & Snyder, 1991). It is technically a form of writing, as messages are entered by key-pressing and are transmitted through visual symbols. On the other hand, it is not typically writing either. Ong (1982) describes writing as a solitary experience, in which the written text is a fixed product. However, CMC text does not have to be solitary, as it could be produced by many participants, and the text is not always fixed on the paper as it frequently scrolls up and disappears from the screen. This seems to be a new form of language, and many interesting characteristics can be explored within this new medium.

From the above observation, I have developed three fundamental questions: (1) How should we handle the issue of language variety? (2) Should CMC be a focus of this language study? (3) What is the appropriate method for the study? The answers to these questions have led to the construction of a framework for this study.

1.2 Three Fundamental Questions

As stated in the previous section, the main interest of the researcher is CMC genres as possible new language varieties. Therefore, it is important to first understand the development of the issue of language variety, and to define the roles of CMC as possible

language varieties. Only when this has been achieved does it become possible to develop an appropriate method of study.

Language variety

The issue of language variety has aroused the interest of many linguistics researchers. For language teachers, the arguments over fluency and accuracy have been an important marker in the concept of variety. In the times of behaviourism, language learning was taken as a process of habit formation, and precise accuracy seemed to be set as the utmost goal. No errors were tolerated, as they might lead to the fossilisation of incorrect habits. Immediate error correction and constant practice were believed to be the keys to mastering the "standard" language.

However, since the 1970s, there has been a trend in English language teaching (ELT) that has swayed away from the behaviourism theory, positing instead that communication, above others, should be the main focus for learners learning English (Canale, 1983; Canale & Swain, 1980). The concept of *communicative competence*, which was first developed by Hymes in 1972, and which refers to an individual's ability to utilise speech appropriately in a variety of social contexts, is a key factor in this new approach.

Canale and Swain (1980) describe communicative competence as a combination of knowledge of basic grammatical principles, knowledge of how language is used in social contexts to perform communicative functions, and knowledge of how utterances and communicative functions can be combined according to principles of discourse. Widdowson (1990) further reminds us that the central tenets of the communicative approach to language learning is that learners learn to communicate by communicating, rather than by learning about the language system (p. 160). More recently, the issue of *mutual intelligibility* was raised (Jenkins, 1995, 1998, 2000), which stresses the importance of recognising various phonological forms of utterances to facilitate communication in international contexts.

Though some scholars proclaim that the job of language teachers is to teach students to produce standard English, to attempt to define exactly what constitutes standard English is highly controversial. What we can do is to try to develop a better understanding of the characteristics of different varieties in the English language. With the understanding as a reference, we can expect to make good use of each variety in developing English teaching activities and also in developing teaching materials, especially for the countries where English is not used in everyday communication. For

the purpose of communication, we are naturally concerned about the varieties, as much as we are concerned about the awareness of the varieties. In this study, I hope to explore the newly developed computer-mediated communication (CMC) as another context for language varieties.

CMC as a subject of varieties

Computer-mediated Communication (CMC) is a new medium and also a new context of language use. It is becoming rapidly more pervasive in all sectors of society. Of course, we do not believe that CMC will take the place of other media for communication. However, it is an undeniable fact that "the future menu of media will be enlarged" (Murray, 1995, p. 97). With the development of the Internet, the speed at which messages are transmitted and the associated change of communication contexts have made possible a variety of modes of message exchange. These new modes have been found to have a great impact on social relationships and education.

Many studies have been conducted on the changes in social relationships. Optimists look at CMC as a fertile land that will nurture new virtual communities (Rheingold, 1994). People with more sceptical views warn of the danger of the overuse of uncontrolled rude language, namely flaming, and other behaviour that reveals weakened social norms in the CMC (Sproull & Kiesler, 1986, 1991). Whatever positive or negative impact it may have on social relationships, it has been argued that CMC is not just a tool, but has constituted a major part of communication contexts that interact with the participants. This is a unique feature that is not found in any other media.

Since its emergence, CMC has been examined for what it may bring to education. When any new development of knowledge or any new instructional technology is introduced, many teachers feel the necessity to incorporate it into their approach to teaching, because they are ever conscious of John Dewey's warning that, "if we teach today as we taught yesterday, we rob our children of tomorrow" (cited by Berquist, 1999). Ray Clifford (cited by Meloni, 1998) also warns that "technology will not replace teachers ... teachers who use technology will probably replace teachers who do not."

There are many studies on the value of the Internet and CMC in teaching. In the STILE (Students' and Teachers' Integrated Learning Environment) project, which includes four universities in the United Kingdom (Wilson & Whitelock, 1997), it is found that students used the system to learn more about the subject area and about each other, and to sustain or motivate themselves throughout the learning period. In terms of language learning, various studies have been conducted on the pedagogical values to

students of CMC (e.g. Amoroso, MacDonald, Shedletsky & Travers, 1993; Harasim, 1989, 1990; Holden & Mitchell, 1993; Hunter, 1987; Johnston, 1992; Kaye, 1989, 1992; Poon, 1993; Romiszowski & deHaas, 1989). On the other hand, some people worry that the special features of the language often found in CMC have deviated from the generally accepted registers, and are therefore unsuitable as models for language learning (Barson, Frommer, & Schwartz, 1993; Beauvois, 1992; Johnston, 1992; Lundstrom, 1995; Wilkins, 1991).

The several studies cited above are just a fraction of a large number of studies that are concerned with the role of CMC as a factor of communication. A more detailed review of literature will be presented in the next two chapters. For the time being, I still need to mention the importance of CMC as a new set of language varieties.

This new medium of communication is certainly unconventional and comprises the features found in both spoken and written modes. Its existence has made the distinction between spoken and written forms even more vague (Close, 1994; Maynor, 1994; Murray, 1988; Warschauer, 1996b). Walther (1996) even argues that communication in certain types of CMC is, by its nature, *hyperpersonal*, i.e. it surpasses the impersonal or interpersonal interaction found in most other media. The language varieties developed in these contexts must be unique and interesting as objects of study, and this suggests that CMC should be examined in its own right, as a unique category of language contexts.

An appropriate method for the study

The study of language variety is based on the belief that languages vary for different reasons. There is no standard language that stays unchanged (Crystal, 1997). Instead of casting any value judgement or giving prescriptive rules, a better way to study language variety is to posit objective descriptions of the nature of the language. The use of an electric concordancer on a large sample of computer-readable texts is an efficient method of conducting objective linguistic research on a large scale. This is the modern approach of corpus linguistic studies.

The very nature of CMC, i.e. that it is computer text, makes it prefect for computerised corpus linguistic studies. There is already a large number of researchers investigating language by using concordancing to see what different patterns or characteristics exist in language forms across different social contexts (Higgins & Johns, 1984; Howarth, 1995; Johns, 1990; Mindt, 1986; Tribble, 1991). Text data has been collected and compiled from different contexts and the language components examined

with the purpose of obtaining an explicit, objective and descriptive view of how they differed in various social contexts.

I believe this is an approach worth taking. It is not just that the collection of the electronic text samples is perhaps easier. It is also not just that objectivity can be best maintained. I believe that in taking this approach, the method attempted may produce a good model for others who are interested in gaining a better awareness of the new language variety. When we hope to facilitate the teachers' and learners' awareness of language variety, demonstrating an approach that can help develop awareness is certainly more important than bringing to their attention some of the more specific features.

1.3 Framework of the Study

This study first began with the aim of enhancing the quality of English language teaching through an exploration of the new varieties of language that have emerged through the development of computer-mediated communication. To achieve this, I have taken several factors into consideration in developing the methodology of this study, including a choice between a qualitative and quantitative research approach, the objects of the investigation, and the types of texts that could best serve as sample data for this study.

After the approach to the analysis had been decided and the objects and the sample texts chosen, a process of data preparation was undertaken. Then, the statistical procedure was implemented in order to make a comparison between the two varieties of CMC texts themselves. I also attempted to observe the differences between the features of CMC texts as compared with those of non-CMC texts.

The findings I gathered, could, I hoped, be of some reference in pedagogy. I hope that the approach and the findings of this study will arouse teachers' attention to the new medium of CMC and enhance the awareness of EFL (English as foreign language) teachers and students alike to the many possible varieties of the language. I also hoped that by detailing some of the problematic or special linguistic features found in the CMC data, and by making some generalisations about these, the research could be of help to syllabus designers, textbook writers and English teachers.

The methods adopted and procedures undertaken in this study are discussed in more detail in Chapter Four. What follows here is a foreshadowing of the approach and procedures of the study.

1.3.1 Approach and objects of the study

In the study, I adopted the approach of corpus linguistic studies and that of quantitative comparison. By taking these approaches, I hope to gather objective data for further interpretation. The objects of my investigation are a list of 67 linguistic features adopted from earlier studies, which I believe to be comprehensive enough for my research purpose.

1.3.1.1 A quantitative corpus linguistic approach

For the exploration, I chose to apply the approach of corpus analysis and implement some statistical procedures to gather quantitative information on the characteristics of the new language varieties found in CMC. CMC texts are posted and exchanged on the new convenient electronic medium. Every day, billions of words are transmitted through the system. The large quantity of texts stored in the electronic version forms becomes an ideal data pool that can be accessed and directly examined with the help of the computer. This is an important advantage in corpus linguistic studies - a large body of machine-readable sample texts as objects of analysis (McEnery & Wilson, 1996). I hope the quantitative data gathered from the study can provide sufficient objective evidence of the nature of the new variety of the language. Therefore, I decided to take a quantitative approach to the study and use the corpus linguistic method for data analysis.

1.3.1.2 Objects of analysis

Secondly, it was necessary to consider the features to be examined in this study. Many of the previous studies carried out on different varieties of language chose to focus on a certain aspect, for instance, lexical density (Stubbs, 1986; Ure, 1971), clausal complexity (Kress, 1994), word choice (Halliday, 1989; Ljung, 1991; Vande-Kopple, 1995), readability (Botta, 1993; van Hout-Wolters & Schnotz, 1992), and message structure complexity (Andersson, 1974; Laina, 1992; Smeltzer, 1992). Besides, most previous studies on the characteristics of CMC texts (Chapelle, 1990; Kim, 1998; Tella, 1999; Warschauer, 1998) seemed to be focused on rather limited sets of language features and could not provide a comprehensive perspective for me to follow.

As this research was intended as a comprehensive study on the nature of CMC texts, as compared with that of non-CMC texts, I chose to survey a large number of features in the texts, so as to get an overall picture. The set of features chosen for this study was adopted from that which is provided by Biber in several of his studies (1986, 1988a,

1989, 1995; Biber, Conrad & Reppen, 1994, 1998). The total set of 67 features is discussed in more detail in Chapter Four. Biber's studies have been recognised by many researchers of language variation and the set of features he chose, I believe, are very clear, explicit, detailed and unambiguous. Thus, I decided to use these features as the objects of this study.

1.3.1.3 Text samples

Last but not least, I needed to choose the source and size of the text samples for this study. As an English teacher, I am a regular user of several electronic discussion lists on the theme of language learning and linguistic analysis. I have easy access to the archived texts, and I also believe that the CMC texts I can examine in those sample texts are of adequately educated registers.

The sample texts of analysis are chosen from archives of two Internet discussion lists, NETEACH-L and TESL-L, as well as two MOO forums, NETEACH-MOO and Netoric-Moo. The former are asynchronous and the latter synchronous. The theme of these texts is always issues concerning teaching English as a second or foreign language, and almost all writers are English teachers.

The selected text sample in the asynchronous CMC group (abbreviated as CA) consisted of a total of 100,194 words in 503 email texts. The synchronous CMC (CS) text sample had a total of 104,162 words in 17 MOO session logs.

While the major focus of the study is on the CMC texts, the features found in non-CMC texts can also serve as an important reference. To avoid the technical difficulty of converting a large amount of non-CMC texts into computer readable form, I decided to adopt the quantitative data on already-existent non-CMC text samples of similar genres for possible reference. The data were collected from Biber's work (1988a).

Biber (1988a) used 15 written genres from the LOB Corpus (Johansson, 1982; Johansson, Leech & Goodluck, 1978), 6 spoken genres from the London-Lund Corpus (Svartvik & Quirk, 1980), and two types of unpublished letters (Biber, 1988a).

The corpus Biber used has a total of 960,000 words in 481 texts. They are distributed in 23 genres – 17 written and 6 spoken. For some of the linguistic features I examined, I compared the relative standings of the CMC genres with those of the 23 non-CMC genres in an attempt to present the special characteristics of the features apparent in the former. This is further detailed in Chapter Four.

1.3.2 Data preparation and tools used

To handle the large quantities of CMC text data, I would have to resort to computerised tools of concordancers and tagging devices. The tools I selected were: (1) the concordancing programme CLAN (MacWhinney, 1995, 1996a, 1996b), developed at Carnegie Mellon University, and (2) the part-of-speech tagging programme TAGGER (Mason, 1996) developed in the University of Birmingham.

With the help of the tools, the CMC text samples can be prepared and processed through several steps to yield the required results.

1.3.2.1 Concordancing and tagging programmes

CLAN (Child Language ANalysis) is a set of programmes designed to allow users to perform a large number of automatic analyses of transcript data. The main programmes used are COMBO and FREQ, which are required for key-word-in-context searches and their frequency.

The programme TAGGER is an automatic part-of-speech (POS) used to affix tags of the part-of-speech to each word, so that the features can be identified and counted by the CLAN programme. A list of part-of-speech tags is included in Appendix 1.

1.3.2.2 Data preparation and processing

For CMC data downloaded from the Internet, all irrelevant lines, such as the email headers and lines generated by the computer system, were removed. This was done in order to achieve a precise analysis, because these components are not written by the CMC users.

The CMC texts were then sent by e-mail to TAGGER for the part-of-speech tagging. The tagged files, together with their original untagged version, are the objects for analysis later used by the CLAN concordancing programmes.

All of the 67 linguistic features are first depicted in notations in order for the concordancing programme to be able to find their patterns and structures. This is further discussed in Chapter Four. The CMC texts are then processed using the COMBO command in CLAN. Although the jobs of tagging and concordancing are carried out by the computer programmes, it is still necessary to manually proofread and correct the output.

1.3.3 Model of analysis

An adequate model of analysis is the key aspect of the research method. For the most adequate model to emerge, I have undergone a long and painstaking thought and

trial process. Several methods were considered and attempted, but eventually the procedure of Non-Parametrical Statistical Test was decided upon because it was found to be most suitable for the characteristics of the figures generated in the text analysis procedures. The statistical methods considered included: a 2 x 2 Factorial Experiment between CMC and non-CMC texts (Section 4.6.1), a Factor Analysis on CMC texts (Section 4.6.2), and a Non-Parametric Statistical Test of synchronous versus asynchronous CMC texts (Section 4.6.3). Each of these is discussed in Chapter Four alongside the findings and issues in order to examine how the methodology was developed to best fit the purpose of the present study.

1.4 Expected Pedagogical Contributions

There are several channels through which I expect this study to make contributions to linguistics and to English teaching. The most important, however, is to draw to the attention of language teachers and learners to the emerging new varieties of English in CMC. Language occurs in various forms when people communicate with different persons at different times and in different places. Linguistic awareness is an important goal in language education. The patterns found in this study will possibly reveal a new variety of language that is emerging in CMC. My intention, then, is to explore the varieties of language that are emerging in CMC, and to disclose language features that appear significantly more frequently in some text groups and less so in others. More importantly, of course, I hope that this study can arouse the interest of language teachers and students to seek a better awareness of the language varieties, especially those that emerge in the new versatile contexts of computer-mediated communication.

1.5 Organisation of This Research

This study is made up of seven chapters, covering the literature review, the investigation procedures, the discussion of findings, and the suggestions for pedagogical implications. First, Chapter Two is devoted to reviewing the literature that discusses how language is realised in human communication, involving many elements, such as grammar, styles, register, contexts, participants, etc. Chapter Three continues with the literature studies to review the communication elements in the context of computer-

mediated communication. Chapter Four then demonstrates how this research proceeds, including the selection and preparation of text samples, the choice of tools and learning how to use them, how the investigation was performed, and the adjustments carried out in order to obtain the best statistical approach for the study.

Chapter Five is based entirely on a discussion of the findings obtained in Chapter Four. A comparison of the characteristics of CMC texts with those of non-CMC texts is also included in this chapter. The following chapter, Chapter Six, is dedicated to suggestions, based on the findings and the approaches implemented in this study, of how language teaching practitioners can utilise this study for English language teaching. My hope is that it can be of help to English teachers, syllabus designers, textbook editors and publishers of language materials. Finally, Chapter Seven provides a summary to the whole study, and some concluding remarks on how this study could be evaluated and what directions future researchers could take.

Chapter 2: The Study of Language Variations

2.1 Introduction

For decades, ESL/EFL teachers have been devoted to the cause of helping their students master the target language in the most effective and efficient ways possible. From time to time, theorists and practitioners also develop new assumptions and theories, as well as techniques and methods in language teaching to provide teachers with the support necessary to achieve this goal. The development of teaching approaches from grammar-translation, the natural approach and audio-lingual, to the contemporarily very popular communicative approach (Canale, 1983; Canale & Swain, 1980), and the hundreds of innovative teaching methods that blossom from time to time in different places all mark the unceasing effort of researchers and practitioners in seeking new breakthroughs.

One of the major trends in contemporary language teaching methods is the resort to authentic language materials. As the "style of language in dialogues of the sort found in textbooks may not always be used consistently" (Rivers & Temperley, 1978), it has been argued that authentic language should be used in instruction wherever possible (Hadley, 1993; Purgason, 1991), and the use of authentic language has often been taken as a characteristic of communicative language teaching (Brown, 1994; Geddes & White, 1978). Moreover, Melvin and Stout (1987) say that "using authentic materials ... allows students to experience early in their study the rewards of learning a language. Students who work with authentic materials have an interest in the language that is based on what they know it can do for them" (p. 55).

There are many reports that advocate the use of authentic language materials and that describe different experiences of using authentic materials in various circumstances (Graham, 1980; Jones, 1984; Lado, 1988; Medley, 1984; Morton, 1999; Peretz, 1988; Porter & Roberts, 1981; Short, 1999; Smutek, 1986; Swan, 1985; Thomas, 1979; White, 1975). Among them, White (1975) points out that the concept of register in language should be considered as a basis for selecting authentic language materials. To expand this point further, Lado (1988) argues that "it is inadequate to turn out students who can use English (more or less) in only one register, one unified style, and one genre; that

humourless, opinionless, colourless, errorless conversation that consists of calling a pencil a pencil, a pen a pen, and an eye an eye" (p. 150). Other researchers also discuss the importance of grading and selecting authentic materials so that they meet the interests and needs of students of various levels (Furmanovsky, 1996; Graham, 1980; Johns, 1993; Maingay, 1980; McKnight, 1995; Rivers, 1987).

While many language teachers believe in the value of authentic materials, there are also different opinions. Freeman and his colleagues (Freeman, Fennacy, & Freeman, 1993) describe the resistance of some teachers, administrators, and parents to their use. Cook (1997) also challenges the belief in contemporary English language teaching that students should be exposed to authentic or natural language focused on achieving practical purposes. To make his point, he draws some lessons from the classroom and the way in which young children play with language (Cook, 1997).

A discriminating position seems the most favourable. As argued by Hadley (1993, p. 82), "the use of authentic materials does not imply that we should abandon the use of materials created for instructional purposes. Rather, a blend of the two seems more appropriate." Kuo (1992) further compares authentic and simplified materials. Swan (1985) also suggests integrating both authentic and specifically prepared materials.

The interest of many language teachers in using authentic materials in language teaching illustrates a belief that the language taught and learned in a second or foreign language setting should be compatible with the varieties of language used in real life by native speakers. There is no doubt that the ultimate goal of foreign language learning is to foster the learners' competence in communicating in the language in authentic situations. The students are expected to be able to use the language in real communication contexts, and not just to acquire the language system as a fixed set of knowledge. In this sense, there is definitely a need for the teacher, as well as the students, to better understand the features of a variety of styles of English as an authentic language, not just the structures depicted in textbooks.

Since the 1980s, the personal computer use has been pervasive. It has also been used as a means for personal communication since the early 1990s. Via a computer network, people, even children, communicate with each other using functions such as chat, talk, email, and file transfers. As a result, the computer network has become one of the major medium of communication in the modern world, and the role of computer-mediated communication (CMC) may be of growing importance in the coming years.

Many innovative language activities are also made possible with the help of CMC. Activities such as key-pals through email exchange, bulletin board systems (BBS) and

online conferencing, to name a few, are very common practice nowadays in language teaching. Many teachers encourage these activities and even design syllabuses for their students to engage in authentic use of the language on the Internet (Bowers, 1995; Kitao, 1998; Howell-Richardson & Bish, 1997; Vilmi, 1996; Warschauer, 1996a, 1996d). The convenience of the computer network also makes higher interactivity in distance education possible. Many online courses offer such a convenience to students who are unable to attend the physical site of a classroom for various reasons, such as having to work, time conflicts, disability and so on (Morgan, 1999; Pincas, 2000; Schaeffner, 2000; Vilmi, 1994, 1996; Wilson & Whitelock, 1997).

The benefits of gathering students from different areas and various backgrounds in a virtual classroom, besides extending the opportunity of learning to more people, is that the exchange of information and opinions is no longer bound by time or space and, with adequate planning and encouragement, the frequent constructive interaction of all of the participants may very well enrich the content and insight of the communication. In addition to innovations in language teaching applied by teachers, the development of the computer, the new medium for communication in the current decade, has resulted in the emergence of new types of language use (Barnitz & Speaker, 1999; Baron, 1984; Carey, 1980; Ferrara, Brunner, & Whittemore, 1991; Maynor, 1994; Patterson, 1996; Spitzer, 1986).

Under such circumstances, as Barnitz and Speaker (1999) point out, language diversity flourishes, as children develop many electronic and non-electronic literacies. Under this new form of communication, people use language with patterns or lexicons more or less different from that which they normally use in traditional forms of communication, such as telephones, face-to-face conversations, or written letters. These differences, revealed in some studies, include the frequent exchanges of messages, deviations on vocabulary, variations in structure, and so on.

Many scholars have been attracted by these differences, and are curious about how this new form of communication has influenced language use. They have researched the possible prominent elements of language used in this electronic medium. Issues on which they have focused include, for example, textual analysis (Collot, 1991; Collot & Belmore, 1996; Howell-Richardson & Mellar, 1996), gender (Herring, 1992, 2000; Herring & Lombard, 1995), discourse and identity (Pincas, 1999; Warschauer, 1996b; 2000), and many more. Findings from these studies do unveil some diversities in the new language form. However, a close investigation of the overall language pattern under this new medium has not been fully exploited. Hence, there is a need to fill this gap with a

detailed examination, so that researchers and interested parties have a clearer understanding of how language use is different in this area.

This study intends to examine the use of language in the CMC, with a further hope that the findings reached would benefit education practitioners, offering improved insights into how new styles of language are generated in the computer medium. It is also hoped that the findings could serve as a set of descriptive principles to better facilitate language teaching when teachers incorporate the authentic use of CMC into English classes.

Since CMC texts form an emerging range of uses of language, they can be approached against a background of wider existent understandings of language varieties. In examining the CMC texts, we may well begin with a review of the studies on various perspectives of language variety. The views from several linguists on grammar, language style, stylistics, register, genre and context, as well as the context of situation will be discussed to seek an understanding of why language use varies in different contexts. Then in later sections, a further review on the difference between written and spoken language will be presented before a major theme of this research is established - a multi-dimensional approach to distinct language variety.

It is hoped that previous studies on the varieties of language can be of reference here in setting up the significant features to investigate in the present study of the language used in the medium of CMC. Together with this approach, some early studies will also be reviewed. After reviewing these, I will propose a rationale about the language used in a new medium, i.e. CMC. As more and more people communicate with others using CMC, and as English is used as an international language, mutual intelligibility is a key issue in this new form of communication. An understanding of selecting an appropriate language variety is also important for teachers and learners in English language teaching (ELT). Finally, I will close this chapter with a short summary of some studies that focus specially on the CMC for ELT, which will be further discussed in the next chapter.

2.2 Background to the Study of Language Variation

Language variety has been an important field in linguistics studies for many years. It can be traced back to the early 20th century, when Sapir (1915) found that, in Nootka, an Indian tribe in the coastal region of British Columbia, Canada, separate linguistic forms were used in speaking to or about children, fat people, dwarfed people,

hunchbacks, etc. He also found that, in Yana, an Indian tribe in northern California, some linguistic variations were correlated simultaneously with the identity of the person spoken to and that of the person speaking (Sapir, 1929). That is to say, when women most often use "female speech," men also use this variety in talking to women.

Besides the identity of people speaking or spoken to, social and cultural contexts are also found to correlate with language variations. For example, language forms may be considered formal versus informal and, in some cases, the two varieties of formality become sharply complementary to each other in a situation termed as "diglossia" (Ferguson, 1959).

More studies have continued to research this issue, looking at why different linguistic forms occur within a language, even though it is recognised as "one language" (Bernstein, 1970; Biber, 1988a; Brown & Yule, 1983; Halliday, 1978; Hymes, 1974; Kress, 1994). Whether the language forms possess a nature of formal, or informal, or intimate, or academic communication, it is revealed that within a single language, there are several variants, and each variant is to refer to a specific social context or function. With regards to this phenomenon, Henry Widdowson concludes that the use of different language variants, or varieties, is for the purpose of group identity: "Language is naturally used to define social identity; and conformity to the norms of a particular language variety is an expression of group membership" (1994).

Crystal, in his language encyclopaedia, gave a definition to the *variety* of a language: "a situationally distinctive system of linguistic expression" such as for legal documents, or formal speeches (Crystal, 1987, p. 432). He elaborates later in the *Penguin Dictionary of Language*, defining it as "any system of linguistic expression whose use is governed by situational variables, such as regional, occupational, or social class factors" (Crystal, 1999). Other researchers also try to give definitions from the point of view of their studies. Novak indicates that language varieties can be "systematic variants of the usage of a language determined by individual and/or group characteristics" (Novak, 2000). Crenn states that language variety is

> a set of variations within a language that are either used by an individual or shared by a group of speakers and are either the product of the social or demographic characteristics of the speaker or speakers, or are associated with different situations in which the speaker or speakers find themselves. (Crenn, 1996)

and when language variation is determined by the situation in which the speaker finds himself or herself, it is a so-called "situational variation" (Crenn, 1996). From the above

definitions, we can summarise that language variety is linguistic forms that are used by individuals in specific social contexts in a speech community.

In the study of language varieties, linguists take different approaches in classifying the dimensions of investigation. Bright (1960) names such variations "social dialects", arguing that they should be studied in minute detail. He proposes three basic dimensions of sociolinguistics, namely, the social identity of the sender, that of the receiver, and the setting of communication, which account for most linguistic diversity (Bright, 1966). He also clarifies that these dimensions are not mutually exclusive, but commonly intersect to condition a particular type of sociolinguistic behaviour. Other researchers examine language variation from different perspectives such as grammar, style and stylistics, register and genre, and the context of the situation, which will be further discussed in the following sections.

2.2.1 Grammar

People use language to communicate (Brown & Yule, 1983; Cook, 1989; Crystal, 1969; Gumperz, 1982; Halliday, 1976b, 1989). Language used in a linguistic community usually has its own conventional rules. The set of rules, known as grammar, is shared by the members of the same community. In a narrow sense, grammar means the syntactic rules of the language. In a larger sense, however, it is an explicit system dealing with possible structures and the cultural and social functions of the community, and is characterised as a "form-meaning rule mechanism" in communication (van Dijk, 1977, p. 3).

The grammar contains components of, among other things, the meaning of the words in a text or conversation, components of the reference to which the words and phrases refer, and the knowledge that the linguistic community shares (van Dijk, 1977). Thus, a system of grammar must be able to account for the ability to construct "correct" utterances and texts, and to use such utterances and texts adequately in some communicative situation (van Dijk, 1977, p. 13).

This view is also shared by other linguists. Hymes refers to the ability to use utterances and texts adequately as a "communicative competence" (1972), while Halliday describes it as a prerequisite before embarking upon any study of language use (1973). In summary, grammar can systematically account for the different forms with different meanings or different functions of language used in a linguistic community (van Dijk, 1977, p. 5).

There was previously a period in which grammar was considered a set of prescribed

rules governing the use of language (Jespersen, 1922). However, more contemporary views take the view of grammar as a theoretical reconstruction of the rule system of a language, thus the term descriptive grammar. Grammar is believed to mean not "correct use of language" but "the way a language works" (Dillard, 1972). Grammarians collect texts from written and spoken forms and describe their typical features to accumulate a set of descriptive rules about language use. They hope that these descriptive rules could account for the language use in most situations (Labov, 1972; Quirk, Greenbaum, Leech & Svartvik, 1972). They seek some objective rules that suggest that most languages have been written or spoken in such a way. These descriptions definitely make a tremendous contribution to the language study. However, it is also found that many language uses cannot be accounted for solely by grammar rules.

Grammar rules cannot fully account for examples such as, "That one's over there in that but it isn't there," given by Brown and Yule (1983, p. 57). This is an initial response from a five-and-a-half-year-old girl when she is asked to say how two pictures are different from each other. In the two pictures, the teddy bear is on the chair in the first picture, but it is not in the second. This represents a problem of referents by using deitic terms in a context-dependent conversation, in which prescriptive grammar rules cannot interpret well. Another example is, "You, you, but not you, all come here!" In this context, grammar rules are unable to interpret well because there is a referent problem, too. This sentence needs more contextual information about the persons at which the speaker is pointing at the moment of utterance.

Linguists also find that literature texts deviate from grammar rules. Those rules cannot explain sufficiently why certain types of literature appear differently when they are written by different writers. This gives rise to the study of styles and stylistics (Crystal & Davy, 1969).

Grammar is but one of the several constraints on language use. Among the billions of sentences that are grammatically acceptable, a speaker must choose to utter ones that make sense, "those which correspond to the events he wishes to discuss and those which are suitable to the situation in which he speaks" (Burling, 1970, p. 5). Burling also proposes other extra-linguistic constraints on language use. These include reference, situation, and personality. The rules of reference mean the definitions of the referents of terms, which provide explicit criteria for deciding between terms such as *father* and *mother*. The rules of the situation govern the choice of terms based on the condition in which an utterance is made, such as between *father*, *dad*, *daddy*, *pop* and *pappa*. The rules of personality are the linguistic choices governed by the idiosyncrasies of individual

personalities (Burling, 1970).

In discussing the choice between the *–ing* and *–in* forms as verb endings in different language situations, Fischer (1964) marks the socio-symbolic variants of speakers, the addressees, the formality and the specific verbs used. He also argues that "people adopt a variant not because it is easier to pronounce ... or because it facilitates some important distinction in denotational meaning, but because it expresses how they feel about their relative status versus other conversants" (p. 486).

In summary, we can see that language use is constrained by several factors, including the participants involved, their moods and backgrounds and the settings of communication. These factors interact with each other to determine if certain sets of sentences are acceptable. To understand the mechanism of language uses, we need to further explore these concepts

2.2.2 Style and stylistics

The English language, like any other, is a complex of many different "varieties" of language that is in use in various situations throughout the world (Crystal & Davy, 1969, p. 3). These are clearly varieties of one language. Though they have much more in common among themselves, they are not a single homogeneous phenomenon. Various varieties may develop out of the same language as a result of differences in gender, age, education, social status, as well as region of residence. Different registers and genres of the same language may develop in different contexts of situation, too. Within them, words, phrases and structures used in one form are not exactly the same as in the other.

When used with a sense of evaluation, style is assigned with a more restricted meaning. It refers to the overall effectiveness of expression in language use, such as "saying the right thing in the most effective way," or "good manners" (Crystal & Davy, 1969, p. 10). It has long been allied mainly with literature with a focus on literary critic, which is partially evaluative and descriptive, and as a "characteristic of good, effective, or beautiful writing" (p. 10). Thus, we are led to the widespread use of the word "style" to refer solely to literary language.

In a broarder sense, however, stylistics contributes weight in language variation by studying "certain aspects of language variation," and searching for the "linguistic features which could characterise the main varieties of English" (Crystal & Davy, 1969, p. 9). By linguistic features, we mean any part of speech or writing that a person uses from his/her flow of language and discussion.

Stylistics attempts to reach a commonness of language use in many occasions. It

includes the study of language habits from one person, either in part or all, which is usually referred to as "style" in stylistic discipline. Style also refers to some or all of the language habits shared by a linguistic community at one time, or over a period of time. Examples of this study include the style of the Jane Austen novels, sixteenth century renaissance English literature, as well as the style in which government documents are written, or that of public-speaking during a political campaign.

As the aim of stylistics is to analyse the language habits of a person, or a group of work, in trying to identify language use from a general mass of linguistic features, we can see how these common features might be bound up with certain social contexts. If so, stylistics could possibly explain why such features, instead of others, have been used. Stylistics also tries to classify features into categories based upon their functions in the social context (Crystal & Davy, 1969). It studies the internal structure of sentences, comparing the lexical items in a text with their meanings and distributions in relation to the frequent use of others. Hence, the more important stylistic feature in a text will be that which occurs more frequently within that text, and such a feature is shared less by other texts or varieties of language. One example given by Crystal (1969, p. 21) is the use of the passive in scientific English texts. The passive is a distinctive feature in this variety, i.e. scientific texts, as it has a greater frequency of occurrence than in most other varieties. It predicts a great stylistic importance, because of its uniqueness. As Crystal claims, a different unique feature which occurs more frequently in a text could be graded higher on the stylistic scale (1969).

From Crystal's point of view, the task of stylistics incorporates three directions. It first should be able to recognise the complete range of linguistic features that people feel to be stylistically significant, and be able to specify an accurate way to discuss them. It should then be capable of outlining analysis methods that permit researchers to organise the significant features in a systematic way so that they can compare any one use of language with any other. Thirdly, stylistics should be competent enough to determine the function of these significant features, by classifying them into categories based on the different linguistic research (Crystal & Davy, 1969).

Crystal uses the high frequency of certain linguistic features as a measure by which to scale the style of texts. At the same time, other researchers raise different issues on language use, such as register, genre, context, the context of situation and factor (Brown & Yule, 1983; Halliday, 1973, 1976b; Hymes, 1962; Lyons, 1977; Sadock, 1978; Wootton, 1975). As we will see in the next sections, those researchers, in proposing different terms to describe the varieties of language styles, also largely associate the

varieties of language with the relative frequency of certain linguistic features.

2.2.3 Register and genre

Register is raised by Halliday and his colleagues to distinguish a variety of a language "according to use" (Halliday, McIntosh, & Strevans, 1964, p. 87). The crucial criteria of a register are related to its grammar and lexis (p. 88). If two pieces of language samples from different situations show no differences in grammar or lexicons, these two samples can be assigned to one register, and the same register. For instance, reportage, editorial comments and feature writing are in one register, i.e., the register of journalism, as words that qualify as names, places, and political events are more likely to be used within the texts.

Register specifies the "range of meanings that is activated by the semiotic properties of the situation" (Halliday, 1975, p.126). It is the set of meanings, the configuration of semantic patterns that are "typically drawn upon under the specified conditions, along with the words and structures that are used in the realisation of these meanings" (Halliday & Hasan, 1976, p. 23). In other words, register is closely related to a situational context.

Register is also a reflection of the context of the situation in which language is used, and the ways in which one type of situation may differ from another. Types of linguistic situation differ from one another in three respects (Halliday, 1978). First, "what is actually taking place;" next, "who is taking part;" and finally, "what part the language is playing" (p. 31). These three variables determine two things, first, the range within which meanings are selected, and second, the forms which are used for their expression. In other words, these three aspects determine the register.

According to Halliday and his colleagues, the above is best defined by three dimensions of discourse, namely, *field*, *mode*, and *tenor*, which are the formal properties of any given language event with which they have always been associated. The field of discourse is the subject matter talked about in the language use, while mode of discourse is the medium of language use, such as the written or spoken form. The tenor of discourse is the style used by language users in relation to the other participants in a linguistic community (Halliday, 1974; Halliday et al., 1964).

All language events are closely related to these three elements. For instance, a language event that comprised of a lecture on chemistry in a science college would be categorised as in a field of *scientific biology*, a mode of an *academic lecture*, and a tenor of *teacher-to-student style*. In other words, the linguistic features that are typically associated with a configuration of situational features - with particular values of the field,

mode and tenor - constitute a register (Halliday & Hasan, 1976, p. 22).

Halliday (1974) points out that we always listen and read with expectations. This is because the notion of register is really a theory about these expectations, providing a way of making them explicit at the time of listening and reading (p. 53). Alternatively, register refers to the fact that "the language people speak or write varies according to the type of situation" (Halliday, 1978, p. 31). As Halliday claims, register attempts to unveil the general principles that govern this variation, in trying to understand "what situational factors determine what linguistic features" (p. 31). For instance, given that we know (1) the situation, and (2) the social context of language use, we can predict a great deal about the language that will occur, with reasonable probability of being right (p. 32).

In register, Halliday et al. (1964) further develop the notion of genre. Take the language of literature as an example. Literature is a single register; but various genres can be assigned within it, such as prose fiction, light verse, sonnet, etc. The journalism mentioned earlier as a register can be further divided into genres of reportage, editorial comments, feature writing, and so on (p. 92). In summary, there seem to be complementary ways of describing the relationship between text-types and situation-types. Register specifies the "range of meanings that is activated by the semiotic properties of the situation" (Halliday, 1975, p. 126). It is rather as if specific lexical items and syntactic structures are primed by one's reading of the situation in terms of field, tenor and mode, so that, in speaking or writing, these items would be more likely than others to be selected at the point of utterance or writing. Genre, by contrast, describes the way in which texts are constructed sequentially of different elements in an organisation which stages texts. For example, in standard reportage and editorial comments, the arrangement of the organisation is in a cultural convention that has been found to be functionally and rhetorically effective in the situation, i.e. reportage and editorial comments in journalism.

In contrast with the concept of *styles* as used by Crystal (1969), *registers* and *genres* are used by Halliday and his colleagues (Halliday, 1974, 1975, 1976b, 1978; Halliday & Hassan, 1976; Halliday et al., 1964) to describe the varieties of a language. A comparison of the two approaches reveals that where Crystal takes a more general view of the language varieties, Halliday et al. (1964) attempt to look at them on two hierarchical levels, with register on a higher level, embedding a certain number of genres. Besides, the term *style* used by Crystal seems more compatible with the term *register* used by Halliday et al. Both terms are defined as characterised by the relatively frequent use of certain language features. *Genre*, as used by Halliday et al., is determined more by the

text construction than by the frequency of language features.

It must be pointed out, however, that these terms are also used differently in other studies.

Whatever terms we take for the varieties of language, they are clearly associated with the features used in the language, and in the use of the language, they are largely determined by variables such as participants, topic and situation of communication, which we will examine in the next section as the context in which language varieties develop.

2.2.4 Context, co-text and context of situation

Use of language can be divided into a range of varieties. They are referred to as social dialects (Bright, 1960), styles (Crystal & Davy, 1969), or registers and genres (Halliday, 1974, 1975, 1976b, 1978, 1989; Halliday & Hasan, 1989; Halliday et al., 1964). These varieties are chosen in relation to many elements, like the speaker, the person spoken to, the topic, the means and settings of communication, etc. (Brown & Yule, 1983; Cook, 1989; Halliday, & Hasan, 1976, 1989; Halliday et al., 1964; McCarthy & Carter, 1994). These extra-linguistic variables, which strongly correlate with the varieties of language, are usually considered as elements in the context of communication.

Since the beginning of the 1970s, linguists have become increasingly aware of the importance of context in the interpretation of language discourse. For instance, Hymes claims that context supports the intended interpretation as well as limiting the range of possible interpretations (Hymes, 1962, cited in Wootton, 1975, p. 44). Brown and Yule use the concept of co-texts to alert us to the fact that contexts always change as discourse continues (1983). Halliday also emphasises that language study must begin with its social context, as language is realised as an activity carried out by people in a certain situation (Halliday 1974; Halliday et al., 1964).

The basic concept of this "context of situation" is first proposed by Malinowski (1923), and later expanded by Firth (1935). The essential point here is that "language comes to life only when functioning in some environment" (Halliday, 1974, p. 28). Language is not used in isolation, but is always related to actions and events associated with persons and people's background knowledge to make things said and done. This is referred to as the "situation," so language is said to function in "contexts of situation" (Halliday, 1974, p. 28). In his argument, Halliday proposes that the types of linguistic situation differ in three ways: what takes place, what part language plays, and who takes part. These three variables together determine the forms used and meanings selected for

language use (p. 32). These variables have been further termed by him as *field*, *mode* and *tenor*, (previously referred to as *style* in his early studies), and are used to determine registers and genres (Halliday, 1974, 1978; Halliday & Hasan, 1976; Halliday et al., 1964).

In context, there are three basic elements: participants, topic and setting (Brown & Yule, 1983; Hymes, 1964a). Addresser, addressee, and audience are all among the participants. They include: (1) the speaker or writer who produces the utterance; (2) the hearer or reader who receives the utterance; and (3) possible over-hearers. Topic, the second element of the context, is the contents or subject-matter about which the speaker talks during the communication. Setting is the place and time in which the language event is situated, with inclusions of gesture and the facial expression of the participants (Brown & Yule, 1983; Firth, 1957).

Hymes (1964a) also discusses other features in his work on context, such as *channel*, *code* and *event*. In his view, *channel* is the way in which participants contact each other, either by speech, writing, signing or smoke signals. *Code* is the form intended in the language event, like chat, debate, fairy-tale, sonnet, etc., while *event* refers to a sermon or prayer in a church service, or an annual report in a commercial firm. Based on these similar criteria, Brown and Yule present an example to show how knowledge about a context can help people to interpret the discourse: a young scientist (addresser) is responding to questions asked by a language researcher (addressee) in the researcher's office, with a tape-recorder on, during the late 1970s (1983, p. 39). In this example, Brown and Yule point out that the *channel* is speech, the language *code* is English, while the message-form performed is conversation, and the *event* embedded is an interview.

In attempting to interpret the functions of contextual elements in language varieties, Brown and Yule (1983) propose several principles. For instance, the term that an inferior uses to address his/her superior is definitely different from what he/she uses to address his/her friends or wife/husband, i.e. *sir* versus *hey/darling*. This shows how the social roles of the participants affect the language styles. Brown and Yule also argue that it can be assumed that in a specific social context, only one role is taken by an individual at any one particular time (Brown & Yule, 1983, p. 54), and the same individual may hold different roles in different social contexts (p. 56).

Other elements included in the context discussed by other researchers are the *referent* and *knowledge* in the discourse. It is believed that the participants must have certain shared knowledge in the context, so that meanings can be adequately conveyed (Dahl, 1976; Grice, 1981; McCawley, 1979).

For knowledge in a communication context, van Dijk (1977) claims that a knowledge set would include: (1) knowledge of the "worlds in which the utterance is interpreted;" (2) knowledge of the "various states of the context;" (3) knowledge of the "language used, i.e. of its rules and of possible uses of rules as well as knowledge of other systems of interactional conventions." Without this knowledge, the utterance "cannot be processed properly," and hence cannot be produced and interpreted, and without this knowledge, the participants "do not know what is spoken about or why there is any speaking at all" (van Dijk, 1977, p. 194).

Besides, in conversation, a general agreement of co-operation between participants, called *Cooperative Principle* by Grice (1981), is that each participant expects the other to conform to certain conventions in speaking, in terms of quantity, quality, relation and manner. By these, the speaker speaks as informatively, truthfully, relevantly and clearly as the situation requires.

Using an example given by Sacks (1972) to explain how Grice's principle works in a language context: "The baby cried. The mother picked it up", it is demonstrated that, though there seems to be no physical connection linked between these two sentences, the situation will direct the hearer to build a context just good enough to reach an interpretation: a baby cried probably in a living room, his/her mother heard his/her cry and immediately came to pick him/her up. Brown and Yule (1983) further elaborate this with their *principle of local interpretation.* They claim that it will restrict any hearer to construct a context larger than the above, such as a baby has cried for one year in one city, and his/her mother knew of it in another city and flew over to pick him/her up one year later. Another element by Brown and Yule relating to the referent and knowledge in context is that, the hearer always assumes that everything remains as it was before in determining interpretation - the shared knowledge, unless the speaker gives specific notice to change some aspects of the discourse. This is referred to as the *principle of analogy* (Brown & Yule, 1983).

It can be seen from the discussion above, therefore, that in a communicative context, the participants would choose the adequate styles and varieties based on their status and relations, and they must cooperate with each other with their shared knowledge about the context, and, to successfully achieve the task of communication, conform to the conventions according to the many elements found in the context.

All of the contextual elements have been discussed by various researchers using different terms. In the following table, Table 2.1, I will attempt to categorise these elements as discussed by three groups of researchers.

Table 2.1
Elements in Context

Hymes (1964a) Brown and Yule (1983)	Halliday (1964, 1974, 1975, 1976, 1978, 1989)	Burling (1970)
Topics	field	referent
Channel, code, event	Mode	
Participants	Tenor	personality
Settings		situations

It is not surprising that different groups have different terms for the elements that effect language styles and varieties. While they are not exactly equal to each other, we can assume some parallel meaning among the terms used in each row in the above table. For instance, what Hymes and Brown and Yule call "topic" may be similar in concept to what Halliday calls "field," and may be quite related to what Burling calls "referent." In this case, we can see that the contextual elements (or extra-linguistic elements) that affect language styles may be roughly categorised as four groups: the content, the medium, the participants and the settings.

Besides, it is worth noticing that in each row in the table, the elements do not function separately, but in conjunction with each other. For instance, in row one, a topic of communication certainly cannot dictate the language style by itself unless the elements of participants and others are considered. For the topic of some scientific knowledge, the text of a professor's class lecture would certainly be different from that of some students' discussion in the lab.

From the above discussion on language variation, we can conclude that, in understanding language variation, more weight has been placed on context than on the form of language itself, as it has been agreed that language is realised in social contexts. Sub-categories may be important but difficult. They may cover items like the speaker, the hearer, the communication channel, the language code, the message-form, the language event, the changing of context, the referent, the knowledge possessed or shared, as well as the minimal and local interpretation within the context. There are many questions of *who*, *whom*, *when*, *where*, *how* and *why*, and they must be answered in seeking the elements that effect certain language varieties.

In addition to the influence of the context, the concept of *co-text* has also been proposed by Brown and Yule (1983) to account for language variation. As contexts can change from time to time, one more dimension has been added to the variation of language. Van Dijk (1977) says that a context is not a world-state, but a sequence of

world-states, within which situations do not remain identical in time, but change. Hence, a context has "an initial state, intermediary states and a final state." Brown and Yule (1983) point out that our inferences are usually determined and constrained by the several contexts occurring in discourse. This sequence of contexts is referred to as *co-text.* Co-text, in some sense, can nest contexts, indicating the complexity of the context which hearers and readers are capable of interpreting. It is one additional dimension of elements that make language styles even more changeable and versatile.

From the above discussion, we can see that language varies mainly from the linguistic forms used when writers/speakers adjust themselves to cope with the different social contexts in which they find themselves. This would result in language varieties within a single language. Then, the next question would be, how does language vary across different contexts? Moreover, for this study, how does language vary when the mode (written or spoken) is different?

Further studies of language variations between spoken forms, as considered from the different points of view proposed by researchers in past years, are carried out in the following section, Section 2.3. Most of the studies have been concerned substantially with the elements co-ordinated in the social contexts discussed in this section. Another school of study raises the issue of the factor approach in language use, which will be briefly dealt with briefly in later section, with more discussion in Section 2.4.

2.3 Studies on Written and Spoken Language Variations

In the previous sections, it has been argued that the varieties of language may be correlated with many elements in the communication contexts, such as participants, settings, topics and channels of communication. As this study will attempt to examine the nature of language as used in the innovative channel of computer-mediated communication (CMC), it is of interest to explore some previous studies on the various channels of language. It is quite clear that most studies on the channels of communication have been on the modes of speech and writing.

The similarities and differences between written and spoken modes of language have been a focus of interest to many scholars (Chafe & Danielewicz, 1987; Halliday, 1989; Kress, 1994). In some studies, dichotomous differences are sought between orality and literacy (Finegan, 1988; Havelock, 1986; Ong, 1982). Spoken language, uttered by speech organs, was considered spontaneous and fragmented (Blankenship, 1978; Blom &

Gumperz, 1968); while written language, produced with a pen or a typewriter, was viewed as prepared and highly organised (Bjorklund & Virtanen, 1991).

The variation between the two modes, as an element in a communication context, does not just function independently, but rather interact closely with other contextual elements in effecting different language styles. There have been some studies on such interaction, which I will present in this section.

In addition to that, in discussing the variation between speech and writing, the language features and styles that are often associated with either of these forms should also be discussed.

I would also like to discuss the argument presented by Halliday that the spoken form of communication is not necessarily a simple, fragmented form, as generally believed (Halliday, 1989). Finally, this section will close with the argument that variations between the spoken and written forms may appear on several dimensions and cannot be accounted for by a simple formula.

2.3.1 Modes of language and other contextual elements

Historically, the two forms of language: speaking and writing, have been believed to serve different purposes and make different kinds of contribution (Halliday, 1989). The variation between speaking and writing, basically a difference in the channel of communication, has shown a close correlation with other elements in the language context, such as settings, participants, and topics (Hymes, 1962, 1972, 1974).

In the development of human civilisation, the spoken and written forms of a language serve different functions, and they have made different contributions in different stages of development in human history (Halliday, 1989). In a non-literate society, spoken language performs all the functions that language is called upon to serve; and there is nothing lacking. In a literate society, however, the functions of language are shared between speaking and writing; there is some overlap, but by and large they fulfil different roles. They are both forms of a language – it is the same linguistic system underlying both, but they exploit different features of the system, and gain their power in different ways (pp. 99-100).

Now that our society has considered written language the most valuable asset and disregarded the spoken form, Halliday (1989) gives some thoughtful points:

> In a literate culture, we tend not to take the spoken language seriously. This is not surprising, since not only has writing taken over many of the high prestige functions of language in our society, but also our highly valued texts are now all written ones. Written

> records have replaced oral memories as the repositories of collective wisdom and of verbal art. (p. 97)

Facing the seemingly superior status of the written form, we should not neglect the important roles of speech. In human learning, spoken language has indeed played an important role, especially for young children before learning how to write (Kress, 1994). Among the several thousand languages in the world, there are many more that are based solely on an oral tradition than those which also incorporate a written form (Crystal, 1997). Besides, in a world of electronic communication in which sound and graphic symbols attract a large amount of people's attention, the power of speech should never be overlooked.

2.3.1.1 Setting: The constraint of time

In typical writing, a large amount of information is frequently packed into relatively few words because the writer can construct a carefully packaged text under lower time constraints (Halliday, 1989; Kress, 1994). Likewise, the reader can read as quickly or slowly as he/she pleases and is able to take advantage of a highly integrated text (Halliday, 1989; Halliday & Hasan, 1989; Kress, 1994, 1995). On the contrary, typical speech cannot be highly integrated with information because it is produced and understood online within time constraints (Halliday, 1989; Kress, 1994).

In the above studies, the fact that reading and writing is carried out under fewer time constraints than speaking and listening seems to have been taken for granted. However, we must note that in the above research, only the language modes in "typical" conditions are examined. As a matter of fact, there are definitely cases of exception. For instance, in note-taking, the writer is certainly under some time constraints, and is therefore unable to carefully package the text. In the delivery of a prepared speech, also, the speaker is certainly under much fewer time constraints than during an impromptu speech or daily conversation, and the information may be highly packed. The association of time constraints with either the spoken or the written mode does need some adjustment.

As will be discussed in Chapter Three, the proliferating use of CMC has provided many communication functions which have blurred the difference of time constraints on the modes of language. This presumably will cause more changes to the language features, which is the object of this study.

2.3.1.2 Participants

Another element that often plays an important role in a communication context is the participants. It is not just the personal features, such as age, gender, education and occupation that will affect one's language style. The relationship between the addresser and the addressee will also exercise influence on the nature of the language used (Brown & Yule, 1983).

In a typical speech condition, the degree of involvement between speaker and listener is a critical factor determining the nature of language. This does not exist in written form of communication, as the writer and reader typically do not interact at the same time (Brown & Yule, 1983). Due to this interaction, speakers often make direct reference to the listener by use of second person pronouns, questions, imperatives, etc., and they concern themselves with the expression of their own thoughts and feelings, e.g., marked by use of first person pronouns, affective forms such as emphatics, amplifiers and cognitive verbs such as *think* and *feel*. As a result, Biber (1988a, p. 43) argues that speech often has a distinctly non-informational and imprecise character, such as marked by hedges *something like*, *more or less*, *maybe*, pronoun *it*, and reduced forms '*coz*, *howdi*. These features can be considered together as the characteristics of speech in a context with interaction and involvement between discourse participants. In contrast to involvement, detachment refers to the characteristics which are common in writing, which results from the fact that the writer and the reader usually do not directly interact. Therefore the written mode is considerably marked by agentless passives and nominalisations.

Again, while it is interesting to examine these variations that are found typically between the spoken and written modes, we are tempted to explore whether or not such variations also exist in the new mode of communication, the CMC.

2.3.2 Typical characteristics of the modes of speech and writing

The two modes of communication, in interaction with other contextual elements such as time constraints and participants, are found to be characterised by a number of features.

In some studies, the dichotomous differences seem to be sought between orality and literacy (Finegan, 1988; Havelock, 1986; Ong, 1982). However, in recent years, it is widely accepted that, instead of being dichotomous forms, i.e. either written or spoken, with no other forms in between, the written mode and spoken mode may differ only in degree along some continuum (Bhatia, 1993; Derewianka, 1996; Littlefair, 1991).

Previous research reveals plenty of diversities. Comparative studies between the two

are usually done along one or several chosen dimensions, such as studies of lexicons, word choices, reading readabilities, clausal structures, etc. Researchers may also use word choice (Halliday, 1989; Ljung, 1991; Vande-Kopple, 1995), readability (Botta, 1993; van Hout-Wolters & Schnotz, 1992), and message structure complexity (Andersson, 1974; Laina, 1992; Smeltzer, 1992) to measure the differences between written and spoken forms of language. Some major differences are summarised below.

2.3.2.1 Lexical density versus grammatical complexity

Among the characteristics that mark the difference between speech and writing, Ure (1971) and Stubbs (1986) use lexical density as a measure to distinguish written from spoken messages, pointing out that written forms use words of heavier lexical meaning than spoken forms. Kress (1994), on the other hand, puts more weight on clause structures, arguing that written language has more clausal complexity, as verbs tend to be nominalised.

Halliday also calls for attention to the fact that sentence complexity is an important distinction between the two forms. In his view, the complexity of written language is static and dense, with high lexical density, "solid like that of a diamond formed under pressure," while that of spoken language is dynamic and intricate, "liquid like that of rapidly running river," with grammatical intricacy. As far as clausal complexity is concerned, "the highly information-packed, lexically dense passages of writing often tend to be extremely simple in their grammatical structure" (Halliday, 1989, p. 87).

2.3.2.2 Linguistic features that mark speech/writing variation

Many researchers attempt to explore the characteristic differences between the modes of speech and writing. For instance, Halliday (1989) finds that the written mode tends to have fewer verbs than spoken, but more nouns than spoken, giving as an example, "applause followed the announcement," versus "after the announcement people applauded" (p. 93). Likely, features used to integrate information into the written form, as Biber (1988a, p. 43) points out, include attributive adjectives, prepositional phrase series, phrasal coordination and careful word choice.

In a summary presented by Biber (1988a, p. 47), it is reported that the following items of difference, at least, existed between written and spoken forms.

1. [Writing is] more structurally complex and elaborate than speech, indicated by features such as longer sentences or T-units and a greater use of subordination (O'Donnell, Griffin & Norris, 1967; O'Donnell, 1974; Kroll, 1977; Chafe, 1982;

Akinnaso, 1982; Tannen, 1982, 1985; Gumperz, Kaltman, & O'Connor, 1984);

2. [Writing is] more explicit than speech, in that it has complete idea units with all assumptions and logical relations encoded in the text (DeVito, 1966, 1967; Olson, 1977);
3. [Writing is] more decontextualised, or autonomous, than speech, so that it is less dependent on shared situation or background knowledge (Kay, 1977; Olson, 1977);
4. [Writing is] less personally involved than speech and more detached and abstract than speech (Blankenship, 1974; Chafe, 1982; Chafe & Danielewicz, 1987);
5. [Writing is] characterised by a higher concentration of new information than speech (Stubbs, 1980; Kroch & Hindle, 1982; Brown & Yule, 1983); and
6. [Writing is] more deliberately organised and planned than speech (Ochs, 1979; Rubin, 1980; Akinnaso, 1982; Brown & Yule 1983; Gumperz et al., 1984).

However, viewing the above, we can see that most of the researchers analyse linguistic variation using just one or a few parameters, and many tend to treat the linguistic variation in some specific contexts. The problem is that most of the findings do not give a comprehensive view of the variation between the two modes and cannot be readily generated to account for language varieties found in other contexts. It is no wonder then, that some of the findings in studies of characteristics of the spoken and written modes are contradictory to those found in each other. For instance, a general view about written language is that it is elaborated, complex, formal and abstract in a complicated structure. Spoken language, in contrast, is concrete, context-dependent and structurally simple (Akinnaso, 1982; Backlund, 1986; Tannen, 1982; Tottie, 1983).

Some studies, though, have found almost no linguistic differences between speech and writing in terms of complexity (Blankenship, 1962, 1974; Blass & Siegman, 1975), while others actually claim that speech is more elaborate and complex than writing (Halliday, 1989; Kroll, 1977; O'Donnell, 1974). In terms of sentence length, Blankenship (1962, 1974) finds that it is nearly the same in speech and writing, while O'Donnell (1974) claims that sentence length is considerably longer in writing. Kroll (1977) finds that writing has more subordination than speech, while Blass and Siegman (1975) find little differences on that front. Facing such contradictory findings, we are convinced that, to obtain a more comprehensive view of the characteristics of the language modes, it is necessary to take into consideration a large number of linguistic features as they occur in a large number of communication contexts.

2.3.3 Multiple dimensions of variables

It is clear now that many the contradictory findings on the characteristics of the two language modes, speech and writing, result from the limited number of linguistic features and the limited scope of language contexts examined. Some researchers (Brown &

Fraser, 1979; Ervin-Tripp, 1972; Hymes, 1974) give warnings that it is misleading if linguistic variation is analysed only within some specific, isolated linguistic markers, without taking into account the sets of co-occurring features in the whole text. Comparative studies between the two modes of language should be carried out along several chosen dimensions, such as studies of lexicons, word choices, reading readabilities, clausal structures, etc. It is indeed desirable to take a multi-dimensional approach to the study.

In addition to the multiple dimensions of linguistic features that are clearly involved in determining the styles of the two language modes, socio-linguistic elements such as participants and settings should also be taken into account in order to obtain a more comprehensive picture of the phenomenon. It is clear that the linguistic features and the variables in the social context all interact to result in the styles in a mode of language. This is why Biber, through his several studies (1986, 1988a, 1988b, 1989, 1995; Biber, Conrad, & Reppen, 1994), argues for a multi-features/multi-dimensions approach to the study of characteristics of the modes of speech and writing.

2.4 Multi-Feature/Multi-Dimension Approach

Study of the linguistic characterisation of speech and writing has a long history. Through a variety of approaches applied, most researchers have identified specific linguistic features that distinguish written and spoken language. The findings from these studies usually display linguistic characterisations of speech and writing from different perspectives. However, as most of them examine only particular variables by picking up instances only from particular samples, it is impossible to view the whole profile of the nature of the language modes without a comprehensive investigation on a large number of linguistic features with a large sample. To achieve such a comprehensive view on a complex phenomenon, some sophisticated statistical techniques are required, such as multivariate analysis. One commonly used multivariate analysis technique in linguistic research is *factor analysis*, which can serve the purpose of describing the inner complexity residing within the data.

2.4.1 Factor analysis

Factor analysis is a mechanical, mathematical procedure for examining the structure of inter-correlation of elements within a large body of data. Used in linguistic studies, it

can effectively norm the correlation of linguistic features into resultant clusters. It uncovers patterned variations of linguistic elements within, normally, a group of texts of the same genre, or across different genres. The linguistic elements investigated can be items of social contexts, as discussed in Section 2.3, such as topics, channel, participants, settings (Hymes, 1974; Brown & Yule, 1983), field, mode, tenor (Halliday, 1974, 1975, 1976b, 1978, 1989; Halliday et al., 1964), or referent, personality, situations (Burling, 1970). To go along with the social context, more items can be included, such as sentence length, T-units, idea units, sentential logical relation, explicitness, decontextualisation, new or given information, planned or unplanned text, shared knowledge, etc.

Applying factor analysis on these numerous elements will provide a patterned variation that some elements will cluster together, or correlate to other elements. This is because the point of the factor analysis is to summarise a large set of variables, i.e. linguistic elements in this case, in terms of a smaller set on the basis of statistical similarities between the original variables. At the same time, it loses the minimal amount of information about their differences.

In procedures, factor analysis first gathers the occurrences of every element or feature in language as a linguistic variable across different text samples. Then, an intercorrelation matrix is calculated from the cross-tabulation. This is in an attempt to summarise the similarities between the variables so as to obtain a smaller number of reference factors. The hypothesis of factor analysis is that the many variables that appear in the original frequency cross-tabulation are in fact masking a smaller number of variables that can help explain better why the observed frequency differences occur. Each of the clusters is referred to as a factor. In a factor, each variable (linguistic feature or element) receives a loading on the factor, signifying its closeness to that factor. For example, in analysing a set of frequencies of different linguistic features across several texts, one might find that certain features received high loadings on one factor, whereas others loaded highly on another factor.

If we want to investigate a large group of linguistic features, for instance, say seventy features, from a large group of texts, factor analysis can derive a set of factors that can underlie the patterns structured within the texts. This set of factors, probably five or ten, is dramatically smaller in number than the original seventy features. The use of this technique can identify factors that underlie large-scale variation of linguistic features across texts, and interpret those factors as linguistic constructs.

To conceptualise this idea, Bainbridge (1986) uses an interesting metaphor. We can imagine the final resultant clusters of variables as clusters of stars in the sky. Each factor,

i.e. every resultant cluster, can be considered as a star cluster – a group of stars that are close to each other in distance. Each cluster usually has a centre of gravity. Within this star cluster, we can see that some stars are near the centre of their cluster, while others are closer to the periphery. In factor analysis, there is some weight assigned to each star, called factor loading, which can be taken as the measure of distance between a star and the centre of gravity of its cluster. The higher the factor loading of a star, the closer this star is to the gravity centre. In other words, each star receives a loading on each cluster that signifies its contribution to that cluster. This is also the main idea of a factor analysis used in many academic disciplines (Kruskal, 1978; Kruskal & Shepard, 1974).

Using this technique, it is possible to extract from a large number of linguistic features some key factors of language variation. With these factors, we can observe and analyse language in its different modes. Biber (1986, 1988a) adopts the method in analysing the variation between speech and writing, and he refers to this a multi-feature/multi-dimension approach.

2.4.2 Previous studies

As the linguistic elements investigated by the factor analysis are usually many, and factor analysis itself is a computer-oriented approach, studies applying factor analysis usually use computers to perform the analysis. Moreover, with the help of computer software for statistics, many linguists have attempted to use statistical methods of linguistic research in trying to refine possible and objective linguistic patterns from numerous language data.

Juhan Tuldava is one of the first linguists since the early 1970s to use factor analysis to carry out systematic analysis (Andreev, 1997). His works are devoted to working out a general approach and methods that could be used in linguistic data analysis, and at the same time, to define possible ways of interpreting the results – the groups, or clusters, of linguistic objects and characteristics. After that, other researchers conduct more studies related to writing style, English verb quality, metaphor, word use and sense, etc. (Andreev, 1997; Bainbridge, 1986; Biber, 1988a, 1993; Koiranen, 1995).

Following Tuldava's use of factor analysis for linguistic patterns, Bainbridge (1986) attempts to examine how authors of scientific fiction relate their ideological characteristics to their styles of writing within the four literature types, i.e. hard science, new wave, fantasy and classic; and the meaning of the authors' writing style to readers. He investigates the responses of 595 respondents on author preference, rating 140 for well-known authors. As the data being collected is big and unstructured, he makes use of

factor analysis. From his findings, he describes what the authors' style means to readers, how the authors' ideological characteristics are revealed in their writing, and why people who like science fiction written in the 1960s and 1970s overlap with those who like the science fiction written in the 1920s the 1930s.

In addition to a series of his work devoted to the multi-dimensional study of language, Biber (1993) uses factor analysis to identify the relationships between word senses and uses of *right* by examining the collocation around words. Four factors have been suggested to best account for the text taken from the Longman-Lancaster corpus. Factor 1 are those words which give high loadings to collocations, such as *right hemisphere*, *right sided*, *right hander*, thus appearing to identify the locational sense of *right*. Factor 2 gives high loadings to collocations such as *right now*, *right away* and *right here*, thus identifying the sense of *immediately* or *exactly*. Factor 3 has high loadings for collocations such as *that's right*, *you're right*, *not right*, indicating the sense of *correct*. Factor 4 appears to mark a somewhat less clearly defined stylistic usage of *right* at the end of a clause. In Biber's research, factor analysis takes a large number of different collocations, and reduces them to a much smaller number of reference factors, with loadings indicating the degree of association of each collocate with each factor.

Metaphor, another interesting linguistic element, is part of everyday language. Metaphors are also the processes by which people view the world. A study about the metaphor and concept of entrepreneur and entrepreneurship is conducted by Koiranen (1995) on a speech made by a group of North-Europeans. He finds that this figurative language is a kind of metaphoric expression of how people think and learn, and influences our lives and actions. For example, some entrepreneurial expressions are not real metaphors, but nearer to literal statements, such as *investor*, *owner*, *innovator*. Some metaphors have been used in speech so often that in normal language, people do not often think about their metaphoric nature and origin. He also finds that different respondent groups have very similar views and ideas of entrepreneurship and entrepreneurs.

Also taking advantages of factor analysis in compressing and reflecting structures, Andreev (1997) uses it to classify a number of formal characteristics of 8,375 English verbs that are listed in the seventh edition of the *Concise Oxford Dictionary of Current English* (Sykes, 1987). The formal characteristics he investigates are phonetic, morphemic, derivational and syntactic aspects of the verbs. He finds that these characteristics are clustered into two main groups, one that is manifested within the word base, and the other whose manifestation takes place outside the word base. In other words, the verbs' characteristics reveal a nature of internal or external bases. The former

group includes phonetic and morphemic, while the latter includes derivational and syntactic characteristics.

Other linguistic studies using factor analysis include those by Dewaele (1993), Des Brisay and her colleagues (Des Brisay, Duquette, & Dirir, 1993), and Warschauer (1996c). Dewaele (1993) identifies the situational, social, psychological and sociobiographical variables that underlie the observed linguistic variation. She unveils that interstylistic and interindividual variation occurs with level of formality, fluidity and complexity of speech among Dutch-speaking students. Des Brisay and her colleagues (1993) analyse two sets of students' test data to build up the patterns of learners' language ability and learning experiences. Warschauer (1996c) investigates the effects on student motivation of using computers for writing and communication in the language classroom. It was found that the students overall had a positive attitude towards using computers and that this attitude was consistent across a number of variables, including gender, typing skill and access to a computer at home. More studies applying the same analysis also attempt to define the class of associated words (Schmidhuber & Prelinger, 1993), to analyse metaphors to restructure how people think and speak (Koiranen, 1995), and to examine students' language test performance (Streiff, 1983).

2.4.3 Biber's approach on seven dimensions

In his study on the variation between speech and writing, Biber (1988a) examines 67 linguistic features, ranging from past tense and perfect aspect verbs, to synthetic negation and analytic negation. The sample texts he used were from a corpus of about 960,000 words, collected from the LOB corpus, London-Lund Corpus and some unpublished letters. The text samples are of 23 genres, covering 17 written and 6 spoken. The frequency of occurrence of each feature in each different genre is first obtained through computerised calculation. Then, the technique of factor analysis is implemented to generate several factors out of the 67 linguistic features. Biber then argues that each of the factors is a dimension, and the variation of the two language modes, speech and writing, should be observed and compared along all the dimensions together.

The seven factors Biber extracts from the 67 linguistic features include:

Factor 1: informational vs. involved
Factor 2: narrative vs. non-narrative
Factor 3: explicit vs. situation dependent
Factor 4: overt expression of persuasion
Factor 5: abstract vs. non-abstract
Factor 6: online informational elaboration

Factor 7: academic hedging

As expected, genres differ from each other in their relative position on the dimensions as the frequencies of the language features are measured. However, the interesting finding is that the differences between the genres are of different patterns on different dimensions. On the dimension of Factor 1, for instance, the genres of face-to-face conversation, spontaneous speech and personal letters are similar to each other in that they all stand very high scores on this dimension (Biber, 1988a, pp. 127-135). On the dimension of Factor 3, however, it is the genres of fiction, conversation and personal letters that show close similarities (pp. 142-148). That is to say, the variation of the speech/writing modes of language cannot be examined from just one dimension, but requires a comprehensive view of all of the dimensions.

From the discussion above, it can be seen that Biber's study has made a strong case that the variation between the two modes of speech and writing is not just across a few language features or on one simple dimension, and that any further study on different language modes may also benefit from a multi-feature/multi-dimension approach like this. As my interest is in exploring the linguistic features as they occur in the new mode of language, the CMC, the MF/MD approach used by Biber is an important reference in this research.

2.5 My Rationale for the Current Study

As discussed in the previous sections, we are aware that there are many varieties of language used in different social contexts. These varieties are conditioned by groupings of speakers according to factors like field, mode and tenor, as raised by Halliday (1989); or topics, channel, participants and settings, as suggested by Hymes (1964a) and Brown and Yule (1983). However, while some language varieties are naturally developed out of these factors, some are deliberately developed by language users. Moreover, some language varieties may become so different from others that they develop into unique languages themselves, while some may only keep a flavour of accent, vocabulary and grammar of their own, though they remain perfectly intelligible to users of other varieties of the same language. The phenomenon of these language variations is further discussed below.

2.5.1 Language varieties and social identity

Language is a major tool that human beings use to communicate. Its forms occur differently in different social contexts because human beings modify their ways of speaking and writing due to the different settings of contexts in which they interact. These diverse linguistic forms vary, hence, various varieties of a language develop.

Some major language variants are, to name a few, jargon, sub-language, new language and new varieties. Jargon is "used for communication in limited referential domains, e.g., trade, labour recruiting" (Romaine, 1988, p. 117). A typical example of jargon is the language full of law terms used by a group of lawyers in a law meeting. This is because the context shapes the communication and the participants here need a set of law terminology to precisely convey their meaning to their audience on the specific topics, i.e. the law matters and law issues. This kind of language variety is naturally developed out of the social factors mentioned by Halliday (1989), Hymes (1964a), Brown and Yule (1983).

However, some varieties are deliberately developed by language users, for instance, the sub-language. In a community, teenagers frequently use a subset language and slang, which is created by themselves and that is only intelligible to the members of their group. In other words, people who are outside of this teenage group might have difficulty in understanding the language used (Rosen & Burgess, 1980). The language used by this teenage group is a kind of sub-language as they still use the *same* language, but a *different* variety. This sub-language and the jargon mentioned earlier exhibit the fact that people use the *same* variety of language to show their identity as part of a social group when they communicate within their own speech community.

In other cases, some language varieties may become so different from others that they develop into unique, new languages themselves. A classic example would be that several modern languages, i.e. *French*, *Italian*, *Romanian* and *Spanish*, descended from *Latin* several hundred years ago, while Latin also earlier descended from another language, *Italic* (a variety of Indo-European language used several thousands years ago), too (Anderson, 1998). Usually, these changes take place so gradually that the people speaking the language will not notice the changes. But after hundreds of years, the changes become apparent. The languages newly emerged here are the result of language change. However, another type of new language that has emerged is very interesting, as it does not follow this changing process. An artificial language, *Klingon*, is a typical example of this new emerging language.

Klingon was originally designed by Marc Okrand in 1985 for Paramount Pictures

Corporation, to serve as the language of the Klingons in the second *Star Trek* movie and all subsequent Star Trek (Grune, 1997; Hermans, 1999). However, this artificial language was later used by a group of people in real life. They translate the Bible, Shakespeare and many other works into Klingon, and write new literature in this newly created language. The designer of the language even wrote a Klingon dictionary. Most of Klingon's forms are mainly based on components of general American-Indian languages with no gender, no case, no adjective and no verb tense. Its writing system looks like English and Dutch, and its word order of the subject, verb and object has two major types: SOV and SVO (Hermans, 1999).

Though Klingon is created artificially by borrowing many components from American-Indian language, it has become a different language and is spread and used in a real speech community by a group of real people who are big fans of Star Trek (Hermans, 1999). In Hermans' study, this group of users claim to be able to speak this artificial language, despite the fact that it was artificially designed to serve as a language used in films, only with the intention of creating reality. As the language was not designed to be used in the real world, the users learn it and use it among themselves, simply because they enjoy the group identity that they can sense when using the language.

One more phenomenon of language varieties is that some develop different accents, vocabulary and grammar of their own. Though they may still be intelligible to others of the same language, they have apparently become new varieties. Crystal (1997), following Kachru's view (1992), mentions that English language is divided into three circles: inner, outer, and expanding circles. The inner circle is comprised of the varieties of English, as mother tongue, in which the linguistic forms of sound, word order and meaning are close to each other. This circle includes British English, American English, Canadian English, Australian English, New Zealand English and South African English.

The outer circle is the varieties of English, as second language, whose linguistic forms deviate far more from each other. This circle includes the English spoken in India, Pakistan, Bangladesh, Sri Lanka, West Africa, East Africa, Caribbean and Singapore, where countries were colonised by Britain in the 17th Century. All of these varieties have deviated from British English and have developed a different accent, vocabulary and grammar of their own. However, they are still intelligible to the speakers of other English varieties. All of the Englishes in the above inner and outer circles belong to the varieties of English, and the speakers in a speech community of a variety use that variety for communication to also show their identity with that social group.

Moving into the electronic age, people now use computers to transfer messages for

communication. As the new context is different from the old, where there is no computer, and time and space are no longer such a big barrier to communication, naturally, new varieties would emerge from this CMC. But is the language performed here a completely new variety? Or is it similar to some registers/genres that already exist? In this research, I intend to investigate the linguistic forms of CMC by examining the language performed by a group of professionals in an academic context. I want to know whether the language used here emerges as a new variety, and if yes, what patterns it has. From my observation of several CMC discussion lists over the past years, I predict that the language form at least deviates from the current forms of spoken and written English. It is the possible differences in this new variety that I plan to examine.

2.5.2 Intelligibility in international English

In a speech community, human beings use language to communicate. When they communicate, they use the "same variety" of the language because it gives them a sense of identity, a feeling that they belong to the same group (Crystal, 1997). This variety makes them different from other speakers who speak other varieties either of the same language, or of other languages. Hence, language varieties provide "identity to the group which own them" (p. 134). Crystal further says that language can show one's identity "If you wish to tell everyone which part of a country you are from, ... the most convenient solution ... is ... you ... speak with a distinctive accent and dialect" (p. 134). It is therefore easy to identify many varieties of a language developed in different contexts. This is true in different social classes, and we are not surprised to find that engineers, political activists and teenagers, for example, use their own variety to communicate with each other.

While people maintain their group identity by adhering to certain varieties of language, they also need to communicate with people of other groups. We can think of many occasions in which people have to switch codes between varieties and even between languages in order to successfully communicate with others. In a context in which speakers need to switch codes because of the new factors involved, i.e. the changing context *co-text* (Brown & Yule, 1983; van Dijk, 1977), we are faced with more important consideration – intelligibility.

In the context of an English speaking country, for example, "Most people are already 'multidialectal' to a greater or lesser extent" (Crystal, 1997, p. 137). They use an *informal* variety at home, or when talking to other members of their local community. They use a *more formal* variety when they are away from home, travelling to different

parts of their country or interacting with others at their place of work. Those who are literate have learned the variety of written *standard* English, "which (apart from a few minor differences, such as British versus American spelling) currently unites the English-speaking world" (p. 137).

It can be seen then that the choice of the use of varieties is quite natural and more and more people have developed an awareness of the varieties and the ability to choose the most appropriate to use in different occasions. In the choice, two important, if not exclusive, factors to consider are group identity and intelligibility. How do people achieve mutual intelligibility without losing their group identity? This is what Crystal describes as an attempt "to have their cake and eat it" (Crystal, 1997, pp. 134, 138).

The variety of English that can serve the function of international communication, according to Crystal (1997), is "World Standard Spoken English" (WSSE). Crystal argues that WSSE is still in its infancy. "Indeed, it has hardly yet been born" (p. 138). However, he predicts that it will receive more influence from American than from British English in terms of grammar, spellings and lexicon. This is due largely to the influence of media, i.e. worldwide magazines, radio broadcasting, TV programmes, and in the recent time, CMC.

Personally, I find the use of the term "World Standard Spoken English" a little bit risky. I do not think most linguists would agree that there is a "standard" variety of spoken English, which implies all the others are non-standard. As language is changing (Anderson, 1998), it is hard to tell what is standard and how long the standard can hold its prestige status. The use of this term would involve a value judgement, and I would rather avoid touching upon it here.

However, I do agree that Crystal's concept of WSSE does not replace any national language variety. Rather, it supplements it. "People who can use both are in a much more powerful position than people who can use only one" (Crystal, 1997, p. 138). Whatever name it has, and whatever linguistic characteristics it involves, there should be a language variety which can guarantee international communication and mutual intelligibility, when people need it.

As introduced in Chapter One, mutual intelligibility is a key issue in international communication where speakers are from multilingual countries (Crystal, 1997; Jenkins, 1995, 2000). In Jenkins' view, phonological forms are very important to intelligibility when multilingual speakers converse with each other in English (Jenkins, 1995, 2000). I believe that the intelligibility has the same high degree of importance in CMC, as more and more people communicate via a computer network worldwide. I also believe that the

mutual intelligibility in CMC would mainly, if not exclusively, lie on the understanding of the nature and characteristics of this new variety used in this new medium. I hope to discover how this new language variety looks, and I believe that my investigation will contribute to the mutual intelligibility of international English.

2.5.3 Language awareness in the CMC age

Language varieties are numerous and only a proper one is good for a specific context. For example, an *informal* variety at home, and a *more formal* variety in an academic context is appropriate. It is also suitable in a social gathering to use an *informal* variety when chatting to close friends, but a *more formal* variety when giving a public speech. The awareness of such variety selection is an important objective of language teaching. Nowadays, in a world of electronic communication, English has been taken as a means for international communication (Kehoe, Pitkow, Sutton, Aggarwal, & Rogers, 1999). As a major language for international communication among users in different contexts across regions, ethnic groups and social status, the English language is undergoing changes and new varieties are emerging. I expect to find new varieties of English emerging in the CMC medium, which serve both the functions of personal identity and international intelligibility, and I believe it is of vital importance to call the attention of English teachers and learners alike to this development. As discussed in Chapter One, English language teaching (ELT) has moved away from the old mechanism towards the communicative approach, and I am confident that my investigation on the CMC varieties will support this theoretical movement in ELT.

2.6 Summary

This chapter began with a review of some basic elements of linguistic forms used when people communicate with each other in a speech community. These include grammar, style and stylistics, register and genre, context, co-text and context situation. In each element, some discussion was given to how each of the features influence the language that people use in different social contexts.

Further discussion focused on modes of language, contextual elements, the characteristics of speech and writing forms. In contextual elements I have talked about setting and participants. In discussing the characteristics of two language forms, I elaborated the lexical density and grammatical complexity with other linguistic features

that mark the variation between speech and writing. Then I carried on to propose a solution that would give a better answer to the distinction of speech and writing forms, i.e. the multiple dimensions of variables used to examine the language varieties. Following this, a thorough demonstration detailed how multiple dimensions can be reached, and how they were implemented in the language analysis. After discussing these, I tried to sum up with the rationale that language used in a CMC medium would result in new varieties, as the context within the new medium is different from the context without it in the past.

The main interest of this study is to examine how the new variety is used when people use CMC for communication. As CMC is a kind of international communication, and English is an international language, how an intelligibility could be achieved across multilingual speakers would depend largely on the understanding of this new variety. Moreover, the understanding of this variety and of a proper selection of language variety is very important in English language teaching.

Language varieties develop out of many social and regional circumstances. In addition to the modes and contexts discussed in this chapter, another possible factor that influences the characteristics of a language variety is the medium of communication. The focus of this study is the language variety that emerged in computer-mediated communication (CMC). More literature on the development of the medium and its influence on language and language teaching will be reviewed in the next chapter, Chapter Three, Observation and Description of CMC Contexts, Language Variation, and Language Teaching.

Chapter 3: Observation and Description of CMC Contexts, Language Variation and Language Teaching

3.1 Introduction

The use of computer-mediated communication (CMC) is becoming rapidly widespread. Six years ago, 17,000,000 people were using 8,000,000 computers to distribute information by interactive TCP/IP (Transmission Control Protocol/Internet Protocol) services (Liu, Peek, Jones, Buus, & Nye, 1994), while 26,000,000 users of 10,000,000 computers were accessing information through interactive TCP/IP services. In the same period, 39,000,000 people were using email (Matrix Information & Directory Services, 1995). In a survey carried out in 1996 (Agre & Schuler, 1997), it was reported that Internet access had grown at a rate of 100% per year since 1988. Moreover, at the beginning of the new millennium, eTForecasts reported that the number of people worldwide using the Internet had grown to up to 414,000,000 (2001). This growth supports the belief that much of our society's life will soon be conducted through the mediation of computing and networking.

Communication via a computer is commonly referred to as computer-mediated communication (CMC). Using technology to achieve such communication has been commonplace in the commercial world since the 1960s. However, using technology in an educational context has been isolated, confined mainly to university faculty, who use it to communicate with colleagues via specialised networks (Holden & Wedman, 1993). No application has focused specifically on researching, teaching or collaborative learning.

In the past few years, advances in hardware and software, including the availability of more powerful and less expensive microcomputers, have enabled increased accessibility to telecommunications technology (D'Souza, 1992; Royar, 1994). The development of a common TCP/IP, central to the operation of what we now refer to as the "Internet," has further increased accessibility (Lai, 1996).

The proliferation of CMC has an immense impact on human life. It means not just a new tool of communication, but the emergence of a new context for human interaction. Many CMC-related studies have been conducted from the perspectives of sociology, education and linguistics, as CMC is believed to exercise influence in those fields. It offers new opportunities to develop a more convenient and efficient way of life, and also

presents challenges to almost all human beings.

In the following sections, we review the literature in the context of CMC and the language variations developed in CMC, with particular emphasis on language learning and teaching in the era of CMC.

3.2 Context of Computer-Mediated Communication and Its Messages

As the use of computer-mediated communication had increased dramatically, it has entered the life of many people and changed the way in which we communicate with each other. With the help of the Internet, some people believe that virtual communities have emerged, in which participants interact with each other in ways no less versatile than in traditional social communities. In this section, different theories about the social contexts in which CMC is used for communication are reviewed. The advantages and disadvantages that the medium offers for communication are also covered.

3.2.1 Virtual community or social vacuum?

The idea of a virtual community was made popular by Howard Rheingold (1994), who describes CMC communities as "cultural aggregations that emerge from the net [Internet] when enough people carry on those public discussions long enough, with sufficient human feeling, to form webs of personal relationships in cyberspace" (p. 5). He thinks that the virtual communities in public CMC networks arise out of people's need to replace disintegrating traditional communities and he describes his own experience of involvement in such a community.

> The idea of a community accessible only via my computer screen sounded cold to me at first, but I learned quickly that people can feel passionately about email and computer conferences. I've become one of them. I care about these people I met through my computer ... . (Rheingold, 1994, p. 1)

3.2.1.1 Online Utopia

The word "community" comes from the Latin *communis* – *com* meaning together and *munis* meaning bound – and is often used to describe people collected in one place (Little & Sanders, 1989). The traditional definition of community includes a group of people, sharing common interests and a common locality (Tonnies, 1957). The modern meaning of the word encompasses more than just the co-location of a group of people. Traweek defines *community* as a group of people with a shared history, with distinct behaviours and with expectations for a shared future (1992). Falk further refines the

definition to include common ideals and a commitment of personal resources and energy (1995). Costello (1993) defines it as a group of people with similar interests, noting that it is characterised by the ways in which the group interacts, shares, participates and experiences fellowship. Wendell Berry (1992) considers community to be a social organiser. He says that a community is identified by mutual interests and operates by "the common virtues of trust, goodwill, forbearance, self-restraint, compassion and forgiveness" (p. 120).

In educational CMC literature, the term community is also used to describe a sense of group identity, which reduces social isolation, encourages interactive mental engagement, and provides a social context for conversation and dialogue (Grabowski, Pusch, & Pusch, 1990; Harasim, 1987).

It is clear, then, that a community exists and functions only with a group of participants who have shared interests and trusts, and who interact with each other within that shared identity. With the help of far-reaching communication technologies, a shared physical location may no longer be a prerequisite for a community to develop.

Morgan (1997) asserts that the Internet can provide specific communities with a venue where they can easily connect and do business. Internet systems serve as an electronic city hall, newspaper, church, etc., tying together a dispersed community. "Whether online communities were formed for business, fantasy or relationship purposes, they offer two-way interpersonal and mass communication" (p. 4).

Rheingold (1994) also describes people in virtual communities as using the words they type on screens to

> ... exchange pleasantries and argue, engage in intellectual discourse, conduct commerce, exchange knowledge, share emotional support, make plans, brainstorm, gossip, feud, fall in love, find friends and lose them, play games, flirt, create a little high art and a lot of idle talk. (p. 3)

3.2.1.2 Suspicion about the virtual community

While researchers find more proof for the emergence of virtual communities on the Internet, there are people who find the concept of virtual communities doubtful. Patterson (1996) says, "Anecdotal accounts abound with tales of the utopian or dystopian effects of this new medium of communication. Organisational studies portray a hostile or negative environment, while Internet studies paint a playful one." Furthermore, Mizrach (2000) states that critics usually argue that cyberspace is not the real space in which people live their lives; their actions in cyberspace are not real actions, formed by expectations of real consequences, and there are no such things as virtual communities.

Both Patterson and Mizrach believe that people interacting in cyberspace do not have the webs of real dependencies and interchanges that those in "true" communities have. Thus, observing the discourse and "behaviour" of people in cyberspace tells us nothing about their "real" lives. For instance, Kiesler (1984), Stoll (1996) and others have found CMC to be an inadequate way for people to share emotional content, let alone develop meaningful, long-lasting relationships. Jones (1995) also observes that not every collection of people who happen to talk (or write) to each other form the sense of trust, mutual interest and sustained commitments that automatically deserve to be labelled as communities.

Comments on the lack of socio-emotional content in cyberspace may be made on different bases. For instance, Kiesler's criticism is based on the assumption of reduced social-contextual cues in CMC (Kiesler, 1984), an issue with which we deal in the next section. This criticism may also come from the task-oriented impersonal nature of the CMC texts examined. Stoll's criticism of the concept of a virtual community (Stoll, 1996), however, seems to be based on his unpleasant experience of using the CMC technology.

We will discuss further the ideal of social contextual cues in the medium in the next section. Presently, however, we will focus on just two issues: the source of CMC texts examined and the transparency of the technology. We will also examine their relation to the perception of CMC as a communicative context.

Source of the CMC texts examined

Murray (1995) and many others argue that, whether written or spoken, language serves two basic functions: transactional and interactional. The former is task-oriented, for the transfer of information, while the latter is socio-emotional, for the maintenance of social relationships. Of course, many discourses involve both functions. Murray finds that these two functions operate differently in CMC than in communication without computers. It is reasonable to assume that CMC texts in organisational settings tend to serve a more transactional than interactional function.

In a survey carried out by Patterson (1996), it is found that the traditional study of computer-mediated communication and groups has relied on data from organisational case studies (Finholt & Sproull, 1990; Hiltz, 1984; Rice, 1984; Sproull & Kiesler, 1991), field and laboratory experiments (Ahern, 1994; Eveland & Bikson, 1988; Fuller, 1994; Hiltz, Johnson, & Turoff, 1986; Straus & McGrath, 1994; Walther & Burgoon, 1992; Weisband & Kiesler, 1996), educational settings (Bonham, Cifuentes, & Murphy, 1995;

Hesser, Hogan, & Mizell, 1992; Hiltz, 1995; Hudspeth & Maxey, 1995) and computer conferencing systems (Nunamaker, Dennis, Valacich, Vogel, & George, 1993).

It is therefore argued that early research on CMC, mostly carried out in organisations or laboratory experiments, led to its being characterised as inherently impersonal, lacking in normative reinforcement and with less socio-emotional content than other forms of communication (Hiltz, Johnson, & Turoff, 1986; Rice, 1984; Sproull & Kiesler, 1991).

Visibility and transparency of technology

In an interview by Hart (1996), Clifford Stoll describes his resistance to the use of CMC.

> For example, if I want to use email, I know I would have to invest many hours of my time just to learn how to use it. I have to figure out how to use a computer. I have to learn how to log in to the Network. I have to learn what's changed, what's busy, what's novel about things. All of these things are an investment of my time and energy. I could equally well invest that time and energy in talking to some friends around a table at the coffeehouse. I could spend that time and energy fooling around with my cat, Milo. What I'm worried about is that people advertise email as being revolutionary. "It'll change the way our lives work." I doubt it. I doubt it. (Hart, 1996)

The assumption that Stoll makes is that interacting with participants in CMC is not "really" interacting with "people." He assumes that to make and have "friends," you must interact in "real life," in a "real physical world," where you can "be with people" rather than investing time and effort in "this virtual world called the Internet" (Hart, 1996).

The interesting thing is Stoll's complaint that CMC, in this case, email, is taking time away from real life human interaction, tuning us out of the "real world." When he complains about the time and effort he has to "waste" on learning to handle the technology (Hart, 1996), we can see a kind of technology phobia. This is similar to people preferring to stand in long lines at bank counters rather than using Auto Teller Machines (ATMs).

As a matter of fact, the user-friendliness of the CMC technology is an important factor in determining the effectiveness of participating in the virtual community. The terms "visibility" and "transparency" are used to describe the degree of effort the user has to make in using the technology (Kollock & Smith, 1994). When visible, the computer is often an impediment to communication; when invisible, the computer no longer exists as a separate artifact in the activity, but becomes an extension of the person, referred to as the "functional organ" in activity theory (Kaptelinin, 1996). This point is further

elaborated by Patterson (1996).

> When invisible, the computer no longer exists as a separate artifact in the activity or the delivery tool, but becomes an extension of the person. After learning to use the various software, editor and quotation system, newsreader and email, participants are no longer aware of the existence of the computer. The computer becomes as invisible as a window pane. The computer becomes visible when problems arise. But the invisibility of the computer does not negate its participation in the interactions. It actively structures the interactions.

In a study on the BBS-based distance education programme, Johanson (1996) also observes the distraction of the students, which is caused by some technical difficulties.

> The difficulties that many of the students had with getting connected, getting class access and/or mastering the functions of the BBS or their own telecommunications software, had a negative impact on the feelings of comfort and satisfaction of the whole group. The students were distracted from content-related interaction by technology-related interaction for at least the first three weeks of the term.

3.2.1.3 Virtual community revisited

While some doubt about the ability of CMC to serve socio-emotional functions or to foster a virtual community may be fostered by the reasons presented above, a rich social environment in cyberspace is still likely, provided the unfavourable conditions are excluded.

In terms of the apparent contradiction between the transactional and interactional functions, Chenault (1998) finds that, in fact, there is a high degree of socio-emotional content observed in CMC (e.g. Rheingold, 1994; McCormick & McCormick, 1992; Ord, 1989; Rice & Love, 1987), even in organisational and task-oriented settings (Lea & Spears, 1991). The Freeman and Freeman study (as cited in Hiltz, 1984, p. 176) of scientists connected to a CMC network revealed that friendships developed online over time. By the end of seven months of interaction, all of the participants reported that each was either a friend of the others or the friend of a friend. The participants also reported that these friendships were valuable when they needed help or information from their online colleagues. Johanson confirms this and says that with the passage of months, the network changed from a clique structure of small groups of *friends* to a *genuine* community (1996). Another study by Chesebro (1985) also claims that around 30 per cent of the postings in a random sample of computer bulletin board messages were interpersonal in nature.

In his criticism of email communication, Stoll says that it "denies the sense of who you are and where you are," and that it leaves out the "most important things about you,"

including appearance and personality. He asks, "Might it be that the nature of electronic mail limits you to only that which you wish to show to other people?" (Quoted in Hart, 1996) However, Lea & Spears (1991) find that even first-time users form impressions of other communicants' dispositions and personalities on the basis of their "communication style."

Many studies examine the interpersonal relationship. Walther (1990) proposes the elapsed time and frequency hypotheses about relational communication in CMC, speculating that closer social relations may be developed in proportion to the time and frequency of interaction. However, other studies (e.g. Gunawardena, 1994) show that the length of the engagement and message frequency did not seem to influence the relational tone of users' messages. Nevertheless, most researches are consistent in the conclusion that people find ways to express immediacy and compensate for the lack of physical intimacy in CMC.

Mizrach (2000) argues that cyberspace allows people to conceal many of their "real life" contexts - e.g. gender, race, culture, ethnicity, etc. Nevertheless, this does not eliminate them, and often creates norms and values of its own in their place - "netiquette." Thus, CMC does carry emotional and impression-forming content.

As for the problem of the medium's visibility in the communication context, as soon as the users are accustomed to the technical operation of the system and the technology becomes a true extension of man, living in the virtual community should be as natural as any other form of human community. In Johanson's experiment (1996), he found that once the technical difficulties were overcome, the students worked comfortably with the technology. It became the tool that allowed them time and place independence, and as such it ultimately had a positive impact on student satisfaction.

The emergence of virtual social communities on the Internet seems a natural tendency of human nature. Wellman and Gulia (1995) find that even when online groups are not designed to be supportive, they tend to be. As social beings, those who use the Internet seek not only information but also companionship, social support and a sense of belonging. Rheingold affirms this by saying "Any difficulty is harder to bear in isolation. There is nothing to measure against, to lean against. Typing out my journal entries into the computer and over the phone lines, I found fellowship and comfort in this unlikely medium" (1994, p. 20).

A famous quote from one of John Donne's sermons in 1624, *For Whom the Bell Tolls* – "No man is an island, entire of itself; every man is a piece of the continent, a part of the main" (Abrams, 1987) – reveals that people are social beings, and they tend to form

communities by all possible means. It should not then be a surprise that the convenience of the Internet provides them with the means to develop new forms of communities, which more and more people will get used to and enjoy.

3.2.2 CMC and social contextual cues

A set of frequently cited research on the characteristics of CMC is that of Sproull and Kiesler (1986, 1991), who have identified a complete lack of expressive (non-verbal) behavioural cues, such as the nuances of a conversation created with nods, smiles, eye contact, distance, tone of voice and other non-verbal behaviours, in CMC contexts. They claim that, in CMC, traditional forms of communication become "mystery" variables. Hence, much misunderstanding arises as a result of perceived meaning derived from context and the tone automatically attributed by a human's imagination (Metz, 1994).

However, others argues that, due to the lack of the paralinguistic and non-verbal cues in CMC, people experience a sense of anonymity, which gives them liberty to behave freely, even irresponsibly, in interacting with each other (Kiesler, Siegel, & McGuire, 1984; Reid, 1991; Sproull & Kiesler, 1986). This results in the weakened social norms in CMC, including unregulated use of language (flaming), but also a more democratic and liberal environment of communication.

3.2.2.1 Theoretical framework

The assumption of social contextual cues is based on several theories, the most important of which is the social presence theory and the media richness theory (Wolski, 1999). Social presence theory (Short, Williams, & Christie, 1976) is the most widely cited theory in CMC literature. Essentially, it focuses on the perceptions of the media's ability to translate the social presence of another. Short, Williams, and Christie define social presence as "the degree of salience of the other person in the interaction and the consequent salience of the interpersonal relationships" (p. 65), meaning the degree to which each person perceives the other as a *real person* and their interaction as a *personal relationship*. The media is ranked on the degree of social presence it allows, with face to face (FtF) being richer than audio channels, and text-based CMC at the lower end of the continuum.

Media richness theory, like social presence theory, also takes into account perceptions of the medium (Daft & Lengel, 1984, 1986; Trevino, Lengel, & Daft, 1987). However, instead of making predictions about the message content or outcome based on the perceived or actual ability of the media to translate social presence cues, media

richness theory makes predictions about the choice of media to match task requirements. The ability of the media to transmit non-verbal information is measured by the number of *channels* a medium provides for the transmission of various kinds of data (Walther, 1990). In the face-to-face context, communication occurs along multi-sensory channels that receive and/or transmit audio, visual, tactile and even olfactory information. It is, therefore, a "rich" media best suited for a complex and ambiguous task. On the other hand, CMC is limited to the dimensions that can be expressed only in visual channels transmitting text and simple graphics, so it is most preferred for simple and unambiguous tasks (Levinson, 1990; Wolski, 1999).

The hypothesis of reduced social contextual cues by Sproull and Kiesler (1986) focuses not on the perceptions of the technology, but on the restrictions the technology imposes on communication. Researchers have different estimations for the weight of non-verbal cues in communication. Lapakko (1997) cites 65 per cent, while Mehrabian (1981) finds that 93 per cent of meaning in face-to-face communication comes from non-verbal information (Feenberg, 1989; Mabrito, 1989). At any rate, according to the hypothesis of reduced context cues, all media restrict some amount of social context cues. These cues (gesture, expression, kinetics, voice quality, and all of the other components of "speech acts"), normally thought to constitute discourse, are believed to be almost totally filtered out in the text-only CMC context, which therefore tends to be very low in social presence (Walther, 1990).

This perspective predicts that as social cues are weak or absent, communication becomes more self-centred and unregulated, more extreme and more impulsive (Sproull & Kiesler, 1986). This also explains the occurrence of "flaming" – the using of offensive and hostile language (Walther, 1992). On the other hand, the lack of social cues also provides the condition for a more liberal and democratic environment.

3.2.2.2 Anonymity and weakened social norms

In 1993, New Yorker magazine published a famous cartoon of a dog at a keyboard saying, "On the Internet, nobody knows if you're a dog" (Kling, 1996). This reveals that the assumed anonymity of participants in CMC contexts is believed to have both negative and positive effects.

On the negative side, anonymity is believed to be associated with isolation. Assuming CMC operates in a social vacuum, electronic communication appears as an environment in which the personal identity of the sender and of possible receivers tends to become blurred. That is, a condition of de-identification occurs, in which people lose

their sense of personal responsibility and of the respect due to social norms. The lack of social context cues are read as an obscuring of the boundaries, which delineates the forms of behaviour that are acceptable or unacceptable.

Reid (1991) finds that "with little regulating feedback to govern behaviour, users behave in ways that would not generally be acceptable with people who are essentially total strangers." This may result in weakened social norms and unregulated behaviour, such as the use of crude insulting language, i.e. flaming (Siegel, Dubrovsky, Kiesler, & McGuire, 1986). The term *flame* has now entered common usage, and episodes of intolerance expressed through CMC usually denote rage and frustration (Mantovani, 1996).

Though flaming is generally defined as "rude, impulsive behaviour, and the use of crude or even insulting expressions (Mantovani, 1996), Thompson (1992) proposes to define flaming by the absence of particular message components. According to Thompson, a flame is expected to have more copula deletions (omitting the "to be" verb) compared to a non-flame message. He also measures user awareness of their omission of copulas in their email messages, as well as their perceptions of how often they are flamed. Though copula deletion was not found to correlate with the perception of a flame, this research is important in that it shifts our thinking about flaming to a linguistic perspective (Wolski, 1999).

3.2.2.3 Compensation for lack of social contextual cues

The assumption that the lack of social contextual cues results in anonymity in CMC, and consequently in weakening social norms, such as flaming, has been challenged widely. Lea and his associates (Lea, O'Shea, Fung, & Spears, 1992) state that flaming is relatively rare and is usually caused by "frustration arising from the communicational inefficiency of the medium" (p. 98), especially when the person concerned is under stress, rather than by any lack of proper social and normative contexts in CMC (Mantovani, 1996). In Thompson's study (1992), 58 per cent of the participants report that they have not been flamed. Recently, Kayany finds that the frequency of flaming differs between the newsgroups examined, as well as within newsgroups according to the general topic under discussion (Kayany, 1998). This suggests that social context, but not the medium itself, is the primary determinant of online uninhibited behaviour.

Walther (1990) challenges that the research on the interpersonal interactions of CMC groups has been flawed by using previously unacquainted communicants and short-lived groups. Later, he also suggests that, given time, CMC participants will employ

compensatory techniques to establish social presence and immediacy, and that they ultimately engage in group dynamics as intense and rich as in face-to-face relational communication (Walther, 1990, 1992, 1994).

Reid (1991) also states that users must compensate for the lack of these contextual cues. CMC users have managed to develop "ways of sending computerised screams, hugs and kisses." Here, she refers to the icons made up of a simple combination of computer keystrokes that are used to express emotions. They are known as "emoticons" or "smileys," such as ":), ;), :-), ;-)" for smile faces, or ":(, :-(" for sad or angry expressions, just to name a few.

Overall, it is estimated that there are well over 500 of these icons. Such creativeness implies that computer users have the ability to adapt to limited communication situations, and even substitute graphic symbolic cues to replace conversational shortcomings (Metz, 1994). This characteristic particular to CMC texts is further discussed in Section 3.3.

Sproull and Kiesler's concept of a lack of social contextual cues in CMC is based on the assumption that visual text messages are the only channel of message conveyance in CMC. They ask, "How do people treat one another when their only connection is a computer message?" (Sproull & Kiesler, 1991, p. ix), which reflects their peculiar approach to the CMC communication. In reality, as Mantovani (1996) argues, people are never connected "only" by means of a computer message. They are linked by an organisational network that parallels the technological, and also by a set of partly shared expectations, needs and goals, which are to some extent reciprocally recognised. Moreover, CMC generally does not work in isolation as an alternative to other media, but fits into the pre-existing media with which it constantly interacts (Kling, 1980, 1994, 1995; Kling & Scacchi, 1982).

Mantovani (1996) criticises the concept of reduced social context cues, stating that it has reduced the social world merely to a direct physical connection between people. However, "It is not enough to be physically alone in a room to be removed from the social world. Social context is not something outside or above people; it is both around and inside them, an integral part of their identity" (Mantovani, 1996). Lea and Spears also mention that CMC users resort to other facilities to compensate for this absence of social cues.

> In fact, it could be argued that an absence of social cues from other interacting individuals, together with the resulting uncertainty, forces people to resort to default social norms to guide their behaviour. (Lea & Spears, 1991, p. 286)

From the above discussion, it can be seen that, while the assumption about social contextual cues does give us an inspiring point for study, CMC is certainly not a simple text-exchange environment, but a complex context that deserves more detailed examination.

3.2.2.4 Democracy and liberty

The lack of social contextual cues in CMC, along with the anonymity that it effects, is not without positive effect on some participants. Sproull and Kiesler (1991) consider CMC as capable of letting people express themselves more openly because it isolates them from social contexts.

> People interacting on a computer are isolated from social cues and feel safe from surveillance and criticism. This feeling of privacy makes them feel less inhibited with others. It also makes it easy for them to disagree with, confront, or take exception to others' opinions. (pp. 48-49)

Reid (1991) also finds that a reduction in social context cues allows "the shy and socially ill-at-ease" a way of learning social skills in a non-threatening environment. People who are shy, who feel guilty or disturbed by conflict with others, may find protection in the presumed isolation of CMC. For those with relational difficulties, the apparent isolation of CMC is only one aspect of the complex dynamics governing communication in groups and organisations (Mantovani, 1996). The lack of social context cues also makes it easier for the shy participants to move in and out of conversations in a CMC format (Davis & Brewer, 1997).

Lee (1996) also finds the use of email for communicating to be special in its sense of immediacy and informality. "Email's immediacy and democratic power have their nearest parallels in oral communication, but email is not an oral medium; rather, email converts correspondence into a virtual conversation" (p. 279). Thus, as claimed by Waddell, it allows for individualisation not possible in verbal conversations (Waddell, 2000).

Morgan (1997) asserts that the Internet creates new norms and rules for interaction, and intranet systems provide the workplace, within an institution, with a more accurate way of organising and controlling a vast amount of information accurately. Morgan believes that the online system serves as an equalising basis for community members who have access to the Internet. Waddell further proposes that the online system is expected to play the role of "the great equaliser" (2000).

The sense of democracy and liberty found by many researchers among CMC

participants is not a universal phenomenon. In certain cases, CMC does allow people to be more open and able to give free rein to their self-expression, because they feel protected by the relative anonymity of CMC, and because they feel (mistakenly) that CMC is "ephemeral" in character (Sproull & Kiesler, 1991). However, as Mantovani (1996) says, whether electronic environments do in general make people feel less inhibited is highly questionable. For the "Internet equality" claimed by Morgan (1997), there are also challenges raised by different researchers.

As a matter of fact, years before Morgan's claim of a democratic equal community, contrary opinions were stated in many works. A series of studies conducted by Tora Bikson and her colleagues (Bikson, Eveland, & Gutek, 1989) find that email is effective in overcoming physical barriers, though not necessarily, but not social barriers, as these technologies primarily enhance existing interaction patterns. Child and Loveridge's (1990) studies, undertaken in different European countries, find that CMC is usually designed precisely to support ongoing hierarchical relations. Reviewing these studies, Mantavani (1996) concludes that whether or not CMC engenders democracy actually depends on the social contest of CMC use, on the organisational history, and on the rules determining CMC's application.

Gender equality is an important structure examined by researchers in the CMC settings. Among the studies, Herring (1994) reports

> Despite the democratising potential ... male and female academic professionals do not participate equally in academic CMC. Rather, a small male minority dominates the discourse both in terms of amount of talk, and rhetorically, through self-promotional and adversarial strategies. Moreover, when women do attempt to participate on a more equal basis, they risk being actively censored by the reactions of men who either ignore them or attempt to delegitimise their contributions. Because of social conditioning that makes women uncomfortable with direct conflict, women tend to be more intimidated by these practices and to avoid participating as a result ... [the] conditions for a democratic discourse are not met: although the medium theoretically allows for everyone with access to a network to take part and to express their concerns and desires equally, a very large community of potential participants is effectively prevented by censorship, both overt and covert, from availing itself of this possibility. Rather than being democratic, academic CMC is power-based and hierarchical.

The studies by Murray (1996), Savicki, Lingenfelter, and Kelly (1996), and Barrett and Lally (1999) all report similar findings: that men outnumber women as participants in online discussion groups.

Other than the issue of gender, researchers also find other points of inequality in CMC settings. Schneider (1996) notes that very few of the participants account for an extraordinarily high proportion of the content. Scott and Easton (1996) also question the claim of equal influence in CMC forums, because historically high- and low-influence

individuals are still perceived as such in CMC contexts.

Moreover, Wolf (1998) disagrees with Morgan's conception of CMC as "the great equaliser." She finds that the typical Internet user is a white male, who is well educated and has an above-average income. Besides, women are often not on the Internet, due to constraints such as time and financial difficulties. They are often working mothers with too little time to attend to too many priorities. She concludes that "the elite's access to the Internet increases rather than decreases inequality, since the ability to deploy one form of power often tends to facilitate access to other forms of power" (p. 26).

In a study on the social identity model of de-individuation (SIDE) effect of CMC, Postmes, Spears, and Lea, (1998) conclude that, "Although CMC gives us the opportunity to traverse social boundaries, paradoxically, it can also afford these boundaries greater power, especially when they define self- and group-identity."

It seems then that the CMC as a medium itself does not guarantee anything. It takes the whole context to decide the nature of a certain CMC environment. That is why Fulk and her associates (Fulk, Schmitz, & Schwarz, 1992) argue for the development of situation-specific hypotheses rather than generalisable theory. More researchers agree that the CMC context includes the interaction setting, the message receiver, the communicative function of the message, the sender's individual psychological state and the medium itself (Matheson & Zanna, 1990; Murray, 1988; Sherblom, 1988). All of these components impact on language choices in that the context influences what and how something is said, while what is said will influence the context (Wolski, 1999). We will discuss context further in Section 3.2.3.

3.2.3 CMC as a communication context

As can be seen from the review in the above sections, although the traditional CMC research, conducted on organisational sites, tends to depict it as inherently impersonal, the recent Internet-centred studies characterise CMC as a lively social space. On the other hand, as Patterson (1996) argues, though CMC may encourage wide participation and candour, with an emphasis on contribution merit, these are not inherent characteristics of the medium, and neither is anonymity, egalitarian social organisations or flattened social hierarchies, as reported earlier (Hiltz, Johnson, Turoff, 1986; Nunamaker, Dennis, Valacich, Vogel, & George, 1993; Rice, 1984; Sproull & Kiesler, 1991.)

CMC, as it has been argued, is not just a tool for communication, but a rich communicative context, combining aspects of oral and written discourse (Ferrara, Brunner, & Whittemore, 1991), and which actively encourages participation in the

interaction (Metz, 1994). In this section, I intend to examine the formation and functions of this context. The discussion begins with the related structuration theory and social identity theory, then moves on to the participants and settings of CMC.

3.2.3.1 Structuration theory and social identity theory

Structuration theory is proposed by Giddens (1979), stating that human behaviour is influenced by context, but the same behaviour also influences and changes the context. As mentioned earlier, the context is found to include the interaction setting, the message receiver, the communicative function of the message, the sender's individual psychological mood and the medium itself (Matheson & Zanna, 1990; Murray, 1988; Sherblom, 1988). All of these context components have an impact on language choices, which influences what and how something is said, and at the same time what is said will also influence the context.

According to this theory, there is a reflexive nature of social life, i.e. the way in which the structure of activity is created and recreated by the very activities of which it is constituted. This image has particular applicability in the context of CMC. We cannot study the effects of CMC upon the participants without at the same time studying the role of the participants in shaping and reshaping the context. Giddens says, "Reflections on social processes (theories, and observations about them) continually enter into, become disentangled with, and re-enter the universe of events they describe" (1979, p. xxxiii). Because the actors in this process are self-aware, theories developed and disseminated through the study of the medium can result in the use of that theory by the participants to further modify their communicative environment (Lawley, 1992).

Researchers supporting the theory of social identity believe that individuals carry the social world and its rules within themselves (Tajfel & Turner, 1986; Hogg & Abrams, 1988). They believe that people, on the one hand, have unrepeatable identities as unique individuals, and that they, on the other hand, also have social identities as members of one or more social groups, no matter whether or not the groups are real or fictional. Belonging to one or more reference groups is part of an actor's identity. In their view, the contexts are created between actors' situated interests and the specific situations they face, and the contexts also emphasise the social or individual poles of personal identity.

Another model of CMC interaction, the SIDE model (social identity de-individuation), was proposed by Lea and Spears (1991, 1992), and Spears and Lea (1992). It explains the group polarisation effects in CMC by taking the context of the interaction into account. This model looks at who the message is being sent to, the

individual's role in the group and the displayed attitudes of the group majority. Its predictions are based on the notion that if the self-identity is high, then the attitude polarisation will not occur. This means that the participants' attitudes will become more extreme. On the other hand, when the group identity is high, the attitude polarisation will occur. Wolski (1999) comments that this theory applies to a context that is quite narrow in scope, though it can be helpful in adding to our knowledge about group social influence processes.

3.2.3.2 Participants

Based on the theory of social identity, individuals participating in any kind of communication, including CMC, can never be anonymous. Lea and Spears reformulate the concept of the de-individuation in a sense that anonymity, if associated with immersion in a group, does not negatively affect social norms, but amplifies the prominence of the group and its norms (Lea & Spears, 1991; Spears & Lea, 1992). Therefore, whether CMC makes the social identity of the participants more or less salient, as Mantovani comments (1996), depends on the type of pre-existing group and the tenor of the ongoing communication. The social world is not only outside, but also inside people, as part of their identity, which functions even when they sit - physically alone - in front of their computer screens.

In Lawley's study (1992), it is found that the nature of the relationships of the participants and their levels of involvement developed according to their individual styles, time commitments and autonomous choices. Moreover, the new expectations and norms developed in CMC may change the technology itself. Therefore, Lawley concludes

> Computer-mediated communication (CMC) has been shown to have significant effects on the patterns of communication among its participants. Although some of these patterns are a specific response to the constraints of the medium, others are the result of participant influence on the medium (Lawley, 1992).

In reviewing the field study carried out by Saunder and her associates (Saunders, Robey, & Vaverek, 1994), Fuller (1994) also finds that gender and status account for the differences in interaction quantity and style. This is further confirmed by the experiments testing the relationship of anonymity and equalisation of participation and status studied by other researchers (Jaffe, Lee, Huang, & Oshagan, 1995; Selfe & Meyer, 1991; Weisband & Kiesler, 1996).

An earlier work that also addressed the role of participants in shaping their communication environment is Myers' (1987). He claims that in the BBS examined,

"meanings are continuously negotiated through symbolic interaction. Such an interactionist perspective has seemed out of place in one-way mass communication systems, but finds new life in heavily interactive, computer-based media" (p. 251).

The assumption of the lack of social contextual cues in CMC, as proposed by Sproull and Kiesler (1986, 1991), conceives CMC as a process of uni-directional influence from technology, which inevitably results in anonymity, de-individuation and loss of social norms. After examining the social identity theory and the performance of CMC participants, more researchers (e.g. Mantovani, 1996) accept the idea that the relation between CMC and participants may be seen as multiple, context-dependent and circular. The experience of people who communicate by CMC depends on the type of groups and contexts.

3.2.3.3 Setting of CMC

There are many issues related to the setting of CMC. We will discuss the difference between synchronous and asynchronous settings, the open and closed CMC settings, and the special discourse patterns found in some synchronous and asynchronous settings.

Synchronous versus asynchronous CMC

Many early researchers of CMC, Sproull and Kiesler (1986) for example, limit their scope of study to email. Perhaps this is due to the immaturity of most other types of CMC technologies at the time of the study.

However, it is now clear that CMC settings can be of two types: synchronous and asynchronous. The former include functions such as online talk, Internet Relay Chat (IRC), multi-user dimensions (MUD), and even, more recently, online video conference in which participants must get online at the same time to exchange messages and interact with each other. The latter are functions such as email, voice mail, discussion lists and discussion forums. Users can choose to send or receive messages any time at their will.

Since the two types of CMC settings are very different, Metz (1994) argues that researchers must be clear about which setting they are working on, especially when they attempt to make comparisons. He argues that asynchronous CMC, such as email, and the findings resulting from studies of such material cannot explain the occurrences of culture within Reid's (1991) analysis of Internet Relay Chat, a synchronous form of CMC.

Open and closed CMC settings

There are many types of Internet networking software for both synchronous and asynchronous communication, which enhance different natures in online communication, i.e. openness and closeness. Some computer-mediated networks with software like LISTSERV, ListProc, Majordomo, or JISCmail are open to the public, so every one who is interested can access them directly, with at most the requirement to first go through some very simple subscription formality. This is the case of most commercial and academic networks, including the TESL-L and NETEACH-L discussion lists on the theme of English teaching, which are used as the sources of samples in this study, asdiscussed in Chapter 4: Methods and Procedures of Investigation. CMC settings that have been established by certain organisations or institutions for a limited scope of users, such as employees of a company or students in a distance education programme also exist. This open or closed nature of a CMC setting may have an effect on the discourse patterns (Pincas, 2000).

Message organisation in asynchronous CMC

One special characteristic of asynchronous CMC settings is the message organisation. Reading messages is not the same experience in such settings. In an email discussion list, the postings are usually sent to subscribers regularly, so the messages are listed chronologically as they are received by the list server. A typical page of email messages from several discussion lists is illustrated in Figure 3.1.

In the list of incoming mail, the user can see the messages from two different discussion lists arranged in the order of their arrival time. The user can simply type the number of the message he or she would like to read and check its content.

In discussion forums, where the user does not receive the messages by email, but goes online to a specific web page to check and read messages, the messages are often listed in topic threads. A topic thread is a group of messages with the same subject line, meaning that all messages that reply to a previous posting will be collected together. This allows the user to check messages of the same topic more easily as they are clustered together. A sample of this threaded, hierarchical format is shown in Figure 3.2.

From the above illustrations, we can see that messages can be read sequentially as they arrive or in threads every day. However, reading messages in bulk, once a week or so, is not the same as reading daily. This is discussed further below.

```
You have mail.
mentor% mail
"/var/mail/teemsht": 23 messages 23 unread
>U  1 owner-tesl-l@CUNYV Sun Dec 31 16:28   44/2350  Bilingual Dictionaries
 U  2 owner-tesl-l@CUNYV Mon Jan  1 04:27   49/2550  Urging Students to Read
 U  3 owner-tesl-l@CUNYV Mon Jan  1 11:40   33/1415  happy new year
 U  4 owner-tesl-l@CUNYV Mon Jan  1 14:17   42/1806  Should we teach Ergative
 U  5 owner-tesl-l@CUNYV Mon Jan  1 22:24   48/2494  Re: Bilingual Dictionarie
 U  6 owner-NETEACH-L@ra Tue Jan  2 01:49   48/2185  Online learning and Telne
 U  7 owner-NETEACH-L@ra Tue Jan  2 09:25   54/2402  Online Learning Materials
 U  8 owner-NETEACH-L@ra Tue Jan  2 11:40   52/2405  Re: Online Learning Mater
 U  9 owner-tesl-l@CUNYV Tue Jan  2 14:56   69/3846  Pre-writing
 U 10 owner-tesl-l@CUNYV Tue Jan  2 14:57   61/2970  Re: should we teach ergat
 U 11 owner-NETEACH-L@ra Tue Jan  2 16:53   92/3714  RE: Online Learning Mater
 U 12 owner-NETEACH-L@ra Tue Jan  2 18:02   48/2052  Re: Online Learning Mater
 U 13 owner-NETEACH-L@ra Tue Jan  2 18:33   71/2576  Re: Online Learning Mater
 U 14 owner-NETEACH-L@ra Tue Jan  2 19:02   94/3278  RE: Online Learning Mater
 U 15 owner-NETEACH-L@ra Tue Jan  2 19:07   68/2811  RE: Online Learning Mater
 U 16 owner-NETEACH-L@ra Tue Jan  2 19:12   89/3383  RE: Online Learning Mater
 N 17 owner-NETEACH-L@ra Tue Jan  2 20:20   52/1853  Solar Eclipse Slide Show
 N 18 owner-NETEACH-L@ra Wed Jan  3 17:32   37/1410  RE: unsubscribe
 N 19 owner-NETEACH-L@ra Thu Jan  4 03:02   62/2810  Salaries for teaching at
 N 20 owner-NETEACH-L@ra Thu Jan  4 10:21   39/1429  Transfer audio cassettes
? z
>N 21 owner-NETEACH-L@ra Thu Jan  4 14:08   95/3721  RE: Salaries for teaching
 N 22 owner-NETEACH-L@ra Thu Jan  4 14:33  129/4848  RE: Salaries for teaching
 N 23 owner-NETEACH-L@ra Thu Jan  4 17:43  118/5200  Your ideal LL environment
?
```

Figure 3.1 A Linear Structure of Message Organisation

- Corpora: 3 Ph.D. positions in Tilburg Harry Bunt (Tue Sep 12 2000 - 16:44:30 MET DST)
- Re: Corpora: Question about a Brown Corpus tag Dirk Ludtke (Thu Sep 14 2000 - 09:22:41 MET DST)
  - Re: Corpora: Question about a Brown Corpus tag E S Atwell (Thu Sep 14 2000 - 12:37:44 MET DST)
  - Re: Corpora: Question about a Brown Corpus tag Frank Henrik Mueller (Thu Sep 14 2000 - 12:57:40 MET DST)
    - RE: Corpora: Question about a Brown Corpus tag Mark Lewellen (Thu Sep 14 2000 - 15:42:25 MET DST)
      - RE: Corpora: Question about a Brown Corpus tag Miles Osborne (Thu Sep 14 2000 - 16:15:03 MET DST)
      - RE: Corpora: Question about a Brown Corpus tag Mark Lewellen (Fri Sep 15 2000 - 18:21:17 MET DST)
      - Re: Corpora: Question about a Brown Corpus tag Tylman Ule (Thu Sep 14 2000 - 16:22:06 MET DST)
      - RE: Corpora: Question about a Brown Corpus tag E S Atwell (Thu Sep 14 2000 - 16:32:21 MET DST)
      - Re: Corpora: Question about a Brown Corpus tag Atro Voutilainen (Thu Sep 14 2000 - 16:57:12 MET DST)
- Corpora: Exciting Job Opportunities in NLE Linda Page (Thu Sep 14 2000 - 10:30:13 MET DST)
- Corpora: AiML-ICTL: Final Call for Participation Advances in Modal Logic (Thu Sep 14 2000 - 23:42:36 MET DST)
- Corpora: spoken/multimodal dialogue jobs at FTW Vienna Gregor Erbach (Fri Sep 15 2000 - 11:49:53 MET DST)

Figure 3.2 A Hierarchical Structure of Message Organisation

Email messages accumulate until they are read. Threaded messages read in bulk can present a more coherent conversation, as Rheingold points out, because larger numbers of messages are grouped in the thread. However, reading this "multilogue" in bulk loses the dynamism of the message exchange that is experienced by frequent readers and posters (Rheingold, 1994).

Reading messages daily, whether threaded or sequential, can be a more disjointed experience because fewer messages are in a thread (Tannen, 1989). When one reads messages sequentially, he or she has to construct the conversation beyond conversational involvement to recapture an idea of what the interaction is about, who the post came from, what their intentions may be and where the post fits in the context of the newsgroup interactions. Patterson points out an interesting phenomenon in reading email messages daily. As many messages in the forum are follow-ups, it may be more difficult for the reader to grasp the overall picture of the discussion, because some follow-up messages may precede the original message and may not include the text of the original (Patterson, 1996).

At any rate, the threaded organisation of messages in many asynchronous CMC settings is a unique characteristic, which differs from non-CMC settings. It would be interesting if further studies could be conducted on the effect of such organisation on reading habits.

Hyper-personal synchronous CMC

Though non-CMC interaction generally proceeds sequentially, it is not often the case in CMC. The processing of messages in threads in asynchronous CMC discourse, as discussed earlier, is a characteristic hardly found in non-CMC settings.

As a matter of fact, the non-sequential multi-threaded discourse, as shown in Figure 3.2, is also likely to occur in synchronous CMC settings. In IRC, MUD or MOO, though the participants are interacting synchronously, they may not take their turn and follow each other's messages in a strict sequence, such as when people are talking on the phone. The main reason may be that there is often some time lag between when a participant is waiting for his message to get through and when his counterpart responds. During the time lag and the waiting, it is tempting to pick up another topic and send out some other messages.

Though this is only my personal experience, it may be an interesting topic for some further research. Moreover, this may also be the reason why Erickson (1999) says

> ... it is disquieting to see how drastically conversation changes when it takes place in the digital medium. Computer-mediated conversation (CMC) is quite different from its face-to-face namesake. It is cumbersome, particularly among more than a few participants. Conversations often seem to lack coherence, tending towards drift and dissolution. In CMC many of our finely honed skills become irrelevant.

An excerpt of a MOO conversation (Figure 3.3) and its discourse structure (Figure 3.4) can be of further reference to this "non-sequential" discourse. In the former, there is the partial sample text from a MOO meeting, with 18 utterances included (Younger, 1996). These utterances are presented sequentially by the computer network system at the time at which it receives the messages sent by all participants around the world. All of the participants are teachers of English, using the Internet for their classes. Among them, Gregor, Frizzy, Millie and Puzzle are from the USA, while Yoshi is from Japan and Chrissy (the researcher of this study) is from the UK. Gregor was the host who organised the meeting to discuss what Internet activities could be successfully applied in ELT.

We can see the discourse structure more clearly in Figure 3.4. Though at a first glance it seems to be very chaotic, the structure of the utterances does reveal a systematic order. In this short conversation, Gregor made utterances U1, 2, 9, 11 and 17, while Mullie made U3, 6, 13 and 18. Frizzy communicated U8 and 15, and Yoshi U4, 7, 12, 14 and 16. Other participants only made one utterance in this initial stage of the conversation (Chrissy: U5, and Puzzle: U10).

From the beginning, Gregor started with U1 and 2, inviting talks from other participants. However, Millie threw out her question (U3) about the proficiency of talking in the MOO environment, which immediately initiated a sub-topic. Yoshi wanted to help Millie (U4), and Chrissy agreed that she was a newcomer, too (U5).

In trying to help Millie, Yoshi suggested that Gregor show an OHP covering the basic MOO commands (U7), while Millie requested for some patience from the others (U6) (to tolerate her not-so-skillful MOO commands that might hinder the communication).

To respond to Gregor's call for the talk, Frizzy gave her experience of using the Internet in her classroom (U8). While Gregor continued his host role (U9), he also replied to Yoshi's request (U11). At the same time, Puzzle felt confused about Gregor's words (U9), and wanted him to elaborate (U10). Delayed by trying to help Millie, Yoshi then gave his experience of using the Internet in ELT (U12), but he still kept in his mind Millie's problem (U13), and further offered to help her (U16) while agreeing with Gregor's decision not to show the basic commands (U14) in the meeting.

Utterance	Part of the MOO Log from a Meeting
	...
U1	Gregor says, "So, let's start by sharing anything we have already tried with students here..."
U2	Gregor nudges Frizzy.
U3	Millie [Guest] says, "I am a newbie...HELP!!""
U4	Yoshi [to Millie [Guest]]: OK, what can I do?
U5	Chrissy [Guest] is a newbie, too.
U6	Millie [Guest] says, "Just be patient...or maybe I should and just take it all in!""
U7	Yoshi [to Gregor]: Show them the basic commands OHP sheet, will ya?
U8	Frizzy says, "I've used schMOOze to hold conferences with students. I logged them (with permission, of course), then sent the log to the students via e-mail."
U9	Gregor says, "I have briefly experimented with students for collaborative brainstorming... Iwant to expand my use of that."
U10	Puzzled [to Gregor]: "you could expand on what you mean.
U11	Gregor [to Yoshi]: I'm going to skip the basic commands for newbies today... they know how to talk, and that's enough for today! ;o)
U12	Yoshi says, "I offered an online class at Frizzy's FUN, but it didn't turn out very well."
U13	Millie [Guest] says, "How about cloosing off the systyem just for your purposes so that no one else can access it...Is that possible?""
U14	Yoshi nods to Gregor.
U15	Frizzy says, "This summer, I had a group of students do role play at schMOOze. They had a political debate (one was Bob Dole, another was Clinton, and one student logged on as God!)."
U16	Yoshi says, "Page me if you have trouble: page yoshi "
U17	Gregor [to Puzzled]: I mean having two (or more) students in the same virtual room brainstorming a topic and logging what they say. They can then look at their conversationlater and pice together what they need to make a next draft.
U18	Millie [Guest] says, "Frizzy, that sounds like fun!""
	...

Figure 3.3 A MOO Conversation Sample

Finally, Gregor then spoke to Puzzle (U17), and Millie returned to the meeting topic by praising Frizzy's Internet activities for ELT (U18) after noting Yoshi's offer of help.

In this short synchronous CMC interaction, it is interesting to see that two adjacent pairs are so far away from each other, which is very different from the norm in face-to-face conversations. The first adjacent pair is U10 and U17, in which Gregor explained to Puzzle what he meant in U9, "I want to expand my use of that." Gregor, acting as the host of the meeting, needed to take care of the sub-topic discussed between Millie, Yoshi and Chrissy, and it was this concern that made him delay his response to Puzzle.

The second "far away" adjacent pair is U8 and U18, in which Millie praised Frizzy's idea. We can see why U18 is so far away from U8, simply because Millie was afraid that her primitive MOO skills might slow down the meeting. Hence, she first asked for help (U3), then for others' patience (U6 and U13). She returned to the main-stream topic to praise Frizzy's idea (U18) only after receiving Yoshi's assurance to help her (U16) whenever she encountered technical problems. This supports the fact that participants within the context of a different medium, though still maintaining social rules

for interaction, perform a different pattern of linguistic structures. In sum, conversing over CMC is different from talking face-to-face.

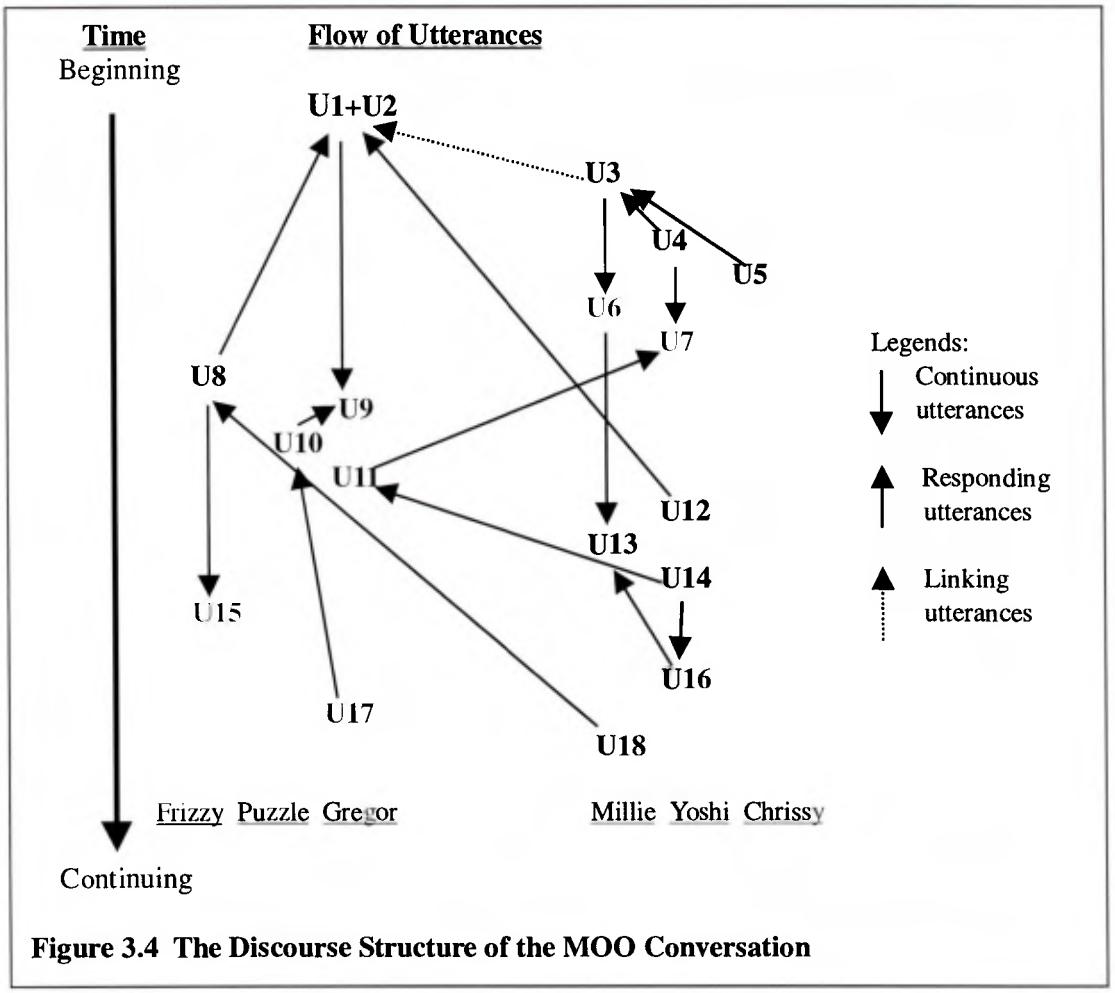

Figure 3.4 The Discourse Structure of the MOO Conversation

Another interesting thing about synchronous CMC setting is the use of "poses" and "emoting" to compensate the lack of non-verbal cues. The command for emoting is ":" or "emote." For instance, if a user named Peter types ": happily jumps up and down," other participants will see the line "Peter happily jumps up and down." The system automatically includes the modification on the subject and presents the sentence in a third person perspective to the other participants. Poses, sometimes referred to as "emoticons," are the ASCII graphics mentioned earlier, such as " :), ;), :-), ;-) " for smile faces, or " :(, :-(" for sad or angry expressions.

Rheingold comments on the new dimension that emoting gives to MUD and IRC conversations: "Instead of replying to a statement, you can smirk. Instead of leaving the room, you can disappear in a cloud of iridescent, bubble-gum-flavoured bubbles" (Rheingold, 1994, p. 148). This "new dimension" is a way of adding socioemotional content to the CMC setting, and is hardly available in other media (Chenault, 1998).

Indeed, Walther consolidates the various CMC theories into a theory of impersonal,

interpersonal and hyperpersonal interaction (1996). The extensive use of poses and emoting in more recreational situations such as MUDs and IRC has enabled CMC to surpass the level of affection and emotion of parallel face-to-face interaction, and to reach a level of hyperpersonal communication.

These studies demonstrate that people who use CMC for enjoyment manipulate the text to enhance their communication, develop relational skills and express individuality and originality. Reid (1994) emphasises that MUDing interweaves the play with social rules and expectations, that "descriptions, communicative commands and specialised language and textual forms enable the social understandings which link people together and allow the evolution and transmission of social norms."

3.3 Language Variations Found in CMC

In the past few decades, several forms of electronic communication developed, such as email, the bulletin board system, usenet, discussion lists, talk, chat, World Wide Web and video conferencing. The development of the new technology has brought with it new styles of language and aroused the interest of many researchers in this field.

In terms of the varieties of language found in CMC, most researchers found the following characteristics: (1) the simplified register, (2) special linguistic and paralinguistic features, and (3) a hybrid of written and spoken varieties.

3.3.1 The simplified register

One special type of characteristic identified in most CMC texts is simplified registers (Murray, 1995, p.15). The simplified registers may result from particular aspects of the context in which the language is used. One type of simplified register results from the speaker's perception that the addressee is incompetent in the language. Caretaker talk (e.g., Ferguson, 1977), foreigner talk (e.g., Ellis, 1985) and teacher talk (e.g. Henzl, 1973; Gaies, 1977) have been identified as such simplified registers. The goal of this simplification is to facilitate comprehension. Another possible context that results in simplified registers is limitations in time or space. This is often found in telephones, newspaper headlines (e.g. Straumann, 1935), advertising (e.g. Leech, 1966), or note-taking (Janda, 1985), where a simplified register is chosen to achieve more efficient communication.

Some typical simplification devices in CMC include *u* (you), *r* (are), *re* (regarding), *btw* (by the way), *asap* (as soon as possible), and *imho* (in my humble opinion). They are used for the purpose of saving time and space, and have become a unique characteristic of CMC texts.

3.3.2 Special linguistic and paralinguistic features

The use of special linguistic and paralinguistic features has often been identified as a characteristic of CMC. Early in 1980, Carey observed the use of "electronic paralanguage" in CMC, which included the use of intentional misspelling, lexical surrogates for vocal segregates, spatial arrays, grammatical markers, absence of corrections and capitalisation (Carey, 1980).

Reid stated there were four types of such features (Reid, 1991). First, there is the common practice of simply verbalising physical cues. For example, it is difficult to reply to a humorous comment with genuine laughter, so the literal typing of "hehehe" is considered acceptable. Secondly, emphasis and physical actions taken by CMC users are typically described within two asterisks, such as *Jefferson wanted a university *without* grades*. The third form includes the terms addressed in all capital letters, so the stress of the word is to be perceived as near-yelling, suc as, *I am a newbie...HELP!* The fourth feature is emoticons (emotive icons), a "shorthand" for the description of physical condition. Emoticons are called "smileys" by Reid (1991), but "relational icons" by Asteroff (1987), as she claims that they serve to establish the interpersonal relationship. Some examples are described by Reid (1991):

:-) or :) a smiling face, as viewed side-on
;-) or ;) a winking, smiling face
:-(or : (an "unsmiley": an unhappy face
:-} or : } someone having a so-so day
:-(*) someone about to throw up
8-) someone wearing glasses
:-P someone sticking out their tongue
>:-O someone screaming in fright, their hair standing on end
:-& someone whose lips are sealed
@}-`-,-`-- a rose

Looking at the icons 90 degrees counter-clockwise, we can see how wittily each simple combination of a few keys expresses certain emotions, and we cannot but help admire the talent and creativity of the people who designed these icons. According to the estimation of Metz (1994), over 500 of these emoticons are in use.

The extensive use of the linguistic and paralinguistic cues in CMC texts marks the efforts of CMC participants to compensate for the reduced social contextual cues, which was previously recognised as a defect in CMC settings (Sproull & Kiesler, 1991).

3.3.3 A hybrid of written and spoken varieties

In addition to the styles of simplified register and extensive use of paralinguistic cues, CMC texts are also found to exhibit a blend of written and spoken varieties.

Spitzer (1986) cites remarks from his colleagues in email form in which he found this new variety of language to be in use, which he described as "talking in writing," "writing letters which are mailed over the telephone," and "a panel discussion in slow motion." He himself spotted that participants "must use language as if they were having a conversation, yet their message must be written" (Spitzer, 1986, p. 19).

In a study on the interactive discourse that takes place in simultaneous terminal-to-terminal typed dialogues, Ferrara, Brunner, and Whittemore (1991) discovered a reduced register and a hybrid language variety that displayed characteristics of both oral and written language. The co-occurrence of these characteristics very much blurred the difference between speech and writing styles in CMC texts.

On the same topic of email messages, Maynor (1994) observes that the features distinctive to email are a lack of capital letters, simplified spellings, clippings, icons, punctuation, etc. He also noted that, except for the clippings, these features seem to share many similar features to those of speech. The use of these particular lexical items, Maynor believes, is associated with an effort, from the point of view of email users, to imitate spoken language.

Uhlioova (1994) examines the textual properties of a corpus of email messages and concludes that email contributes significantly to the development of language use, which offers "new writing strategies in the frame of new constraints and requirements of the medium." However, with written substance, email messages are no less interactive than speech in some respects, and this obscures the boundary of writing and speaking. However, the existence of many features of spoken registers, in my point of view, does not necessarily mean that CMC texts take the typical spoken registers.

In *Spoken and Written Language,* Halliday (1989) proposed the concept of lexical density as a criterion to differentiate betweem spoken and written forms, arguing that spoken forms generally have a lower lexical density. In comparing the discourse and participation of ESL students in face-to-face and electronic discussion, Warschauer (1996a) studies the lexical complexity by using the type-token ratio to measure word

repetition. He found that discussion in the electronic medium is significantly more complex than in the face-to-face group. His findings for syntactic complexity are similar, when he uses the coordination index to measure the ratio of subordinate clauses to coordinate clauses. Though Warschauer's formula to calculate the lexical density is different from that of Halliday's, who used entries listed in Roget's Thesaurus (Roget, 1982) as the criteria to count lexical words, the findings still suggest that electronic discussion is of a language variety not typically found in ordinary face-to-face conversation.

As argued by Collot and Belmore (1996), "electronic language" cannot be strictly labelled as spoken messages, since the participants neither see nor hear each other; nor can they be considered strictly written since many of them are composed directly online.

3.3.4 Situational features of CMC

The CMC texts do exhibit special characteristics not typically found in any traditional medium. Many of the studies mentioned above focus on different forms of CMC. The relational and situational constraints on the communication participants can be very different when the forms of communication vary. Though it is hard to generalise about common situational features in all forms of CMC, we can generally agree that the limits of time, compensation for the reduced contextual cues, and the intimacy that most participants intend to maintain with others largely account for the unique language characteristics found in many forms of CMC texts.

3.4 CMC and Language Learning and Teaching

The use of media to enhance the quality of teaching has been attempted by many teachers. Blackboard, flashcards, charts and tables, models, realia, tape recorders, video, films and overhead projectors are some of the most common instructional aids used in educational settings. For language teaching, the language laboratory – a specially developed medium – is used, which is meant to provide intensive audio-lingual training to the students. Historically, it can be seen that, each time a new technology comes into being, there are enthusiastic teachers who try to incorporate it into their teaching.

The computer is certainly no exception. In the 1970s, a period in which computer technology was very primitive by today's standards, there were herculean efforts, such as PLATO, to "computerise" whole sets of primary and secondary school curriculums. In

the 1980s, when micro-computers became popular, more stand-alone CAI packages (computer-assisted instruction) were developed for students' self-study. Since the 1990s, CMC has attracted the attention of many people, so much so that it has became a widely tested and discussed medium in language learning.

In this section, I will first review some studies on the relations between CMC and learning, and continue to discuss CMC in language teaching in particular.

3.4.1 CMC and learning

As CMC boasts almost unlimited accessibility over time and space, it is often used in programmes of distance education (Buckland & Dye, 1991; Lever, 1992). While this is a reasonable use of the medium, it does not mean that CMC is only used for transmitting knowledge to students in a remote context. Small-scale CMC settings are also set up to offer students various types of learning contexts.

In most of these projects, the results are reported to be satisfactory (Bonham, Cifuentes, & Murphy, 1995; Hesser, Hogan, & Mizell, 1992; Hiltz, 1995; Hudspeth & Maxey, 1995; Scardamalia & Bereiter, 1993; Williams & Merideth, 1995), without regard to the negative portrayal of CMC. The findings are usually that learning did take place, that faculty and students worked harder, and that satisfaction with the course was good (Patterson, 1996).

3.4.1.1 CMC attributes in learning

What we are more interested in are the possible attributes that lie behind the reported positive results in the CMC-based learning programmes. Harasim (1990) proposes five attributes of CMC instruction, which not only set it apart from other modes of education, but also support learning in a unique domain for learner-centered instruction. These five attributes are: (1) many-to-many communication, (2) place-independent communication, (3) time-independent communication, (4) text-based communication and (5) computer-mediated interaction.

Many-to-many communication

A primary component of CMC instruction is the computer conference, which supports and encourages interactive group communication. In a sound virtual conference, communication is no longer solely teacher-initiated, and students are no longer just receivers of information, as sometimes in many traditional classes. All of the students can also easily initiate new topics or questions and the interaction is multi-directional.

This sense of participation is beyond the capacity of any other distance learning system and may substantially enrich the content of the course, which is otherwise impossible. Harasim (1990) contends that, "with careful attention to [curriculum] design—computer conferencing supports and facilitates active learning collaborations" (p. 43).

Place-independent communication

CMC students are not limited to the learning opportunities, resources and experts available in some fixed geographic locations. As long as they can get access to the CMC system, they can learn and teach each other, at school, at home, or in any other location.

Time-independent communication

The asynchronous mode of CMC gives the students as much time as they need to read, understand and respond to information. This is considered an important valuable of learning through CMC (Hiltz, 1986; Chesebro & Bonsall, 1989). On the other hand, we should not neglect the value of synchronous conferences, which allows the participants a chance to interact more intimately.

Text-based communication

The lack of non-verbal and paralinguistic information in CMC has often been charged as a defect, that is, the so-called lack of social contextual cues, in communication. However, some learning theorists propose that the process of writing may be more helpful to learning than face-to-face talk or lectures because it is heuristic. In the process of writing, thoughts are manipulated by the simultaneous actions of synthesising, imaging and graphically recording a representation of thought in words (Emig, 1977). The mainly text-based CMC conference system is really at an advantage in serving as a learning environment.

Computer-mediated interaction

Harasim (1990) argues that CMC interactions are "revisable, archivable, and retrievable," and they give the user an exceptional ability to "present, receive, process and manage information" (p. 51). Unlike in other educational settings, the process of CMC automatically creates an electronic transcript of all the interactions. This allows participants unprecedented control over the way they participate in collaborative and discursive activities. They can sort, scan and/or ignore the contributions of the others in their group. The transcript also allows participants repetitive access and information

retrieval. This attribute facilitates both retrospective analysis and enables a critical review of the interaction.

3.4.1.2 Value of the CMC attributes in learning

The five CMC attributes described above can be of substantial assistance to the quality of education, as they are compatible with some learning theories.

Collaborative learning

Harasim (1989) describes the process of collaborative learning as the construction of knowledge by the engagement of students, instructors and experts in interactive discussions and activities. Bruffee (1986) has defined collaborative learning theory based on a model of the learner-centred environment: (1) students actively participate in the learning process; (2) students are invested with responsibility for knowledge acquisition; (3) the instructor role shifts from knowledge transmitter to knowledge facilitator, enabler, coach, model and guide; and (4) the environment facilitates peer interaction, evaluation and cooperation.

Collaborative learning has been an ideal of many teachers, but implementation is somewhat difficult in ordinary class situations sometimes. However, with the help of the unique attributes of CMC, the realisation of the ideal is much more feasible.

Davie (1988) reports a study of two CMC courses which required collaborative writing assignments. It is found that all students contributed to the final project in some way, and the quality of the writing was judged superior to the products of similar small groups in face-to-face settings.

Free from teacher dominance

In education research, message act analysis is used to explore the patterns of message function (Mehan, 1978). It is found that the sequence of *teacher initiation, student response and teacher evaluation* (IRE) is the typical message pattern in ordinary classrooms. Using a message act analysis, Levin, Kim, and Riel (1990) found the initiation, evaluation and reply type messages were more evenly distributed among teachers and students in CMC. This clearly suggests that students participate more actively in a CMC class. They also found two other patterns of message function: a star pattern, where the messages are a series of replies to a single initiation, and the thread pattern, where the messages are more linear in a chain of replies following a similar thread. The CMC messages were found to be more complex than in the ordinary class

interactions, as they were less dominated by the instructor (Levin, Kim, & Riel, 1990).

The study, carried out from the perspective of message act analysis, suggests a change of roles between the teacher and students. We may not say that the CMC setting is necessarily an ideal student-centred learning environment. However, it is safe to say that such an environment can be more easily fostered with the help of the certain attributes of CMC.

Writing as a creative thinking process

It should, by now, be agreed that writing does not involve simply taking down verbal messages as they are. In the process of writing, thoughts are mediated by the simultaneous actions of synthesising, imaging and graphically recording a representation of thought in words. Vygotsky (1962) refers to this as the "deliberate structuring of the web of meaning" (p. 100).

Kenneth Bruffee (1984) views writing as the resocialisation of internalised conversation and the collaborative conversation. The writing process is a way of "demonstrating to students that they know something only when they can explain it in writing to the satisfaction of the community of their knowledgeable peers" (p. 652).

Emig (1977) states that, "because writing is often our representation of the world made visible, embodying both process and product, writing is more readily a form and source of learning than talking" (p. 124). The dependence on writing can indeed make CMC a better environment for learning.

In this section, we reviewed the possible values of CMC in learning from the perspectives of collaborative learning, media act analysis and text-based learning. It can be seen that, with sound planning and management, a CMC setting can be an efficient learning environment. It has the potential to move distance education from a *mass production* model to one that meets the ideal of a student-centred collaborated learning context.

3.4.2 Development of computer-assisted language learning

The history of computer-assisted language learning (CALL) is along the same lines as the development of computer-assisted instruction. Warschauer (1996e) divides the development of CALL into three stages: behaviouristic, communicative and integrative. CALL started in 1960s, when some software was developed to teach languages in a behaviouristic approach. The assumption was to have the computer serve as an all-knowing, patient tutor, delivering instructional materials and guiding the learner through

individualised practice. Based on the behaviouristic approach, repetitive drills and practice were emphasised.

CALL software designed in this approach resembles in the model of *programmed instruction*. It assumes that for the content of teaching to be better conceived and digested by the learner, it should be presented in small pieces. Immediate feedback based on the individual response is also essential in guiding the learner through the learning process. The approach, though it may work for the teaching of some domains of knowledge, may not be appropriate for the teaching of language. As language is such a sophisticated, not-always-logical system, nobody has a thoroughly systematic way of presenting a language in details, and no teacher can claim to have a complete teaching plan of a language lesson that prepares for all of the possible responses of the learner. The idea of a sound programmed instruction type of CALL package is hence criticised as unrealistic (Chou, 1985b).

Furthermore, Underwood (1984) sarcastically criticises the behaviouristic CALL software as taking a *wrong-try-again model*. It lacks flexibility, creativity, or any semblance of communication. He advocates the development of communicative CALL software, such as games, simulations, text manipulation programmes and text generation programmes, and also encourages the use of email and multimedia in English learning settings.

Other than open-ended, personalised activities on the computer, the features of communicative CALL also include interpersonal communication, which is stimulated by the software, and the use of the computer as a tool, such as for word processing or desktop publishing. For these purposes, it is interesting to see that there is no definite reliance on software specifically designed for language learning purposes (Yuen & Foley, 1998).

The development of the integrative CALL was along the same lines as the rapid technological development of hypermedia and the Internet in the 1990s. The hypermedia environment may provide comprehensible input to the students in sound, text, graphics and video segments in a format to which they can have easy access at their will. The Internet is the benchmark to an unlimited amount of resources, and also a highly efficient environment for authentic interperson communication. With the help of the Internet, language learners are no longer are confined to simulated communication with the computer, or with the next-seat neighbours in the same class. They can experience the true pleasure of discussion, questioning, explaining, consoling and debating with counterparts from every corner of the world, which was previously never possible.

In the next section, I will review some of the research projects conducted on the use of the Internet to facilitate the learning of English as a second or foreign language.

3.4.3 English learning activities using CMC

With the rapid development of the Internet, there are increasing opportunities to study English using a wide variety of Internet resources. Among many other resources, there are discussion lists where students can discuss various topics with other students; places where students can find keypals, i.e. pen pals who communicate using a keyboard; online writing labs initially established in Purdue University, and later popular in many universities, where students can practice writing and get advice; plus an endless number of sources of reading materials, word games and grammar lessons. For English language students in countries where English is a foreign language, these resources are particularly important.

Much research has also been conducted on using the Internet to facilitate students learning English in the past decade, and their results have shown that using technology can provide them with a sense of empowerment and develop their communicative abilities. In West Europe, Ruth Vilmi (1994) of Finland is one of the pioneers to conduct many CMC projects for computer-assisted language learning. In England, the computer network, *National Grid For Learning*, has been in operation since 1997 specifically for education, i.e. for all primary and secondary schools and universities nationwide within the UK (Cole, 1998). In North America, many professionals from universities in New York, Virginia and Colorado, as well as many regional schools/universities in Canada, all go on teaching with the assistance of the Internet (Tillyer, 1997, 1998; Corio, 1996; Younger, 1997, Turbee, 1995, 1996, 1997, 1998; Holmes, 1996). In East Asia, Randall Davis and Kenji Kitao in Japan, Susan Moody in Hong Kong, C. T. Chou and Hsien-Chin Liou in Taiwan all use CMC to facilitate the learning of English (Chou, 1985a, 1996, 1998; Davis, 1999, 2000; Kitao, 1998; Liou, 1997; Moody, 2000).

Researchers and teachers have begun to discuss the advantages of CMC as a facilitator to language learning. Many researchers demonstrate through empirical studies how their students become better motivated in learning English after gaining access to the Internet. The Internet is believed to provide an efficient and economical means of communication, an unlimited range of authentic texts for reading, a psychologically safe environment to communicate in anonymity and a real target audience for communication (Chou, 1996).

In an attempt to demonstrate that computer-networked class discussion is an

effective method for increasing interactive competence, Chun (1994) shows that networking not only helps students' writing skills, but also increases their spoken proficiency. She finds that using CMC provides students with opportunities to generate and initiate different kinds of discourse structures and speech acts, and that the writing competence gained from CMC can gradually be transferred to the students' speaking competence as well.

In comparing the discourse and participation of ESL students between face-to-face and electronic discussion, Warschauer (1996a), using a formula that measures the number of words of discourse per student, finds that student participation in the electronic group is double that in face-to-face discussion. He infers that the reason behind this could be that his Japanese students are largely quiet in nature and silent in face-to-face interaction, but they participate much more regularly in the electronic mode.

Other than learning English through asynchronous CMC, students are believed to benefit from synchronous CMC, too. A popular Internet activity, which has a great deal of potential for English language students, is chatting. In doing so, the users get on line at the same time and communicate in real time with each other from different corners of the world. A more sophisticated type of online chat is MOO (Multi-user Object Oriented). MOO is a virtual environment in which the users not only chat with each other, but also envision a common physical setting in which the communication takes place. Despite being a text-based environment, there are virtual objects in the setting that can be moved and manipulated. More experienced users can even cooperate to construct and modify the setting. An excerpt of MOO samples discussed earlier and shown in Figure 3.3 and 3.4 can be used as further reference.

SchMOOze University and the *Lingua MOO* are two typical MOO environments on the Internet for language learners. In 1994, Falsetti, along with Schweitzer, established the *schMOOze University* as the most well known MOO site for ESL/EFL students and teachers (Falsetti, 1995). *Lingua MOO* was co-founded by Haynes and Homevik in 1995 for the purpose of writing classes for learners of Spanish (Haynes & Holmevik, 1998). Since the establishment of these two sites for CMC synchronous communication, the use of educational MOOing has grown rapidly.

By reviewing the research carried out prior to 1996 on using CMC within a class, Conrad (1996) finds that learners possess greater motivation and produce more language. Nevertheless, he also warns about the fossilisation effect due to the consistent interlanguage input. The solution he puts forth is the use of synchronous Internet communications, such as MOO and IRC, to bring non-native speakers into contact with

native speakers.

Pica, Lincoln-Porter, Paninos, and Linnel (1996) study the effect of language learning in MOO. They focus on the theme of negotiated interaction, which is composed of input, output and feedback of language learning. One of their findings supports the view that non-native speakers can supply only part of their own linguistic needs for movement along the interlanguage continuum. Beyond that, native speakers are needed for necessary interaction. This supports the use of synchronous Internet communications to bring non-native speakers into contact with native speakers.

Using MOO for language learning, Hall (1998) points out that, although students in the traditional classroom might learn *about* a particular grammatical form, they would be hard pressed to use it correctly. On the other hand, in the situation of immersion in the target language culture, the students would not only learn how to use the grammatical structure, they would also learn the cultural norms of using the target language because they are in contact with native speakers in MOO sites. Learners learn from native speakers how to start conversations, how to ask for information, how to persuade, how to end conversations, etc. in the target language.

However, Hall warns that a MOO site, based on learning theory, could function as a constructivist environment only if the teacher allows it to. Hall notes the tendency of a growing number of traditionally minded teachers to use technology to recreate traditional forms of pedagogy. Real life classrooms, embedded in the culture of traditional education, which frequently does not use language as a tool to further cognitive apprenticeship, could be turned into virtual traditional classrooms. Students would be transported to virtual chairs and desks, reading information being presented on virtual blackboards. Such situations would lead to widespread virtual failure, just as they do in real life classrooms. Instead, Hall calls for student-centred methodologies, where teachers would be facilitators and guides. It would be the students who would ultimately do the learning, and they would be most successful if they were placed in situations of authentic language with a maximum amount of control over their own learning.

The development of chat and MOO has resulted in innovative environments for foreign language learning. In terms of synchronous CMC, Frizler (1995) applies a case study of conversations carried out in a university-level ESOL (English for speakers of other languages) writing class via CMC and the Internet. Her findings suggest that such an application has an innovative impact on learners' accomplishment. Based on her findings, Frizler draws conclusions and makes recommendations for maximising the educational benefits and minimising the limitations of using the Internet in the ESOL

writing classroom to develop the confidence and writing ability of students. She finds that, with the Internet as a teaching tool to supplement the traditional classroom, "elements of the virtual classroom can indeed be beneficial to both students and teachers. In fact, as the world around us evolves technologically, learning online may better prepare our students for the thinking and tasks they will be expected to accomplish not only in the work place, but across all aspects of their lives" (Frizler, 1995).

However, Frizler also suggests that using the Internet as a virtual classroom in which to learn writing should be an option, but not a complete replacement for the traditional classroom, except for distance learning students, who have little choice but to use this format.

Meunier (1997), performing a meta-analysis of the various affective elements of CMC learning in empirical studies, reports that CMC procedures in second/foreign language classes results in high level situations, task motivation and a generally positive attitude on the part of the students. Learners are more intrinsically and socially motivated online than in face-to-face conversations.

The above discussion shows that as CMC develops to become a common means of communication, it brings about new changes and problems to language learning. The high efficiency and low cost of CMC has made possible immediate communication with real living counterparts using authentic language. This used to be an ideal beyond the capability of most foreign language learners. However, the students can now rather easily be engaged in both synchronous and asynchronous forms of CMC to practice the target language. It seems certain that more and more foreign language learning activities will be attempted and experimented with by teachers and students on the Internet, and this will cause a tremendous change not only to the methods of learning, but also to the relationship among all those involved in the CMC environment – the teacher, the students and other users.

3.5 Rationale of the Present Study

The discussion on the context of CMC, the language variations developed in CMC and the language learning and teaching in CMC contexts has formed a background of reasons for further studies into the newly emerging language varieties in CMC.

3.5.1 New communication contexts and emerging language varieties

Though there is no unanimous agreement on the issue, it is now generally accepted that CMC is not just a tool, but actually forms new contexts of communication. With the many factors of participants, settings and purposes involved, the CMC contexts can be extremely versatile, and the operation in the contexts can be very sophisticated.

The CMC contexts may be a rich ground of social and emotional interaction. In both synchronous and asynchronous CMC settings, new behavioural patterns are developed so that the people can achieve their goals, and, more importantly, enjoy the pleasure of communication even though there may be certain limitations (e.g. time, non-verbal cues, etc.), which make some aspects of interaction less convenient than in traditional settings.

In the many new virtual communities, language registers exhibit changes. This is a logical consequence of the change of the contexts. In the review, we find the generally observed CMC characteristics to include the simplified register, special linguistic and paralinguistic features, and a hybrid of written and spoken varieties.

Understanding the complexity of various CMC contexts, the above observation still cannot give us a more detailed picture of the specific language patterns used in certain CMC settings. More empirical research in this field may be needed so as to help depict the CMC contexts more fully.

3.5.2 Facilitation of English language learning

Generations of teachers have made the effort to use instructional technologies to assist learning. Since the coming of the computer era, much has been done to incorporate computer technology into all kinds of teaching programmes. Moreover, since the appearance of early CAI software, which claimed to offer an individualised interactive learning experience to the students, there are several lines of thought on the role of the computer in different teaching programmes.

In terms of computer-assisted language learning (CALL), the trend has developed from behaviouristic and communicative to the presently mainstream integrative approach. CMC is a major context of integrative CALL. Nowadays, learners can enjoy boundless resources and unlimited opportunities to communicate with other people. This used to be little more than an impossible dream of many language teachers due to the limitation of time and space. It now allows a comprehensible input which fits the interests and needs of the students, and endless chances for communication with authentic purposes.

While we can expect to find more language learning activities conducted in CMC settings, a natural concern is the kind of language to which the learners are exposed. We

understand that new varieties of the English language are forming in CMC, and we should not neglect the importance of these new varieties. A very important objective of language teaching is to foster the linguistic awareness of the varieties of language (Donmall, 1985; Hawkins, 1987; Wolfram, 1998). All language learners will face different varieties of the language, and it is agreed that instead of being given a prescriptive set of rules to follow, they should be nurtured with the ability to observe and analyse the varieties of language as they change in different contexts.

In this study, I plan to examine linguistic features in certain CMC texts. My intention is not just to present the findings and suggest that these *are* the CMC features to remember and to use. The method in which I gathered and analysed the data to reach the findings is, namely, that of corpus analysis. It is an efficient instrument. Together with Biber's multi-dimension examination, corpus analysis can best serve in detecting various forms of language varieties in which human beings communicate in different contexts. I hope to demonstrate this approach to language teachers, learners and perhaps syllabus designers alike, so that more of them can make use of similar methods to better understand the new varieties of the English language in the era of CMC.

3.6 Summary

In this chapter, I reviewed the literature on the context of CMC, the language variations developed in CMC, and learning and teaching in CMC contexts.

Some people believe that, with the help of the Internet, virtual communities have emerged in which participants interact with each other in ways no less versatile than in traditional social communities.

However, critics like Kiesler (1984), Stoll (1996), and others argue that cyberspace is not the real space in which people live their lives, and that CMC is an inadequate way for people to share emotional content, not to mention developing meaningful, long-lasting relationships. The suspicion of the lack of socioemotional content in cyberspace may have developed on several bases: the assumption of a lack of social contextual cues in CMC, the mostly task-oriented impersonal nature of the CMC texts in the study samples, and the difficulty, for some users, of manoeuvring in cyberspace. After examining a larger variety of CMC types and settings, it is argued that, as long as human beings are social creatures, they will not neglect any opportunity to exchange their emotions with each other through any possible channel, including CMC. The concerns mentioned above

will not, therefore, hinder the development of virtual communities.

The lack of social contextual cues was proposed by Sproull & Kiesler (1986, 1991) to argue that, as social cues are weak or absent in CMC, communication becomes more self-centred and unregulated. The negative effect can be weakened social norms, and the use of offensive and hostile language, i.e. flaming. On the more positive side, we can expect a more democratic and liberal context of communication. As a matter of fact, more researchers (e.g. Reid, 1991) find that CMC users often develop ways to compensate for the lack of these contextual cues. In doing so, the interpersonal relationship during CMC is not impersonal, and may even be hyperpersonal, which is not possible through any other media (Walther, 1996). For the argument of more equality and democracy in CMC contexts, we also find that the statement is not necessarily true, as there are more elements, especially the participants and the purposes of communication, that determine the nature of a communication context.

In terms of the varieties of language found in CMC, most researchers found the following characteristics: (1) the simplified register, (2) special linguistic and paralinguistic features, and (3) a hybrid of written and spoken varieties. These are developed out of some of the situational features of CMC, including the time constraints, compensation for the reduced contextual cues and the intimacy that most participants intend to maintain with others.

In terms of the relationship between language learning and CMC, we first examined these two variables in general. It was found that collaborative learning, freedom from teacher dominance and writing as a thinking process are the ideals in which CMC contexts can help assist in the learning process. As for the relationship between CMC and language learning, we first reviewed the three stages of the development of computer-assisted language learning (CALL), i.e. behaviouristic, communicative and integrative (Warschauer, 1996e), and found that CMC provides the necessary prerequisites for integrative CALL. We also cited many cases of language learning programmes in synchronous and asynchronous CMC settings.

From the review, we conclude that CMC, as a rich context of communication, may be the grounds on which we can find the nurturing of new forms of social relations, new chances for learning and new varieties of language use. In terms of language learning, we should not fail to seek a better awareness of the important new varieties. As English teachers and students, we cannot afford to fail, because CMC is becoming a major means of communication in the new century.

Chapter 4: Methods and Procedures of Investigation

4.1 Introduction to the Overall Research Framework

As mentioned in the previous chapters, one of the main aims of this study is to seek a better awareness of the new varieties of the English language that have developed in the context of a particular new technology, namely that of Computer-mediated Communication (CMC).

Through the literature review of previous chapters, we find that the use of language varieties depends on many sociolinguistic factors, and that the choice also reflects the social identity of the language users. As language learners learning a language, students should be helped to develop an awareness of the different varieties so that they can both better understand others and also express themselves better in different contexts. When this mature, open-minded awareness prevails, linguistic parochialism may gradually fade, making international intelligibility easier to achieve.

CMC as a means of communication has grown rapidly in the past few years and is rapidly turning into a major medium in our daily life. The new medium not only signifies a new way of message conveyance, but also involves new types of human relationship and new forms of language. As a new context for communication, CMC is believed to form new virtual communities, involving different participants, settings, and purposes. Such communities are a rich breeding ground for the development of new behaviour patterns and new language varieties. So far, most studies have found that language varieties in CMC generally contain the characteristics of a simplified register, special linguistic and paralinguistic features, and a hybrid of written and spoken varieties. However, more empirical research is needed to help us better understand the specific language features of CMC.

This understanding is valuable not just for linguistic research, but also for language teaching. The proliferation of CMC technology has aroused the interest of language learners who are keen to use it to facilitate their learning, especially of English. This makes us even more concerned about the language varieties that often appear in CMC contexts. There is an argument that the special characteristics of the language often found in CMC have deviated from the generally accepted registers and are therefore unsuitable as models for language learning (Lundstrom, 1995). However, if we agree with the

importance of language awareness, we have no choice but to develop a better understanding of the varieties, simply because they exist and are important, despite any value judgement.

Most English learners will have more and more opportunity to use CMC, and as such, it is unthinkable that they should be ignorant of the language varieties that prevail in this context. A preliminary study, such as this, into the language features of CMC genres can reveal some important characteristics of the varieties not previously discussed. Besides, as it is believed that language learners' ability to observe and analyse the varieties of language should be nurtured, it is hoped that the methods of analysis adopted in this study can also be of some reference to language teachers and students.

In terms of the framework, this study started with an aim of enhancing the quality of language education, proceeded through corpus analysis and quantitative statistical procedures, and will end with a discussion on how the use of CMC texts can effect and perhaps facilitate English teaching. This chapter is mainly concerned with the methodology applied for this research, and it begins with a background of the overall research framework and the considerations that were made in choosing a research approach. I then describe the considerations undertaken in developing the methods with which the study can be expected to serve pedagogical purposes in English language teaching. In the later sections of the chapter, I also detail the procedures expedited for the analysis, i.e. how the text data were prepared; how the notations for the linguistic features were designed; how the computing processes were carried out; and how the statistical comparisons were implemented. Finally, before going into the detailed findings and discussion in the next chapter, a general introduction on the findings resulting from the statistical comparison on the text data is discussed briefly.

4.2 Factors to Be Considered for the Linguistic Analysis

To achieve the purposes of the study, many factors had to be taken into account in the development of the methodology of this study. In terms of linguistic analysis, considerations were necessary at least in three aspects. They included the choice between a qualitative or quantitative research approach, the objectives of the language investigation, and a choice of texts that could best serve as sample data for this study.

First of all, an approach to the study must be decided. In text analysis, researchers often take one of the two approaches: qualitative and quantitative. In the qualitative

approach, researchers would examine a certain amount of text data in depth to detect some unclear but significant clues. They should also consider the context of communication, such as the relationship of participants and the context of communication, so as to make a correct judgement on the texts. In the quantitative approach, researchers can resort to the help of computer technology to examine a large amount of text rather objectively to induce general patterns of language features. The latter is the approach of corpus linguistics, which is receiving growing attention with the development of computer technology.

CMC texts, a major object of analysis in this study, are of a rather new medium. For a person subscribed to computer discussion lists, these texts often come in large quantity. But the observer can usually have access only to the texts, not the context in which the texts are produced. In other words, CMC texts are different from non-CMC texts in that they are more abstract and of larger quantity. Therefore, they may not be as suitable for qualitative analysis as non-CMC texts. Discerning this, I decided to take a quantitative approach in the study and used the corpus linguistic method for data analysis.

Secondly, there was a need to consider the features to be examined in this study. Previous studies have been done on various aspects of the language, though these all hail from different points of view. These included genres (Bhatia, 1993; Brainerd, 1972; Halliday, 1990; Littlefair, 1991; Marckworth & Baker, 1974), registers (Besnier, 1986; Janda, 1985), styles (Blankenship, 1962; Carroll, 1960; Kochman, 1981; Vachek, 1979), and similarities and differences between written and spoken language (Akinnaso, 1982; Cayer & Sacks, 1979; Tottie, Altenberg, & Hermeren, 1983). Most of these studies also chose to focus on a number of features.

As this research was intended as a preliminary comprehensive study on the nature of CMC texts as compared with that of non-CMC texts, I did not want to limit the study to only a few features. Rather, I chose to survey a large number of features in the texts so as to get an overall picture. Instead of compiling a list of features on my own, I found the 67 features used by Biber in several of his studies (1986, 1988a, 1989, 1995; Biber, Conrad, & Reppen, 1994, 1998) a good set to adopt. I made no pre-selection of them, mainly because most previous studies of CMC texts (Chapelle, 1990; Kim & Sag, 1995; Tella, 1999; Underwood, 1984; Warschauer, 1998) had concentrated on rather limited sets of language features and could not provide a comprehensive perspective for me to follow. Though some studies (Collot & Belmore, 1996; Halliday, 1989; Lundstrom, 1995; Maynor, 1994; Tribble, 1998) concluded that the CMC text was a kind of text bridged

between the written and spoken styles, I felt that this could be just a general vision, instead of a complete portrait. Moreover, there was a worry that I might not be able to claim the reliability of my findings at the end if such a pre-selection was made before my full understanding of its whole nature. I consider it a safer strategy to work on a larger number of features, lest I should miss some features that might be found, by others, to play a prominent role in the CMC text. Therefore, I finally decided to stick to this set of 67 features with a hope that I could see the whole picture better and more clearly.

The 67 features, partly discussed in Chapter 2's The Study of Language Variation (pp. 50-51), are very clear, explicit, detailed and unambiguous. Thus, they form a whole set of language features ideal for the comprehensive comparison intended in this study.

Thirdly, it was necessary to decide the source and size of the text samples for this study. There was already a huge amount of CMC texts in existence, from which I must choose a sample. The primary concerns in sample choosing included what texts could be accessed and which would be more appropriate for this analysis. With my own background as an English teacher, I hoped the CMC texts examined in the study to be of adequate educated registers. It seemed then that discussion lists with ELT teachers as participants were the most appropriate source of text samples. As I hope the findings would serve some pedagogical functions in English teaching, the choice of ELT-related discussion lists as samples was also justified.

With regard to the size of the text samples, there is no criterion claiming that a certain number of words is necessary (Cain, 2000; Klotz, 2000; Lopes, 2000; Luk, 2000; Rome, 2000). Even a small collection, such as three poems, can serve the purpose for a language analysis (Rome, 2000). However, Lopes (2000) suggested 5,000 words as a minimum text, as it was an optimal number for him when he went through several studies. Many (Cain, 2000; Klotz, 2000; Lopes, 2000; Luk, 2000; Rome, 2000) agreed that the general rule is that the sample texts should be as large as possible, so that the researcher can always get enough data for most, if not all, of the investigation. As I wanted to be confident with my results obtained on the 67 features, I arbitrarily decided that my texts would be no less than 50,000 words in each text group, i.e. to have an equal size of sample texts. I hoped with this quantity, my texts could mirror the facts of language used in the CMC medium on the one hand, and avoid any bias results generated in the later stage on the other.

Finally, since I planned to take a quantitative approach, the concordancing and statistical software for analysis and comparison needed to be reviewed. I needed to

consider several possible alternatives in choosing the concordancing instruments, and made the selection. In order to get a clear idea how the language features differ between several text groups, the statistical software had to be implemented. It should be able to compare groups of text data in relation to multiple features in each text. The selection of a specific statistical approach for the comparison, such as a t-test, Chi-square, Analysis of Variables (ANOVA), or others had to be taken into careful account. In summary, I had to select carefully a method that was suitable for comparing four groups of data in two variables, which would provide, at the same time, maximum reliability and minimum constraints.

Above are the linguistic concerns that this study attempts to examine. In education, using CMC has become a major practice in English language teaching. Through this study, it is hoped that some contribution can be made to English teachers with a better awareness and understanding in applying this technology in their teaching career.

The following sections deal with data preparation (Section 4.3), tools for data preparation and analysis (Section 4.4), calculation of the linguistic features (Section 4.5), and statistical analysis (Section 4.6). In addition, the actual procedure by which the study was conducted will also be described in detail.

4.3 Data Preparation

4.3.1 Language features examined

A set of 67 language features was chosen as the object investigating analysis among the four groups of CMC texts in the present research. These 67 features, as discussed in Chapter 2, were used by Biber (1988a) in his study to identify the variation across speech and writing. They are arranged in 16 categories and shown in Table 4.1.

Various researchers may choose different language features as their objects of analysis. In the literature review made in Chapters 2 and 3, I have reviewed many of the previous studies on language use, each of which may focus on a certain number of features. As the present study was intended as a detailed investigation of the language variations between CMC and non-CMC texts, as a researcher, I believe it is necessary to include as many systematic features as possible so as to avoid a channel vision in the investigation, based on a single feature or two. In collecting the features, however, it was found that researchers of previous studies all used their own terms to label the features being examined, and these terms often differed from or overlapped with those used by

others. The 67 features used by Biber had been found as the largest set of mutually exclusive language features presented in a single study. This was the main reason for adopting the 67 features as the objects of analysis.

Table 4.1
67 Language Features to Be Examined in This Study

(A) *TENSE AND ASPECT MARKERS*
past tense (F01)
perfect aspect verbs (F02)
present tense (F03)
(B) *PLACE and TIME ADVERBIALS*
place adverbials (F04)
time adverbials (F05)
(C) *PRONOUNS AND PRO-VERBS*
first person pronouns (F06)
second person pronouns (F07)
third person pronouns (F08)
pronoun IT (F09)
demonstrative pronouns (F10)
indefinite pronouns (F11)
DO as pro-verb (F12)
(D) *QUESTIONS*
WH questions (F13)
(E) NOMINAL FORMS
nominalisations (F14)
gerunds (F15)
nouns (F16)
(F) *PASSIVES*
agentless passives (F17)
BY passives (F18)
(G) *STATIVE FORMS*
BE as main verb (F19)
existential THERE (F20)
(H) *SUBORDINATION*
THAT as verb complements (F21)
THAT as adjective complements (F22)
WH clauses (F23)
infinitives (F24)
present participial clauses (F25)
past participial clauses (F26)
past participial WHIZ deletions (F27)
present participial WHIZ deletions (F28)
THAT relatives clauses on subject position (F29)
THAT relatives clauses on object position (F30)
WH relatives clauses on subject position (F31)
WH relatives clauses on object position (F32)
WH relatives clauses pied pipes (F33)
sentence relatives (F34)
causative adverbial subordinators: *because* (F35)
concessive adverbial subordinators: *although, though* (F36)
conditional adverbial subordinators: *if, unless* (F37)
other adverbial subordinators (F38)
(I) *PREPOSITIONS, ADJECTIVES AND ADVERBS*
prepositions (F39)
attributive adjectives (F40)
predicative adjectives (F41)
adverbs (F42)

Table 4.1
67 Language Features to Be Examined in This Study (Continued)

(J) *LEXICAL SPECIFICITY*
type/token ratio (F43)
word length (F44)
(K) *LEXICAL CLASSES*
conjuncts (F45)
downtoners (F46)
hedges (F47)
amplifiers (F48)
emphatics (F49)
discourse particles (F50)
demonstratives (F51)
(L) *MODALS*
possibility modals (F52)
necessity modals (F53)
predictive modals (F54)
(M) *SPECIALISED VERB CLASSES*
public verbs (F55)
private verbs (F56)
suasive verbs (F57)
SEEM/APPEAR (F58)
(N) *REDUCED FORMS AND STRUCTURES*
contractions (F59)
THAT deletion (F60)
stranded prepositions (F61)
split infinitives (F62)
split auxiliaries (F63)
(O) *COORDINATION*
phrasal coordination (F64)
non-phrasal coordination (F65)
(P) *NEGATION*
synthetic negation (F66)
analytic negation (F67)

These 67 features are arranged in 16 categories by Biber. The ordering and the categorisation of the features represents the outcome of a comprehensive comparison of written and spoken CMC texts on a large number of linguistic features (Biber, 1988a) as has been discussed in Chapter 2. I found adopting this set of features a convenient alternative to building a new set of features of my own.

For the purpose of providing concrete ideas to English teachers on how CMC texts are different from non-CMC texts, it was my intention to compare the occurrences of each of these 67 features in the synchronous and asynchronous CMC texts, with a possible comparison between those in non-CMC texts, so as to see how language uses vary between media and between temporalities.

4.3.2 Text samples collected

Four groups of sample texts were collected for the study: asynchronous CMC (abbreviated as CA), synchronous CMC (CS), asynchronous non-CMC (NA), and synchronous non-CMC (NS). The structure of these CMC and non-CMC text samples is depicted in Table 4.2. The sources of these text samples are described in the next few subsections.

Table 4.2
Data of Sample Texts Taken for This study

Sub-corpora		No. of Texts	No. of words
CMC Asyn. Texts	1. NETEACH-L	248	50,028
	2. TESL-L	255	50,166
Subtotal		503	100,194
CMC Syn. Texts	3. NETEACH-MOO	10	52.109
	4. Netoric-MOO	7	52.053
Subtotal		17	104,162
Total		520	
Non-CMC Asyn. Texts (Written)			
	1. Academic prose (LOB)	80	160,000
	2. Professional letters (Biber, 1988a	10	10,000
Subtotal		90	170,000
Non-CMC Syn. Texts (Spoken)			
	3. Face-to-face conversation	44	115,000
	4. Telephone conversation	27	32,000
	5. Public conversations, debates, and interviews (London-Lund)	22	48,000
Subtotal		93	295,000
Total		183	

4.3.2.1 Sources of CMC texts

CMC texts can generally be classified as being of asynchronous temporality and synchronous temporality. Asynchronous CMC (CA) is used in email, discussion forums, discussion lists, etc, where participants are not engaged in simultaneous communication with one another. Synchronous CMC (CS) is typical of the style of talk, chat rooms and, Multi-user dimension Object-Oriented sites (MOO), where all the participants must log-in to the computer network simultaneously to be engaged in synchronous communication.

For this study, I collected CA texts from the discussion lists of NETEACH-L (NETEACH-L, 1996) and TESL-L (TESL-L, 1996). The theme of these two lists is related to academic discussions of language teaching. They are concerned mainly with issues related to teaching English as a second or foreign language, and almost all the

writers are English teachers, who are either native or non-native speakers teaching in primary, secondary, tertiary or adult levels in different countries.

The selected archives were downloaded in electronic form from the servers of NETEACH-L (NETEACH-L, 1996) and TESL-L (TESL-L, 1996). All irrelevant lines in the email message texts, like mail headers or lines generated by the computer system, were removed in order to achieve a precise analysis. Samples with and without mail headers can be further referrenced in Appendices 2 and 3. The CA text group has a total of 100,194 words in 503 email texts. How asynchronous CMC texts were prepared for further analysis is detailed in Section 4.5.1, Text Data Preparation.

The synchronous CMC (CS) texts used in this study were downloaded from the log files archived in the servers of NETEACH-MOO (Younger, 1996) and Netoric-MOO (Netoric-MOO, 1996). The theme of the NETEACH-MOO is to use the Internet to enhance English teaching and learning, while that of the Netoric-MOO is to use the Internet to facilitate English composition in language teaching. The CS text group has a total of 104,162 words in 17 MOO session logs. Its preparation for further analysis mirrors that of the CA texts, which is detailed in Section 4.5.1.

4.3.2.2 Reasons for choosing the CMC samples

At the time of planning this research, there was already a very large amount of text data on the Internet and the amount was still growing rapidly. Choosing the texts as CMC samples for the present study was based on several considerations.

First, being an English teacher myself, I found the topic of English language teaching (ELT) a major interest. This was also the theme of the CMC activities with which I had been most often involved. I believed that I was familiar with the language and manners of communication on this topic, though I still needed the related background knowledge in analysing the result to be generated.

Secondly, I hoped that all the texts being investigated were of acceptable register to educated native language users. As Internet communication is usually open to all, people of different language backgrounds can join discussion lists, newsgroups and MOO sessions to post messages, which may result in the sub-standard usage of language on certain web sites. For instance, there are web sites developed for students of the English language to participate in the discussion, such as Student-L hosted in La Trobe University in Australia. Apparently, the texts posted in these discussion lists are from beginning or intermediate EFL learners, which shows the different interlanguage stages writers have reached. Their

writings may be useful for the study of error analysis, but not for the purpose of this study.

To ensure that the language being analysed was from educated English users, I had to be selective and careful in choosing text samples. The participants in NETEACH-L, TESL-L, NETEACH-MOO and Netoric-MOO are all English teachers, as most of them are native speakers of English or at least native-like, their English proficiency is believed to be high.

Of course, there were other discussion lists with the similar theme of ELT, but the above lists were, of all those which I have access to, the most popular, displaying a large number of postings in 1995-1996, when I attempted to collect the text samples.

In addition, as my aim is to interest the language teacher in the results of this study, I also hoped that the text samples I used were of a similar background. Choosing a discussion list on the theme of chemistry could probably be good, but it would not be as attractive to language teachers.

4.3.2.3 Sources of non-CMC texts

In the study of language use of CMC texts, non-CMC texts can also serve as an important reference. However, due to the technical difficulty of converting a large amount of non-CMC texts into computer-readable form, rather than trying to locate comparable original texts, I decided to adopt the quantitative data on already-existent non-CMC text samples of similar genres for possible reference. The data were collected from the work of Biber (1988a).

Biber (1988a) used 15 written genres from the LOB Corpus (Johansson, 1982; Johansson, Leech, & Goodluck, 1978), 6 spoken genres from the London-Lund Corpus (Svartvik & Quirk, 1980), and two types of unpublished letters (Biber, 1988a), forming a corpus of 960,000 words in 481 texts.

To achieve a fair comparison between CMC and non-CMC text groups, the non-CMC text groups should be of similar genres. I adopted from Biber's statistical figures only those texts from genres comparable to those used in the CMC text groups. Therefore, in terms of the asynchronous temporality, the statistics of two written genres, namely, academic prose and professional letters, were chosen. They were composed of a total of 90 text files, with a total for 170,000 words (Biber, 1995). For the synchronous temporality, three spoken genres were chosen for comparison: face-to-face conversation, telephone conversation, and public conversations, debates and interviews. These were in 93 text files with a total of 295,000 words (Biber, 1995).

4.3.2.4 Reasons for choosing the non-CMC data

As my intention was to show the variation of language use between CMC and non-CMC, it was first necessary to find comparable non-CMC texts. However, building a corpus is a highly technical endeavour and certainly beyond my ability for the present study. Therefore, the study by Biber, which uses a collection of written and spoken genres was deliberately chosen for its analysis of variation between written and spoken genres. It was a most convenient source of non-CMC data for the study.

4.4 Tools for Data Preparation and Analysis

The sample texts from the above CMC archives were converted into appropriated electronic files and further processed using two computer programmes: (1) the concordancing programme CLAN, developed at Carnegie Mellon University (MacWhinney, 1995, 1996a, 1996b); and (2) the part-of-speech tagging programme TAGGER developed in the University of Birmingham since the early 1990s (Mason, 1996).

4.4.1 CLAN – the concordancing programme

CLAN (Child Language ANalysis) is a set of programmes written by Leonid Spektor at Carnegie Mellon University, with design assistance from Brian MacWhinney (MacWhinney, 1995, 1996a, 1996b). These programmes are designed to allow users to perform a large number of automatic analyses of transcript data formatted according to the CHAT system of Child Language Data Exchange System. However, many of the programmes can run on ASCII files of any texts. The two main programmes used in the present study are COMBO (for key-word-in-context search) and FREQ (for frequency count).

COMBO is a small programme in CLAN to search for key words in context across every text in each text group. For instance, to find out how many times *agentless passives* (Feature 17) in the TESL-L text group, a COMBO command with the following notation can be given to search all the occurrences from all the texts in this group:

combo +y +m –w1 +w1 +s@verb_be^*^(/VBN)^*^!by c:\tsl9611t*.* +ff17

Then, COMBO searches and generates the output. An excerpt in Figure 4.1 is the output that COMBO reports for the Text No. LN0008 in TESL-L.

```
COMBO.EXE +y +m -w1 +w1 +s@verb_be^*^(/VBN)^*^!by c:\tsl9611t\LN0008 +ff17
**** line 20; file c:\tsl9611t\LN0008 ****
> John/NP mentioned/VBD that/IN the/DT gift/NN was/VBD brought/VBN here/RB
                                                    1       1
for/IN the/DT party/NN tonight/RB ./.

**** line 20; file p:\tsl9611t\LN0008 ****
> The/DT last/JJ example/NN is/VB a/DT noun/NN clause/NN embedded/VBN
                            1                           1
in/IN a/DT question/NN, an/DT incredibly/RB
Strings matched 2 times
```

Figure 4.1 An Excerpt of COMBO Output for *F17 Agentless Passives*

In Figure 4.1, we can see that COMBO marked "1 1" below the words *was brought*, and *is ... embedded.* At the end, it reported "*Strings matched 2 times.*" This means that it found two occurrences of *agentless passives* in this text, Text No. LN0008. This number of strings became the occurrences of this feature for this text. COMBO continues to search the same feature from other texts in the same text group, and reports back how many strings it found for every text. Respectively, the number of strings for each feature from every text was recorded as its occurrences.

However, the second occurrence in Figure 4.1 is actually not a correct instance for the *agentless passives* feature. This is due to the limitation of the computer capability in linguistic analysis for the time being. This point will be further discussed in Section 4.5.1.

4.4.2 TAGGER – the part-of-speech programme

The programme TAGGER (Mason, 1996) was used to affix tags of the linguistic features to each word, so that the features could be identified and counted by the CLAN programme. TAGGER is an automatic part-of-speech (POS) program developed at the University of Birmingham when John Sinclair and his colleagues edited the COBUILD English Dictionary in 1988. This programme was developed based on the probability of part-of-speech of all the texts collected for the dictionary during the compilation stage. A list of part-of-speech tags has been presented in Appendix 2. This programme was publicly accessible by means of an experimental email tagging service (Mason, 1996). The text could be sent to the TAGGER in the University of Birmingham, and the output files would be automatically sent back. To avoid the difficulty of getting a long text tagged, all the CA, and CS text samples used for this study were kept within 50 kilobytes in each file.

The sentence in Figure 4.1 in Section 4.4.1 also demonstrates how TAGGER affixed

a part-of-speech tag to every word of the sentence. Before tagging, the first sentence was:

(1) **John mentioned that the gift was brought here for the party tonight.**

After tagging, the first sentence became:

(2) **John/NP mentioned/VBD that/IN the/DT gift/NN was/VBD brought/VBN here/RB for/IN the/DT party/NN tonight/RB ./.**

A part-of-speech tag was affixed to each word to give more syntactic information during the tagging process, i.e. */NP* for the proper noun, */VBD* for the past tense verb, */IN* for the subordinate conjunction, etc. (with reference to the tag codes in Appendix 1). This tagging process is necessary for the present study, as many linguistic features investigated involve the words' part-of-speech, such as features like *nouns, adverbs, agentless passives, BY-passives, infinitives,* etc.

4.4.3 Reasons for choosing CLAN and TAGGER

In the corpus linguistic studies, some linguists hold the belief that language researchers need to be able to write their own software, rather than relying on those that already exist, in order to expedite their linguistic analysis (Sampson, 1998; Scott, 1998; Moffat, 1998). This is because every piece of software embodies its own particular theories, and thus might limit the work performed within a certain scope. Hence it would be difficult to suit the software to the other kinds of linguistic search that different researchers wish to accomplish. Moreover, it is believed that only the researcher himself is most clearly aware of the investigation work to be performed. Thus, the researcher is the best person to write the software for his/her study.

On the other hand, other linguists hold a different view. They propose that language researchers do not need to write their own software for linguistic analyses, but can simply get help from that which already exists, or rely on colleagues who are expert in computing technology (Caldas, 1998; Juola, 1998; Mason, 1998; Reetz, 1998; Wolters, 1998). This group also claims that writing a computer programme is one thing, but, testing and proving its correctness and reliability is another. Besides, if there is good software available, there is really no need to re-invent the wheel.

Considering my background and needs, I decided to adopt some existing software to achieve the aim of this study. However, it must be noted that, with the ready-made

software, I still needed to make necessary adjustments on my own so that it would present the work as I required. The measures of adjustment I needed to do included the design of the notations for the language features to be analysed, and the manual proof-reading of the output files from the computer to make sure that the software performed the analysis and calculation as intended. This will be discussed in more detail in the next few paragraphs.

For concordancing, there are already a number of software packages developed and tested by experts, which are readily available and also proved quite reliable (University of Essex, 1999; Granger & Tribble, 1998). These include CLAN (MacWhinney, 1995, 1996a, 1996b), WordSmith (Scott, 1996), MiniConcord (Bernstein, Bourne, & Bernstein, 1997), Monoconc (Barlow & Neumann, 1995), OCP (Micro-OCP, 1988), Sarah (Burnard, 1995), and TACT (TACT, 1990). However, each of these has its merits and limitations.

CLAN was chosen for this study for several reasons. First, it is a powerful tool that places no limits on the length of texts being analysed. Certain other popular concordancing tools usually place a strict limitation on the text length, or are not suitable for the large amount of texts being examined in this study. Besides, CLAN has an excellent online service by which the users can get technical help from the designers very easily. I have used it on several pilot studies (Tsai, 1997a, 1997b, 1997c, & 1999) and gratefully benefited a lot from the online service.

For tagging, TAGGER was chosen, because it claims to have a very high percentage of correctness, around 97 per cent, in labelling the part of speech of each word in a text. The output files generated by TAGGER are in linear format, which is appropriate for the structure searched by concordancing in the later stage. Moreover, its codes of part-of-speech tags can also serve the need of my linguistic analysis for all of the 67 features. For instance, TAGGER distinguishes the preposition *to* from the infinitive *to*; differentiates the existential *there* from the adverb *there*; tags codes differently to the various forms derived from verbs, etc. TAGGER is reliable as it was developed in the University of Birmingham and has been continuously improved in terms of accuracy in labelling the parts of speech since John Sinclair and his colleagues compiled the Collins COBUILD Dictionary in the 1980s. TAGGER is also free for academic use. It has been put online, with its software engineer as the system administrator taking care of the tagging service. It is very easy to use, even for a novice linguistic researcher. I believed that TAGGER was the most appropriate choice for my study. I would also like to demonstrate its use to English teachers during a discussion of its pedagogical implications in a later chapter.

There are other titles of tagging software, such as Amalgam Tagger (University of

Leeds, 1997), Brill Tagger (WinBrill, 1997), EngCG Tagger (Karlsson, Voutilainen, Heikkila, & Antilla, 1995), MBT (University of Essex, 1999) and CLAWS (Lancaster University, 1999). However, many of them generate the output files in a format that is not compatible for the use of CLAN in the following steps, or they have a length limit on the tagging texts. On the other hand, TAGGER is an automatic service provided online. The user can simply send the texts to the server by email and the output file is sent back in a short time. After carefully comparing the relative advantages of different tagging software, I decided to use TAGGER for the study.

The ready-made software was chosen after considering many factors. However, the software packages are just instruments. As the researcher, I would also have to exercise my own creativity and judgement in designing the notations and procedures so that the packages can function as expected. The creativity and judgement, I believed, were the major elements I needed to exhibit in the process of the study.

4.5 Calculation of the Linguistic Features

4.5.1 Text data preparation

The calculation was to be made on sample texts several hundred thousand words long, so it was necessary to resort to the use of the computer and the technique of corpus linguistics. For the language analysis tool CLAN, as introduced in Section 4.4.1, to locate and recognise certain words or linguistic structures, the structures must first be depicted in notations that could be recognised by CLAN. Where we need to examine the occurrence of a specified feature when intact words are not the only elements involved in a structure, the texts have to be first tagged, for example */VBD* (for *verb past tense*) and */VBN* (for *verb past participle*), etc. Such tags must be attached to all words for their respective part-of-speech. This tagging process has been discussed in Section 4.4.2. The researcher took advantage of the automatic TAGGER service at the University of Birmingham and had the CMC text files tagged.

Although the jobs of tagging and concordancing were done by the computer programmes, there was still a need for manual proofreading and correction. The TAGGER programme still has limitations in its power of discrimination. Though it claims to be 97 per cent correct, there is still a three per cent chance of errors. In addition, the notations designed with CLAN looking for the language features can not always elicit

precisely the features needed. This is due to the clear limitations of the computer's capability in linguistic analysis at the present stage.

An example given earlier in Figure 4.1 did show this problem. In Figure 4.1, the target feature was the *agentless passive*. This feature represents that a *verb-to-be* must be followed by a *past participle* either adjacently or remotely, as in the following sentences:

(3) ***The car was fixed yesterday.***

(4) ***The car was perfectly fixed yesterday.***

(5) ***The car was perfectly and nicely fixed yesterday.***

In order to cover all occurrences in the above sentences, a notation needed to be designed with a broader scope of the text in mind. Therefore, instead of using a notation that would retrieve only the adjacent occurrence, i.e. *verb_be^(/VBN)*, the other notation that could retrieve a remote occurrence, i.e. *verb_be^*^(/VBN)*, was applied. Here, the former notation stands for the *past participle* "immediately follows" the *verb-to-be*, while the latter for the *past participle* "immediately or remotely follows" the *verb-to-be*, as that used in Figure 4.1 earlier. The asterisk mark in the latter stands for anything between the *verb-to-be* and the *past participle*.

However, a notation that was good in retrieving the remote occurrence of a feature from a broader context would also trigger unwanted entries. For instance, in the earlier Figure 4.1, COMBO command searched "*is ... embedded*" as an occurrence for the *agentless passive* feature from the sentence:

> The/DT last/JJ example/NN is/VB a/DT noun/NN clause/NN embedded/VBN in/IN a/DT
1 1
question/NN, an/DT incredibly/RB

This was indeed not a correct search, and was an unavoidable result when I wanted to retrieve all the occurrences like that in the Sentences (3), (4) and (5). I had no choice but to tolerate such errors, and leave them for the process of proofreading afterwards. Therefore, the occurrences COMBO reported "*Strings matched 2 times*" was corrected as "1 time" only after the proofreading.

For each of the 67 features, a notation was developed. A few of the notations, and their illustrations, are listed below to show how they would work. For a complete list of the notations, please see Appendix 4.

F17: Agentless Passives ***i.e. The car was fixed yesterday.***

@verb_be^*^(/VBN)^*^!by

This notation represents the occurrence of a *verb-to-be* followed by a word tagged as */VBN* (i.e. a *verb past participle*), which is NOT followed by the word *by*. The occurrence of the ^*^ means there may or may not be any number of words occurring at that position.

F18: BY Passives ***i.e. The car was fixed by John yesterday.***

@verb_be^*^(/VBN)^*^by

This notation is similar to F17, the "*agentless passive*" above, with the exception that the word *by* must follow the *verb past participle* in the sentence.

F26: Past Participle Clauses

@sym.txt^(/VBN)

The @sym.txt is a file of several punctuation marks (,.?!) that serve as delimiters. When a *verb past participle* occurs after one of the punctuation marks, it is counted as an occurrence of this feature.

This notation is supposed to locate sentences like:

He sat there with a smiling face, satisfied with what he had earned.

He couldn't say a word, astonished at what he saw.

In earlier days, built in a single week, the house would stand for fifty years.

Of course, manual proofreading and revising are necessary to distinguish this feature from other sentences like:

The diamond was found, bought, and finally sent to the queen as a present.

F27: Past Participle WHIZ Deletions

/VBN

This notation simply tells the CLAN to look for any occurrence of words with the tag */VBN* (*verb past participle*). However, here it is the intention of the researcher to locate texts like:

The solution produced by the process ...

A situation resulted from this decision ...

To distinguish texts with this feature with other texts simply containing a word with */VBN* tagging, there seems no choice but to resort to manual proofreading after the texts have been concondanced.

F38: Other Adverbial Subordinators

@ad_sub

The @ad_sub means a pre-edited file containing all words that are considered adverbial subordinators. These include *since, while, whilst, whereupon, whereas, whereby*, etc.

F45: Conjuncts

@sym.txt^(@conj)

The @conj is a file of all conjuncts, such as *alternatively, altogether, consequently, else, furthermore*, etc. This notation indicates the occurrence of such conjuncts after a delimiter.

All of the above notations served as part of the commands to COMBO to search the specific features from the sample texts, so the occurrences of that specific feature could be recorded for the statistical comparison in the later stage. The full list of the notations of the 67 features is further demonstrated in Appendix 4.

4.5.2 Calculation of the linguistic features

As a comparison was to be made on the linguistic features found in texts from different sources, the mean (M) and standard deviation (SD) of the frequency of occurrence of each linguistic feature in each text group had first to be calculated. In the statistic process, the mean is the sum of all feature occurrences, divided by the text numbers, representing the centre value of each feature of one of the four CMC groups (Dorak, 2000; Easton & McColl, 2000). Standard deviation is a measure of the spread of the feature occurrences of each CMC group. The more widely the occurrences are spread, the larger the standard deviation. For example, if we have two sets of exam scores each from a class of 30 students, one ranges from 31 to 98, the other from 82 to 93, then the standard deviation would be larger for the results of the first exam.

However, when calculating the occurrences of every feature across different texts in each text group, i.e. the "*Strings matched N times*" reported by concordancing mentioned in Section 4.4.1, the calculation must be made on the same basis in all of the text groups so as to make the comparison fair and justifiable. This is because each text is in different in length, and it is very likely that the longer the texts, the higher the occurrences of the features. A way to justify this is to further adjust the occurrences found for each feature in every text with a formula, shown in Figure 4.2, as if the occurrences are taken from the same length text, which is 1,000 words.

$$C = (A / B) * 1{,}000$$

While A = *Occurrence frequency found for each feature in every text*
B = *Total words of this text*
C = *Feature frequency for later statistical comparison*

Figure 4.2 Adjustments for the Occurrences Found for Each Feature in Every Text

This formula standardised the feature occurrences of each feature for every text as if all of the feature occurrences were from the same text length of 1,000 words, despite whether the text is short or long. It avoided the bias that shorter texts might have fewer occurrences, or longer texts that might have higher occurrences.

The notation, as mentioned in Section 4.5.1, and the calculation, as mentioned in Section 4.5.2, as well as many others in the following stages, needed to be processed on sample texts several hundred thousand words long. Therefore, the need to resort to the use of the computer and the technique of corpus linguistics was necessary. This step involved the calculation of feature occurrences found by the set of notations of the 67 linguistic features.

After each of the texts had been through the process of concordancing and frequency count for each feature, the resulted figures from different pieces of texts were further merged and processed to obtain the mean (M) and standard deviation (SD) of the frequency of occurrence of the linguistic feature in the whole text group. As mentioned earlier, mean is the centre value of each feature in a text group, and standard deviation is how each feature spreads in a text group. They are two of the important data needed, in addition to the number of texts in each text group, for the statistical comparison in the later stage. A full list of the text numbers (N), means (M) and standard deviations (SD) of the occurrences of each feature in each of the four groups of text samples were collected

and arranged in Appendix 5. All numbers in this table had been used in the statistical comparisons to be discussed in the Section 4.6.

4.6 Statistical Analysis

To develop an adequate model of analysis, a long painstaking process of thinking and trial was necessary. The methods considered and tested are described in this section in order to give a picture of the methodology development. They included a 2 x 2 Factorial Experiment between CMC and non-CMC texts (Section 4.6.1), a Factor Analysis on CMC texts (Section 4.6.2), and a Non-Parametric Statistical Test of synchronous vs. asynchronous CMC texts (Section 4.6.3). Each will be discussed with its findings and issues to see how the methodology was developed to best fit the purpose of the present study. Finally, a brief description of the findings of the Non-Parametric Statistical Test will lead to further discussion in the next chapter.

4.6.1 2 x 2 Factorial Experiment Between CMC and non-CMC texts

The first approach used in an attempt to analyse the CMC texts was that of a 2 x 2 Factorial Experiment, also referred to as Two-Way Analysis of Variance (ANOVA).

ANOVA is the measure widely used to compare the means of data value in several groups. When differences are found between groups, it is up to the strict statistical procedure to ascertain that the difference is indeed substantial, and not the effect of some unknown sources or sampling errors. That is the reason for using the statistical method of ANOVA instead of resorting to one's own subjective judgement.

The 2 x 2 Factorial Experiment means that there are two variables examined in the study, and each has two levels. For the present study, the two variables were the media of communication and temporalities of communication, and the two levels were CMC and non-CMC for media, and synchronous and asynchronous for temporalities.

To process the comparison, the following data of each feature are needed, as mentioned in Appendix 5 in Section 4.5.2: mean (M), standard deviation (SD), and the number of the texts (N) in each text group. Here are the data, taken from Appendix 5, of the first feature, *past tense* (F01). The M, SD, and N are depicted below in Table 4.3 for ready reference.

Table 4.3
M/SD/N for *F01 Past Tense*

Text Groups	M	SD	N
CA (CMC Asyn.)	9.51	14.1	503
CS (CMC Syn.)	11.55	8.57	17
NA (non-CMC Asyn.)	20.59	20.36	90
NS (non-CMC Syn.)	33.03	19.47	93

The means of the four groups, i.e. CA, CS, NA, and NS, can then be depicted in a 2 x 2 (4 cells) tablet like the following, with the sums of two related cells shown right under or beside them, as in Table 4.4 for *past tense* (F01):

Table 4.4
M and Sum of Ms for *F01 Past Tense*

A(Media)* \ B(Temporalities)*	Asyn.	Syn.	Sum
CMC	9.51	11.55	21.06
non-CMC	20.59	33.03	53.62
Sum	30.10	44.58	74.68

*: A stands for media of communication; B stands for temporalities of communication

In the above table, A stands for the first variable -- media of communication covering CMC and non-CMC media, while B refers to the other variable -- temporalities of communication covering asynchronous and synchronous temporalities. The number 30.10, in the bottom row is the sum of means 9.51 and 20.59, while 21.06, in the rightmost column, is the sum of means 9.51 and 11.55, etc. 74.68 is the sum of the Ms in all of the four cells of 9.51, 11.55, 20.59, and 33.03.

As the four text groups have a different number of texts (N), i.e. 503, 17, 90, and 93, as shown in Table 4.2, the procedure required in processing the data is somewhat different from that where the number of texts is the same for all the groups. Technically, the procedure for this study was a 2 x 2 Factorial Experiment with unequal numbers of texts in the cells (Broota, 1989, pp. 162-165).

A figure and two tables were produced to summarise the ANOVA findings for *past tense* (F01), as shown in Figure 4.3, and Tables 4.5, and 4.6.

In Figure 4.3, we can see the relative position of the four means of the feature *past tense* (F01) in the four text groups. We can see that synchronous non-CMC (NS:33.03) has a higher frequency of *past tense* (F01) than asynchronous non-CMC (NA:20.59). It also has a higher frequency of F01 (NS:33.03) than synchronous CMC (CS:11.55), etc.

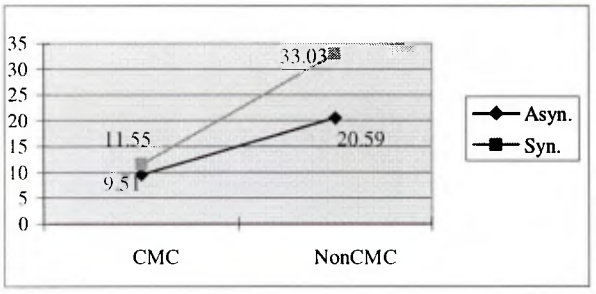

Figure 4.3 Means of Four Text Groups for *F01 Past Tense*
(*Oh, you **took** your car to school? Why **was** that?"* (tesln-107))

However, we cannot be sure those differences are truly significant or just the result of the sometimes inevitable sampling errors. In Tables 4.5 and 4.6 below, the asterisk (*) is used to indicate that the difference for a variable is significant subject to the 5 per cent possibility of error generated by the Two-Way ANOVA Test. That is to say, when the asterisk appears, we are 95 per cent confident that the difference is truly significant.

In the analysis depicted in Table 4.5, the three F values represent the between-group differences of the variables A (media) and B (temporalities), and the interactive effect of A and B respectively. The $F_{.95(1,\infty)}=3.84$ is the criterion to judge if the between-group differences have reached the 95 per cent level of confidence. When an F value meets this criterion, an asterisk is affixed to indicate a significant difference.

Table 4.5
Significant Differences Found by Two-way ANOVA on *F01 Past Tense*

Two-way ANOVA					
	SS	df	MS	F	
A	12823.084	1	12823.08	51.56909	*
B	2536.0676	1	2536.068	10.199	*
AB	1308.249	1	1308.249	5.2612306	*
w cell	173812.2	699	248.6583		
				$F_{.95(1,\infty)}=3.84$	

The findings in Table 4.5 show that the difference between the two levels of variable A, the media of communication, is significant, as an asterisk (*) appears in the cell of F and A. The same is true for the differences caused by variable B, the temporalities of communication, i.e. the cell of F and B. The interactive effect of the two variables A and B is also significant in determining the difference of the feature *past tense* among the four text groups, i.e. the cell of F and AB. As interactive effect of more than one variable is found, the F value of AB interaction becomes the focus of observation and interpretation,

and no attempt is made in observing the main effect of each individual variable, the F value for A or B. The reason why no attempt is needed here is further explained in the following paragraph.

Take the feature *past tense* (F01) in Table 4.5 for example. Though both variables of media (A) and temporalities (B) have a main effect on the differences, the effect from the individual variables is not further examined as the significant interactive effect (AB) takes priority in the study here. Hence, observation and interpretation will be focused on the interactive effect (AB), but not for the main effect (neither A nor for B). This is the primary rule in ANOVA.

When an interactive effect (AB) is found, a further investigation is necessary to see how they interact with each other. Table 4.6 below further examines the simple main effect of the two variables with this task.

Table 4.6
Significant Interactive Difference Found by Two-way ANOVA on F01 *Past Tense*

Simple Main Effect					
	SS	df	MS	F	
A for b1	2969.841	1	2969.841	11.94346	*
A for b2	11161.49	1	11161.49	44.88686	*
B for a1	100.6732	1	100.6732	0.404866	
B for a2	3743.643	1	3743.643	15.05537	*
w cell	173812.2	699	248.65834		
				$F_{.95(1,\infty)}$=3.84	

As mentioned earlier, variable A stands for that of medium and variable B stands for that of temporality. We can see from this table that there is significant interactive effect in three of the four sets of comparison. The set of comparison that does not show significant differences is "B for a1", which means that variable B (temporality) does not show significant difference when the texts are in the first level (CMC) of variable A (medium of communication). In other words, CMC texts (a1) do not show any significant difference between the two temporalities (CA:9.51, CS:11.55, See Figure 4.3 mentioned earlier). In all of the other three sets of comparison, namely "A for b1," "A for b2," and "B for a2," we can find significant differences. That is to say, the feature *past tense* (F01) shows significant difference between the media when the temporality is either synchronous or asynchronous (NA:20.59>CA:9.51; NS:33.03>CS:11.55, referred to in Figure 4.3), and between the temporalities when the medium is non-CMC (NS:33.03>NA:20.59, also referred to in Figure 4.3).

The findings resulting from the 2 x 2 Factorial Experiment showed quite complicated patterns between the CMC and non-CMC, and between the asynchronous and synchronous texts. Basically, the two variables, i.e. medium and temporality, each had its impact on language used in terms of medium and synchronicity, while both of the variables mutually interacted with each other and affected the texts in different and complicated patterns. Details of how significant differences were found with many features among the different text groups is outlined in Appendix 6.

The adoption of the approach of a 2 x 2 Factorial Experiment here is based on the assumption that there are two variables involved in the text samples: media and temporalities. In the variable of media, the two levels are CMC and non-CMC, and in terms of the temporalities, the two levels are synchronous and asynchronous.

By using the two variables, I hoped to provide a broader and also deeper understanding when comparing the language used in CMC in contrast with that used in non-CMC texts. Technically, however, there were some problems that cannot be overcome easily.

For one thing, as the researcher did not have access to a comparable non-CMC corpus composed of similar genres to the CMC texts, she had to rely on the statistical results of non-CMC text analysis as a basis for comparison with CMC data which was gathered firsthand. Though every care had been taken to minimise possible errors, there was still not strong evidence to prove that the procedures of processing of data were the same for both CMC and non-CMC texts. For example, the tool used by Biber in tagging the non-CMC texts was certainly different from the TAGGER, which the present researcher used. Though the effect of the different tools used might be very small, there was no way to prove it. The present researcher did try to contact Professor Biber, who carried out the study on the non-CMC texts, to see if it was possible to have access to at least some of his original data so that the TAGGER could be used to see if the result was differentt. However, as expected, the study by Biber was done over 10 years ago and all the original data were long since lost.

The inaccessibility of the original non-CMC data also caused concern that the texts in those files were not necessarily of comparable genres with the CMC texts. The genres selected from the non-CMC corpus for analysis were academic prose, professional letters, face-to-face conversation, telephone conversation and public conversations, debates, and interviews. The CMC genres were asynchronous and synchronous discussion on language

teaching issues. Without the original texts to serve as evidence, it was hard to argue the text samples on the two media were of compatible genres.

Though there were limitations in the 2 x 2 Factorial Experiment above, the Two-Way ANOVA approach still had its strength in demonstrating the differences found between the CMC texts and non-CMC texts. Among the 67 features investigated, it was found that 15 had no differences across the media and temporalities. It was also found that the differences between the two temporalities, i.e. asynchronicity and synchronicity, mainly existed in non-CMC texts rather than in CMC texts (See Appendix 6). This is in line with some previous findings that CMC is a relatively more neutral medium, which does not exhibit as many differences between temporalities as does the non-CMC medium (Collot & Belmore, 1996; Halliday, 1989; Lundstrom, 1995; Maynor, 1994; Tribble, 1998). In other words, CMC medium blurs the distinction of spoken and written forms of communication.

4.6.2 Factor Analysis on CMC texts

Another approach to comparing the differences between CMC and non-CMC texts was to follow the Factor Analysis approach of Biber (1988a) and undertake a similar analysis on CMC texts such as that discussed in Section 2.4 in Chapter 2: The Study of Language Variation. Unlike other researchers' studies based on the assumption of dichotomous distinction, i.e. written and spoken languages, Biber's approach takes the oral/written distinction as on a continuous scale of variation. Different types of texts in various styles, registers and genres are not the same or are dichotomously different; rather they are "similar, or different," to differing extents with respect to each dimension (Biber, 1988a, p.22). Based on the 67 linguistic features found in a corpus of 960,000 words and 23 written and spoken genres (Biber, 1988a, p. 67), he implemented the method of factor analysis and came up with a total of seven factors, which may serve as dimensions on which texts may be measured as being more speaking-like or writing-like. These include:

Factor 1: informational versus involved production,
Factor 2: narrative versus non-narrative concerns,
Factor 3: explicit versus situation-dependent reference,
Factor 4: overt expression of persuasion,
Factor 5: abstract versus non-abstract information,
Factor 6: online informational elaboration, and,
Factor 7: academic hedging.

Each factor is characterised by some of these 67 features (See Appendix 7). Each of these factors is taken as a dimension on which speech and writing varies. After its publication, these seven dimensions have been quoted by many (e.g. Allomong, 1996; Collot & Belmore, 1996; Ljung, 1991; Tribble, 1997) in the study of language styles and in the design of language teaching syllabus and materials.

A possible approach to a comparison between CMC and non-CMC texts, especially when original non-CMC texts are not available, was to implement the Factor Analysis method on the CMC texts, as was discussed in Section 2.4.1 in Chapter 2. It was hoped that the analysis would result in a set of factors similar to those extracted in the non-CMC texts. Assuming that there would be agreement, then it would be easy to plot where each of the CMC data sets fitted along these dimensions.

To do this, the data of M, SD, and N of all CMC texts were gathered, as listed in Appendix 5, and used in an SPSS Principal Axis Factoring and Varimax Rotation (Lin, 1992, pp. 639, 650), with an Eigen Value chosen larger than or equal to one (Kaiser, 1958). The results of the Factor Analysis were quite out of expectation. Shown from the Appendix 4.7, among the CMC texts, there were 26 factors extracted from the 67 features!

In the structure, it can be seen that the one factor with the highest Eigen Value has eight features in it, i.e. Factor 1 in Appendix 8, six with positive loadings and two with negative loadings. The other factors usually have just less than three features in each. This is clearly very different from the case of non-CMC texts examined by Biber (Appendix 7). Besides, a comparison between the data in Appendices 7 and 8 can show that, in the first four factors, where there is a substantial number of underlying features, the features in each factor are not in any sense related to those features gathered in Biber's seven factors for the non-CMC texts.

It seemed then that Factor Analysis was not appropriate for the present study. Usually, Factor Analysis is hoped to derive a relatively small set of underlying variables, called "factors," from large sets of variables, such as the seven factors retrieved from the 67 linguistic features (Biber. 1988a). This small set of factors is considered to be a guideline in explaining the differences existing in the sample. McEnry and Wilson (1996) in their book, *Corpus Linguistics*, recommended that four factors are the best number to account for the nature of the sample texts. Moreover, from a number of previous studies, the number of factors (Bainbridge, 1986; Des Brisay, Duquette, & Dirir, 1993; Dunning, 1995; McEnery & Wilson, 1996; Warschauer, 1996d) usually ranged from two to five.

Based on these guidelines, the number of the factors found in the present study, as many as 26, has little value in interpreting the essence of the CMC texts.

The result of factor analysis on the CMC texts was quite out of expectation. A very clear phenomenon of the factors gathered in CMC texts was the diversity of the features. The features, when occurring in CMC texts, did not gather into clusters as they occurred in non-CMC texts. This seems to suggest that CMC texts are much more diversified in language use and cannot be easily represented as being of several types.

The diversity of the Factor Analysis in the CMC texts made it extremely difficult to make any comparison between CMC and non-CMC texts.

As neither the 2 x 2 Factorial Experiment nor the Factor Analysis could provide a good instrument with which to make a convincing comparison between the CMC and non-CMC texts, the framework of analysis had to be modified. It was therefore decided to focus only on the CMC texts, because a comparison of the two temporalities of CMC was envisioned to still help foster some insight into the nature of the new medium, CMC.

Though there were restrictions, as mentioned above, for the present study, applying a Factor Analysis did unveil part of the mystery of the CMC text. In the past, many studies (Collot & Belmore, 1996; Halliday, 1989; Lundstrom, 1995; Maynor, 1994; Tribble, 1998) claimed that, either from the researchers' own intuition or from their empirical observations, the CMC text was somewhat a mixture of written and spoken language. Now, from the empirical work done in this study, the results from the Factor Analysis did confirm this claim to some extent. It is true that the CMC text does not have the nature that many linguistic features cluster into a few groups, similar to that in the non-CMC text.

In non-CMC texts, the linguistic features tend to concentrate in major groups, i.e. many of them share similar or opposite characteristics so they form a few groups (Appendix 4.6, Biber's seven factors). On the contrary, the same features do not tend to form groups in the CMC text in the same way (Appendix 4.7, 26 factors for the present study).

In other words, while the non-CMC texts can be compared or contrasted along a few dimensions, CMC texts do not provide us with a controllable number of dimensions. It can be seen that most of the features in CMC texts are quite independent and do not correlate very much with each other. A simple conclusion from the comparison may be that CMC texts are just unconventional. Thus, we cannot compare the synchronicity of CMC as we compare the spoken and written forms of non-CMC texts. The present result

from the Factor Analysis did pave a ground for further work on studying the CMC texts. As the Factor analysis can be used to uncover patterned variation, it would be interesting to investigate further why the CMC texts are so diversified in nature. However, the findings have provided a preliminary piece of evidence that CMC texts are newly emerging genres. CMC texts are new varieties that is quite diversified in nature, and therefore we need to take other approaches in order to make some concrete observation.

4.6.3 Non-Parametric Statistical Test of synchronous vs. asynchronous CMC texts

In the hope of comparing the occurrences of linguistic features between the synchronous and asynchronous temporalities of CMC texts, the first plan was to use the technique of One-Way Analysis of Variance (ANOVA). This is the usual way to compare the means between different groups in the linguistic statistical approach. As the CMC texts were gathered first-hand, the standardised frequency of occurrences, as that discussed in Section 4.5.2, of each feature in each text can be clearly depicted, such as in Table 4.7 below for NETEACH-L and Table 4.8 for NETEACH-MOO.

Table 4.7
A Sample Data of Adjusted Feature Scores in Individual Postings of NETEACH-L

NETEACH-L	Total words	F01 Past tense	F02 Perfect aspect	F03 Present tense	...	F66 Synthetic negation	F67 Analytic negation
No001	776	3.87	2.58	39.95	...	.00	9.02
No002	50	60.00	.00	20.00	...	.00	20.00
No003	117	.00	.00	59.83	...	.00	17.09
No004	46	.00	21.74	43.48	...	.00	.00
No005	79	37.79	.00	12.66	...	.00	12.65
No006	154	19.48	.00	19.48	...	.00	6.49
No007	197	.00	5.08	76.14	...	.00	.00
No008	118	.00	.00	67.80	...	8.47	8.47
No009	39	.00	.00	51.28	...	.00	.00
No010	95	.00	.00	42.11	...	.00	.00
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...

The technique of ANOVA is implemented on the basis that the values of the data in each of the groups being compared are normally distributed, or with normal distribution. Usually, it is believed that when there is a large enough number of data, the values will form a pattern of natural distribution, with most of the values gathered toward the mean,

while fewer and fewer of the values are distributed toward the extremes. Mean has been referred to earlier as the central value of a feature in a text group. How the feature distributes across all texts in the group is represented by distribution. Distribution is the feature frequency distributed on a scale. When the data sample is large enough, its distribution usually appears in a symmetrical, bell-shaped curve, centred at its mean, called *normal distribution* (Dorak, 2000; Easton & McColl, 2000). This is the prerequisite for ANOVA test. However, the concept that it is usually normally distributed is not always true, even though the sample is large enough.

Table 4.8
A Sample Data of Adjusted Feature Scores in Individual Postings of NETEACH-MOO

NETEACH-MOO	Total words	...	F31WH rel. cl. On subj. pos.	F32 WH rel. cl. On obj. pos.	F33 WH rel. cl. Pied pies	...	F67 Analytic negation
Nmo0417	6381	...	.47	.00	.00	...	5.49
Nmo0424	9911	...	.40	.00	.00	...	4.94
Nmo0509	4616	...	.65	.00	.00	...	9.97
Nmo0516n	6691	...	1.20	.00	.00	...	7.17
Nmo0612	4753	...	1.68	.00	.00	...	5.47
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...

A look at Tables 4.7 and 4.8 can lead to the observation that the values in many cells are zero. This would arouse the concern that the values are not presented in a normal distribution. Hence, the idea of carrying out a One-Way ANOVA was not appropriate because of this skewed distribution. A simple distribution chart of the feature F01 *past tense* in NETEACH-L (see Figure 4.4) can show the problem of skew.

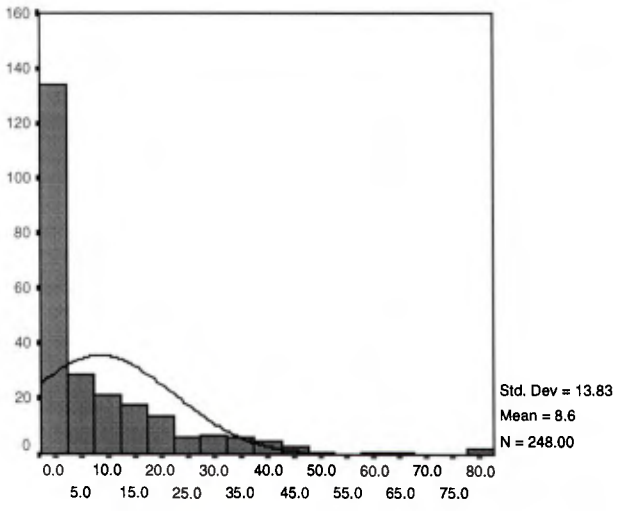

Figure 4.4 ***F01 Past Tense*** **Distribution in NETEACH-L**

From Figure 4.4, it is clear that among the 248 asynchronous CMC texts in NETEACH-L, up to 131 texts that did NOT use the *Past Tense Verb* at all (zero value, the highest polygon in the chart). This results in an obvious skewed distribution of data. Most other features occurring in the sample texts also show similar skewed distribution.

The reason is clearly the large number of cells with the value zero. As there are many texts that are very short in length, a lot of features are just not used at all.

This skewed distribution appearing in the analysis of many features in the text files made it inadequate to apply the more popular process of One-Way ANOVA. Instead, a Non-Parametric Statistical Test was used for comparison, as it is more appropriate to apply the latter when normal distribution criterion is not met. Also, to minimise the skewedness caused by the short length of many texts, it was decided to select, for the purpose of analysis only, the texts that were over a certain length.

The criteria set for the final text-length selection was 200 words per text. This was necessary because this number was close to its original mean of 199.19 words per text, in all of the 503 CA texts, with a total of 100,194 words. By way of a new selection, the number of CA texts was reduced to 203 texts, with a total of 67,527 words, with a new mean of 332.65 words. It was hoped that the extremely skewed distribution could be reasonably reduced in this new selection. Below, Tables 4.9 and 4.10 show the information of the original texts, and that of the new texts after further selection.

Table 4.9
Samples of CMC Texts in Original Length

	Texts	Minimum Words	Maximum Words	Word counts	Mean Words	SD
CA	503	4	1,249	100,194	199.19	148.40
NETEACH-L	248	4	1,249	50,028	201.73	176.53
TESL-L	255	25	649	50,166	196.73	114.94
CS	17	3,369	9,911	104,162	6,127.18	2,018.06
NETEACH-MOO	10	3,369	9,911	52,109	5,210.90	1,964.64
Netoric-MOO	7	5,816	8,896	52,053	7,436.14	1,293.14
TOTAL	503			204,356		

Once the text length had been decided, the Non-Parametric Statistical Test was followed using the Man-Whitney U function. Man-Whitney U is a Non-Parametric Statistical Test used for two independent samples to compare their distribution, just as the

present research is going to compare asynchronous vs. synchronous CMC texts (CA vs. CS).

Table 4.10
Samples of CMC Texts over 200 Words per Text

	Texts	Minimum Words	Maximum Words	Word counts	Mean Words	SD
CA	203	200	1,249	67,527	332.65	145.53
NETEACH-L	89	200	1,249	33,493	376.33	186.44
TESL-L	114	201	649	34,034	298.54	90.09
CS	17	3,369	9,911	104,162	6,127.18	2,018.06
NETEACH-MOO	10	3,369	9,911	52,109	5,210.90	1,964.64
Netoric-MOO	7	5,816	8,896	52,053	7,436.14	1,293.14
TOTAL	220			171,689		

4.6.4 Results from the Non-Parametric Statistical Test

The results from the Non-Parametric Statistical Test yielded a total of 29 features, with significant differences for CA vs. CS (Table 4.11). Among them, 17 features have a mean value in CA larger than CS texts, and 12 features in CS larger than CA. It shows that CA texts generally follow the textual nature characterised in written language, while CS texts follow that of spoken language. However, as Biber (1988a) argues that a dichotomous distinction is not appropriate to categorise language texts, a further discussion on the findings will be explored in the next chapter, Chapter 5: Findings and Discussion.

4.7 Summary

This chapter first gave an introduction to the overall background of the research, with a short review of the impact of CMC on the language use, and of language teaching in countries where English is a second or foreign language. This led to the researcher's interest – how language uses in asynchronous and synchronous CMC communication vary, how the language used in this medium should be introduced into the language classroom, and what language characteristics exist in the CMC medium. The chapter went on to discuss the factors that are involved in language analysis, and its contributions to the pedagogical practices.

Table 4.11
Significant Differences Found from Non-Parametric Statistical Test at $\alpha = 0.005$

FEATURES	CA > CS	CS > CA
(A) TENSE AND ASPECT MARKERS		
F03 Pres. tense V.		*
(B) PLACE AND TIME ADVERBIALS		
F05 Time adv.		*
(C) PRONOUNS AND PRO-VERBS		
(C3) PRO-VERBS		
F12 DO as pro-V.		*
(D) QUESTIONS		
F13 WH question		*
(E) NOMINAL FORMS		
F14 Nominal.	*	
(F) PASSIVES		
F18 BY-passive	*	
(H) SUBORDINATION		
(H1) COMPLEMENTATION		
F22 THAT cl. as adj. compl.	*	
F23 WH clause	*	
(H2) PARTICIPIAL FORMS		
F25 Pres. part. cl.		*
F26 Past. part. cl.	*	
F27 Past part. WHIZ del.	*	
(H3) RELATIVES		
F29 THAT cl. on obj. posit.	*	
F30 THAT cl. on obj. posit.	*	
F32 WH cl. on obj. posit.	*	
F34 Sentence rel.	*	
(H4) ADVERBIAL CLAUSES		
F35 Causative subo.	*	
F36 Concessive subo.	*	
F38 Other adv. Subo.	*	
(I) ADJECTIVES AND ADVERBS		
(I2) ADJECTIVES AND ADVERBS		
F40 Attrib. adj.	*	
(K) LEXICAL CLASSES		
F47 Gen. hedge		*
F50 Discourse part.		*
(M) SPECIALISED VERB CLASSES		
F55 Public verb		*
F58 SEEM, APPEAR	*	
(N) REDUCED FORMS AND STRUCTURES		
F59 Contraction		*
F60 THAT del.		*
F61 Final prep.	*	
F62 Split inf.		*
(O) COORDINATION		
F65 Non-ph. coord.	*	
(P) NEGATION		
F66 Synthetic neg.		*
Subtotal	17 cases significant	12 cases significant
Total	**29 cases significant**	

To achieve the goal of analysing language use in the medium of CMC, a collection of 67 features was adopted from Biber (1988a) as the objects of study in the texts. Then, the researcher collected and processed a certain amount of texts as samples. These were from asynchronous and synchronous CMC contexts of discussion lists and MOO respectively. For the possible need for comparison, the figures resulting from Biber's text analysis (1988a) from non-CMC texts were also collected. To process the CMC texts, notations were designed for each of the 67 features so the concordancer could recognise the features correctly. The device for text annotation, the TAGGER, wase used to prepare the texts, and the concordancing tool of CLAN was used to locate the features in the texts. With the frequencies of the features calculated from the sample texts, the data were to proceed through statistical analysis. The process of statistical analysis would depend on the framework of analysis.

At first, the researcher intended to conduct a Two-Way ANOVA on the data, considering both media and temporalities as variables, in the hope of finding some interaction between the two factors affecting the language use in both CMC and non-CMC texts. However, this approach met with difficulties: the non-CMC data came from the results reported in previous studies and could not provide statistically convincing data as first-hand materials. Therefore, the scope of the research was reduced to only the CMC texts, comparing the two temporalities: synchronous and asynchronous. As there was one variable, the temporalities, involved, the original approach of the Two-Way ANOVA gave way to the procedure of a One-Way analysis. However, as the values of many of the features examined were found to be of skewed distribution in different texts, the procedure of Non-Parametric Statistical Test was chosen, rather than the more commonly known One-Way ANOVA.

The procedure is known by its capability to reduce the errors of skewed distribution of data. The result from the Non-Parametric Statistical Test showed that 29 out of the total 67 features showed significant differences in their frequencies across synchronous and asynchronous texts. Among them, 17 had significantly higher frequency in asynchronous CMC texts, while 12 features exhibited a high frequency in synchronous CMC texts. A detailed discussion of how these 29 features appear differently across CA and CS will be made in the next chapter, Findings and Discussion. After that, some pedagogical implications based on the present research will be presented in Chapter 6: Pedagogical Implications.

Chapter 5: Findings and Discussion

5.1 Introduction and Overall Discussion of the Statistical Findings

5.1.1 Introduction

In Chapter Four, Methods and Procedures of Investigation, a detailed process was set out, explaining how each step of this research has been executed. These steps included the CMC sample text collection, text tagging, text concordancing, proofreading of the concordance output, the adjustment of feature occurrences, the series of statistical procedures tested and finally, the conclusive statistical approach implemented. After this long process, the findings on the 67 linguistic features, which fell under 16 categories, were gathered. The statistical procedure of Non-Parametric Statistical Test was the main instrument used to test the differences between the two temporalities of CMC sample texts: synchronous and asynchronous.

The findings are presented in this chapter under the 16 categories. In each section, the statistical findings of one category are presented in a table and, where necessary, in a number of charts. Some discussion is made based on the findings and related research. Where necessary, cross-reference is made to the content of the different sections in this chapter. In addition to comparing the features between the CA and CS texts, I also present a comparison of the use of each feature between CMC and non-CMC genres. The genres in the non-CMC texts are taken from Biber's study (1988a) on LOB Corpus and London-Lund Corpus. My findings on the CMC and non-CMC texts are not the result of any strict referential statistic, though I believe a direct observation on some simple comparison can still generate some very interesting ideas.

The purpose of the study is certainly not to play with the numbers. It is based on the belief that the frequencies of these features, as used in different genres, may be a good indication of the nature and characteristic of that genre. In comparing the features used in the different temporalities of the CMC texts, and between the CMC and the non-CMC texts, I hope to reveal the special characteristics of the new medium of Computer-mediated Communication. It is also hoped that some pedagogical implications can be found from the results of the study. This element will be presented in Chapter Six.

Chapter Five has 18 sections. The data for each category are presented in individual sections, starting with tense and aspect markers. At the end of the chapter, Section 5.18 summarises of the major findings reached in the study.

5.1.2 Overall findings from the statistical procedures

All of the 67 linguistic features used in the two CMC text groups were analysed using the statistical method of Non-Parametric Statistical Test, which was outlined in Chapter Four. The overall statistical findings are summarised in Table 5.1. Findings on all of the 67 features are examined further under the 16 separate categories - Section 5.2 to 5.17.

The results for each feature are presented in four columns in Table 5.1. The third and fourth columns list the comparison between the CA and CS text groups. Column CA>CS details those features that demonstrate a higher frequency in the CA texts. The numbers listed, such as 5.30>4.03 for the perfect aspect verb (F02), are the means of the adjusted scores of the feature in the two text groups. The adjusted scores stand for the frequency of occurrence of that feature in a text that is 1,000 words. The example of the case in "5.30>4.03 for the perfect aspect verb (F02)" means that the perfect aspect verb has an average number of the adjusted score of 5.30 in the CA texts, and an adjusted score of 4.03 in the CS texts. These average numbers, i.e. 5.30 or 4.03, are termed as "mean" in the statistical processes.

The column CS>CA, is of course just the opposite, detailing the cases in which the CS texts have a higher mean of that feature than the CA texts. For instance, the numbers of 11.55>10.93 for the past tense verb (F01) indicate that the feature past tense verb occurs 11.55 times in every 1,000 words in the CS texts, and 10.93 times per 1,000 words on average in the CA texts.

The "P value" column on the rightmost shows the index of probability of error. As the level of significance is set at α=0.005, we allow only 0.5% of chance for an error based on sampling, that is, we have 99.95% confidence in claiming that the finding is a significant one. If the P value generated from the statistical procedure exceeds 0.005, then we will not consider the difference significant. On the other hand, if the P value is not larger than 0.005, the difference is considered significant and an asterisk will be marked to indicate a significant difference.

The numbers used in type/token ratio (F43) and word length (F44) are exceptions. They do not denote the frequency of occurrence as per 1,000 words, which the other features adopt. Type/token ratio is the ratio of the different types of words in a piece of text, against the total number of words in that piece of text. Word length, as its name suggests, means the average length of words (in the number of characters) in the texts.

Table 5.1
Significant Differences of 67 Features Found from CMC Texts

FEATURES	CA > CS	CS > CA	P Value
(A) TENSE AND ASPECT MARKERS			
F01 Past tense V.		11.55 > 10.93	0.111
F02 Perf. asp. V.	5.30 > 4.03		0.632
F03 Pres. tense V.		102.62 > 42.14	0.000*
(B) PLACE AND TIME ADVERBIALS			
F04 Place adv.	2.30 > 2.02		0.262
F05 Time adv.		4.00 > 2.45	0.000*
(C) PRONOUNS AND PRO-VERBS			
(C1) PERSONAL PRONOUNS			
F06 1st per. pron.	30.10 > 26.00		0.561
F07 2nd per. pron.		11.09 > 8.79	0.010
F08 3rd per. pron.		13.85 > 12.42	0.121
(C2) IMPERSONAL PRONOUNS			
F09 Pron. IT	8.58 > 8.17		0.312
F10 Demonst. pron.		4.47 > 4.46	0.388
F11 Indefinite pron.	2.41 > 2.40		0.148
(C3) PRO-VERBS			
F12 DO as pro-V.		0.65 > 0.33	0.000*
(D) QUESTIONS			
F13 WH question		1.27 > 0.73	0.000*
(E) NOMINAL FORMS			
F14 Nominal.	22.10 > 13.01		0.004*
F15 Gerund	6.38 > 4.15		0.128
F16 Noun	211.00 > 192.10		0.543
(F) PASSIVES			
F17 Agentles pass.	5.94 > 3.34		0.040
F18 BY-passive	0.73 > 0.34		0.000*
(G) STATIVE FORMS			
F19 BE main V.		14.96 > 13.10	0.096
F20 Exist. THERE	1.54 > 1.01		0.010

Table 5.1
Significant Differences of 67 Features Found from CMC Texts (continued)

FEATURES	CA > CS	CS > CA	P Value
(H) SUBORDINATION			
(H1) COMPLEMENTATION			
F21 THAT cl. as V. compl.	2.33 > 1.24		0.230
F22 THAT cl. as adj. compl.	0.37 > 0.10		0.000*
F23 WH clause	0.18 > 0.17		0.000*
F24 Infinitve		12.89 > 12.82	0.490
(H2) PARTICIPIAL FORMS			
F25 Pres. part. cl.		0.35 > 0.13	0.000*
F26 Past part. cl.	0.20 > 0.08		0.000*
F27 Past part. WHIZ del.	0.93 > 0.38		0.001*
F28 Pres. part. WHIZ del.	1.69 > 1.27		0.029
(H3) RELATIVES			
F29 THAT cl. on subj. posit.	0.91 > 0.50		0.000*
F30 THAT cl. on obj. posit.	0.41 > 0.15		0.000*
F31 WH cl. on subj. posit.	1.41 > 0.68		0.011
F32 WH cl. on obj. posit.	0.27 > 0.08		0.000*
F33 Pied pip. const.	0.13 > 0.00		0.406
F34 Sentence rel.	0.33 > 0.15		0.000*
(H4) ADVERBIAL CLAUSES			
F35 Causative subo.	0.58 > 0.41		0.000*
F36 Concessive subo.	0.62 > 0.60		0.000*
F37 Condit. subo.	4.18 > 2.71		0.152
F38 Other adv. subo.	1.28 > 0.76		0.003*
(I) ADJECTIVES AND ADVERBS			
(I1) PREPOSITIONAL PHRASES			
F39 Preposition		76.61 > 73.73	0.363
(I2) ADJECTIVES AND ADVERBS			
F40 Attrib. adj.	30.54 > 21.14		0.000*
F41 Pred. adj.	3.39 > 2.51		0.878
F42 Adverb		38.10 > 37.45	0.777

Table 5.1
Significant Differences of 67 Features Found from CMC Texts (continued)

FEATURES	CA > CS	CS > CA	P Value
(J) LEXICAL SPECIFICITY			
F43 Type/token ratio		0.56 > 0.55	0.588
F44 Word length	4.83 > 4.82		0.668
(K) LEXICAL CLASSES			
F45 Conjunct	2.58 > 1.15		0.810
F46 Downtoner	1.79 > 1.11		0.066
F47 Gen. hedge		1.07 > 0.55	0.000*
F48 Amplifier	2.66 > 1.04		0.828
F49 Gen. emphatic		7.85 > 6.99	0.129
F50 Discourse part.		1.15 > 0.07	0.000*
F51 Demonstrative	4.56 > 3.37		0.560
(L) MODALS			
F52 Possib. modal	7.37 > 6.75		0.975
F53 Necess. modal	1.92 > 1.46		0.028
F54 Predic. modal	7.41 > 5.46		0.490
(M) SPECIALISED VERB CLASSES			
F55 Public verb		44.78 > 6.63	0.000*
F56 Private verb		17.11 > 15.26	0.088
F57 Suasive V.	5.08 > 3.85		0.719
F58 SEEM, APPEAR	1.12 > 0.80		0.000*
(N) REDUCED FORMS AND STRUCTURES			
F59 Contraction		17.51 > 11.95	0.001*
F60 THAT del.		3.53 > 2.30	0.002*
F61 Final prep.	0.75 > 0.53		0.000*
F62 Split inf.		0.14 > 0.13	0.000*
F63 Split auxiliary	3.65 > 1.92		0.080
(O) COORDINATION			
F64 Ph. coord.	4.53 > 2.38		0.044
F65 Non-ph. coord.	1.55 > 1.03		0.001*

Table 5.1
Significant Differences of 67 Features Found from CMC Texts (continued)

FEATURES	CA > CS	CS > CA	P Value
(P) NEGATION			
F66 Synthetic neg.		0.72 > 0.67	0.000*
F67 Analytic neg.		7.86 > 7.50	0.234
Total cases	17 cases significant	12 cases significant	29 cases significant

Remarks: CA: CMC Asynchronous; CS: CMC Synchronous.
*: Significant difference reaches $\alpha = 0.005$.

5.1.3 Significant differences between the CA and CS texts

With the level of significance set at α=0.005, significant differences were found in 29 of the 67 features. As shown in Table 5.1, among the 67 features, 17 were found to be relatively higher in frequency in asynchronous CMC (CA) texts, and 12 features were found to be relatively higher in frequency in synchronous CMC (CS) texts. There are 38 features that do not show significant differences in frequency between the CA and CS texts. The 17 features on which CA texts have a significantly higher frequency are nominalizations (F14), *BY* passives (F18), *THAT* clause as adjective complements (F22), *WH* clauses (F23), past participial clauses (F26), past participial WHIZ deletions (F27), *THAT* clauses on subject position (F29), *THAT* clauses on object position (F30), *WH* clauses on object position (F32), sentence relatives (F34), causative adverbial subordinators - *because* (F35), concessive adverbial subordinators - *although, though* (F36), other adverbial subordinators (F38), attributive adjectives (F40), *SEEM/APPEAR* (F58), final prepositions (F61), and non-phrasal coordination (F65).

The 12 features in which the CS texts have a significantly higher frequency than in the CA texts include present tense verbs (F03), time adverbials (F05), *DO* as a pro-verb (F12), *WH* questions (F13), present participial clauses (F25), general hedges (F47), discourse particles (F50), public verbs (F55), contractions (F59), *THAT* deletions (F60), split infinitives (F62), and synthetic negation (F66).

A detailed presentation and discussion of these differences are made in the following sections. Other than the difference between CA and CS texts, reference will often be made in different sections to the difference between CMC and non-CMC texts by referring to an earlier piece of research on LOB and London-Lund Corpora (Biber, 1988a).

5.2 Tense and Aspect Markers

This section describes and discusses the result of the tense and aspect markers found as a result of the statistical comparison processed in Section 4.6.3 in Chapter Four. This category includes three features: present tense verbs (F03), past tense verbs (F01), and perfect aspect verbs (F02). Only one feature, i.e. present tense verbs (F03), was found to have significant difference between the asynchronous and synchronous CMC texts.

5.2.1 Findings

The statistical findings on the three features in the tense and aspect markers, taken from the more comprehensive Table 5.1 of Section 5.1, are depicted in the Table 5.2, below. Before describing the findings, some examples of these features are listed here to show how they appeared in the CMC texts. Basically, they present actions of the present time, past time or with a continuing result.

Past tense verbs (F01):
*A friend just **suggested** checking out http://www.thinker.org.* (ntcho-014)
*Oh, you **took** your car to school ? Why **was** that?* (tesln-107)
*Every presentation on the Internet or WWW **was** at double the room's capacity.* (rmoo-0402)

Perfect aspect verbs (F02):
*Don't press "enter" till you **'ve typed** in the whole command.* (ntcho-001)
***Has** anyone **tried** this out?* (ntcho-028)
*Barrym **has tried** mud dweller but always returns to raw telnet.* (rmoo-0504)

Present tense verbs (F03):
*What **is** MBU-L?* (ntcho-002)
*I **advocate** a language awareness approach which **explores** all linguistic information in an active and conscious manner.* (tesln-006)
*Everyone **wants** to know about this stuff.* (rmoo-0402)

From Table 5.2, we can see that only present tense verbs (F03) exhibit significant differences between the two temporalities of CMC (α=0.005). The present tense verb occurs significantly more often in the CS texts than in CA (102.62>42.14). This can be seen more easily in the illustration in Figure 5.1, which shows that present tense verbs were used twice as often in the synchronous as in the asynchronous CMC texts.

Table 5.2
Comparison Between Sync. and Asyn. CMC Texts on *Tense and Aspect Markers*

FEATURES	CA > CS	CS > CA	P Value
(A) TENSE AND ASPECT MARKERS			
F01 Past tense V.		11.55 > 10.93	0.111
F02 Perf. asp. V.	5.30 > 4.03		0.632
F03 Pres. tense V.		102.62 > 42.14	0.000*
Total cases	0 case significant	1 case significant	

Remarks: CA: CMC Asynchronous; CS: CMC Synchronous.
*: Significant difference reaches α = 0.005.

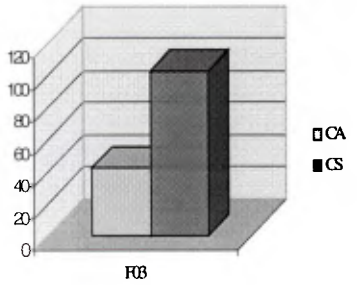

Figure 5.1 Comparison of Feature from F03
F03 *Present Tense Verb* (*Everyone **wants** to know about this stuff.* -rmoo-0402)

5.2.2 Discussion of tense and aspect markers

Several studies claim that tense and aspect markers have the function of distinguishing between speech and written styles (Biber, 1986; Grabe, 1987; Ochs, 1979). Present tense verbs deal with topics and actions occurring in the immediate context of interaction, or the current state of affairs. They are also used in academic writings to remove any temporal focus in order to present the information. Previous studies report that present tense forms serve as a marker of immediate situations (Biber, 1986; Grabe, 1987), and are also frequently used in unplanned speech (Ochs, 1979).

Past tense verbs and perfect aspect verbs describe past events, or past events with continuing results. They have been associated with descriptive texts and with certain types of academic writing and are often involved with the third person pronoun (Feigenbaum, 1978). They focus on the temporal sequence and also serve as primary markers of narrative. Schiffrin (1981) finds that within narratives, the past tense form and the historical present are

used alternatively. Other studies (Biber, 1986; Marckworth & Baker, 1974) report that past tense forms co-occur frequently with perfect aspect forms as markers of narrative.

From Table 5.2 and Figure 5.1, we can see that the synchronous CMC (CS) texts have a significantly higher frequency of use of the present tense (F03) than the asynchronous CMC (CA) texts, which is similar to what happens in traditional face-to-face conversation that usually deals with topics and actions of immediate relevance. The texts, therefore, are of an immediate nature and may involve more use of the present tense verbs. This phenomenon may also indicate that the CS texts focused more, when compared with the CA texts, on the information being presented, removing the focus from any temporal sequencing, as in academic writings.

In Biber's study (1988a), present tense markers are found to be a symbol of involved discourse, while past tense and perfect aspect markers are found to be a symbol of narrative style. The findings here suggest that synchronous CMC texts (CS) are clearly more of the involved style than the asynchronous CMC texts (CA), while the two types of CMC texts do not vary much in terms of the narrative style.

We can further examine this by looking at the comparison of the use of present tense in CMC as well as in non-CMC genres, as shown in Figure 5.2.

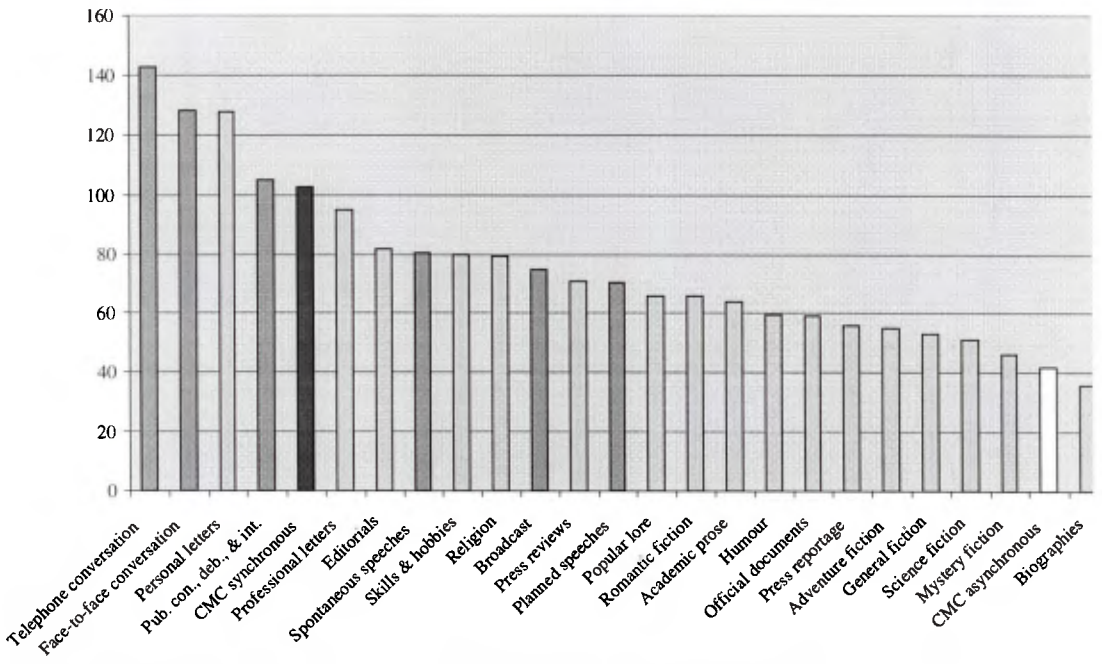

Figure 5.2 Use of *Present Tense Verbs* Between CMC and Non-CMC Genres
F03 *Present Tense Verb* (*Everyone* ***wants*** *to know about this stuff.* -rmoo-0402)

In Figure 5.2, we can see quite clearly that most speech genres generally stand higher in the ranking of frequencies of present tense than do writing genres, with telephone conversation and face-to-face conversation taking up the highest two places. This is in agreement with the general observation that present tense is used more often in speech than in writing.

It is easy to understand that the CS text stands for a more involved style, as participants are engaged in a simultaneous exchange of information. Their common interest is focused on the present, the here and now. On the other hand, the finding that the CS and CA texts do not vary in past tense and perfect aspect suggests that the narrative discourse is not a distinctive nature in the two types of CMC texts. This is different from traditional texts, where narration is often carried out through asynchronous communication in the written and printed forms. The lack of narration in the CA texts may be related to the short length of the postings in the discussion lists. It has been observed that people do not have the patience to read long articles on the computer (Chou, 1996, 1998). When people exchange ideas and information through the computer network, they also have no interest in giving a full account of something, with its full background. They simply point out some of their main ideas or make an announcement. The world of computer-mediated communication is just like our modern fast-food culture, in which speed is of utmost concern.

For past tense verbs (F01) and perfect aspect verbs (F02), it appears that neither shows a significant difference in terms of frequency of occurrence between the CA and CS texts. As both of these features have been associated with descriptive texts, or narrative texts focusing on the temporal sequence, they do not seem characteristic of either the synchronous or the asynchronous CMC texts.

In summary, the feature present tense (F03) is the only one of the three tense and aspect markers being examined that does demonstrate the significant characteristics of the CMC texts. While CS texts have a significantly higher frequency of use of the present tense than the CA texts, the two types of texts do not vary significantly in the use of past tense and perfect aspect.

5.2.3 Summary of tense and aspect markers

This section deals with findings related to the category of tense and aspect markers. Among the three features examined in this category, only the feature present tense verbs (F03) showed any significant difference in the CMC synchronous texts. No difference was noted between the CA and CS texts in terms of the frequency of the past tense verbs (F01)

and the perfect aspect verbs (F02) used. This suggests that CA and CS texts vary only in the use of present tense, making CS significantly more of the involved style than CA. However, the two types of CMC texts, CS and CA, do not vary on the other tense and aspect markers, so that the style of narration between the two types of texts is not found to be distinctive. We can speculate that this is connected to the nature of the short length of texts and the efficiency of communication that is emphasised in the CMC environment.

5.3 Place and Time Adverbials

Place and time adverbials include adverbials of place (F04), such as *aboard*, *behind*, *downstairs*, *east*, *far*, *outside*, and adverbials of time (F05), such as *again*, *early*, *immediately*, *late*, *now*, *today*, etc. In this category, only time adverbials demonstrate a significant difference across asynchronous and synchronous CMC texts.

5.3.1 Findings

The statistical findings on the features of the category of place and time adverbials, taken from the overall Table 5.1 of Section 5.1, are depicted in Table 5.3 below. Some examples of these features are listed here to give an idea of how they were used in the CMC texts. Basically, they provide the information of time and place that the writer or "speaker" wants to express or to which they want to refer.

place adverbials (F04):
Hello all, the message ***below*** *was posted to mbu-l. (ntcho-001)*
Briefly, this approach is centered ***around*** *skill related modules rather than the traditional leveled classes. (tesln-010)*
You can get back to the room by going ***east*** *on the path. (hmoo-0424)*
We need to jump ***ahead****. (rmoo-0402)*

time adverbials (F05):
We'll collect the lists in the cafe and post the final list on the Netoric web site later. (ntcho-001)
Recently *it hit me that if I had enough guest speakers whom I could regularly count on, I wouldn't need a note-taking textbook! (tesln-022)*
Are you familiar with the commands Gregor mentioned ***earlier****? (hmoo-0417)*
Now *we seem to be closing in on the hottest thing in town. (rmoo-0402)*

The statistical findings, as depicted in Table 5.3 below, show that time adverbials (F05) demonstrate significantly higher frequency in synchronous CMC (CS) texts than in asynchronous CMC (CA) texts ($\alpha = 0.005$), while place adverbials (F04) show no significant

difference between synchronous and asynchronous CMC texts.

5.3.2 Discussion of place and time adverbials

Place and time adverbials can be used both for text-internal referents and for reference to places and times outside of the text itself. According to Biber (1988a, p. 110), these forms often serve as deictics that "can only be understood by reference to an external physical and temporal situation."

Table 5.3
Comparison Between Sync. and Asyn. CMC Texts on *Place and Time Adverbials*

FEATURES	CA > CS	CS > CA	P Value
(B) PLACE AND TIME ADVERBIALS			
F04 Place adv.	2.30 > 2.02		0.262
F05 Time adv.		4.00 > 2.45	0.000*
Total cases	0 case significant	1 case significant	

Remarks: CA: CMC Asynchronous; CS: CMC Synchronous.
*: Significant difference reaches $\alpha = 0.005$.

For both text-internal references (e.g. *see above*, *discussed later*) and text-external references (e.g. *over there*, *tomorrow*) in a conversation or a text, the addressee must infer to where and when the adverbials refer to. Both place and time adverbials, therefore, have the characteristics of the situation-dependent reference.

Unlike place adverbials, which show no effect by the temporalities of CMC in this study, time adverbials display significant differences between the asynchronous and synchronous temporalities in CMC texts. Specifically, time adverbials occur more frequently in synchronous than in asynchronous CMC texts (4.00 vs. 2.45), as shown in Table 5.3 above. This feature is typical of the synchronous temporality and this phenomenon can be further observed in Figure 5.3.

That the CS texts display a higher use of time adverbials suggests that synchronous communication is the condition in which time references are more often required than in asynchronous contexts. In synchronous CMC, participants often use time adverbials for both narrative concerns (past time) and non-narrative concerns (present time). This finding is in line with the thinking that synchronous communication is more situation-dependent in terms of time reference (Biber, 1988a, 1988b; Elsness, 1983).

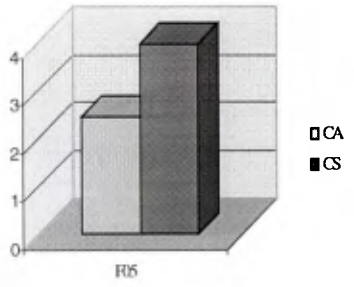

Figure 5.3 Comparison of Feature from F05
F05 *Time Adverbials* (***Now*** *we seem to be closing in on the hottest thing in town.* - rmoo-0402)

We can also look at the comparison of the use of time adverbials (F05) in CMC as well as in non-CMC genres, as shown in Figure 5.4. In Figure 5.4, we can see quite clearly that many speech genres, e.g. broadcast, telephone conversation and planned speeches, generally stand higher in the ranking of frequencies of time adverbials than writing genres. This is in agreement with the general observation that the present tense is used more often in speech than in writing.

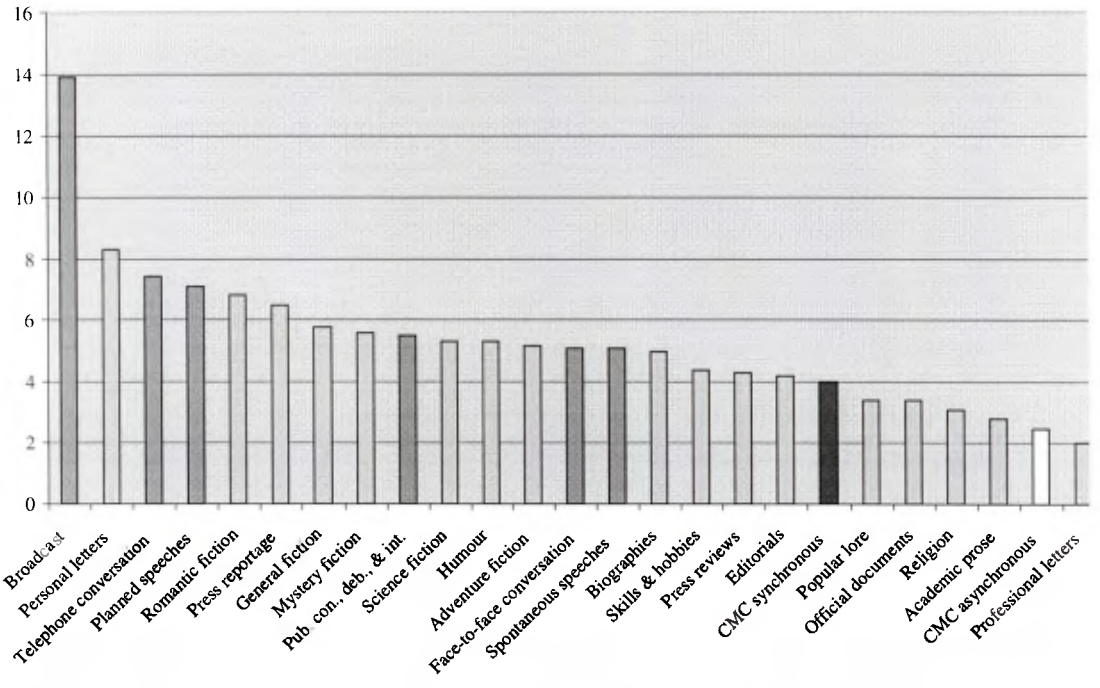

Figure 5.4 Use of *Time Adverbials* Between CMC and Non-CMC Genres
F05 *Time Adverbials* (***Now*** *we seem to be closing in on the hottest thing in town.* - rmoo-0402)

On the other hand, no significant difference in the use of place adverbials is found between asynchronous and synchronous CMC texts. It may be that the CMC participants are

mostly working in the cyberspace and the concept of space is different from the physical space that people are used to when they are engaged in the interaction of traditional written or spoken modes.

5.3.3 Summary of place and time adverbials

From the findings in Section 5.3.1 above, the feature time adverbials (F05) shows a significant difference across the asynchronous and synchronous CMC contexts. They occur significantly more often in the CS text than in the CA text. However, the feature place adverbials (F04) does not show any significant difference between the two temporalities.

The finding that time adverbials occur significantly often in synchronous communication certainly suggests the relative importance of the time referents in the synchronous temporality. On the other hand, it is noteworthy that such relative difference is not found for place adverbials.

As both time and place adverbials are believed to stand for the discourse style of situation-dependent reference, as opposed to explicit reference in traditional written and spoken texts (Biber, 1988a, 1988b; Elsness, 1983), the finding here on the time adverbials is certainly not new. However, the finding that the place adverbials do not demonstrate a significantly higher frequency in synchronous CMC texts seems to suggest a different way of thinking of the space relationship when CMC participants interact in the cyberspace.

5.4 Pronouns and Pro-verbs

The category of pronouns and pro-verbs can be further divided into three sub-categories: personal pronouns, impersonal pronouns, and pro-verbs. Personal pronouns include first person pronouns (F06), second person pronouns (F07), and third person pronouns (F08). The personal pronoun includes all cases such as *I, me, my, he, him, his, she, her,* etc., and numbers, such as *I, we, he, she, they,* etc., of the *pronouns* and the reflective forms *-self* and *-selves.*

The impersonal pronouns include the pronoun *IT* (F09), the demonstrative pronouns like *this, that, these, those* (F10), and indefinite pronouns (F11) such as *anybody, someone,* etc. The final item in this category, the pro-verbs, are words such as *do, does, did,* when they are not used as auxiliaries or in questions.

In the present study, only the pro-verb *DO* demonstrates significant differences across the CA and CS texts. Its details are discussed further in the following sections.

5.4.1 Findings

The statistical findings on the seven features in the pronouns and pro-verbs category, taken from the overall Table 5.1 of Section 5.1, are depicted in the Table 5.4 below. Some examples of these features are listed here to show how they are used in the CMC text.

Personal pronouns:

First person pronouns (F06):
I *shall post a summary within 10 days.* (tesln-009)
My *skills feel like zero.* (hmoo-0417)
We *now pause briefly for introductions.* (rmoo-0402)

Second person pronouns (F07):
Thanks for ***your*** *help.* (tesln-003)
Could ***you*** *introduce yourselves then?* (hmoo-0417)
Create ***your*** *own audience.* (rmoo-0507)

Third person pronoun (F08):
Are ***they*** *formalised with written agreements?* (ntcho-007)
Yoshi takes out a towel and dries ***himself.*** (hmoo-0424)
Gregs wonders why ***his*** *banana bread isn't cooking right tonight.* (rmoo-0402)

Impersonal pronouns:

pronoun *IT* (F09):
It *is not a large list and the traffic is not high.* (ntcho-003)
I used ***it*** *in college writing class with American students (and they liked* ***it****), but* ***it****'s not very difficult and could be used with late intermediate/advanced non-native speakers.* (tesln-005)
It*'ll be on the web page, won't* ***it****?* (hmoo-1217)

demonstrative pronouns (F10):
This *is my new e-mail address.* (ntcho-008)
Hope ***this*** *helps.* (tesln-001)
Which session was ***that****, Douglas?* (rmoo-0402)

indefinite pronouns (F11):
Does ***anybody*** *know of another net talk type application which can record?* (ntcho-038)
Is it because there are ***none****?* (tesln-252)
The look command, without ***anything*** *after it, shows you the room you're in.* (hmoo-0424)

Pro-verbs:
DO as pro-verb (F12):
The other reason is that a good proportion of the students who visit schMOOze on a regular basis ***do*** *so on their own time. (ntcho-043)*
But when I asked her, she just smiled and said she didn't know why she ***did*** *that!* (tesln-090)
I ***did*** *myself on my PC at work and at home.* (hmoo-0417)

Table 5.4
Comparison Between Sync. and Asyn. CMC Texts on *Pronouns and Pro-verbs*

FEATURES	CA > CS	CS > CA	P Value
(C) PRONOUNS AND PRO-VERBS			
(C1) PERSONAL PRONOUNS			
F06 1st per. pron.	30.10 > 26.00		0.561
F07 2nd per. pron.		11.09 > 8.79	0.010
F08 3rd per. pron.		13.85 > 12.42	0.121
(C2) IMPERSONAL PRONOUNS			
F09 Pron. IT	8.58 > 8.17		0.312
F10 Demonst. pron.		4.47 > 4.46	0.388
F11 Indefinite pron.	2.41 > 2.40		0.148
(C3) PRO-VERBS			
F12 DO as pro-V.		0.65 > 0.33	0.000*
Total cases	0 case significant	1 case significant	

Remarks: CA: CMC Asynchronous; CS: CMC Synchronous.
*: Significant difference reaches α = 0.005.

From Table 5.4, it can be seen that, among all seven features, only pro-verb *DO* (F12) exhibits a significant difference between the temporalities of communication, as its P value reaches 0.005, but all other features demonstrate no significant difference between the two temporalities.

Table 5.4 shows that among all the pronouns and pro-verbs, only the pro-verb *DO* shows a significant difference between the synchronous and asynchronous CMC texts (0.65 vs. 0.33). This suggests that synchronous CMC texts have a stronger tendency than asynchronous texts in using pro-verbs. None of the pronouns examined shows any difference in frequency between the two types of CMC texts to the level of significance. Detailed discussion on *Do* as a pro-verb (F12) and other features will be made in the following sections.

5.4.2 Discussion of pro-verbs

Pronominalisation, i.e. personal pronouns, impersonal pronouns, and pro-verbs, is often used across sentence-boundaries and within them to achieve cohesion (Kress, 1994, p. 86). Pro-verbs are words like *do*, *does*, *did*, when not used as auxiliaries or in questions. They

can substitute an entire clause and reduce the informational density of a text. In Table 5.4, Pro-verb *DO* (F12) is the only feature of pro-verb that shows significant differences between the temporalities of communication. Figure 5.5 demonstrates this clearly. We can see that the pro-verb occurs nearly two times more in the synchronous CMC text (0.65 vs. 0.33). This indicates that in synchronous CMC, perhaps due to the pressures of time constraints, MOO participants more often just apply the pro-verb *DO*, i.e. words like *do*, *does*, *did*, to substitute an entire phrase or clause, or to reduce the informational density of their texts.

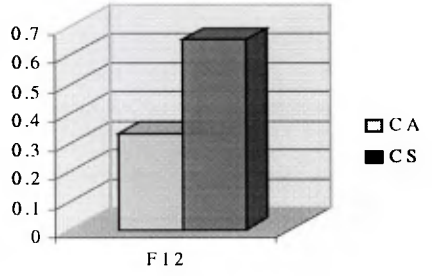

Figure 5.5 Comparison of Feature from F12
F12 *DO as Pro-Verbs* (*I **did** myself on my PC at work and at home.* - hmoo-0417)

We can also look at the comparison of the use of pro-verbs (F12) in CMC as well as in non-CMC genres, as shown in Figure 5.6.

In Figure 5.6, we can see that many speech genres, e.g. face-to-face conversations amd telephone conversations generally stand higher in the ranking of frequencies of pro-verbs than writing genres. Both CMC genres stand quite low in the ranking. This suggests that, though pro-verbs are more of a characteristic for synchronous CMC than for asynchronous CMC, they are used relatively less in either mode of CMC than in most non-CMC settings. This is quite interesting. Pro-verbs are taken as a typical marker for the involved discourse (Biber, 1998a), and they stand for unspecified verbal referents. This causes a "more generalised, uncertain content" (p. 106). A possible reason why they rarely occur in CMC is the limitation on the length of the CMC texts, which makes it impractical to many pro-verbs. In the limited length available, CMC participants may feel the need to be more specific in their choice of words and expressions.

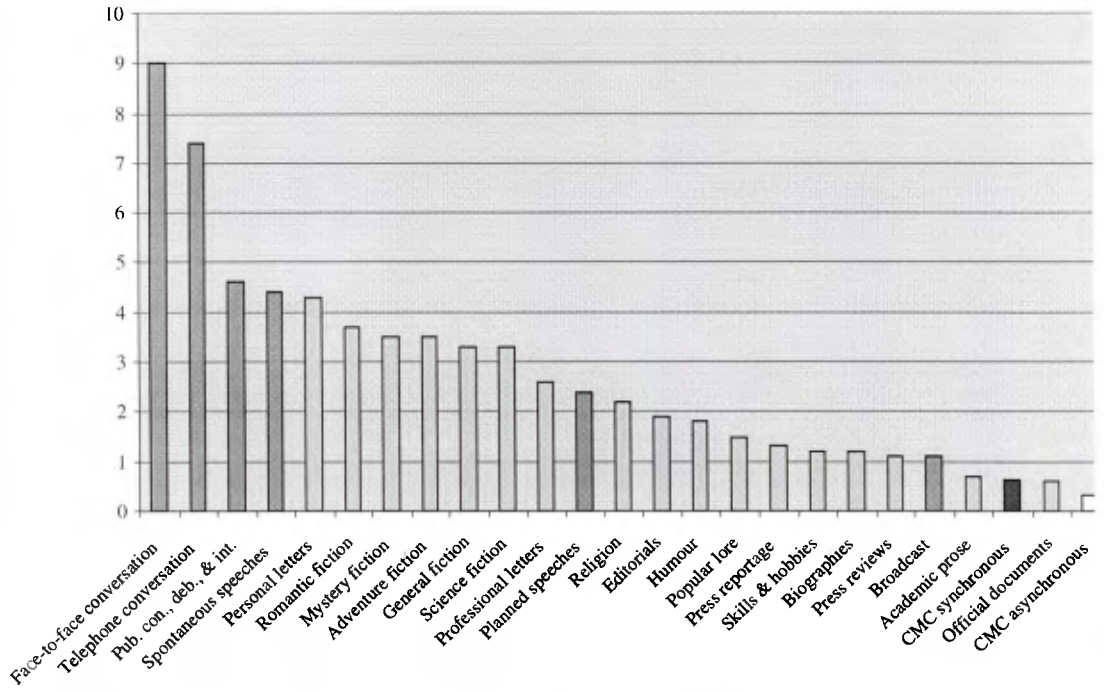

Figure 5.6 Use of *Pro-Verbs* Between CMC and Non-CMC Genres
F12 *DO as Pro-Verbs* (*I **did** myself on my PC at work and at home.* - hmoo-0417)

5.4.3 Discussion of personal and impersonal pronouns

All personal pronouns in pronominalisation are important to serve the cohesion within the text. For example, first and second person pronouns have often been treated as markers of ego-involvement in a text (Biber, 1988a, p. 225). The subjects of cognitive verbs, such as *arrange, describe, choose, analyse, argue,* are found to be usually first person pronouns, indicating that discussion of mental processes is a personal matter often associated with high ego-involvement (Weber, 1985). Moreover, both first and second person pronoun refer directly to the addresser and addressee in a text, and are found to be used frequently in highly interactive discourse (Biber, 1988a). They indicate an interpersonal focus and a generally involved style (Chafe, 1982, 1985).

However, third person pronouns are different. Biber (1988a) argues that third person pronouns co-occur frequently with past tense and perfect aspect forms to serve as a marker of narrative, reported (versus immediate) styles. Actually, narrative discourse depends heavily on third person pronouns to present a description involving specific animate participants.

In discussing the theme of a message, Halliday (1989, p. 72) also says that first and second pronouns are often used in the spoken language, but writing has a more strongly third person orientation.

In this study, though every personal pronoun feature had different frequencies used in the CMC text, none of them reached the significant difference. In other words, all personal pronouns were used across the CA and CS texts without large variation.

For the impersonal pronouns, including the pronoun *IT* (F09), the demonstrative pronouns like *this, that, these, those* (F10), and indefinite pronouns such as *anybody, someone*, etc. (F11), they mark a reduced surface form, a generalised or uncertain presentation of information and a generally fragmented production of text. They usually stand for unspecified nominal referents.

The pronoun *IT* (F09) is the most generalised pronoun, since it can be substituted for nouns, phrases, or whole clauses. Many researchers believe that the pronoun *IT* marks a relatively implicit lexical content and limited amount of information due to strict time constraints in a typical spoken situation (Kroch & Hindle, 1982; Chafe & Danielewicz, 1987; Biber, 1988a). The demonstrative pronouns (F10), like *this, that, these, those,* are found to occur more often in informal, unplanned types of discourse. Their function, as Biber (1988a, p. 113) believes, is to integrate information into idea units. Those demonstrative pronouns that are used without nominal referents are considered errors in speech due to faster production and the lack of editing (Chafe, 1985; Chafe & Danielewicz, 1987). The indefinite pronouns, like *anybody, someone*, etc. (F11) are also markers of generalised pronominal reference.

In the present study, none of the impersonal pronouns presented significant differences across the asynchronous and synchronous CMC texts. The fact that various types of pronouns had been found to be of different characteristics in signifying the different levels of involvement of the participants in traditional written and spoken communication, does not show similar situation in this study. It may be that in either synchronous or asynchronous CMC, the participants had relatively similar levels of involvement, so that their use of pronouns shows no great difference.

5.4.4 Summary of pronouns and pro-verbs

Most of the pronouns and proverbs (with the exception of third person pronouns) can be characterised as verbal, interactional, affective, fragmented, reduced in form, and

generalised in content (Biber, 1988a, p. 105). In the present study, however, pronouns do not show significant differences between synchronous and asynchronous CMC texts. It suggests that CMC texts are of relatively similar characteristics in both temporalities. Only one feature, the pro-verb *DO,* shows a significantly higher frequency in synchronous CMC texts. It is speculated to be a technique of synchronous CMC participants to shorten their utterances under the time constraint. It is also found that both types of CMC use pro-verbs quite rarely as opposed to non-CMC genres. This may be a result of the limitation of time and space that forces the CMC users to use more specific verbs.

5.5 Questions

5.5.1 Findings

The *WH*-questions cover the use of interrogatives such as *when*, *what*, *why*, *which*, *how*, *whoever*, etc. in direct questions. They are the devices that enable people to get more information in an interactive context. Some examples of *WH*-questions from the sampled CMC texts are:

WH-questions (F13):
***What** is MBU-L?* (ntcho-002)
*Just out of curiosity, **why** are these two works on the curriculum?* (ntcho-013)
***Where** would I find one?* (tesln-124)
***How** can I offer a student a coherent rule to go by when the very text they're using isn't consistent?* (tesln-0217)
***When** are the refreshments coming?* (hmoo-0402)

The statistical findings in Table 5.5 below show that significant differences exist in the synchronous CMC text found in the present study. Moreover, in Figure 5.7, the comparison is depicted in a chart to give a clearer presentation of the difference of *WH* questions. It can be seen clearly that *WH*-questions occur much more frequently in CS than in CA texts.

Table 5.5
Comparison Between Sync. and Asyn. CMC Texts on *Questions*

FEATURES	CA > CS	CS > CA	P Value
(D) QUESTIONS			
F13 WH question		1.27 > 0.73	0.000*
Total cases	0 case significant	1 case significant	

Remarks: CA: CMC Asynchronous; CS: CMC Synchronous.
*: Significant difference reaches $\alpha = 0.005$.

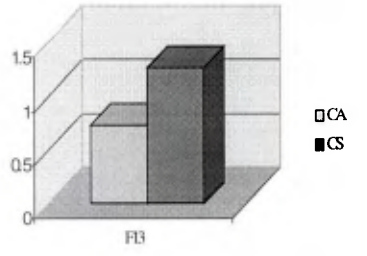

Figure 5.7 Comparison of Feature from F13
F13 *WH-questions* (***What*** *is MBU-L?* -ntcho-002)

5.5.2 Discussion of *WH*-questions

The *WH*-question is believed to be of the style for interactional and involved discourse. They are used primarily in interactive discourse where there is a specific addressee present to answer questions. Moreover, they are used to indicate a concern with interpersonal functions and involvement with the addressee (Marckworth & Baker, 1974; Biber, 1988a). In the present study, the *WH*-question has been found to be of significantly higher frequency in the synchronous than in the asynchronous CMC texts, as illustrated in Figure 5.7.

If we look at Figure 5.8, we can also see the comparison of the use of *WH*-questions (F13) between CMC and non-CMC genres.

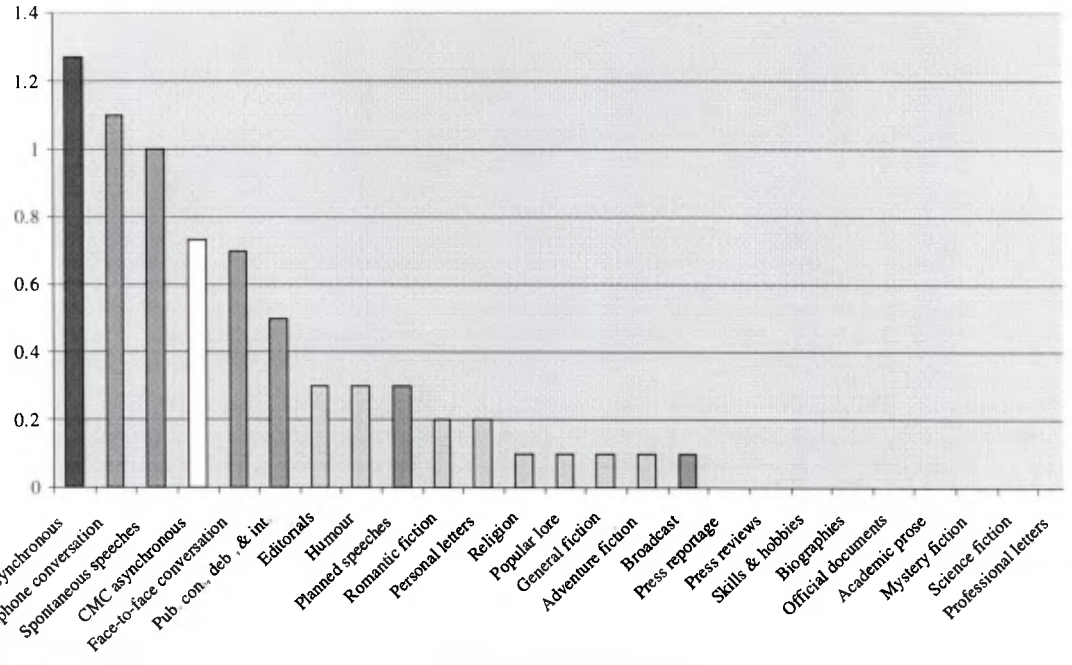

Figure 5.8 Use of *WH-Questions* Between CMC and Non-CMC Genres
F13 *WH-questions* (***What*** *is MBU-L?* -ntcho-002)

In Figure 5.8, we can see that synchronous CMC (CS) has the highest frequency of *WH*-questions. Speech genres like telephone conversations and spontaneous speeches also stand high in the ranking. It is interesting that asynchronous CMC (CA) also stand among the highest in the frequency of *WH*-questions. On the other hand, *WH*-questions never occur in nine writing genres, e.g. press reportage, biographies, academic prose, and personal letters. It is quite clear that, due to the time constraints, CMC users, in either synchronous or asynchronous temporality, use *WH*-questions to raise questions or make requests directly, rather than resorting to euphemism.

5.5.3 Summary of *WH*-questions

Generally speaking, we find that participants in synchronous CMC, more often than in asynchronous CMC, use *WH-questions* as a strategy to enhance interpersonal relationships as well as to get more information. It may be the presence of an immediate addressee that helps foster the context in which direct questions are often raised. It is also interesting to see that people, when under the limitation of time and text length in CMC settings, use the direct question forms more often than in non-CMC settings.

5.6 Nominal Forms

There are three features listed in the category of nominal forms: nominalisations (F14), gerunds (F15), and total other nouns (F16). Nominalisation includes all the cases of the use of suffixes like *-ity*, *-ment*, *-ness*, or *-tion* to form a noun out of a verb or adjective. Gerunds are verbal forms serving nominal functions, such as *going, working, studying,* etc. Total other nouns include all the nouns listed in the dictionary other than those counted as nominalisation and gerunds. In the present study, significant difference between the asynchronous and synchronous CMC texts is found only in nominalisation (F14).

5.6.1 Findings

The statistical findings on the three features in the nominal forms, taken from Table 5.1 of Section 5.1, are depicted in the Table 5.6 below. Some examples of these features are listed below to show how they appeared in the CMC text across asynchronous and

synchronous temporalities. Generally, they are mainly the objects and entities to which people referred in the language context.

Nominalisations (F14):
*Are they formalized with written **agreements**?* (ntcho-007)
*It has a good **explanation** of the four **acquisition** stages of **negation**.* (tesln-001)
*By relevant **action**, do you mean in their community or something like that?* (hmoo-0417)
*You're confusing **randomness** with chaos now.* (rmoo-0402)

Gerunds (F15):
*Please send **greetings** to the ESL students at Oyster.* (ntcho-018)
*Are you addicted to cigarettes but would like to stop **smoking**?* (tesln-082)
*This session's agenda: **Brainstorming**: future topics format for discussions times?* (hmoo-0417)
*He was quite persistent, and went back the next day to finish **typing** in the list.* (rmoo-0702)

Total other nouns (F16):
*If you are interested reading these **reviews**, you may access the **report** on the **WWW** at the following **URL**.* (ntcho-005)
*Our **ESL program** has received an application from a **student** who meets all **entry criteria** and is, therefore, admissible.* (tesln-003)
*My **skills** feel like **zero**.* (hmoo-0417)
*Yeah, the **fringe** is the **centre**.* (rmoo-0402)

From Table 5.6, it can be seen that nominalisation (F14), gerund (F15), and other nouns (F16) all are used more often in the asynchronous CMC (CA) texts than in synchronous CMC (CS) texts, though only nominalisation reaches the level of statistical significance (α=0.005).

Table 5.6
Comparison Between Sync. and Asyn. CMC Texts on *Nominal Forms*

FEATURES	CA > CS	CS > CA	P Value
(E) NOMINAL FORMS			
F14 Nominal.	22.10 > 13.01		0.004*
F15 Gerund	6.38 > 4.15		0.128
F16 Noun	211.00 > 192.10		0.543
Total cases	1 case significant	0 case significant	

Remarks: CA: CMC Asynchronous; CS: CMC Synchronous.
*: Significant difference reaches α = 0.005.

5.6.2 Discussion of nominal forms

Figure 5.9 is a depiction of the comparison between CA and CS on nominalisation (F14), the only set of comparisons in this category where significant difference is found. In

Figure 5.9, it is quite clear that CA and CS texts differ significantly on the frequency of nominalisation.

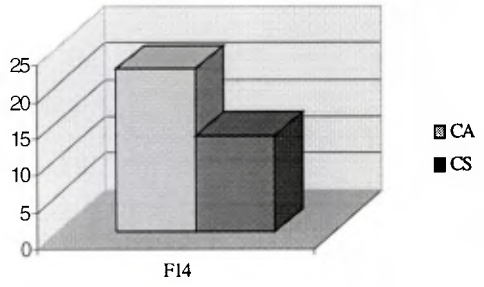

Figure 5.9 Comparison of Feature from F14
F14 *Nominalisations* (*Are they formalized with written* ***agreements****?* -ntcho-007)

The distinction between nominal and verbal styles has been identified as one of the most fundamental distinctions among registers by Wells (1960), and Brown and Fraser (1979). The frequent use of nominal forms is used to indicate an abstract, informational focus and conceptual abstractness (Biber, 1988a, p. 227). The high lexical density achieved by the frequent use of all kinds of nominal forms is a major character of the modern written English (Halliday, 1989).

As the frequent use of nominal forms is considered a characteristic of the rather formal and careful written style, it is not surprising to see that all three types of nominal forms appear relatively more often in asynchronous than in synchronous CMC texts. This seems in line with our general assumption that when writing in the CMC medium, the time constraint off-line is not as high as when writing online. Hence, it enables writers to choose precise noun forms to express their thoughts explicitly. Therefore, the CA texts might be of a more formal register than the CS texts.

A comparison of the use of nominalisations (F14) in CMC and in non-CMC genres is presented in Figure 5.10. Here, we can see that professional letters, official documents, and academic prose are the highest in the ranking of frequencies of nominalisation. Asynchronous CMC (CA) ranks sixth among all the 25 genres and synchronous CMC (CS) ranks 15th, higher than six writing genres and three speech genres. It can be seen that writing genres tend to have higher frequency in the use of nominalisation, but the tendency is not very strong. Besides, both of the CMC genres examined here are relatively higher than non-CMC genres, showing that nominalisation is a feature often applied by CMC users.

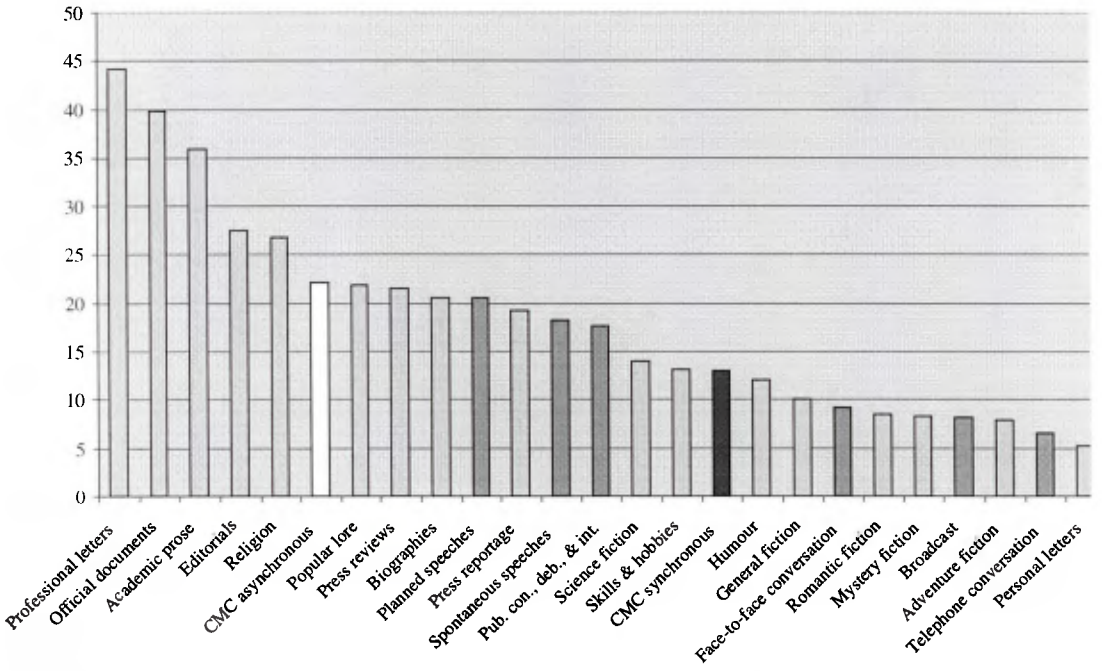

Figure 5.10 Use of *Nominalisations* Between CMC and Non-CMC Genres
F14 *Nominalisations* (*Are they formalized with written* ***agreements****?* -ntcho-007)

The other two features, i.e. the gerunds (F15) and the other nouns (F16), also occur more often in the CA text, although they have not reached the significant level (α=0.0005).

On the one hand, gerunds, the verbal forms serving nominal functions, are closely related to nominalisations in their functions. They have the same form as present participials, but the functions are quite different. On the other hand, the total other nouns (F16) are the primary bearers of referential meaning in a text, and they are used to integrate large amount of information into a text; to present information as concisely and precisely as possible. They are associated with communicative situations that require a high informational focus and provide ample opportunity for careful integration of information and precise lexical choice. As Chafe and Danielewicz (1987) suggest, precise lexical choice is a very difficult production task and is thus rarely accomplished in speech.

The fact that gerunds and other nouns do not show significant enough difference between synchronous and asynchronous CMC texts, whilst the forms of nominalisation do, is quite interesting. As stated earlier, the use of suffixes to result in nominalisation has the function of making information compact by integrating it into fewer words. It is then easy for us to associate the compactness of information with synchronous CMC, which we feel is natural when the time constraint is felt. This is what we have seen in the discussion of pro-

verbs in Section 5.4.3. However, the finding here shows that the asynchronous CMC texts actually have significantly more frequent occurrences of nominalisation than synchronous CMC texts. This seemingly controversial finding leads us to speculate that the formal register associated with nominalisation plays a more important role than the time constraints and need for compactness in determining the relatively more frequent use of nominalisation in asynchronous CMC texts. That is to say, in synchronous CMC texts, nominal forms are not a preferred characteristic, even though it has the effect of compactness. It is not preferred probably because it is too formal in register.

5.6.3 Summary of nominal forms

This section deals with findings related to the category of nominal forms. Among the three features, nominalisations (F14) showed significantly higher frequency in the CA than in the CS text. Gerunds (F15) and other nouns (F16) also occurred more often in the CA text, but had not reached the significant level. The findings in this section largely meet the assumption that nominal forms are characteristic of the asynchronous written style, which carries a higher density of lexical meaning. This finding can be examined in connection with that in Section 5.2 on tense and aspect markers. Besides, the fact that nominalisation is often applied in CMC suggests that CMC users use words of higher lexical density to offset the pressures of time constraints. This can be understood in connection with the discussion on subordination and coordination in Sections 5.9 and 5.16 later.

5.7 Passives

This section presents the results of the passives, with a discussion based on the statistical comparison carried out for the purposes of this study. The passive category includes two features: agentless passives (F17), and *BY*-passives (F18). From the results found in Section 4.6.3 in the previous chapter, only *BY*-passives show a significant difference between the asynchronous and synchronous CMC texts.

5.7.1 Findings

The results from the statistical findings for this category are depicted in the Table 5.7 below, taken from Table 5.1 of Section 5.1. Passives are usually the technique required for

people to switch focus from the agent to the patient of a verb. Here are several examples of the two types of passives taken from the CMC sample texts.

Agentless passives (F17):
ACW-L is a similar list that ***is housed*** *at Texas Tech.* (ntcho-003)
That form ***is posted*** *on my website for TESOL practice teachers and the* URL ***is listed*** *below.* (tesln-019)
I was just pointing out that wp software ***can be networked.*** (hmoo-0402)

BY-passives: (F18)
This week's topic ***was suggested by*** *Claudine Keenan.* (ntcho-022)
Is it worse to ***be done by*** *a native speaker?* (tesln-039)
But terms in one field are radically changed when they ***are used by*** *another field.* (rmoo-0402)

From Table 5.7, it can be seen that the *BY*-passive (F18) has a significantly higher frequency in CA than in CS texts. The agentless passive (F17), though also higher in the CA texts, has not reached the level of significant difference (α=0.005).

Table 5.7
Comparison Between Sync. and Asyn. CMC Texts on *Passives*

FEATURES	CA > CS	CS > CA	P Value
(F) PASSIVES			
F17 Agentles pass.	5.94 > 3.34		0.040
F18 *BY*-passive	0.73 > 0.34		0.000*
Total cases	1 case significant	0 case significant	

Remarks: CA: CMC Asynchronous; CS: CMC Synchronous.
*: Significant difference reaches α = 0.005.

5.7.2 Discussion of passives

Passives have been taken as an important surface marker of the formal written style (Kress, 1994). In passive constructions, the agent of the verb can be shown or dropped altogether, resulting in a static, more abstract presentation of information, thus two types of passives are formed: agentless passives (F17), and *BY*-passives (F18).

Passives are used to give prominence to the patient of the verb, the entity acted upon, which is often an abstract concept rather than a concrete referent. They are often used in procedural discourse, where the agent is pre-supposed and the specific agent of a clause is not important to the discourse purpose. According to Biber (1988a), discourse with very frequent passive constructions is typically abstract, technical in content, and formal in style,

which has a relatively low lexical variety when compared with other types of informational discourse.

From the findings presented in Table 5.7, it can be seen that the two types of passives are both more popular in the asynchronous than in synchronous CMC texts. Between the two, the difference of *BY*-passive (F18) has reached the significant level. It is nearly two times more frequent in the CA than in the CS text. The difference in the *BY*-passive between the CA and CS texts can be examined more clearly when it is individually shown in Figure 5.11.

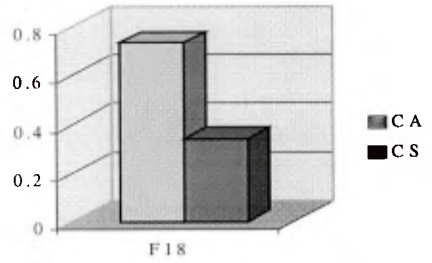

Figure 5.11 Comparison of Feature from F18
F18 *BY-passives* (*Is it worse to* ***be done by*** *a native speaker?* -tesln-039)

As both forms of passive devices are characteristic of the formal register widely used in professional writing on academic subjects, their more frequent occurrence in asynchronous CMC texts seems appropriate. It is often observed that in several cases, such as in the discussion of cases of nominalisation in Section 5.6.2, asynchronous CMC texts tend to be more formal than synchronous CMC texts. Therefore, it is not surprising to find that passives occur more often in asynchronous than in synchronous CMC texts.

The next figure is a comparison of the ranking of all the CMC and non-CMC genres on their frequencies of *BY*-passives. In Figure 5.12, we can see that nine genres standing at the top of the ranking scale for *BY*-passives are all writing genres. Asynchronous CMC texts (CA) stand at 10th, then the speech genres show up, mixed with other writing genres in the lower half of the ranking. The genres that show highest frequency of *BY*-passives are official documents, academic prose, and press reportage. On the other extreme, some writing genres, e.g. science fiction, romantic fiction, and personal letters, and the speech genre of telephone conversation have no *BY*-passives in the sample texts. The two CMC genres, though shown to be significantly different from each other, stand near the middle of the ranks. Basically, the tendency of *BY*-passives is to occur more often in certain types of writing genres, namely those

that are more formal and that deal with serious matters.

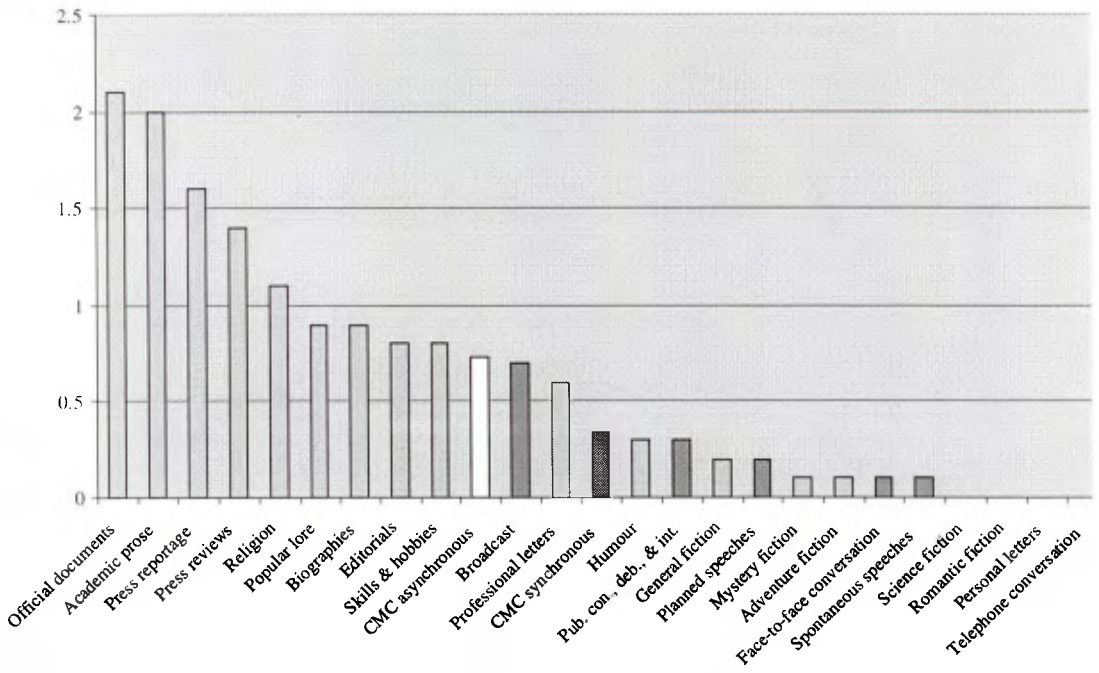

Figure 5.12 Use of *BY-Passives* Between CMC and Non-CMC Genres
F18 *BY-passives* (*Is it worse to* ***be done by*** *a native speaker?* -tesln-039)

5.7.3 Summary of passives

The two passive features both show the tendency to occur more frequently in asynchronous CMC texts. *BY*-passives (F18) show a prominent frequent use in the CA texts, while agentless passive (F17) was also found popular in the asynchronous CMC texts. This is believed to be connected with the more formal register of passive devices that the participants applied in the asynchronous mode of CMC.

5.8 Stative Forms

This section displays the findings of the stative forms proceeded by the test in the previous chapter from Section 4.6.3 Non-Parametric Statistical Test of synchronous vs. asynchronous CMC texts. This category includes two features, namely *BE* as main verb (F19), and existential *THERE* (F20). Though differences were found with the two features, neither of them reached the significant level (α=0.005) across two temporalities of the CMC texts.

5.8.1 Findings and discussion

The two features in this category, *BE* as main verb and existential *THERE*, do not demonstrate significant differences between the two temporalities of CMC texts, from what hava been shown in Table 5.8. Here are some examples of these two features to show how they were used in the CMC texts:

BE as main verb (F19):
The Fine Arts Museums of San Francisco ***is*** *all about art access.* (ntcho-014)
They also state re vocabulary: Lexical differences ***are*** *far more numerous, but many of these* ***are*** *familiar to users of both standards.* (tesln-002)
Plaid Guest, I missed part of your thought. What wheels ***were*** *you referring to?* (rmoo-0402)

Existential *THERE* (F20):
Are ***there*** *any www URLs I could check out for more information?* (ntcho-007)
There *was a lot of loot besides gold bars, and some of it might have been used to rebuild Germany.* (tesln-009)
There *is a classroom here with instructions.* (hmoo-0417)

Table 5.8
Comparison Between Sync. and Asyn. CMC Texts on *Stative Forms*

FEATURES	CA > CS	CS > CA	P Value
(G) STATIVE FORMS			
F19 *BE* main V.		14.96 > 13.10	0.096
F20 Exist. *THERE*	1.54 > 1.01		0.010
Total cases	0 case significant	0 case significant	

Remarks: CA: CMC Asynchronous; CS: CMC Synchronous.
*: Significant difference reaches α = 0.005.

The stative forms, i.e. *BE* as main verb, and existential *THERE*, are considered "stative" because they preclude the presence of an active verb. They are typically used to modify a noun with a predicative expression, instead of integrating the information into the noun phrase itself (Biber, 1988a). Existential *THERE*, in particular, is used to introduce a new entity while adding a minimum of other information (Kroch & Hindle, 1982). Normally, the *BE* as main verb can follow the Existential *THERE*, the pronouns and nouns. It is also formed with a perfect aspect in a passive construction, and with a *WH*-question into a question. All of these constructions serve different purposes, such as to give minimum information, to ask for new information, to add predicative expression to the noun, or to switch the focus to the object of the action instead of the agent. As the spoken style is

believed to have more clauses, but fewer high-content words per clause (Halliday, 1979), it can be partly characterised by the stative verbs, which represent a reduced informational load.

The finding that CA and CS texts do not vary much in terms of the use of stative forms *BE* and existential *THERE* suggests that CMC texts are not different in this aspect.

Though there is no significant difference of the stative forms between the asynchronous and synchronous CMC texts, we may still take a look at the comparison of the ranking of all the CMC and non-CMC genres on their frequencies of stative forms. In Figure 5.13, we can see that genres at the top of the ranks for *BE* as main verb (F19) are mostly speech genres, e.g. telephone conversations, face-to-face conversations. The two CMC genres both stand at the bottom of the ranks, i.e. CS and CA.

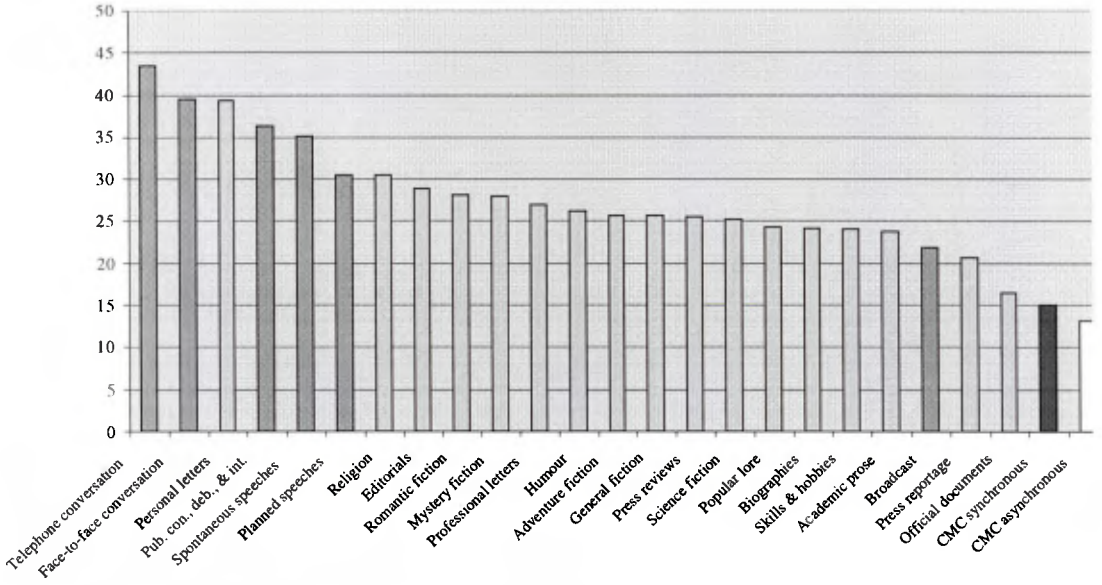

Figure 5.13 Use of *BE as Main Verbs* Between CMC and Non-CMC Genres
F19 *BE as main verb* (*The Fine Arts Museums of San Francisco* ***is*** *all about art access.* -ntcho-014)

In Figure 5.14, the top places of the ranks for existential *THERE* (F19) are also occupied by speech genres, e.g. spontaneous speeches, public conversation, debates, interviews. Most of the writing genres are at the bottom of the ranks. It is interesting to see that synchronous CMC (CS) also stands near the bottom of the ranks.

The ranking of the CMC and non-CMC genres in terms of the frequency of stative forms suggests that these forms are more characteristic of the involved discourse in speech genres. However, it may not be appropriate to jump to the conclusion that CMC genres lack the sense of involvement, as they are low on the ranks of stative forms.

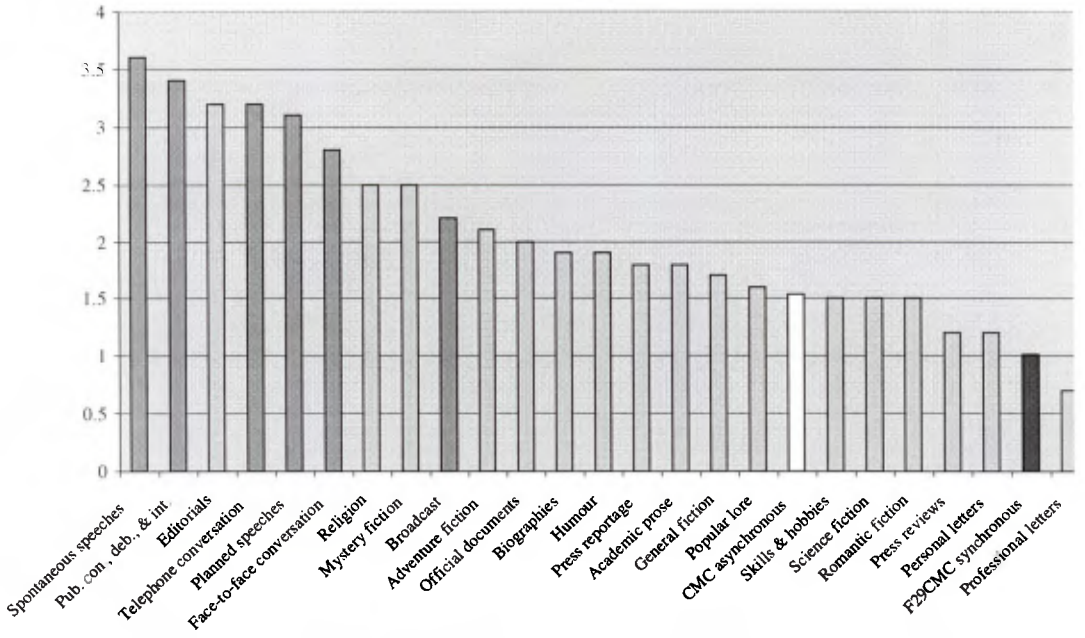

Figure 5.14 Use of *Existential THERE* Between CMC and Non-CMC Genres
F20 *Existential THERE* (***There** is a classroom here with instructions.* -hmoo-0417)

The stative form of *BE* is often associated with a fragmented presentation of information, resulting in low information density. Unlike people involved in speech genres, who feel free to expand the sentence length, thus resulting in lower lexical density, CMC users actually try to increase the information density under the limitation of time constraints. As discussed in other sections in this chapter, i.e. the earlier section that dealt with nominal forms, and the later sections on subordination and coordination, the techniques to which they resort may include the use of more nominal forms (Section 5.6), less subordination (Section 5.9) and coordination (Section 5.16) devices.

5.8.2 Summary of stative forms

Neither of the two features in this category of stative forms shows differences significantly. The *BE* as main verb (F19) seems a more popular feature in CMC texts than the existential *THERE* (F20). However, there is no significant difference between the two temporalities of CMC texts as long as these two features are concerned.

As the use of stative forms is associated with a reduced informational load with "more clauses, but fewer high-content words per clause" (Halliday, 1989), it is usually taken as a symbol of the involved spoken style. The finding that there is no significantly higher

frequency of stative forms in one of the CMC temporalities than the other suggests that CA and CS are not very much different from each other in the aspect of involvement.

5.9 Subordination

Subordination is the focus of this section. A total of 18 features are considered to serve this function of subordination. They are the various types of *THAT* clauses, *WH* clauses, *WHIZ* deletions, participial clauses, sentence relatives, infinitives, and subordinators. They have been sub-categorised by Biber (1988a) as complementation, participial forms, relatives, and adverbial clauses.

Subordination is a concept very often discussed for register comparison. Halliday (1989) argues that subordination is found more often in speech than in writing, with the former characterised by "complex sentence structures with low lexical density." Biber (1986) also finds that many types of subordination devices occur more often in speech than in writing. In contrasting spoken and written styles, Kapoli (1992, p. 312), however, argues that "subordination is most – though not exclusively - relevant and auspicious in written discourse because of the explicitness and specificity of detail required of written course." Kress (1994) also argues "Topics are developed jointly in spoken interaction ... Hence much knowledge may be left implicit, much information may be left unsaid." Thompson (1983, 1984, 1985) emphasises that subordination is not a unified construct and has different functions in different types of discourse. Beaman (1984) finds that there are more finite nominal clauses, such as *THAT*-clauses and *WH*-clauses in speech and more non-finite nominal clauses, such as infinitives and participial clauses, in writing. In Biber's study (1986) as mentioned above, relative clauses and infinitives are found to have a separate distribution from other types of subordination, despite the general tendency of occurring more often in speech. All this suggests that the construct of subordination is a complex one and should be discussed separately.

Many of these features in the present study are found to show a significant difference between CA and CS texts, and most of the comparisons show that these features are significantly more frequent in CA than in CS texts. This study shows that the subordination features are popular in the asynchronous CMC texts. As the types of subordination devices are many, the 18 features in this category are further divided into several small groups based

on its structure nature. As mentioned earlier, these structures are subordination as complementation, as participial forms, as relatives, and as adverbial clauses, thus, the findings of these 18 different features will be presented in these sub-categories in the following sections.

5.9.1 Findings and discussion of complementation

The statistical findings on complementation features are depicted in Table 5.9 below. Some examples of these subordination features from CMC texts are listed here.

THAT *verb complements (F21):*
I have not tried Cool Talk but doubt ***that*** *it is a MOO client.* (ntcho-034)
I didn't know ***that*** *I was in the pool.* (hmoo-0424)

THAT adj. complements (F22):
I was very grateful ***that*** *Julie supported my need for a place to meet my students.* (ntcho-045)
I don't mind the conflation myself, just thought it interesting ***that*** *some comp/rhet people would be such sticklers for sticking to a mathematical concept.* (rmoo-0402)

WH clauses (F23):
Please let me know ***what*** *you think.* (ntcho-051)
I didn't understand ***how*** *you extend the activity up on here though.* (hmoo-1016)

infinitives (F24):
According to L - F and Long, these stages seem ***to*** *occur in all learners despite their L1.* (tesln-001)
Trying ***to*** *avoid Burke if I can.* (rmoo-0402)

Table 5.9
Comparison Between Sync. and Asyn. CMC Texts on *Complementation*

FEATURES	CA > CS	CS > CA	P Value
(H) SUBORDINATION			
(H1) COMPLEMENTATION			
F21 THAT cl. As V. compl.	2.33 > 1.24		0.230
F22 THAT cl. As adj. compl.	0.37 > 0.10		0.000*
F23 WH clause	0.18 > 0.17		0.000*
F24 Infinitve		12.89 > 12.82	0.490
Total cases	2 cases significant	0 case significant	

Remarks: CA: CMC Asynchronous; CS: CMC Synchronous.
*: Significant difference reaches $\alpha = 0.005$.

Complementation is the structure of *THAT* clause, *WH* clause, or *to* infinitive serving as a complement to the verb or adjective in the same sentence. Features examined under this

category included *THAT* clause as verb complement (F21), *THAT* clause as adjective complement (F22), *WH* clause (F23), and *to* infinitive (F24). Among these four features, *THAT* clause as adjective complement (F22) and *WH* clause (F23), reveal significant differences between synchronous (CS) and asynchronous (CA) texts. In both cases, the complementation features appear more often in CA than in CS texts. This is further depicted in Figure 5.15 below.

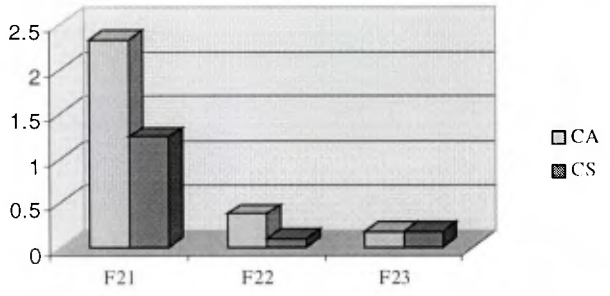

Figure 5.15 Comparison of Features from F21 to F23
F21 *THAT Verb Complements (I didn't know* ***that*** *I was in the pool.* -hmoo-0424)
F22 *THAT adj. complements (I was very grateful* ***that*** *Julie supported my need for a place to meet my students. -ntcho-045)*
F23 *WH Clauses (Please let me know* ***what*** *you think.* -ntcho-051)

Most of the features in this sub-category show the tendency of occurring more often in CA than in CS texts. This is in contradiction with the earlier findings that complementation constructs usually occur more often with interactive features often found in speech (Biber, 1986; Finegan & Biber, 1986b). If we compare the distribution of the frequencies of complementation constructs, especially F22 and F23, in CMC texts with those found in non-CMC genres, we can also find more evidence for this phenomenon.

As can be seen from the Figures 5.16 and 5.17, the speech genres tend to be of higher frequency, though they are mostly scattered among the writing genres in the relative standings of the frequencies of complementation constructs. In Figure 5.16, *THAT* clause as adjective complement (F22) occur highest in planned speeches. In Figure 5.17, *WH* clause (F23) occurs most often in face-to-face conversation. In both figures, however, synchronous CMC texts stand as nearly the lowest in terms of the frequency of these constructs. This marks a very clear difference between non-CMC speech and synchronous interaction on CMC.

When the high frequency of the complementation devices are perceived by Halliday as characteristics of spoken language as affected by the time constraints and "little opportunity

for careful production or revision" (Halliday, 1978), it is quite interesting to find here that asynchronous CMC texts have more of those devices than synchronous CMC texts, as the former is supposed to be associated with more flexible use of time.

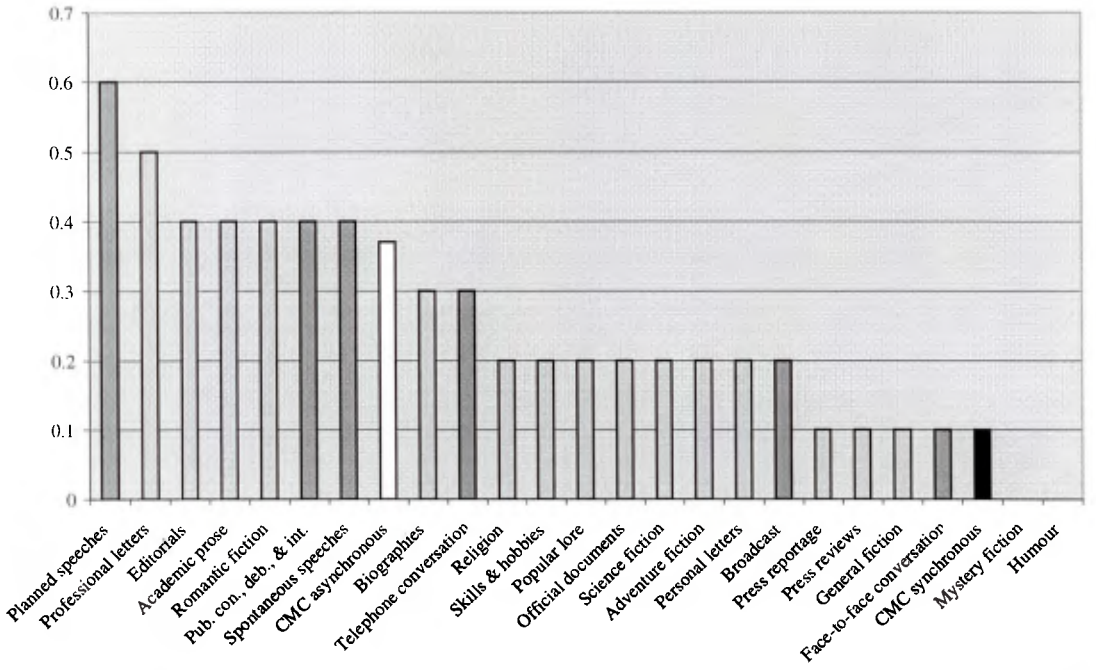

Figure 5.16 Use of *THAT Clauses as Adjective Complement* Between CMC and Non-CMC Genres
F22 *THAT adj. complements (I was very grateful* ***that*** *Julie supported my need for a place to meet my students. -ntcho-045)*

A possible explanation of this phenomenon may still be in the time constraints. It may be that as time constraints are more clearly sensed by the CMC participants, they have to resort to more careful choice of words than those in non-CMC speaking contexts. This can be compared with the discussion earlier in Section 5.6 Nominal Forms and later in Section 5.11 Lexical Specificity, where CMC texts show a relatively higher frequency of nominalisation (F14) and word length (F44) than most speech genres. These features are supposed to result in higher lexical density. The finding in the present sub-section seems to suggest that, under the strict time constraints, synchronous CMC users, especially those of higher educational background in the samples collected for this study, i.e. English teachers, tend to achieve higher lexical density through other means rather than apply complementation devices.

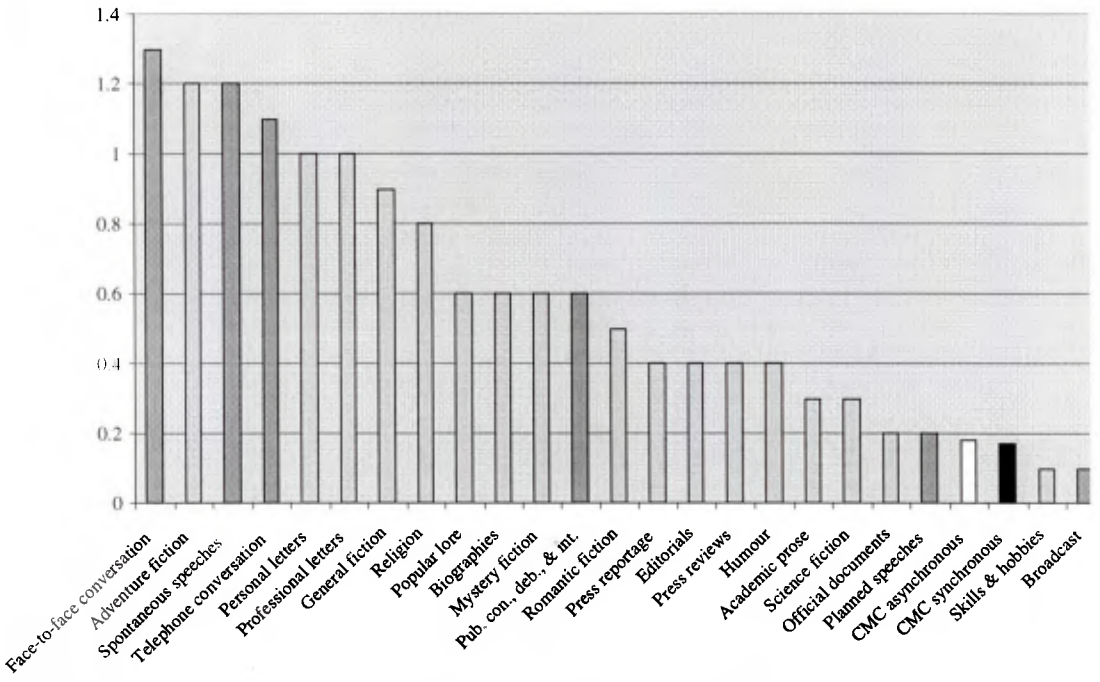

Figure 5.17 Use of *WH Clauses* Between CMC and Non-CMC Genres
F23 *WH Clauses* (*Please let me know* ***what*** *you think.* -ntcho-051)

5.9.2 Findings and discussion of participial forms

Subordination devices examined under this sub-category include present participial clauses (F25), past participial clauses (F26), past participial WHIZ deletion (F27), and present participial WHIZ deletion (F28). WHIZ deletion is the deletion of construction of *which is/are/was/were.* Examples of these five features are listed below.

present participial clauses (F25)*:*
Depending *on the group, this could be a pretty free and open-ended activity.* (tesln-093)
We introduce more and more computer technology into our classes, ***broadening*** *our possibilities and* ***extending*** *the territory into which a writing class can venture.* (rmoo-0910)

past participial clauses (F26):
Compared *to one-level classes, these mixed classes are very hard.* (tesln-171)
What are some of your interests, ***related*** *to NETEACH?* (hmoo-0417)

past participial *WHIZ* deletions (F27):
I offer hands-on workshops to groups then offer individual advising/assistance based on teachers' needs. (hmoo-0516)
It's important to assess the work done in the moo, I think. (rmoo-0402)

present participial *WHIZ* deletions (F28):
The problem of Arabs learning phrasal verbs is a bit difficult. (tesln-013)
I just got a msg from someone looking for other panelists. (rmoo-0402)

In Table 5.10, it can be seen that, among the four types of participial forms, present participial clauses (F25) occur significantly more often in synchronous than in asynchronous CMC texts. Past participial clauses (F26) and past participial *WHIZ* deletions (F27) occur significantly more often in asynchronous than in synchronous CMC texts. Present participial *WHIZ* deletions (F28) also occur more often in asynchronous CMC texts, but the difference has not reached the significance level that we set (α =0.005). Thes comparisons are further depicted in Figure 5.18.

Table 5.10
Comparison Between Sync. and Asyn. CMC Texts on *Participial Forms*

FEATURES	CA > CS	CS > CA	P Value
(H) SUBORDINATION			
(H2) PARTICIPIAL FORMS			
F25 Pres. part. cl.		0.35 > 0.13	0.000*
F26 Past part. cl.	0.20 > 0.08		0.000*
F27 Past part. WHIZ del.	0.93 > 0.38		0.001*
F28 Pres. part. WHIZ del.	1.69 > 1.27		0.029
Total cases	2 cases significant	1 case significant	

Remarks: CA: CMC Asynchronous; CS: CMC Synchronous.
*: Significant difference reaches α = 0.005.

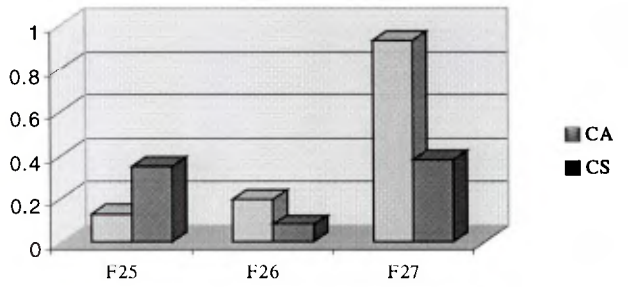

Figure 5.18 Comparison of Features from F25 to F27
F25 *Present Participial Clauses* (***Depending*** *on the group, this could be a pretty free and open-ended activity.* -tesln-093)
F26 *Past Participial Clauses (What are some of your interests, **related** to NETEACH?"* -hmoo-0417)
F27 *Past Participial WHIZ Deletions* (*It's important to assess the work done in the moo, I think.* - rmoo-0402)

That present participial clauses (F25) occur significantly more often in synchronous than in asynchronous CMC texts is quite out of expectation. This feature is found by Biber

(1988a) to characterise the discourse of narration, as present participial clauses have been earlier found by Thompson (1983) to be used to create vivid images in depictive discourse. Moreover, narration is a type of discourse usually associated with writing genres. This can be further discussed by referring to the means of frequencies of this feature in CMC and non-CMC genres, as shown in Figure 5.19.

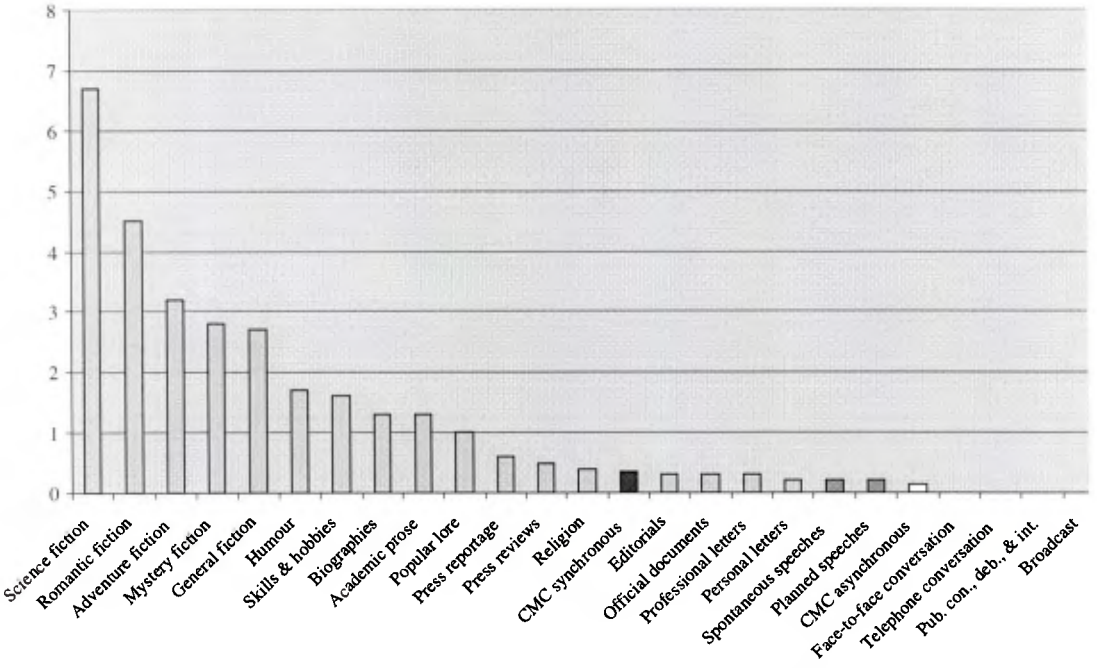

Figure 5.19 Use of *Present Participial Clauses* Between CMC and Non-CMC Genres
F25 *Present Participial Clauses* (***Depending*** *on the group, this could be a pretty free and open-ended activity.* -tesln-093)

In the above figure, Figure 5.19, it can be seen that present participial clauses (F25) usually occur in writing genres and rarely occur in speech genres. In four of the six speech genres examined, present participial clauses never occur. While this feature demonstrates the typical characteristic of asynchronous writing genres in non-CMC, it is interesting that it occurs more in the synchronous than in the asynchronous CMC texts. It seems then that CMC users, when in synchronous temporality, tend to adopt a narrative style more than in asynchronous temporality. This is a very peculiar finding on the subordination features as it is the only one of the 18 subordination features that shows significantly higher frequency in synchronous CMC. However, there is no ready explanation yet for this phenomenon.

In this sub-category of subordination features, past participial clauses (F26) and past participial *WHIZ* deletions (F27) occur significantly more often in asynchronous than in synchronous CMC texts. This is in line with most other subordination features. In Figures

5.20 and 5.21, we can see that writing genres tend to have a clearer tendency of using more of these subordination features. This is more so in Figure 5.20, where none of the speech genres use any past participial clause construct.

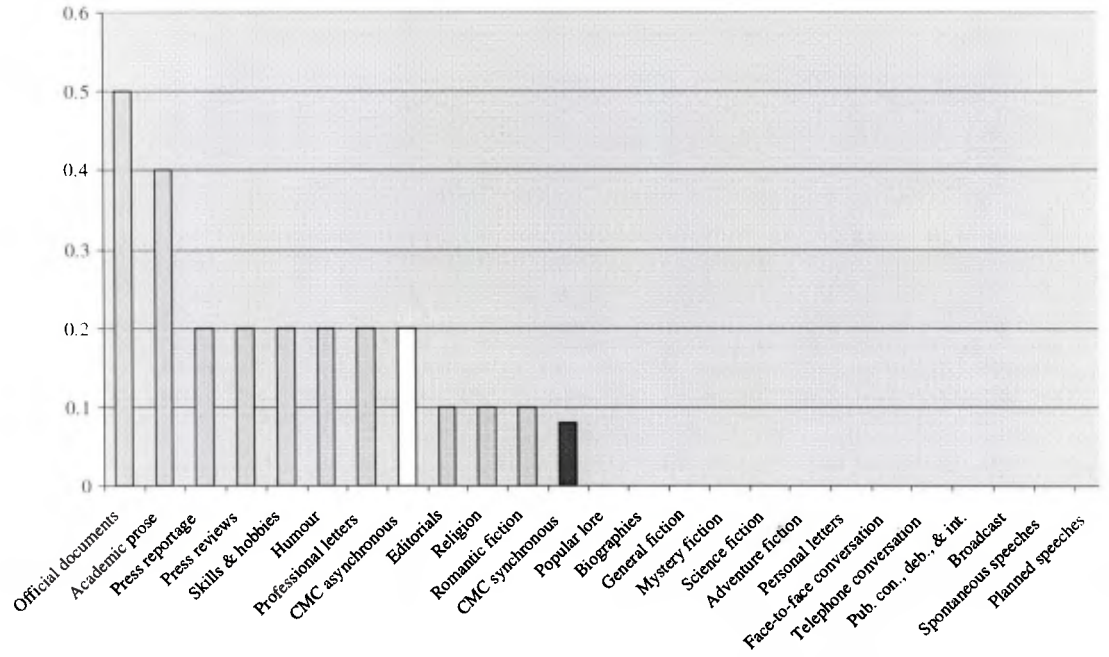

Figure 5.20 Use of *Past Participial Clauses* Between CMC and Non-CMC Genres
F26 *Past Participial Clauses (What are some of your interests,* ***related*** *to NETEACH?* -hmoo-0417)

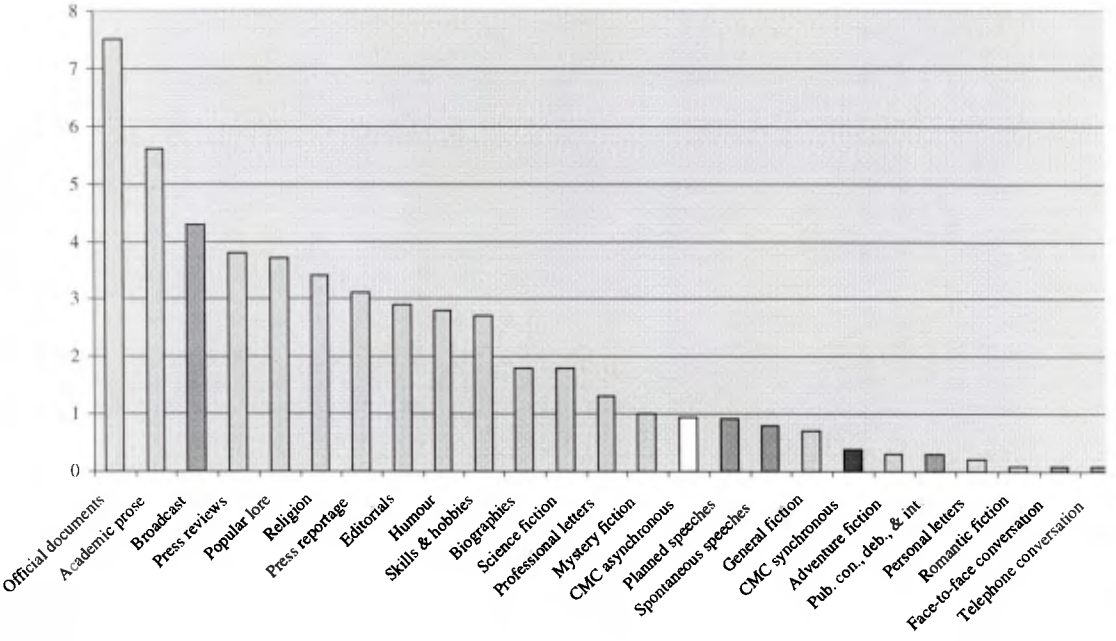

Figure 5.21 Use of *Past Participial WHIZ Deletions* Between CMC and Non-CMC Genres
F27 *Past Participial WHIZ Deletions (It's important to assess the work done in the moo, I think.* -rmoo-0402)

Both Past Participial Clauses (F26) and Past Participial Clauses WHIZ Deletion (F27) are found by Biber (1988a) to characterise abstract discourse, like the use of agentless passives. Both of these features share the nature of being passive, reducing the emphasis on the agent. Passive constructs are often found in procedural discourse, of abstract and technical content and more formal style. It is not surprising that these two features occur more often in asynchronous than in synchronous CMC texts, just as they occur more often in writing than in speech.

5.9.3 Findings and discussion of relative clauses

Relative clauses are an important type of subordination. In this sub-category, there are six features: *THAT* relative clauses on subject position (F29), *THAT* relative clauses on object position (F30), *WH* relative clauses on subject position (F31), *WH* relative clauses on object position (F32), *WH* relative clauses pied pipes (F33), and sentence relatives (F34). Examples of these features are listed below.

THAT relative clauses on subject position (F29):
*I will summarise responses **that** get sent to me directly.* (ntcho-035)
*I think that there are some interesting connections **that** can be made between composing process theory and complex systems theory.* (rmoo-0402)

THAT relative clauses on object position (F30):
*There is a Computers and Writing Reading List in the Tuesday Cafe **that** you can write your contributions on. (ntcho-001)*
*Those URLs are on a note **that** you can read by entering READ SAMPLES. (rmoo-0507)*

WH relative clauses on subject position (F31):
*The person **who** sent me this message has agreed to let me forward it to this list.* (ntcho-021)
*But are * we * the ones **who** determine the definition or is it someone else?* (rmoo-0514)

WH relative clauses on object position (F32):
*Recently it hit me that if I had enough guest speakers **whom** I could regularly count on, I wouldn't need a note-taking textbook!* (tesln-022)
*We're kind of lone rangers here... we're not connected to the U **whose** net access we use.* (hmoo-0718)

WH relative clauses pied pipes (F33):
*I am desperately trying to find GOOD material on CD-ROM **with which** Spanish speaking students can learn French from Beginners to Advanced Levels.* (ntcho-015)
*I'm currently involved in a large email writing training project **in which** I'm working to improve the email writing skills of Unocal's Thai employees.* (ntcho-067)

sentence relatives (F34):
*In my experience the audio quality was pretty insufficient for language learning purposes, **which** leaves the chat and whiteboard.* (ntcho-029)
*Then I shall just publish it on the Web (with my MA thesis, **which** if anyone really wants to read*

it is at http://localsonly.wilmington.net/~eymand/thesis.html. (rmoo-0402)

In Table 5.11 below, we can see that all the six features in this sub-category occur more often in asynchronous (CA) than in synchronous (CS) texts. Among them, the difference in *THAT* relative clauses on subject position (F29), *THAT* relative clauses on object position (F30), *WH* relative clauses on object position (F32), and sentence relatives (F34) has reached the significance level (α =0.005). These four features with significant difference are further depicted in Figure 5.22.

Table 5.11
Comparison Between Sync. and Asyn. CMC Texts on *Relatives*

FEATURES	CA > CS	CS > CA	P Value
(H) SUBORDINATION			
(H3) RELATIVES			
F29 THAT cl. on subj. posit.	0.91 > 0.50		0.000*
F30 THAT cl. on obj. posit.	0.41 > 0.15		0.000*
F31 WH cl. on subj. posit.	1.41 > 0.68		0.011
F32 WH cl. on obj. posit.	0.27 > 0.08		0.000*
F33 Pied pip. Const.	0.13 > 0.00		0.406
F34 Sentence rel.	0.33 > 0.15		0.000*
Total cases	4 cases significant	0 case significant	

Remarks: CA: CMC Asynchronous; CS: CMC Synchronous.
*: Significant difference reaches α = 0.005.

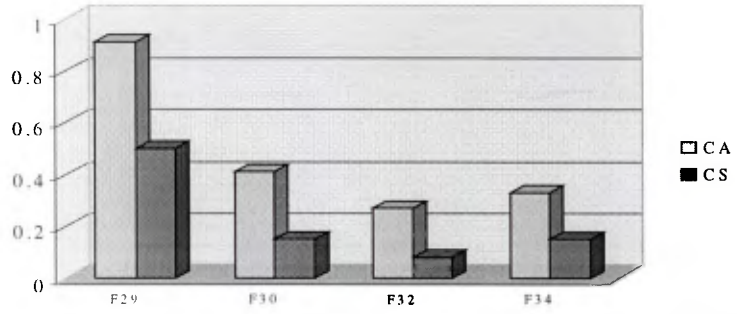

Figure 5.22 Comparison of Features from F29, F30, F32 and F34

F29 *THAT* Rel. Cl. on Sub. Posit. (*I will summarise responses **that** get sent to me directly.* -ntcho-035)

F30 *THAT Relative Clauses on Object Position* (*Those URLs are on a note **that** you can read by entering READ SAMPLES.* -rmoo-0507)

F32 *WH Relative Clauses on Object Position* (*We're kind of lone rangers here... we're not connected to the U **whose** net access we use.* -hmoo-0718)

F34 *Sentence Rel.* (*In my experience the audio quality was pretty insufficient for language learning purposes, **which** leaves the chat and whiteboard.* -ntcho-029)

The relative structure involved in the four features in the CA texts, F29, F30, F32, and F34, is usede to achieve explicitness, specificity, and formality. "Relative clauses ... have been said to be characteristic of written discourse and are assumed to characterise formal language" (Kapoli, 1992, p. 300). As they are found to be more frequent in CA than in CS texts, we can assume that asynchronous CMC texts are more of the written style than synchronous CMC texts.

The findings in relative clauses as subordination suggest that subordination is a style largely associated with asynchronous CMC texts. There is no exception to this general rule.

However, when we compare the frequencies of these features in CMC texts with those in non-CMC genres as shown in Figure 5.23 through 5.26, we can see something interesting. That is, for most of these features, writing genres do not always stand high in the ranking of frequencies, as what is generally suggested. With perhaps the exception of the case of sentence relatives (F34), there does not seem to be any dichotomous distinction between writing and speech genres in their relative standings on the ranking of the frequencies of the features used. This suggests that relative structures are much more too complex to serve as a distinctive feature between the writing and speech genres, though in this study, it does show higher frequency in asynchronous than in synchronous CMC texts.

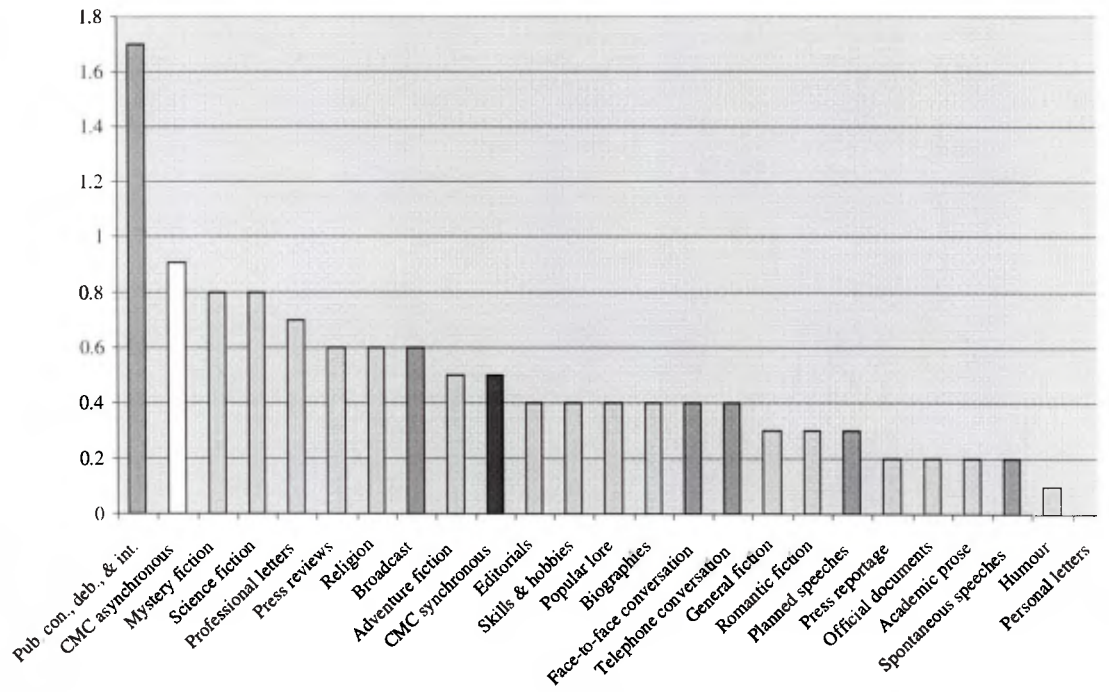

Figure 5.23 Use of *THAT Relative Cl. on Subject Position* Between CMC and Non-CMC Genres
F29 *THAT* Rel. Cl. on Sub. Posit. (*I will summarise responses* ***that*** *get sent to me directly.* -ntcho-035)

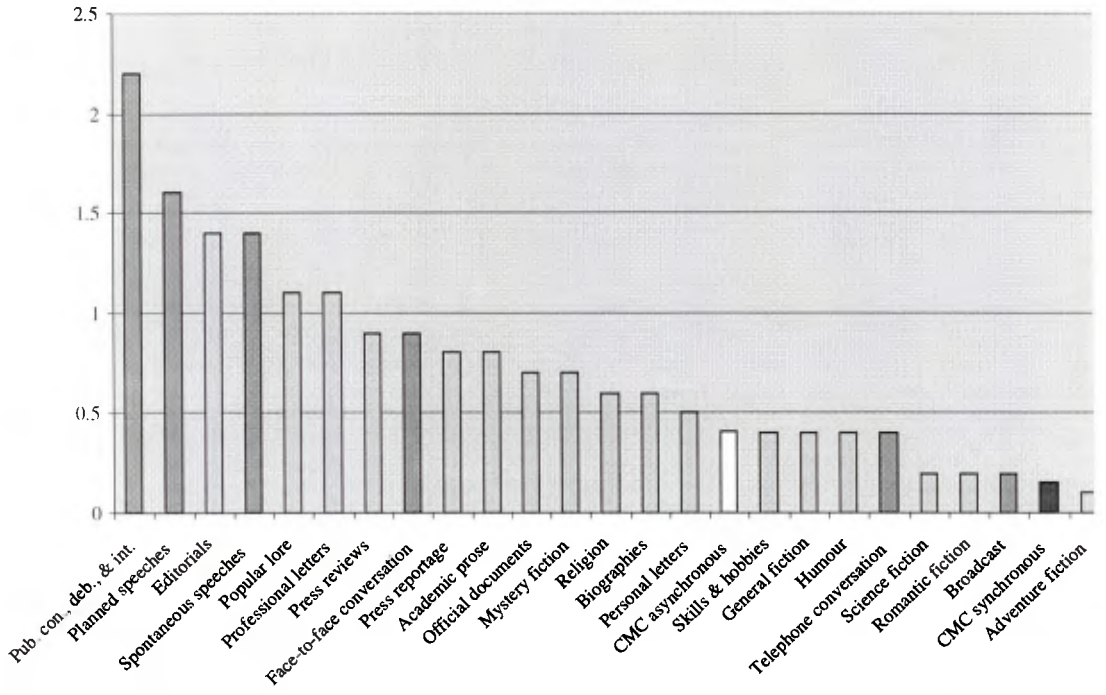

Figure 5.24 Use of *THAT Relative Cl. on Object Position* Between CMC and Non-CMC Genres
F30 *THAT Relative Clauses on Object Position* (*Those URLs are on a note* ***that*** *you can read by entering READ SAMPLES.* -rmoo-0507)

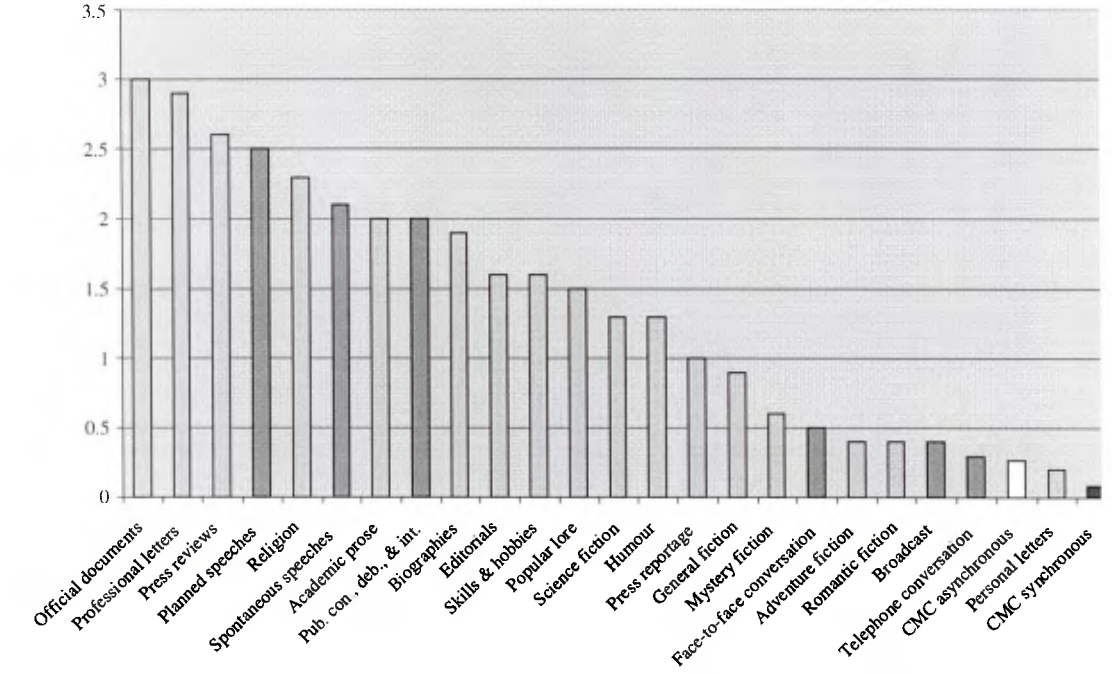

Figure 5.25 Use of *WH Relative Cl. on Object Position* Between CMC and Non-CMC Genres
F32 *WH Relative Clauses on Object Position* (*We're kind of lone rangers here... we're not connected to the U* ***whose*** *net access we use.* -hmoo-0718)

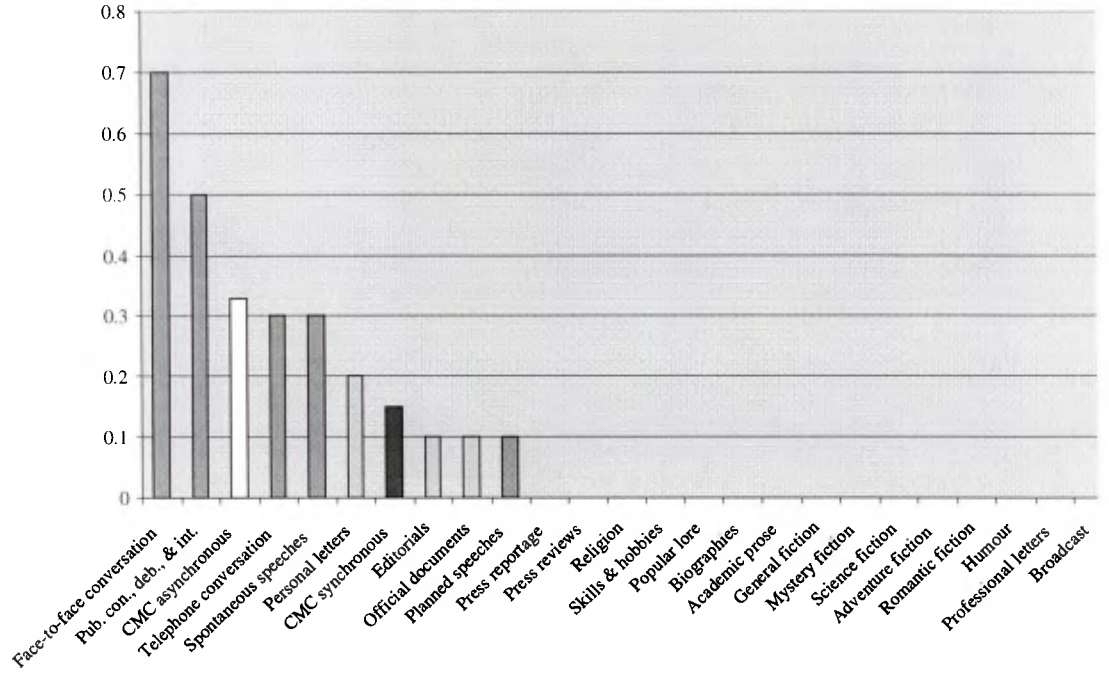

Figure 5.26 Use of *Sentence Relatives* Between CMC and Non-CMC Genres

F34 *Sentence Rel.* (*In my experience the audio quality was pretty insufficient for language learning purposes, **which** leaves the chat and whiteboard.* -ntcho-029)

5.9.4 Findings and discussion of adverbial clauses

There are four types of adverbial clauses examined in this sub-category, including causative adverbial subordinators (F35), concessive adverbial subordinators (F36), conditional adverbial subordinators (F37), and other adverbial subordinators (F38). Among them, F35, F36, and F38 show significantly higher frequency in asynchronous (CA) than in synchronous CMC (CS) texts. Following are the example sentences and the table, Table 5.12, of the comparison of the means of these features.

Causative adverbial subordinators (F35):

*We teach it in our program **because** no one else does.* (hmoo-0424)

*I have a story about that... got so pissed at my son's homework last night I said, go get your father to help you **because** I think this is too stupid to be conscionable.* (rmoo-0402)

Concessive adverbial subordinators (F36):

*Check it out, **although** I'm not always in favour of ms.* (ntcho-029)

***Although** I've not used a moo in class (hope to very very soon), I would imagine that one problem would be some students wouldn't write much (say to an open-ended question).* (hmoo-1016)

Conditional adverbial subordinators (F37):

*Again, **if** you accidentally enter a mess, just start over – we'll erase the mistakes later.* (ntcho-001)

***If** we finish in the next ten minutes, I can finish logging.* (hmoo-0471)

Other adverbial subordinators (F38):
It is just a practice we have used ***since*** *I became coordinator.* (tesln-019)
Unfortunately... I'm still at home... and have to leave for school soon... I will reconnect ***as soon as*** *I can.* (hmoo-0417)

Table 5.12
Comparison Between Sync. and Asyn. CMC Texts on *Adverbial Clauses*

FEATURES	CA > CS	CS > CA	P Value
(H) SUBORDINATION			
(H4) ADVERBIAL CLAUSES			
F35 Causative subo.	0.58 > 0.41		0.000*
F36 Concessive subo.	0.62 > 0.60		0.000*
F37 Condit. subo.	4.18 > 2.71		0.152
F38 Other adv. subo.	1.28 > 0.76		0.003*
Total cases	3 cases significant	0 case significant	

Remarks: CA: CMC Asynchronous; CS: CMC Synchronous.
*: Significant difference reaches α = 0.005.

In the following figure, Figure 5.27, the significant differences found between the features in CA and CS texts are depicted in graphic forms. Three features, namely causative adverbial subordinators (F35), concessive adverbial subordinators (F36), and other adverbial subordinators (F38) occur significantly more often in asynchronous CMC (CA) than in synchronous CMC (CS) texts.

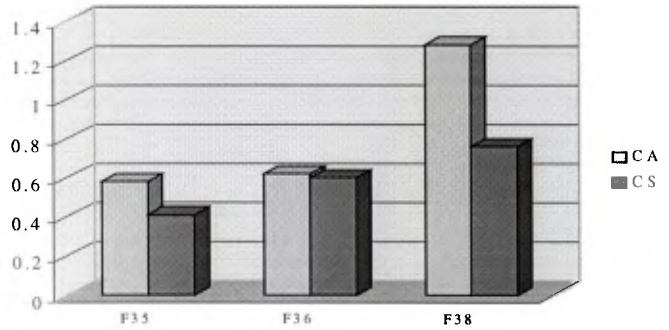

Figure 5.27 Comparison of Features from F35, F36 and F38
F35 *Causative Adverbial Subordinators* (*We teach it in our program* ***because*** *no one else does.* - hmoo-0424)
F36 *Conc. Adv. Sub.* (*Check it out,* ***although*** *I'm not always in favour of ms.* -ntcho-029)
F38 *Other Adverbial Subordinators* (*It is just a practice we have used* ***since*** *I became coordinator.* - tesln-019)

Many researchers found more frequent use of the causative adverbial subordinator *because* (F35) used more often in speech than in writing (Beaman, 1984; Tottie, 1986). In Biber's study (1988a), *because* is also found to be an important marker of an involved discourse typically found in speech. The frequent use of subordinator *because* in speech is assumed to be associated with the "expression of information under real-time production constraints, when there is little opportunity to elaborate through precise lexical choice" (Biber, 1988a, p. 107). In the present study, the causative adverbial subordinator *because* is found to occur more often in asynchronous CMC (CA) than in synchronous CMC (CS) texts. I would speculate that time constraints also account for the difference here. While CMC participants write in the asynchronous mode, they are already under considerable time constraints to make a relatively loose presentation of information, rather than carefully choosing precise words. In synchronous CMC, the time constraints are even harder and maybe the participants simply break the otherwise complex sentences into simple ones to effect rapid exchange of information. This may be the reason why the causative adverbial subordinator *because* is found less in CS than in CA texts.

Two other types of adverbial subordinators in this sub-category also show significant difference between the asynchronous and the synchronous CMC texts. They are concessive adverbial subordinators (F36, CA:0.62>CS:0.60) and other adverbial subordinators (F38, CA:1.68>CS:0.76). The differences found may be for the same reason as that for of the causative adverbial subordinator *because* (F35).

The fact that most of the adverbial coordinators examined here occur more in the asynchronous than in the synchronous temporality is also in line with the argument of Kress that the development of a written sentence "demands the development of planning, deciding which is to be the main clause (the main idea?) and how subsidiary clauses (and ideas) are to be integrated with the main clause" (1994, p. 37). In the asynchronous temporality of communication, the writer has relatively more time for the planning and preparation in producing each sentence. In synchronous CMC, on the other hand, time constraints may be so tense that the user cannot afford the time to connect his clauses into coordinated sentences.

We can also compare the frequencies of the adverbial subordinators in CMC texts with those in non-CMC genres, as shown in Figure 5.28 through 5.30.

From the above figures, Figure 5.28 through 5.30, we can see that asynchronous CMC (CA) tend to use subordinators more often than synchronous CMC (CS). However, these subordinators occur differently in non-CMC genres. The causative subordinator *because*

(F35) occurs most frequently in speech genres, but concessive subordinators *though, although* (F36) and other subordinators (F38) mainly occur more often in writing genres. Moreover, there is not a general pattern showing whether subordinators stand higher in speech or writing genres.

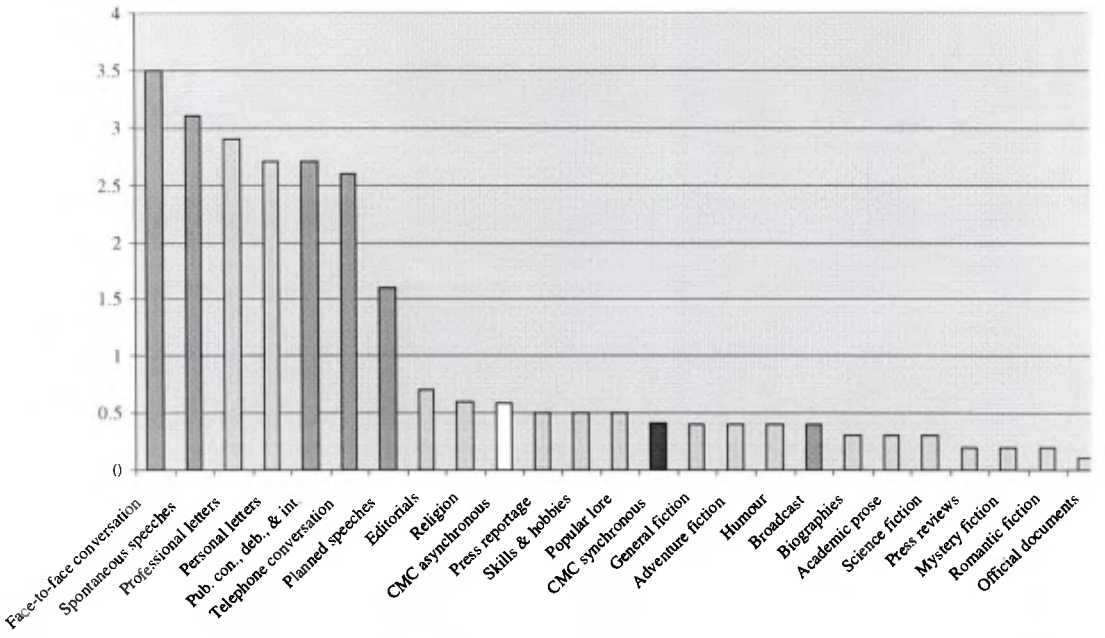

Figure 5.28 Use of *Causative Adv. Subordinators* Between CMC and Non-CMC Genres
F35 *Causative Adverbial Subordinators* (*We teach it in our program* ***because*** *no one else does.* -hmoo-0424)

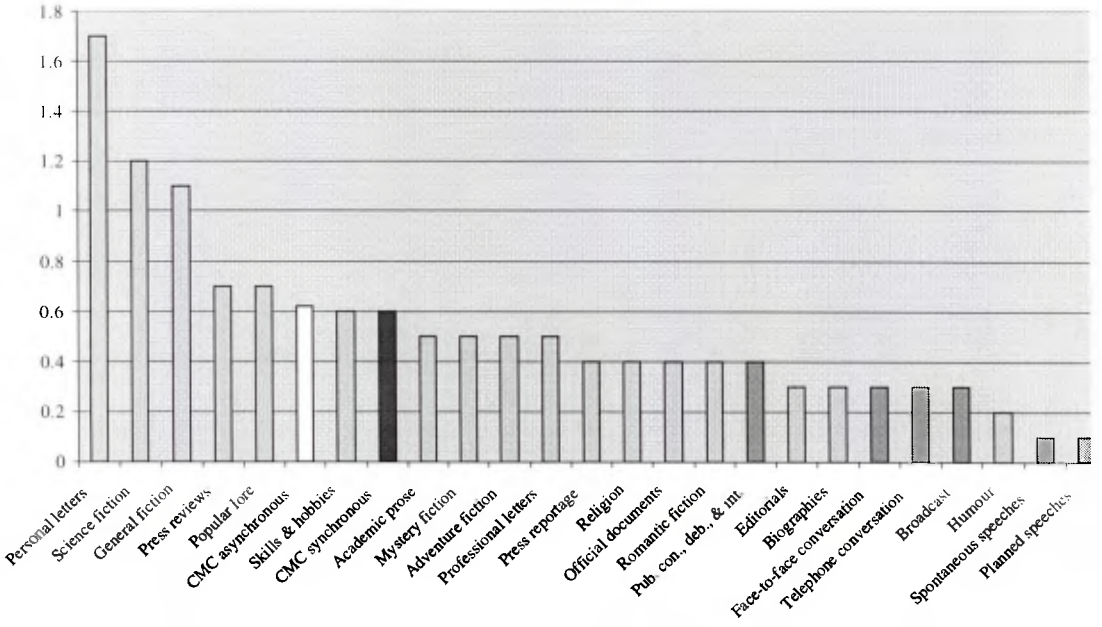

Figure 5.29 Use of *Concessive Adv. Subordinators* Between CMC and Non-CMC Genres
F36 *Conc. Adv. Sub.* (*Check it out,* ***although*** *I'm not always in favour of ms.* -ntcho-029)

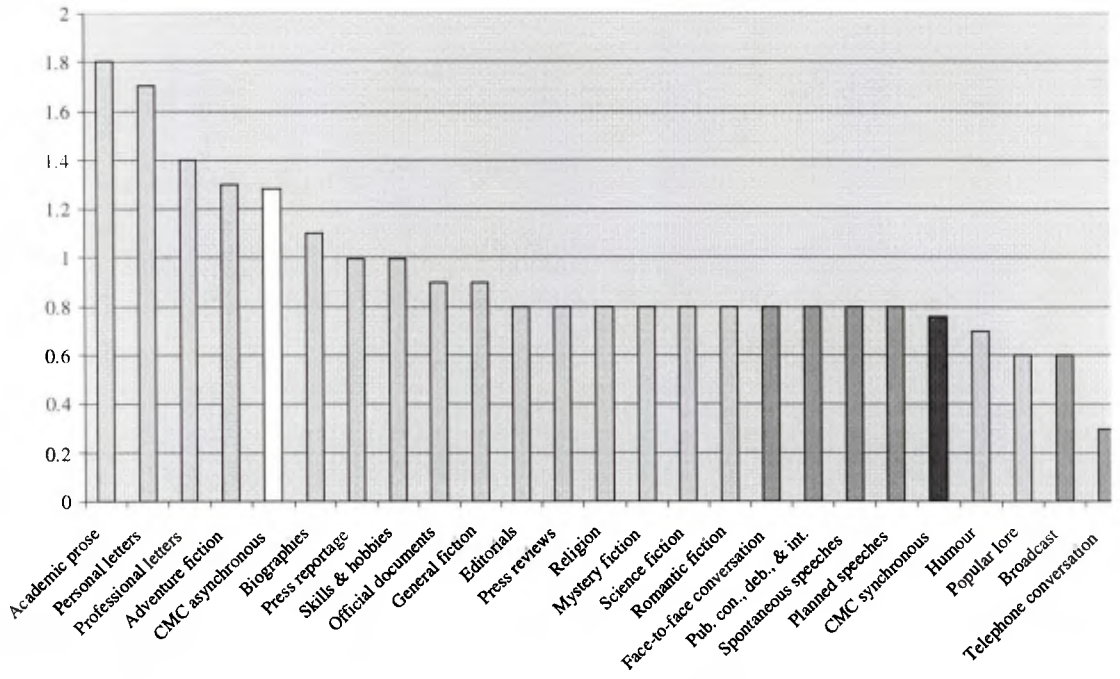

Figure 5.30 Use of *Other Adv. Subordinators* Between CMC and Non-CMC Genres
F38 *Other Adverbial Subordinators* (*It is just a practice we have used* ***since*** *I became coordinator.* - tesln-019)

5.9.5 Summary of subordination

A total of 18 features of subordination are discussed in this section under four categories: complementation, participial forms, relatives, and adverbial clauses.

Subordination is believed by some scholars to mark greater elaboration and be characteristic of informational discourse typical in writing genres (O'Donnell, 1974; Kay, 1977). Conversely, others (e.g. Halliday, 1989) believe that subordination is associated with the production constraints characteristic of speech. In the findings in this section, we can see that participial forms (F25 through F28) occur more often in writing than in speech (e.g. Figure 5.19 for F25, Figure 5.20 for F26, Figure 5.21 For 27). Concessive adverbial subordinators (F36) and other adverbial subordinators (F38) also follow the same pattern (e.g. Figure 5.29 for F36, Figure 5.30 for F38). However, sentence relatives (F34) and the causative adverbial subordinator *because* (F35) occur more often in speech than in writing genres (e.g. Figure 5.26 for F34, Figure 5.28 for F35). Other subordinators, on the contrary, do not show a clear tendency of frequency in either writing or speech genres. Observing from these, it seems inadequate to claim subordination as a unified construct in different genres.

Among the 18 features, there are 12 demonstrating significant differences between CA and CS texts, and 11 out of the 12 occur significantly more often in the CA texts. Only one feature, present participial clause (F25), appears significantly more often in CS than in CA texts (0.35 vs. 0.13). For this peculiar phenomenon, there is no ready explanation why this feature is unique in its performance.

The relatively lower frequency of most subordination devices used in synchronous CMC is interesting. According to Halliday, the time constraints are the cause of the higher frequency of subordination devices used in speech (1989). Synchronous CMC is also exercised under very tense time constraints - why do CMC participants use less subordination devices in this mode than in the asynchronous mode? The possible explanation I would propose for this particular finding is still time constraints. It may simply be that the time constraints are even more tense in synchronous CMC, so that the participants have to resort to (1) a more careful choice of words than those in non-CMC speaking contexts, and (2) shorter and simpler sentences instead of complex sentences with low lexical density.

For the first assumption, the findings on nominal forms, discussed earlier in Section 5.6, and lexical specificity, being discussed later in Section 5.11, can be referred to. In these sections, CMC texts show a relatively higher frequency of nominalisation (F14) and word length (F44), characteristics of high lexical density, than most speech genres. It is believed that, synchronous CMC users, especially those of higher educational background whose texts are sampled in this study, tend to achieve higher lexical density under the time constraints rather than apply subordination devices.

For the second assumption, that participants in synchronous CMC tend to break the otherwise complex sentences into simple ones to facilitate rapid exchange of information, similar findings can be found in the discussion of coordination devices to follow in Section 5.16. There, the word *and* as a clausal coordinator is also found to be less frequent in synchronous CMC (CS) than in asynchronous CMC (CA). This, of course, further proves that the synchronous CMC texts are shorter in general, unlike the spoken texts frequently prolonged by speakers with subordination devices (Halliday, 1978, 1989).

Therefore, for this point, it can be concluded that the participants in the CS texts have to communicate synchronously, but with a slow means of message production, i.e. typing. Due to the limitation on the speed of typing, the general practice of using subordinate and coordinate sentences in speech is not practical in synchronous CMC. To avoid the

synchronicity of communication being hindered, CMC participants tend to use less devices of subordination in the synchronous mode than in the asynchronous mode.

5.10 Adjectives and Adverbs

In this section, the findings of the category of adjectives and adverbs is discussed on the basis of the results generated from the Section 4.6.3 Non-Parametric Statistic Test in Chapter Four Methods and Procedures of Investigation. This category includes four features: prepositional phrases (F39), attributive adjectives (F40), predicative adjectives (F41), and adverbs (F42). Among the four features, only attribute adjectives (F40) show a difference up to the significant level (α=0.005) across the two temporalities of the CMC texts.

5.10.1 Findings and discussion

There are a total of four features examined in this category of adjectives and adverbs. All the findings are displayed in the Table 5.13. Some examples of these features from the CMC texts are:

Prepositional phrases (F39):
*See you **at the cafe!*** (ntcho-001)
*Thanks **for your help**.* (tesln-003)
*I'll draw this **to a close**.* (hmoo-0417)

Attributive adjectives (F40):
*It seems you had an **errant** period.* (ntcho-011)
*I think this is a **serious** problem.* (tesln-018)
*You already passed the **first** lesson!* (hmoo-0424)

Predicative adjectives (F41):
*It will also be **possible** to participate in the conference on-line.* (ntcho-012)
*It is **nice** seeing you all.* (hmoo-0417)
*Everything is **due** on April 19th.* (rmoo-0402)

Adverbs (F42):
*Any information you provide will be **greatly** appreciated.* (ntcho-007)
*I am **usually** on at night.* (hmoo-0417)
*We **merely** asked that students provide a link from their homepage to their Webfolios.* (rmoo-0507)

The comparison of these four features is shown in Table 5.13, from which we can see that only one feature, attribute adjectives (F40), demonstrates a significantly higher occurrence in the CA than in the CS texts (30.54 vs. 21.14).

Table 5.13
Comparison Between Sync. and Asyn. CMC Texts on *Adjectives and Adverbs*

FEATURES	CA > CS	CS > CA	P Value
(I) ADJECTIVES AND ADVERBS			
(I1) PREPOSITIONAL PHRASES			
F39 Preposition		76.61 > 73.73	0.363
(I2) ADJECTIVES AND ADVERBS			
F40 Attrib. adj.	30.54 > 21.14		0.000*
F41 Pred. adj.	3.39 > 2.51		0.878
F42 Adverb		38.10 > 37.45	0.777
Total cases	1 case significant	0 case significant	

Remarks: CA: CMC Asynchronous; CS: CMC Synchronous.
*: Significant difference reaches $\alpha = 0.005$.

Prepositional phrases and adjectives and adverbs are grouped in the same category as they are all believed to have the quality of expanding and elaborating the information presented in a text.

Prepositions pack logical and nominal kinds of information into texts, and are believed to have the function of packing high amounts of information into academic nominal discourses (Chafe, 1982, 1985; Chafe & Danielewicz, 1987; Biber, 1988a). Biber (1988a) also argues that prepositions "tend to co-occur frequently with nominalisations and passives in academic prose, official documents, professional letters, and other types of written discourse" (p. 237).

Adjectives and adverbs are also devices used for idea unit integration and expansion, but they tend to convey more descriptive kinds of information rather than the logical, nominal kinds of information presented by prepositional phrases. Among adjectives, there are studies that distinguish between attributive and predicative. While attributive adjectives are found to be more integrative in their functions, predicative adjectives are believed to be more fragmented (Drieman, 1962; O'Donnell, 1974; Chafe, 1982, 1985).

The significantly higher frequency of attributive adjectives (F40) in the CA than that in the CS texts (30.54 vs. 21.14), as further illustrated in Figure 5.31, suggests that people in the

asynchronous temporality of CMC may be more apt to integrate information in the texts they produce. It may be due to the relatively lower levels of pressure they experience from the time constraints. Hence, they are freer in searching for appropriate lexical entries with proper adjectives to express their thoughts clearly.

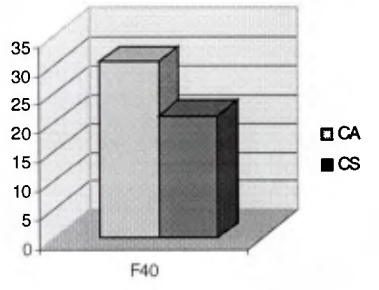

Figure 5.31 Comparison of Features from F40
F40 *Attribute Adjectives* (*You already passed the **first** lesson!* -hmoo-0424)

In contrast with attributive adjectives, predicative adjectives and adverbs do not show significant differences between the synchronous and asynchronous temporalities of the CMC texts. This is interesting, as adjectives and adverbs are both believed to serve the function of integrating descriptive details in the texts. The fact that only attributive adjectives show the differences between CA and CS may be due to their relatively more integrative nature.

Compared with attributive adjectives (F40), predicative adjectives (F41) are considered more fragmented (Biber, 1988a). They are also frequently used for making a stance as heads of *that* or *to* complements (Winter, 1982), such as *I am glad that you managed to come, I am happy to see this.*

Adverbs, which are also believed to serve the function of integrating descriptive details in the texts, come with verbs, unlike attributive adjectives, which more often come with nouns.

A provisional explanation of the significantly higher frequency of attributive adjectives in CA than in CS texts may be their association with nominal forms, which are a characteristic of asynchronous communication.

We can further explore this category by looking at the comparison of the ranking of all the CMC and non-CMC genres on their frequencies of attributive adjectives (F40) and adverbs (F42). In Figure 5.32, we can see that writing genres generally have higher frequencies of the use of attributive adjectives than speech genres, with the genres of press reviews (82.3) and official documents (77.9) standing at the highest ranks.

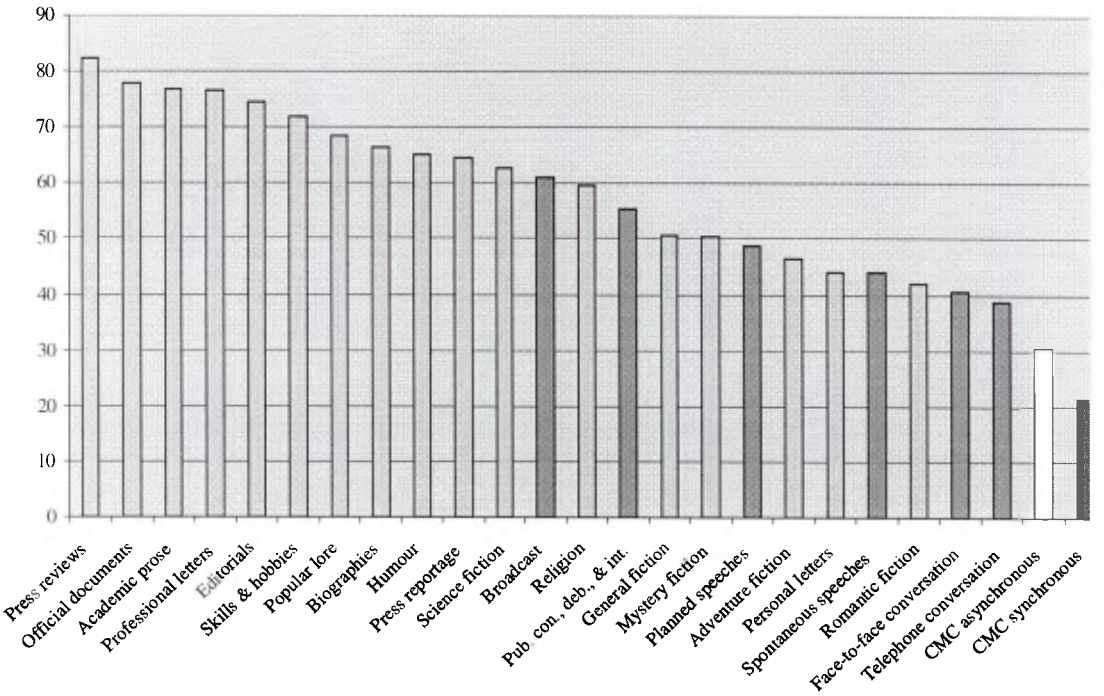

Figure 5.32 Use of *Attribute Adjectives* Between CMC and Non-CMC Genres
F40 *Attribute Adjectives* (*You already passed the **first** lesson!* -hmoo-0424)

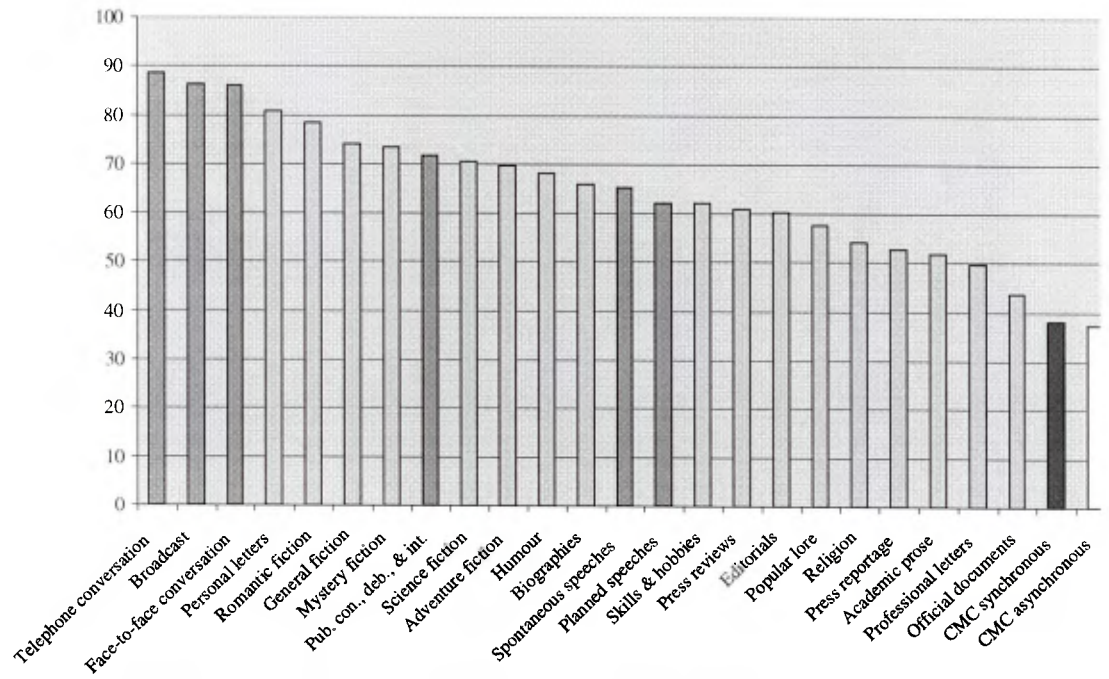

Figure 5.33 Use of *Adverbs* Comparison Between CMC and Non-CMC Genres
F42 *Adverbs* (*I am **usually** on at night.* -hmoo-0417)

In Figure 5.33, however, it shows that speech genres generally stand higher on the ranks, the highest being telephone conversation (88.5) and broadcast (86.3). The interesting thing is that, in either of these figures, the two CMC genres both stand at the lowest ranks. These two figures suggest that the use of attributive adjectives and adverbs is quite rare in CMC in comparison with in non-CMC settings.

While adjectives and adverbs serve to integrate descriptive details in the texts, the finding here may suggest less integration in CMC than in non-CMC texts. However, we should not make this over-simplified conclusion. As a matter of fact, as we find in other sections of the discussion in this chapter, CMC users adopt certain features to achieve higher information density, such as the more frequent use of nominalisation (F14, in Section 5.6.2), higher type/token ratio (F43, in coming Section 5.11.2), and longer words (F44, in coming Section 5.11.3). The features of attributive adjectives and adverbs just seem not to be their favourite measures for integration of information.

5.10.2 Summary of adjectives and adverbs

This section deals with findings related to the category of prepositions, adjectives and adverbs. These four features are all used to integrate or elaborate information, but only attributive adjectives showed significant difference in frequency between CA and CS texts. It is speculated to be related to the association of attributive adjectives with nominal forms. As prepositions, adjectives, and adverbs all tend to pack a large amount of information intp texts and to expand and elaborate the information, the findings in this section suggest that CA and CS texts do not vary much in packing and integrating information, except for in the case of attributive adjectives. Moreover, as it is found that they are used much less in CMC than in non-CMC texts, we can conclude that CMC users do not tend to use these features to achieve integration and high information density in their texts. The features more typically used in CMC for text integration purposes are nominalisation (F14), higher type/token ratio (F43), and longer words (F44).

5.11 Lexical Specificity

This section describes and discusses the result found in the category of lexical specificity in the statistical comparison processed in Section 4.6.3 in Chapter Four. There are

two features included in this category: type/token ratio (F43), and word length (F44). Though they both showed some difference between the CS and CA texts, no significant difference is found in either case.

5.11.1 Findings

The results of the comparison for the type/token ratio (F43) and word length (F44) are shown in Table 5.14. From the table, we can see that neither of them carries any significant differences between the CA and CS texts (α=0.005).

Table 5.14
Comparison Between Sync. and Asyn. CMC Texts on *Lexical Specificity*

FEATURES	CA > CS	CS > CA	P Value
(J) LEXICAL SPECIFICITY			
F43 Type/token ratio		0.56 > 0.55	0.588
F44 Word length	4.83 > 4.82		0.668
Total cases	0 case significant	0 case significant	

Remarks: CA: CMC Asynchronous; CS: CMC Synchronous.
*: Significant difference reaches α = 0.005.

However, if we compare the use of these features among all CMC and non-CMC genres, we find something interesting.

In Figure 5.34, we can see that almost all the writing genres have higher type/token ratio than speech genres, the highest being press review (0.565). In Figure 5.35, there is a similar situation, with the genre of official documents having the longest word length (4.9 characters per word). It is interesting that, in either of the figures, both CMC genres are among the highest on the ranking. In terms of the type/token ratio, as shown in Figure 5.34, synchronous CMC (CS) texts have the second highest number of 0.56, and asynchronous CMC (CA) texts also have a high ratio of 5.5, while the highest type/token ratio found among the speech genres is only 0.497 for broadcast. In terms of word length, as shown in Figure 5.35, the numbers for asynchronous CMC (CA) and synchronous CMC (CS) texts have reached 4.83 and 4.82, only next to the 4.9 found in official documents.

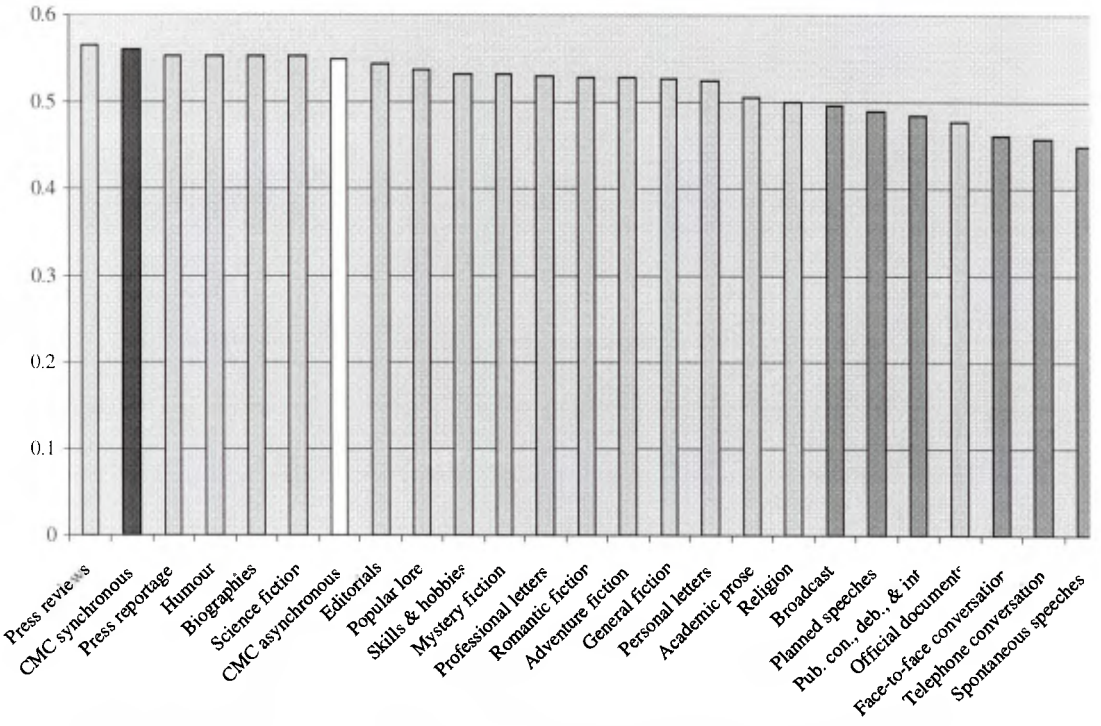

Figure 5.34 Use of *Type/Token Ratio* (F43) Between CMC and Non-CMC Genres

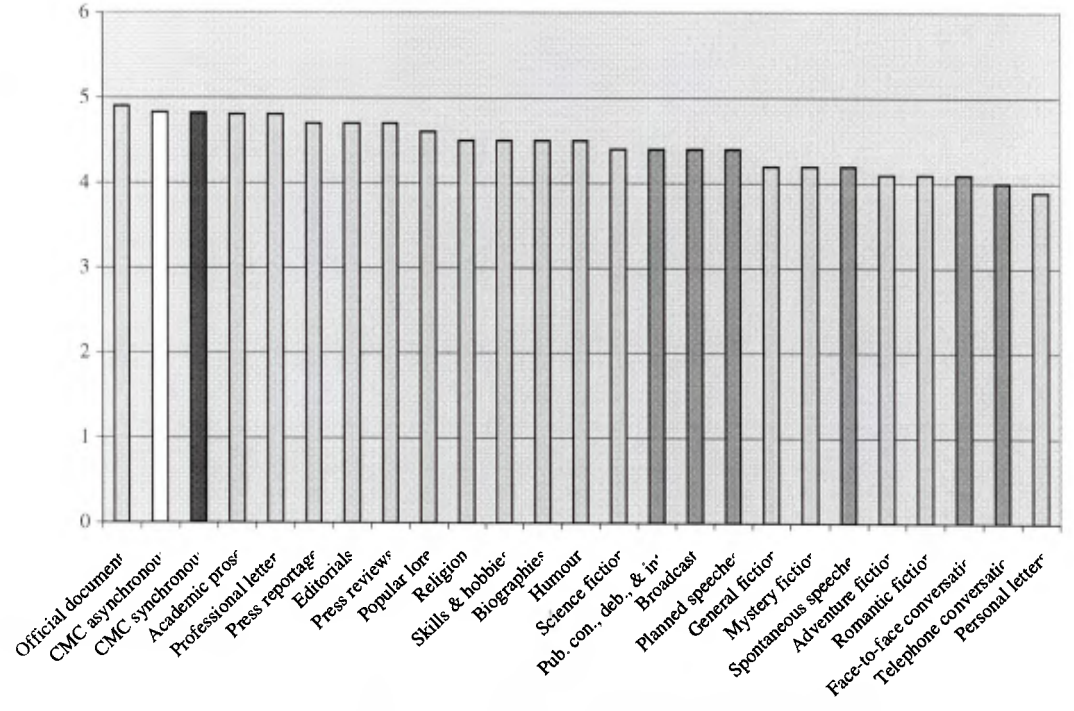

Figure 5.35 Use of *Word Length* (F44) Between CMC and Non-CMC Genres

5.11.2 Discussion of type/token ratio

Type/token ratio refers to the proportion of different word types used in relation to the total number of words in a text. For example, a 500-word article has 500 running words but it may only use 300 different words. The type/token ratio of this article would be 0.6 from the number of 300 divided by 500. The more different words used in an article, the higher the ratio would always be.

The type/token ratio of different texts can be compared only if the length of texts are approximately the same. In Biber's study (1988a), the first 400 words of each text were used and the ratio calculated on the basis of these 400 words. In my first comparison using a 2 x 2 Factorial Experiment, in order to achieve a fair comparison with findings on the non-CMC texts, I also chose those CMC texts that were longer than 400 words and calculated the type/token ratio based on the first 400 words. Among the 503 pieces of asynchronous CMC texts collected, 460 pieces of texts short and contain less than 400 words. After these were disgarded, only 43 pieces of CA texts, each with over 400 words, were analysed for this feature. These 43 pieces were also the ones used in the final statistical comparison, i.e. Non-Parametric Statistical Test. Because of this 400-words criterion of selection, an operational definition of type/token ratio in this thesis is the number of word types on the number of total running words based on the first 400 words in each CMC text.

As mentioned earlier, the more different words used in an article, the higher the ratio would be. So, a high type/token ratio would mean a large variety in the choice of words, which Chafe and Danielewicz (1987) believe is rarely accomplished in speech, when participants are too busy giving immediate responses to think carefully of a large variety of words.

In this section, we just made two observations on the type/token ratio: (1) all the writing genres have higher type/token ratio than speech genres, and (2) both CMC genres (0.56 and 0.55) are among the highest on the ranking of type/token ratios.

For the first observation, the CS texts are mainly interaction at real-time, like that among "speakers," but CA is done asynchronously. Therefore, it is quite surprising that the CS texts have the same high type-token ratio as the CA texts, if we accept the argument of Chafe and Danielewicz (1987) that high type/token ratio is a main characteristic of the writing genres. The finding suggests that when people interact via the CMC medium, either synchronously or asynchronously, they are trying their best to use a good variety of words to pack more information into the texts, even during the busy real-time interaction. In other

words, they are very careful in editing words to express their thoughts, hence, their texts reach an effect similar to the genres of press reportage and press review in the traditional written texts. I assume that a part of this particular effect is due to the high education level of the participants in the sampled CMC data. This is because they are all English teachers who usually use English language as a model for their students to follow.

For the second observation, the reasons may be two-fold: The different means of message conveyance and the different types of time constraints.

The means of message conveyance is different in these two media, i.e. the CMC and non-CMC. In non-CMC conversation, whether face-to-face or over the phone, participants simply talk to each other, whereas in CMC, people have to type on the keyboard. The physical load of typing is much heavier than speaking, and the speed of typing is much lower too. Under such circumstances, non-CMC participants can easily repeat the same idea again and again, not worrying about the load that is involved in the redundancy. CMC participants, on the other hand, must be more careful in choosing the most appropriate words to reduce the load of typing. This may be the first reason for the higher type/token ratio in CMC than in synchronous non-CMC texts.

Time constraint is the other factor we must consider in discussing the variety of words used. When people are engaged in real-time communication, they are usually under a time constraint to give an immediate response to each other so the interaction can be carried on smoothly. This is so in face-to-face or telephone conversation, when a pause over a few seconds would be considered unnatural. Under the time pressure, the precise choice of words is not a strict requirement as the speaker may rephrase his or her ideas again and again to get across the ideas. On the other hand, in synchronous CMC contexts, the main time pressure usually comes from the length of time allowed online. Usually, the online CMC time is limited, and sometimes a fee is charged based on the time the computer is connected. Unlike in synchronous non-CMC contexts, where time constraints come from the pressure for immediate interaction with the counterpart either face-to-face or over the phone, CMC participants do have the advantage of being invisible to each other. That is to say, some time lag is still allowed in synchronous CMC so the participants feel easier to search for more precise words to express their ideas.

It may be, then, that the time constraints felt in synchronous CMC are different from those felt in non-CMC contexts. In CMC, the time pressure comes from the total time allowed online, but not from the need to give immediate response. Therefore, the

participants are inclined to search for a variety of more precise words to use, which results in the high type/token ratio. In synchronous non-CMC, the time pressure usually comes from the need for simultaneous interaction, rather than the total time allowed, so the participants feel more urge to give an immediate response than to choose precise words with a limited length of communication.

5.11.3 Discussion of word length

The other feature discussed in this category is word length (F44). Word length is the average length (number of letters) of the words in a text. Words can range from one-letter words like *I*, *a*, to long words that have more than 15 letters, such as *interdisciplinary, responsibilities*, etc. Several studies on the role played by the length of words in language styles have been carried out (Blankenship, 1962; Drieman, 1962; Osgood, 1960). In Zipf's study, it is also found that the more academic, professionally-registered the text type is, the longer word length the text has (Zipf, 1949).

In the CMC texts examined in this study, there is little difference between the word length (F44) in CA (4.83) and CS (4.82) texts. On the other hand, in non-CMC texts, most writing genres have longer-than-average word length than speech genres, with the genre of official documents having the longest word length (4.9 characters per word).

Again, two observations can be made here. First, CMC participants do not vary significantly in the length of words they use in either synchronous or asynchronous temporalities. Second, CMC participants, in either synchronous or in asynchronous temporalities, use word lengths similar to quite formal written texts and much longer than casual speeches.

For the first observation, that both CA and CS texts have quite high average word lengths, we can suggest that in both synchronous and asynchronous temporalities, CMC participants keep a style similar to that of formal written documents. This may partly be due to the nature of the sample texts used in this study. They are all discussions on language teaching issues and the participants are mostly English teachers, so that even in the synchronous interaction, words of higher length are still used. Another reason for the similar length of words used in both synchronous CMC (CS) and asynchronous CMC (CA) texts may be the time constraints on the texts in CS temporality. It is speculated that, with the need to give a full account of a topic within a certain time limit, and to save the load of

keyboard typing, the addressers in CS temporality, just like those in CA temporality, must choose words that are more specific in meaning.

For the second observation, where both CA and CS texts have longer word lengths, we can also suggest that CMC addressers are not under the same type of time constraints as speakers in synchronous non-CMC. For most CMC addressers, the time constraints come from the need to complete the message conveyance within a limited time. However, in synchronous non-CMC contexts, such as face-to-face conversation and phone conversation, the time pressure is mainly from the need to give an immediate response to one's counterpart. Therefore, CMC addressers must be careful to choose words carrying more specific meanings, while speakers in synchronous non-CMC contexts are more concerned with giving an immediate response to each other and therefore have to use words that are shorter and more general.

5.11.4 Summary of lexical specificity

Both type/token ratio (F43) and word length (F44) serve to mark the density of information. The general belief is that these two features are more characteristic of rich content and higher lexical density. And both of them should occur with higher frequency in asynchronous rather than synchronous contexts. The findings in this section, however, indicate that synchronous and asynchronous CMC texts are similar to each other with regard to these two features. It was also found that CMC texts tend to have a higher type/token ratio and higher word length than synchronous non-CMC texts, i.e. face-to-face conversation and telephone conversation. The above findings suggest that, though CMC users do not vary much in type/token ratio and word length when they are in the synchronous or asynchronous mode of communication, they use a relatively larger variety of vocabulary and longer words than when in non-CMC settings. This is interesting because it shows how CMC users try to incorporate more dense information within the limitation of time and space. The reasons for the findings may include the following: (1) the specific topic and the certain background of participants of discussion in the sampled CMC texts, and (2) the time constraints and the physical workload of typing on the CMC users under which they must give a complete account of a topic. We know that synchronous CMC users type on the keyboard to communicate and are therefore much slower in producing utterances than those who speak. Besides, asynchronous CMC users also write much shorter responses, perhaps because they know that their counterparts will not have much interest in reading long texts (Chou, 1996, 1998). It is therefore logical that CMC

users, in either the synchronous or the asynchronous mode, attempt to achieve higher information density in the texts by resorting to higher type/token ratio and longer words. This finding can be referred to along with those on nominalisation (F14, in Section 5.6.2), subordination (F21-F38, in Section 5.9), and coordination (F64-F65, in coming Section 5.16).

5.12 Lexical Classes

This section discusses the category of lexical classes, which includes seven features of conjuncts (F45), downtoners (F46), general hedges (F47), amplifiers (F48), general emphatics (F49), discourse particles (F50), and demonstratives (F51). In this category, there are two features, general hedges and discourse particles, which demonstrate significant differences across the CA and CS texts.

5.12.1 Findings

The statistical findings on the seven features of this category are depicted in Table 5.15 and Figure 5.36 below. Some examples of these features are also listed below to show how they were used in the CMC texts.

Conjuncts (F45):
Are there any exchanges besides with students, ***for example*** *faculty members or sister cities?* (ntcho-007)
What's happening. Who ***else*** *is in the kitchen?* (hmoo-0417)
Academic prissiness, ***in other words****.* (rmoo-0402)

Downtoners (F46):
Almost *all will be delighted for the free publicity.* (ntcho-052)
Give them, at the first stage, phrasal verbs that have ***only*** *one meaning.* (tesln-013)
Not knowing a huge amount about linguistics, language acquisition etc, they have seemed to me to be ***somewhat*** *retro, traditionally determined areas.* (hmoo-0516)

Hedges (F47):
I am looking for CD-rom courses for Portuguese, if you run across ***something like*** *that let me know.* (ntcho-016)
I'm glad to be here, though I'm ***kind of*** *mixed up.* (hmoo-0417)
It ***sort of*** *sounds like it was.* (rmoo-0910)

Amplifiers (F48):
I've found Gapmaster from Wida software ***very*** *good for the types of applications that you want.* (ntcho-004)
Of course, the use of these items is ***highly*** *idiomatic, and that's what causes the headaches.* (tesln-017)

*They're doing a **perfectly** adequate job without net.* (hmoo-0509)

Emphatics (F49):
*It's pretty easy to do, but if you try it early and have trouble, **just** show up to the cafe and we'll all help you get your items added to the list.* (ntcho-001)
*I **did find** Kombucha, a discussion about using and consuming the Kombucha beverage.* (ntcho-002)
*One **more** command, and then THAT'S ENOUGH!* (hmoo-0424)

Discourse particles (F50):
***Anyway**, if you want to contact me about the specifics of the work involved, feel free to email or call me.* (ntchn-083)
***Now**, the word 'bitch' used to describe someone is a curse word to many, and definitely an insult.* (tesln-045)
***Well**, our classes will be coming in a few weeks when we do our lesson.* (hmoo-0417)

Demonstratives (F51):
*At **this** week's cafe, let's compare lists.* (ntcho-001)
*Try **these** phrases out on your class and see what happens!* (tesln-017)
*Nick's comment, actually caused **that** fall ....* (rmoo-0507)

Table 5.15
Comparison Between Sync. and Asyn. CMC Texts on *Lexical Classes*

FEATURES	CA > CS	CS > CA	P Value
(K) LEXICAL CLASSES			
F45 Conjunct	2.58 > 1.15		0.810
F46 Downtoner	1.79 > 1.11		0.066
F47 Gen. hedge		1.07 > 0.55	0.000*
F48 Amplifier	2.66 > 1.04		0.828
F49 Gen. emphatic		7.85 > 6.99	0.129
F50 Discourse part.		1.15 > 0.07	0.000*
F51 Demonstrative	4.56 > 3.37		0.560
Total cases	0 case significant	2 cases significant	

Remarks: CA: CMC Asynchronous; CS: CMC Synchronous.
*: Significant difference reaches $\alpha = 0.005$.

The statistical findings, as depicted in Table 5.15 and Figure 5.36 above, show that two features, general hedges (F47) and discourse particles (F50), demonstrate significantly higher frequency in the synchronous CMC (CS) texts. General hedges occur more often in the CS texts than in the CA texts (1.07 vs. 0.55), while discourse particles also occur more often in the CS than in the CA texts (1.15 vs. 0.07). All of the other five features show no significant difference between the two temporalities.

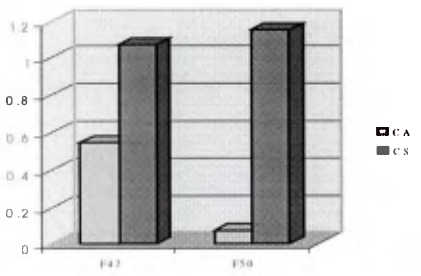

Figure 5.36 Comparison of Features from F47 and F50
F47 *General Hedges* (*It **sort of** sounds like it was.* -rmoo-0910)
F50 *Discourse Particles* (***Well**, our classes will be coming in a few weeks when we do our lesson.* -hmoo-0417)

5.12.2 Discussion of hedges and discourse particles

Certain words and phrases are believed to show characteristics of either the writing or the speaking style, and they are all listed in this category. Among them, hedges (F47) are words and expressions like *more or less*, *maybe*, *something like*, etc. to indicate probability or uncertainty. Hedges simply mark a proposition as uncertain, while downtoners (F46) indicate the degree of uncertainty. Chafe and Danielewicz (1987) argue that the use of hedges in conversation indicates an awareness of the limited word choice. Biber (1986) finds that hedges tend to co-occur with interactive features like first and second person pronouns. He (Biber, 1988a) also finds that hedges are an important indicator of the involved style.

In the present study, general hedges are found to be used more in the CS texts (1.07) than in the CA texts (0.55). This confirms that hedges are often taken as being characteristic of conversational discourse, "to flag uncertainty or lack of precision in the presentation of information" (Biber 1988a, p. 106).

When we look at the comparison of the use of general hedges in CMC and non-CMC genres in Figure 5.37, we can see further evidence of this. Among the 23 non-CMC genres examined in his study, personal letters have the highest frequency of hedges (3.3). Other features with relatively higher frequency of hedges include telephone conversation (2.7) and face-to-face conversation (2.1). Genres with the lowest frequency of hedges include religious documents (0.0), official documents (0.0), biographies (0.1), humour (0.1), and romantic fiction (0.1). The distribution of general hedges in all CMC and non-CMC genres, as seen in Figure 5.37 below, shows that speech genres tend to use more of them than writing genres. And the two CMC genres, though different from each other in the frequency of

hedges used, are both relatively higher than most non-CMC writing speech genres. This also suggests that CMC users tend to use a more casual style than in writing in non-CMC genres.

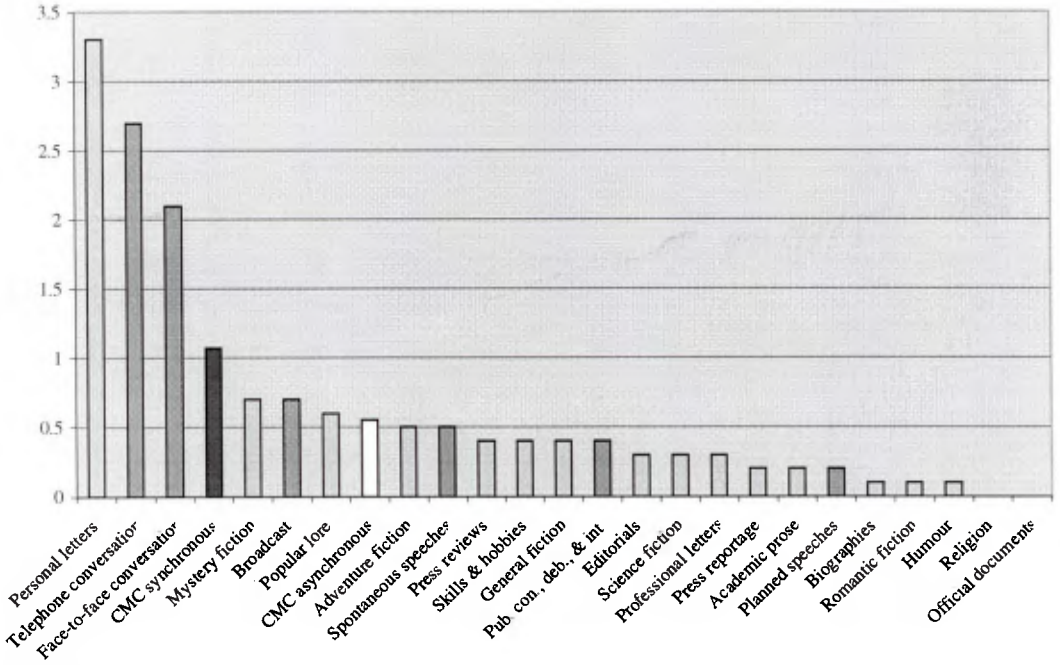

Figure 5.37 Use of *General Hedges* Between CMC and Non-CMC Genres
F47 *General Hedges* (*It **sort of** sounds like it was.* -rmoo-0910)

Hedges, as defined in this study, include words and expressions like *at about*, *something like*, *almost*, *maybe*, etc. They are informal, less specific markers of probability. It is not surprising that they occur more often in MOO than in asynchronous CMC texts as the former is similar to the context of speech communication. Hedges, as defined in this study, are different from features that characterise academic hedging, i.e. downtoners (F47) like *almost*, *barely*, *hardly*, and the feature *seem/appear* (F58). Downtoners and *seem/appear* are often found as more formal markers of uncertainty in academic writing. They will be discussed later in this section and in the coming Section 5.14.

The other feature that shows significant difference between the two temporalities is discourse particles. Discourse particles (F50) are expressions like *well*, *anyway*, *now*, etc. used after a "tone unit" boundary. They are often found in involved discourse to monitor the information flow (Chafe, 1982, 1985), and are rarely found outside the conversation genre (Biber, 1988a). In the present study, discourse particles occur extremely more often in CS texts than in CA texts (1.15 vs. 0.07). This is in line with earlier findings that discourse

particles are rarely used outside the conversation genre (Biber, 1988a). Moreover, Biber, (1988a) argues that discourse particles, like hedges, are an important indicator of the involved style.

Among the 23 non-CMC genres examined, as seen in Figure 5.38 below, telephone conversations have the highest frequency of discourse particles (6.6). Other features with a relatively higher frequency of hedges include face-to-face conversation (3.9) spontaneous speeches (3.6), and interviews (3.0). Genres with lowest frequency of discourse particles include press reportage (0.0), biographies (0.0, official documents (0.0), academic prose (0.0), official documents (0.0), and romantic fictions (0.0). In the present study, discourse particles are found to be used more in the CS texts (1.15) than in the CA texts (0.07). The discourse particles in CS texts are not as frequent as those in speech genres. However, they are significantly more frequent than those in CA texts. This suggests that CS addressers are more inclined than CA addressers to use discourse particles to achieve the effect of coherence, but the inclination is not as strong as that of speakers in non-CMC contexts.

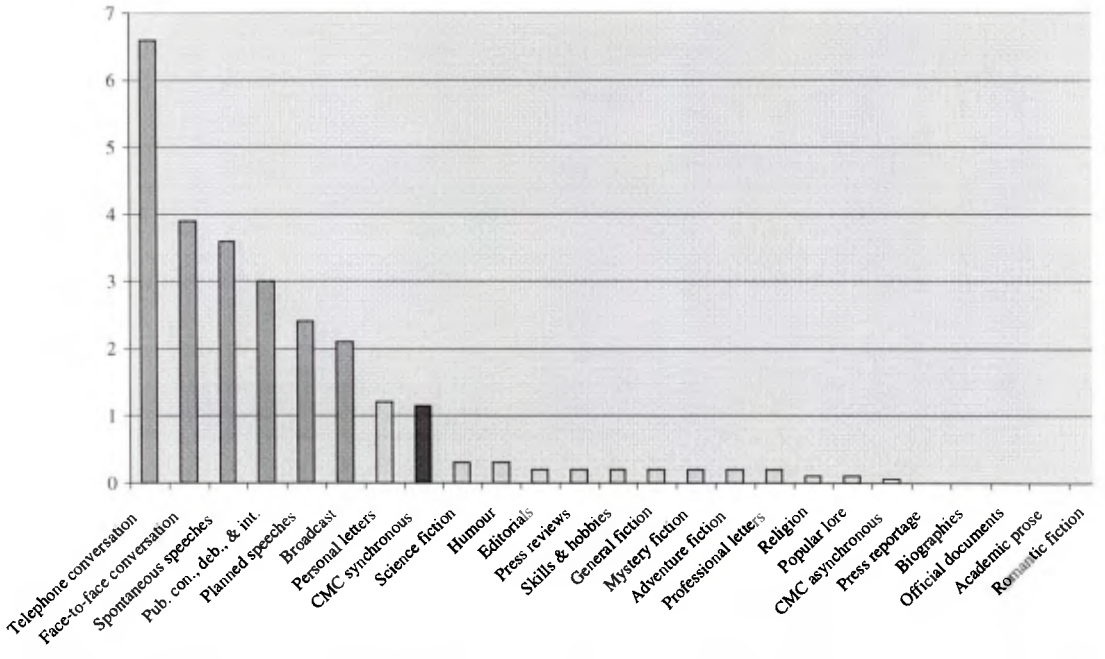

Figure 5.38 Use of *Discourse Particles* Between CMC and Non-CMC Genres
F50 *Discourse Particles* (***Well**, our classes will be coming in a few weeks when we do our lesson.* -hmoo-0417)

Lexical classes tend to add or lower the forces of the verbs. It is expected that they may be crucial in CMC texts, where non-verbal cues such as facial expressions and gestures rich in face-to-face conversation are unavailable (Chou, 1996, 1998). The above findings do

suggest that hedges and discourse particles are verbal messages that people often use to add or lower the forces used in synchronous CMC to make up the insufficient non-verbal cues in face-to-face conversation. However, it is also interesting to notice that in synchronous non-CMC contexts, such as face-to-face conversation and telephone conversation, speakers use even more hedges and discourse particles than in synchronous CMC. It seems that the use of these lexical classes is a popular practice in synchronous CMC, but not as popular as in synchronous non-CMC. This again may be due to the time constraints on the CMC users, as discussed earlier in Section 5.11.

5.12.3 Discussion of other features in this category

Other features in this category, like conjuncts, downtoners, amplifiers, general emphatics and demonstratives, do not show significant differences between the two temporalities.

Conjuncts (F45) are words and phrases like *alternatively*, *e.g.*, *instead*, *in particular*, *by contrast*, etc. They have the function of marking logical relations between clauses and usually show an informational focus. Quirk and his colleagues (Quirk, Greenbaum, Leech, & Svartvik, 1985) even argue that conjuncts have the functions of listing, summative, appositive, resultive, inferential, contrastive, and transactional. Moreover, conjuncts are found to be more formal and therefore more common in planned discourse (Ochs, 1979). They are also found to occur more in writing than in speech (Altenberg, 1986), and to "co-occur with passive forms to mark the complex logical relations among clauses" (Biber, 1988a, p. 112).

Downtoners (F46), the other feature that occur higher in the CA than CS texts, are words like *almost*, *barely*, *only*, *somewhat*, etc. They have a lowering effect on the force of the verb. They are believed to occur more often in academic writing to indicate the lower degree of probability (Chafe & Danielewicz, 1987). Holmes (1984) also believes that, besides marking uncertainty, these forms can also mark politeness.

Amplifiers (F48), the other feature that is also popular in the CA text, are words like *absolutely*, *entirely*, *strongly*, etc. They are the opposite of downtoners in that they boost the force of the verb (Quirk et al., 1985). They indicate the degree of certainty of a proposition. In addition to marking certainty or conviction, they are believed to signal solidarity (Holmes, 1984).

Demonstratives (F51) are the last feature that occurs high in the CA text. They are the four words *this*, *that*, *these*, and *those* when not used as demonstrative pronouns (as F10 discussed earlier), relative (as F29 and F30 discussed earlier), complementizsr (as F21 and F22 discussed earlier), or subordinator (as in F60, which will be discussed later). They are believed to mark the referential cohesion in a text (Halliday & Hasan, 1976). Ochs (1979) finds that they are used more often than articles in unplanned discourse.

From Table 5.15 and Figure 5.36, it is found that of the above four features, all occur more often in the CA than in the CS texts though the level of significant difference has not been reached. The tendency, though, is somewhat in line with previous findings on non-CMC texts.

The feature that occurs more often in the CS texts, but that does not reach the significant difference, is general emphatics (F49). General emphatics include expressions like *for sure*, *such a*, *most*, etc., or structures like *so* followed by an adjective (*so + ADJ*), and *do* followed by a tenseless verb (*DO + V*), etc. Instead of indicating the degree of certainty, they simply mark the presence of certainty. They are believed to characterise informal, colloquial discourse (Chafe, 1982, 1985). Biber (1986, 1988a) finds them to co-occur with hedges in conversational genres. It is logical to find that this feature occurs more often in the CS texts, though it does not reach the level of significance.

5.12.4 Summary of lexical classes

Among the features examined in this category, two have reached the significant differences. They are general hedges (F47) and discourse particles (F50), and both of them occur more often in the synchronous than in the asynchronous CMC texts. The differences found in the other five features do not reach the significant level. The general hedges, as defined here, means informal expressions like *at about*, *something like*, *almost*, *maybe*, etc. They are informal, less specific markers of probability and different from features that serve the function of academic hedging. The more formal expressions of hedges, like *seem, appear*, will be discussed in Section 5.14 later. Here, it is believed that general hedges and discourse particles are a popular practice in synchronous CMC as a means to mark uncertainty or facilitate coherence in a colloquial way, just as like in speech.

5.13 Modals

In language, modals serve many functions in communication, such as (1) predicting something will happen, (2) indicating something that is believed, (3) expressing uncertainty in academic writing, (4) serving as a form of politeness of request, (5) showing the degree of uncertainty that the speaker claims for something, and (6) showing the degree of speaker's commitment towards requests from others.

Modals are examined in this study in three groups, namely, possibility modals (F52), necessity modals (F53), and predictive modals (F54). Predictive modals, like *can*, *could*, *may*, are used to give direct procurements, though with some uncertainty, that certain things may happen. Necessity modals, like *should*, *must*, *ought*, pronounce obligation or necessity. Possibility modals like *will*, *shall*, *would*, address the possibility or ability for certain things to happen. Chafe (1985) argues that possibility modals mark reliability and necessity modals mark some aspects of the reasoning process. Though all three show higher occurrences in the asynchronous CMC texts, none of the differences reaches the significant level. The following sub-sections describe and discuss the result of the three features found in the Non-Parametric Statistical Test processed in Section 4.6.3 in Chapter Four.

5.13.1 Findings and discussion

The statistical findings on possibility modals, necessity modals, and predictive modals are depicted in Table 5.16, in which, we can see that all three features show somewhat higher frequencies in the CA than in the CS texts, but none of them appears to be significantly different across the asynchronous and synchronous CMC texts. Before going into detailed discussion, some example sentences of these features are listed below for reference.

Possibility modals (F52):
*Everyone **can** write on the list, even if you're using a guest character to log onto Mediamoo.* (ntcho-001)
***Could** someone suggest some resources I might consult to research the morphological/syntactical principles of this negation.* (tesln-001)
*Folks **might** have been better off with you!* (rmoo-0402)

Necessity modals (F53):
*That someone **should** have something for me by then.* (ntcho-017)
*I **must** go supermooers, see you at the next session.* (hmoo-0417)
*Maybe the question **ought** to be what are the problems of taking grad students to the moo?* (rmoo-1008)

Predictive modals (F54):
We'll collect the lists in the cafe and post the final list on the Netoric web site later. (ntcho-001)
*I **shall** post a summary within 10 days.* (tesln-009)
*Do you think student access **would** be a good Neteach discussion topic?* (hmoo-0417)

From Table 5.16, we found that all the three features occur more often in CA than in CS texts: possibility modals (F52, 7.37 vs. 6.75), necessity modals (F53, 1.92 vs. 1.46), and predictive modals (F54, 7.41 vs. 5.46). However, none of the feature reaches the significant level.

Table 5.16
Comparison Between Sync. and Asyn. CMC Texts on *Modals*

FEATURES	CA > CS	CS > CA	P Value
(L) MODALS			
F52 Possib. modal	7.37 > 6.75		0.975
F53 Necess. modal	1.92 > 1.46		0.028
F54 Predic. modal	7.41 > 5.46		0.490
Total cases	0 case significant	0 case significant	

Remarks: CA: CMC Asynchronous; CS: CMC Synchronous.
*: Significant difference reaches α = 0.005.

Modals are believed to serve the purpose of overt-expression of persuasion. In the factor analysis by Biber (1988a), all the three kinds of modals are found to mark overt expression of persuasion in discourse. They indicate,

> "the degree to which persuasion is marked overtly, whether overt marking of the speaker's own point of view, or an assessment of the advisability or likelihood of an event presented to persuade the addressee." (Biber 1988a, p. 111)

The concept of "modality" is used in a very general way to cover features of texts which "expresses speakers' and writers' attitudes towards themselves, towards their interlocutors, and towards their subject-matter (Fowler, Hodge, Kress, & Trew, 1979, p. 200). It is expressed by modal auxiliary verbs like *may, might, must, should, can, can't, ought,* but also by various other formal features including adverbs and tense (Fairclough, 1989, p. 127).

Modals are a kind of construct to indicate modality. For any proposition, one may make categorical assertions about it or deny it, there are also available various less

categorical and less determinate degrees of commitment to it or against it. Fairclough (1992, p. 158) gives the example that between the statements "the earth is flat" and "the earth is not flat," there may be statements like "the earth may be/is probably/is possibly/is sort of flat." This is the sphere of modality, and is what Hodge and Kress (1988) mark as the "degree of affinity." Therefore, in language, any proposition has the property of modality.

Modals share, to some degree, meanings of possibility and necessity, which justifies the use of the term. It expresses the speaker's assessment of probabilities and predictability. As Halliday observes (1976d, p. 204), modality is external to the content and is therefore not subject to variation or constraints of tense or polarity. Therefore, modality has no tense, voice or polarity of its own, but combines freely with all values of these variables in the clause.

Modality in grammar was traditionally associated with the modal auxiliary verbs (*must, may, can, should* and so forth). However, the "systemic" approach to grammar which Hodge and Kress (1988) draw upon has stressed that modal auxiliaries are not the only one modality feature (Halliday, 1985, pp. 85-89). According to Fairclough (1992, p. 159), tense, "modal adverbs" such as *probably*, *possibly*, *obviously*, and *definitely*, and hedges such as *sort of*, *a bit*, or *something*, intonation patterns, speaking hesitantly, and so forth, all may serve the function of indicating modality. Fairclough (1995, p. 27) further argues that modal auxiliaries are typical features of modality, but choices of pronouns, speech acts, and many others, are also included within the scope of modality.

According to Halliday (1976a, 1976b, 1976c, 1976d), modality is a form of participation by the speaker in the speech event and derives from the interpersonal function of language. This, further confirmed by Turner and Pickvance (1971), is the function through which the child learns to participate, as an individual, and to express and develop his own personality and his own uniqueness.

Fairclough (1989, p. 127) maintains that there are two dimensions to modality, depending on what direction in which authority is oriented in. Firstly, if it is *relational modality*, it is a matter of the authority of one participant in relation to others. Secondly, it is *expressive modality,* i.e. the modality of the speaker/writer's evaluation of truth. Halliday (1976d, p. 211) makes the distinction between modality and modulation by arguing that modality is an attitude towards the content that is being expressed., while modulation is a characterisation of the relation of the participant to the process – his ability, to carry it out.

All of these studies show how complex this topic may be. In the present study, as I use only the three types of auxiliary modals for analysis, I intend to limit the discussion only to these modals. In the present study, there are no significant differences of the frequency of these modals found between the synchronous and asynchronous temporalities of the sampled CMC texts. This suggests that CMC users do not vary much in making overt persuasion in different temporalities. CMC texts are roughly the same in the use of the frequency of these modals.

5.13.2 Summary of modals

The category of modals includes three features, namely possibility modals (F52), necessity modals (F53), and predictive modals (F54). Modality is an essential characteristic in any proposition, which can be expressed by a number of structures, including modals. In the present study, the scope is limited to the three types of modals and none of the features was found to demonstrate a significant difference between synchronous and asynchronous CMC texts, though they generally occur more often in the asynchronous CMC temporality.

5.14 Specialised Verb Classes

This section discusses the category of specialised verb classes that include public verbs (F55), private verbs (F56), suasive verbs (F57), and *SEEM/APPEAR* (F58). In this category, there are two features, public verbs and *SEEM/APPEAR*, which demonstrate significant differences between CA and CS, while private and suasive verbs show no significant differences across the CA and CS texts.

5.14.1 Findings

The statistical findings on the four features of the specialised verb classes, taken from the overall Table 5.1 of Section 5.1, are depicted below in the Table 5.17 and Figures 5.39 and 5.40. Some example sentences are listed below showing how these features are used in the CMC texts.

Public verbs (F55):

Your help will be ***acknowledged*** *in the publication unless you prefer to remain anonymous.* (ntcho-007)

Could someone ***suggest*** *some resources i might consult to research the*

morphological/syntactical principles of this negation. (tesln-001)
*Do you know the commands I **mentioned**?* (hmoo-0417)

Private verbs (F56):
*I am **supposed** to say something sensible about them to my pupils and I feel more than helpless.* (ntcho-008)
*I **think** we have more teachers than students.* (hmoo-0417)
*What did you **see** that troubled you?* (rmoo-0402)

Suasive verbs (F57):
*If a student doesn't need a particular module ("need" being based on test score) they are not **required** to take it.* (tesln-010)
*Let me go over just a couple commands with our new players to **allow** us to move on to our discussion.* (hmoo-0417)
*I do want to **recommend** that EJC issue I mentioned earlier.* (rmoo-0402)

SEEM/APPEAR (F58)
*It **seems** there's a mistaken trailing slash in the url.* (ntcho-006)
*I don't wish to **appear** paranoid but I think it would be a good idea to post some kind of logging prohibition sign - perhaps at the connection stage (or does such a reminder already exist and I have not noticed it?)* (ntcho-063)
*It **seems** a bit unfair doesn't it - you have to achieve this style to pass but I won't show you how!* (hmoo-0725)
*Here are the contents of the samples note -- there **appears** to be some writing on the note.* (rmoo-0507)

From Table 5.17, we can see that two of the features, public verbs (F55) and *SEEM/APPEAR* (F58), exhibit significant differences between the temporalities of communication, as their P values reach 0.005, while the other two features demonstrate no significant differences. Figures 5.39 and 5.40 give a further depiction of the comparisons.

Table 5.17
Comparison Between Sync. and Asyn. CMC Texts on *Specialised Verb Classes*

FEATURES	CA > CS	CS > CA	P Value
(M) SPECIALISED VERB CLASSES			
F55 Public verb		44.78 > 6.63	0.000*
F56 Private verb		17.11 > 15.26	0.088
F57 Suasive V.	5.08 > 3.85		0.719
F58 SEEM, APPEAR	1.12 > 0.80		0.000*
Total cases	1 case significant	1 case significant	

Remarks: CA: CMC Asynchronous; CS: CMC Synchronous.
*: Significant difference reaches $\alpha = 0.005$.

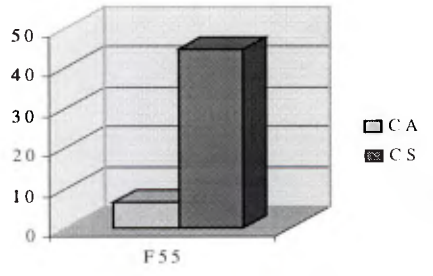

Figure 5.39 Comparison of Feature from F55
F55 *Public Verbs* (*Do you know the commands I* ***mentioned****?* -hmoo-0417)

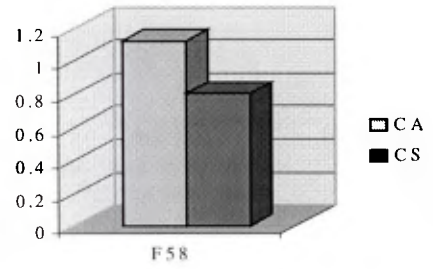

Figure 5.40 Comparison of Feature from F58
F58 *SEEM/APPEAR* (*It* ***seems*** *there's a mistaken trailing slash in the url.* -ntcho-006)

5.14.2 Discussion of public verbs and *SEEM/APPEAR*

The category of specialised verb classes is composed of four classes of verbs: public verbs, private verbs, suasive verbs, and *SEEM/APPEAR* (Biber, 1988a, p. 242). The statistical findings indicated in Table 5.17 show that two of the features, public verbs (F55) and *SEEM/APPEAR* (F58), exhibit significant differences across asynchronous and synchronous CMC texts. Public verbs (F55) occur significantly more often in CS than in CA texts (44.78 vs. 6.63). *SEEM/APPEAR* (F58) occurs significantly more often in CA than in CS texts (1.12 vs. 0.80).

Public verbs (F55) are those words like *acknowledge*, *assert, claim*, *insist*, *say*, etc. (Quirk, Greenbaum, Leech, & Svartvik, 1985) that can be observed publicly. They are primarily speech act verbs "used to introduce indirect statements" (Biber, 1988a, p. 242). In Biber's study, public verbs were found to function as markers of indirect, reported speech (p. 109).

In the present study, the finding that public verbs occur significantly more often in synchronous CMC texts than in asynchronous texts is quite beyond expectation. The

extremely high frequency of public verbs found in synchronous CMC texts (44.78) is much higher than the figure found in non-CMC texts. In the non-CMC data of Biber, the several non-CMC genres exhibiting the highest frequency of public verbs are press reportage (12.0), general fiction (10.3), and adventure fiction (10.0).

In synchronous non-CMC texts, public verbs are found in face-to-face conversation (8.8) and telephone conversation (6.2). The extremely high frequency of public verbs found in synchronous CMC texts is a result of the special format of MOO, in which each of the sentences entered by each participant is automatically displayed with a heading "[So and so] says..." (See Appendix 9). This is believed to be the cause of the extremely large number of public verbs found in synchronous CMC texts.

SEEM/APPEAR (F58) are perception verbs, which mark evidentiality with respect to the reasoning process (Chafe, 1985). From Table 5.17, *SEEM/APPEAR* are shown to occur significantly more often in CA than in CS texts. In Figure 5.40, there is a clear distinction between the two temporalities in the CMC text. Biber (1988a) finds that these two words serve as a marker of academic hedging, "to qualify the extent to which an assertion is 'known' in academic discourse" (p. 114).

In the non-CMC data cited by Biber, the genres that used *SEEM/APPEAR* most often are personal letters (2.2), press reviews (1.4) science fictions (1.2), general fictions (1.1), and academic prose (1.0). The score of *SEEM/APPEAR* is 0.4 in face-to-face conversations, and 0.7 in telephone conversations. This does indicate an effect of uncertainty. That these two words appear significantly more often in CA than in CS texts confirms that CMC users also have an inclination to use *SEEM/APPEAR* more often in asynchronous conditions, most likely to sound modest and reserved in their discussion.

5.14.3 Discussion of other features in this category

The other two features in this category are private verbs and suasive verbs. Private verbs (F56) are verbs of cognition (Carroll, 1960; Weber, 1985). Examples include *anticipate*, *discover*, *feel*, *remember*, *think*, etc. (Quirk et al., 1985). They are used to express intellectual states or non-observable intellectual acts. From the CMC text, we found that private verbs appear more often in CS texts than in CA texts, though the difference does not reach the significant level (17.11 vs. 15.26).

Based on Biber's factor analysis, private verbs are the feature that has the highest loading (0.96) on the factor of involved and non-informational focus (See Factor 1 in

Appendix 8). That means they are the most evident mark of the involved style in texts. In Biber's findings, the non-CMC genres that had the highest scores of private verbs are telephone conversations (35.6), face-to-face conversations (35.4), personal letters (27.2), and interviews (23.7). On the other hand, private verbs occur least in official documents (7.8), broadcasts (10.1), and press reportage (10.4) (See Figure 5.41). The insignificant difference between synchronous and asynchronous CMC texts in terms of the frequency of private verbs suggests that CMC users use private verbs rather moderately, not as often as in telephone and face-to-face conversations, nor as little as in official documents.

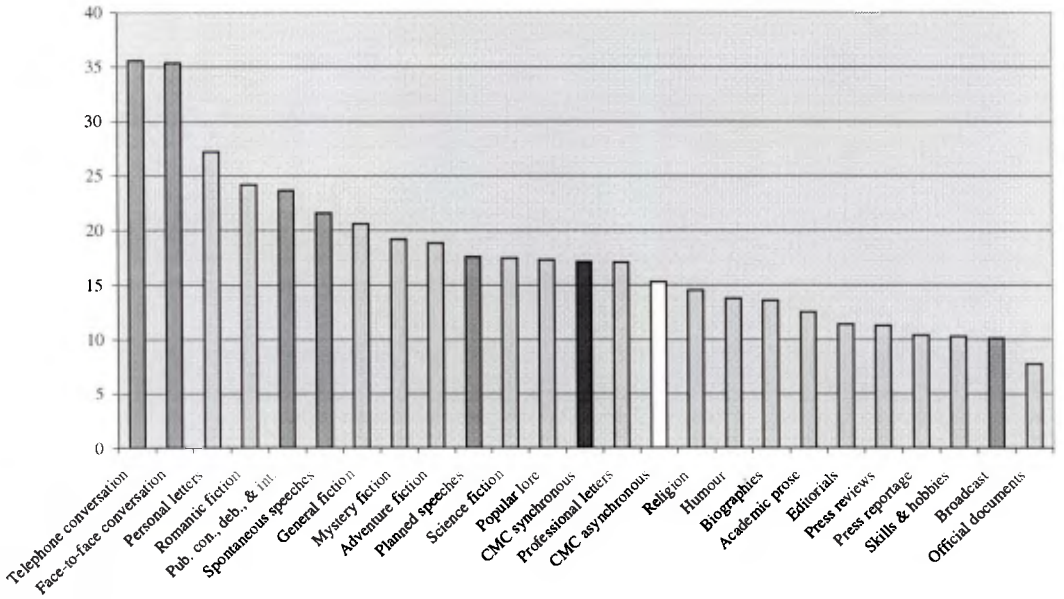

Figure 5.41 Use of *Private Verbs* Between CMC and Non-CMC Genres
F56 *Private Verbs* (*What did you **see** that troubled you?* -rmoo-0402)

Suasive verbs (F57) are those that tend to bring about some changes, such as *agree*, *beg*, *decide*, *instruct*, etc. The verbs are used to overtly mark persuasion. In this study, suasive verbs occur more often, but not significantly so, in CA than in CS texts (5.08 vs. 3.85). As the difference is small, we may assert that the CMC temporalities have no big impact on the use of suasive verbs.

In Biber's factor analysis, suasive verbs are a feature that help constitute the factor of over-persuasion. In Biber's findings, the non-CMC genres that had the highest scores of private verbs are official documents (5.2), press editorials (4.0), and academic prose (4.0). On the other hand, suasive verbs occur least in personal letters (1.0), face-to-face conversation (1.5), and broadcast (1.9) (See Figure 5.42).

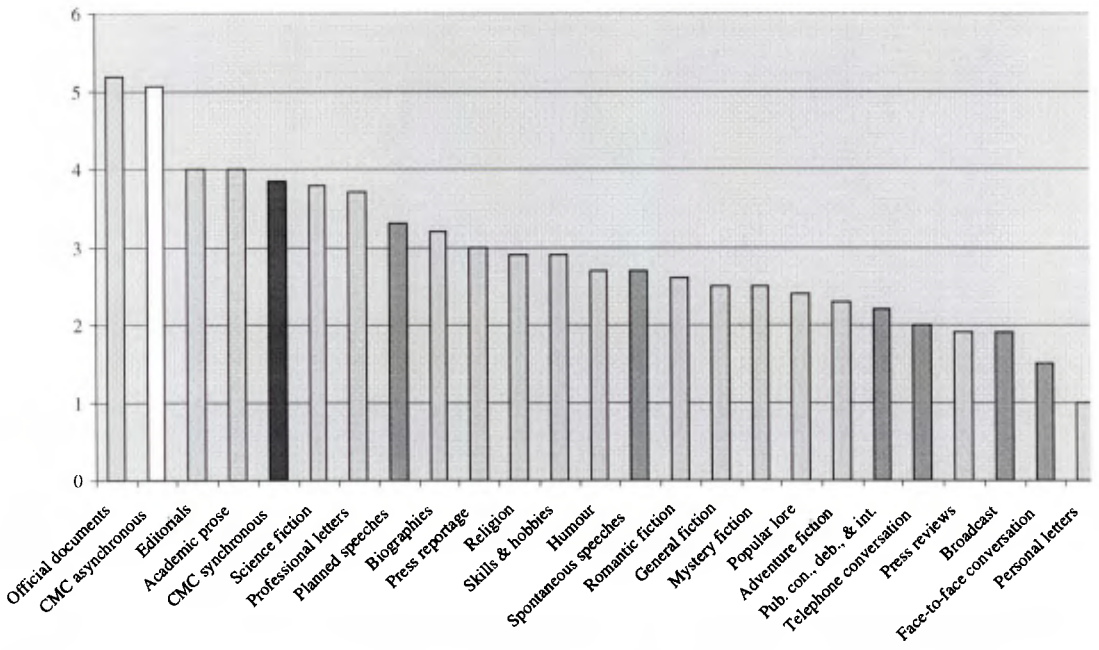

Figure 5.42 Use of *Suasive Verbs* Between CMC and Non-CMC Genres
F57 *Suasive Verbs* (*I do want to* ***recommend*** *that EJC issue I mentioned earlier.* -rmoo-0402)

While there is no significant difference between synchronous and asynchronous CMC texts in terms of the use of suasive verbs, we can see that both CMC genres stand very high in the ranking of the use of suasive verbs. This leads us to believe that CMC, whether synchronous or asynchronous, is a context for more direct expression of ideas. If this is the case, we may also conclude that it is time constraints that force the CMC participants to avoid the more time-consuming indirect ways of euphemism.

5.14.4 Summary of specialised verb class

This section deals with four features in the category of specialised verb class. There are two features, i.e. namely public verbs (F55) and *SEEM/APPEAR* (F58), which show significant difference between the asynchronous and synchronous CMC texts. Two other features, i.e. suasive verbs (F57) and private verbs (F56), do not show any significant differences in the comparison.

The finding that public verbs are associated with the synchronous CMC texts is to be interpreted carefully. As the synchronous CMC texts downloaded from the MOO logs are filled with computer-generated line headings, for examples, "[So and so] says...," the large number of public verbs found in the CMC data do not represent the choice of the addressers

themselves. We should therefore not interpret the finding as to marking a significantly higher use of public verbs in synchronous CMC texts. On the other hand, the finding that *SEEM/APPEAR* are associated with the asynchronous CMC texts is of more significance. It suggests that users in asynchronous CMC communication try to sound more modest and careful than those in the synchronous condition when discussing topics in which they are interested.

For private verbs and suasive verbs, no significant difference was found between synchronous and asynchronous CMC texts. However, we did find that susaive verbs are generally used more often in CMC than in non-CMC texts. This is suspected to be an effect of the limited online time of CMC users, who tend to use suasive verbs more often to overtly persuade others.

5.15 Reduced Forms and Structures

There are five features in the category of reduced forms and structures, namely, contractions (F59), *THAT* deletions (F60), final prepositions (F61), split infinitives (F62), and split auxiliaries (F63). They are all reduced forms used by people mainly as a result of time pressure or other social reasons. All of these five features, except split auxiliaries (F63), demonstrate significant differences across the asynchronous and synchronous CMC texts.

5.15.1 Findings

Four out of the five features in this category show a significant difference between synchronous CMC (CS) and asynchronous CMC (CA) texts, as depicted in Table 5.18. Below are some example sentences from the CMC texts, which show how they were used in the both temporalities.

Contractions (F59):
I've found Gapmaster from Wida Software very good for the types of applications that you want. (ntcho-004)
It seems like we've done it again. (tesln-036)
Hi, everyone. I'm glad to be here. (hmoo-0417)
That's a nice position! (rmoo-0402)

Subordinator-*THAT* deletions (F60):
I do not know where else to turn with my request for help and I know ***[that]*** *I run the risk of being "flayed" now, but necessity knows no law and I need help pretty desperately, so*

here goes. (ntcho-008)
*My husband, who had spent a lot of time in England, said [**that**] it was quite common amongst what we would call upper crust people.* (tesln-024)
*I think [**that**] it's important to *show* others the value.* (hmoo-0509)
*Backy is saying [**that**] navigation is part of presentation ....* (rmoo-0507)

Stranded prepositions (F61):
*How do genres evolve & where do their conventions and standards come **from**?* (ntcho-051)
*At the two Colleges I have taught **at**, students have to pay $ to use the I-net.* (hmoo-0417)
*I see the plan, but it's a tough collab to start **with**.* (rmoo-0507)

Split infinitives (F62):
*I'd like **to** publicly **thank** him for his dedication to NETEACH-L.* (ntcho-114)
*In my experience, mixing (crosslinguistic confusion) is likely to occur when we try **to** artificially **isolate** languages from each other in our minds.* (tesln-006)
*Let's get a good set of ideas for future discussions, and come back in **a** week or so **to** really **hammer** them out!* (hmoo-0424)

Split auxiliaries (F63):
*It **will** also **be** possible to participate in the conference on-line.* (ntcho-012)
*We **might** then **conclude** that the origin of such incorrect phrasing is to be found in their contrastive or counter-intuitive nature.* (tesln-008)
*When seated at one of the tables we **can** only **hear** others at the table.* (hmoo-0417)

From Table 5.18 and Figure 5.43 we can see that contractions (F59) show a significantly higher frequency in CS than in CA texts (17.51 vs. 11.95), as do *THAT* deletions (F60) (3.53 vs. 2.30), and split infinitives (F62) (0.14 vs. 0.13). In contrast, final prepositions (F61) demonstrate a significantly higher frequency in CA than in CS texts (0.75 vs. 0.53).

Table 5.18
Comparison Between Sync. and Asyn. CMC Texts on *Reduced Forms and Structures*

FEATURES	CA > CS	CS > CA	P Value
(N) REDUCED FORMS AND STRUCTURES			
F59 Contraction		17.51 > 11.95	0.001*
F60 THAT del.		3.53 > 2.30	0.002*
F61 Final prep.	0.75 > 0.53		0.000*
F62 Split inf.		0.14 > 0.13	0.000*
F63 Split auxiliary	3.65 > 1.92		0.080
Total cases	1 case significant	3 cases significant	

Remarks: CA: CMC Asynchronous; CS: CMC Synchronous.
*: Significant difference reaches $\alpha = 0.005$.

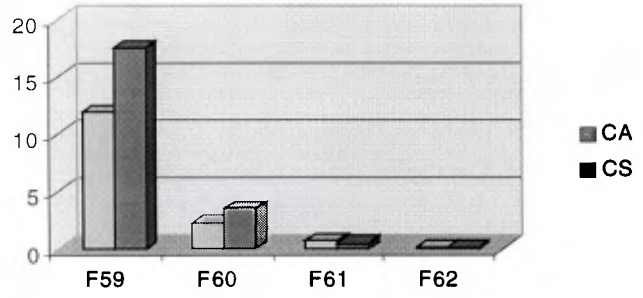

Figure 5.43 Comparison of Features from F59 to F62
F59 *Contractions* (*Hi, everyone. I'**m** glad to be here.* -hmoo-0417)
F60 *Subordinator-THAT Deletions* (*I think [**that**] it's important to *show* others the value.* -hmoo-0509)
F61 *Stranded Prepositions* (*I see the plan, but it's a tough collab to start **with**.* -rmoo-0507)
F62 *Split Infinitives* (*I'd like **to** publicly **thank** him for his dedication to NETEACH-L.* -ntcho-114))

5.15.2 Discussion of reduced forms and structures

Some linguistic features, e.g. contractions, stranded prepositions, and split infinitives, are not encouraged in serious writing. While such prescriptive discouragement is often considered arbitrary, there is an argument to say that those grammatical and rhetorical prescriptions are indeed systematic in that they disprefer those features that involve a mismatch between the surface form and the underlying representation, resulting in either a reduced surface form or a weakened isomorphism between form and meaning (Finegan, 1980, 1987; Biber, 1988a). Biber (1986, 1988a) also argues that those reduced structure features often co-occur with interactive features and with certain types of subordination. However, contractions are believed to be a device for power and solidarity. The contraction is used more often among people who are close to each other, in contrast to a full form which shows power and distance.

Two features in this category involve the surface reduction of forms. They are contractions (F59) and subordinator-*THAT* deletions (F60). In the present study, both contractions (F59) and subordinator-*THAT* deletions (F60) were found to occur significantly more often in the synchronous CMC texts than in the asynchronous context. Contractions demonstrate a high frequency in the CS texts compared with those used in the CA texts (17.51 vs. 11.95).

Contractions are a reduced form that is mainly considered a typical spoken feature as a consequence of fast and easy production. However, Chafe and Danielewicz (1987) find that there is no absolute difference between speech and writing in the use of contractions. On the

contrary, Biber (1987) finds that contractions are used more often in American writing than in British writing.

The deletion of subordinator-*THAT* is another important, though less mentioned, dispreference in edited writing. Few explicit prescriptions prohibit it, but the deletion is argued to be very rare in writing (Beaman, 1984; Finegan & Biber, 1986a; Frawley, 1982; Weber, 1985). In the present study, subordinator-*THAT* deletions, in a similar way to contractions, occur significantly more often in the CS than in the CA texts (3.53 vs. 2.30). It further confirms that the subordinator-*THAT* deletions do occur less in the CA texts, which is similar to the traditional writing context.

The three other features of this category involve a weakened isomorphism between form and meaning. They are final prepositions (F61), split infinitives (F62), and split auxiliaries (F63).

Final prepositions (F61), alternatively named "stranded prepositions," are the feature in which the relative pronoun and the preposition, which should belong to the same phrase in the deep structure, get separated in the surface structure. Chafe (1985) argues that this feature is an indication of spoken "errors." If we examine its frequency along with non-CMC genres in Figure 5.44, we can see quite clearly that all speech genres, with no exceptions, do stand highest on the ranking of the use of stranded prepositions, with public conversation at the highest (5.4), face-to-face conversation at the second (4.8) and so on.

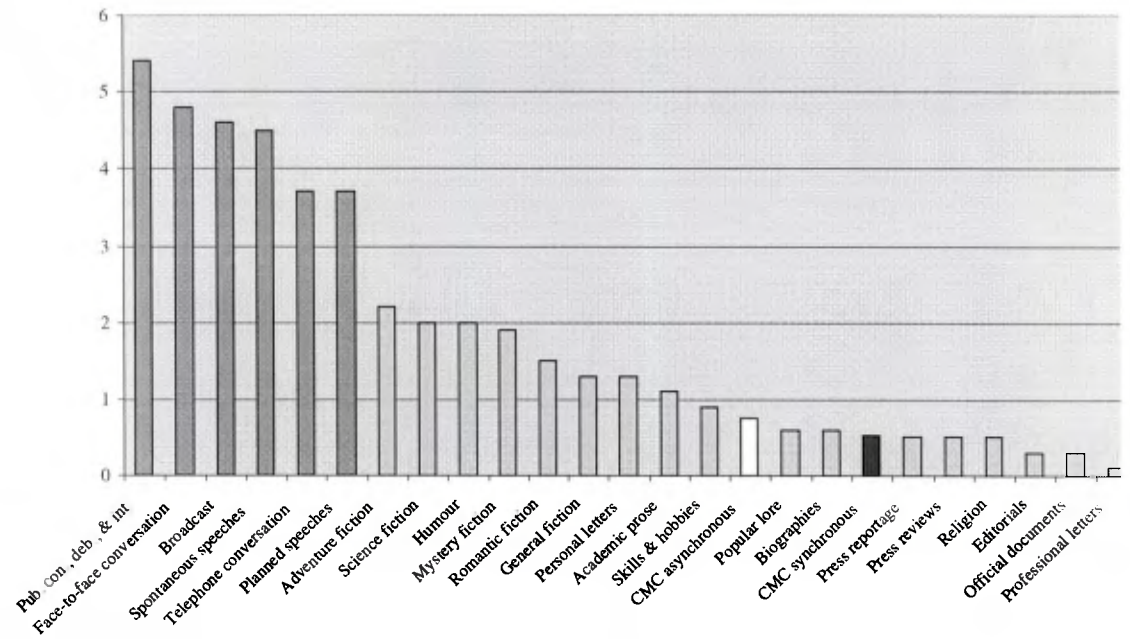

Figure 5.44 Use of *Stranded Prepositions* Between CMC and Non-CMC Genres
F61 *Stranded Prepositions* (*I see the plan, but it's a tough collab to start* ***with***. -rmoo-0507)

However, it is interesting to see that final prepositions do occur more, and even significantly more often in CA texts than in CS texts (0.75 vs. 0.53), while the latter are sometimes considered a very similar context to public conversation. This is indeed a very unusual characteristic found in CA texts, and it is worth further examination. For the time being, it can only be speculated that this may be affected by the low frequency of *THAT* relative clauses on object position (F30), as discussed earlier in Section 5.9.3.

The split infinitive (F62) is another widely discussed condition of mismatching between the surface and the underlying structures. In non-CMC texts, there appears to be no difference between speech and writing, as it is uncommon and reduced in both of these forms (Chafe, 1984). In the present study, however, this feature did occur with significantly higher frequency in the CS than in the CA texts (0.14 vs. 0.13), as shown in Figure 5.43. This finding is not in agreement with that of previous studies (Biber, 1986; Chafe, 1984), in which the feature of split infinitives is found to be equally uncommon in both spoken and written genres. This is illustrated in Figure 5.45, where the frequencies of the use of split infinitives in CMC and non-CMC genres are listed. No plausible explanation is available at the moment for this finding, though it does appear to be a very interesting phenomenon for further exploration.

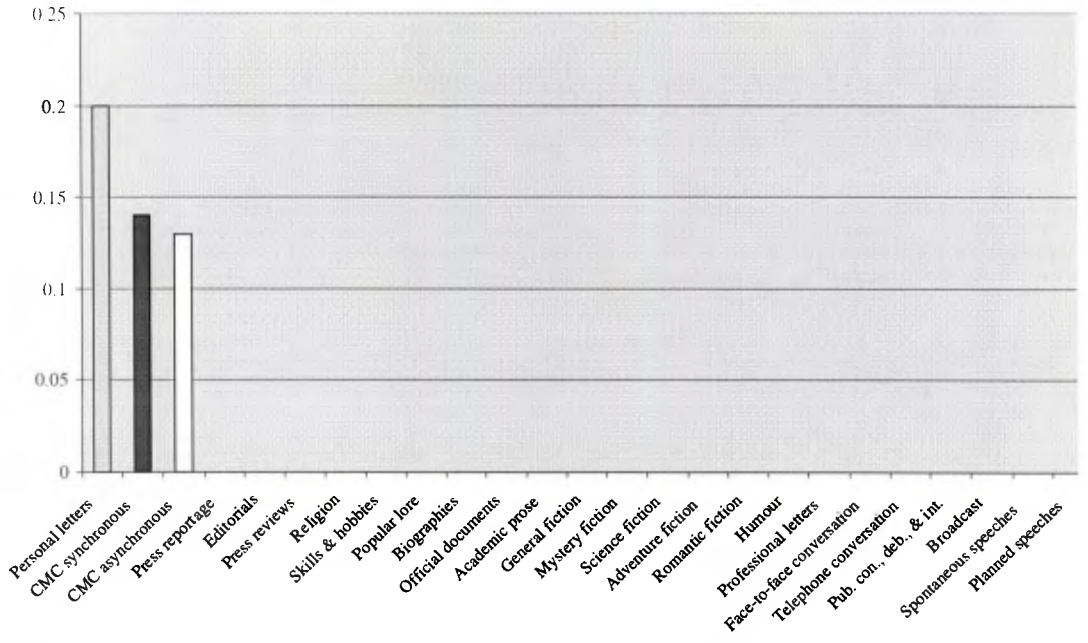

Figure 5.45 Use of *Split Infinitives* Between CMC and Non-CMC Genres
F62 *Split Infinitives* (*I'd like* ***to*** *publicly* ***thank*** *him for his dedication to NETEACH-L.* -ntcho-114)

Split auxiliaries occur when adverbs are placed between the auxiliaries and their main verb. This feature is found to exist more often in certain written genres than in typical conversation (Biber, 1988a). However, in the present study, this feature did not show any significant difference, though it did occur more often in the CA text (3.65 vs. 1.92). In comparing this feature in all of the CMC and non-CMC genres in Figure 5.46, we can also see that it is generally more popular in writing than in speech genres.

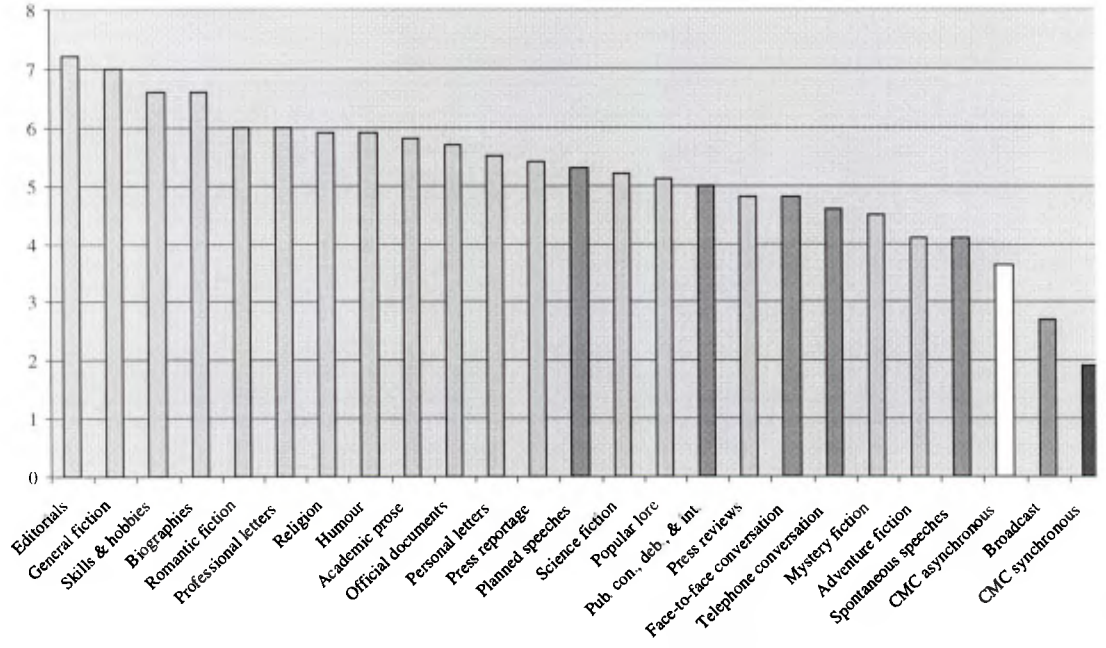

Figure 5.46 Use of *Split Auxiliaries* Between CMC and Non-CMC Genres
F63 *Split Auxiliaries* (*It **will** also **be** possible to participate in the conference on-line.* -ntcho-012)

5.15.3 Summary of reduced forms and structures

This section deals with findings related to the category of reduced forms and structures. Among the five features, only split auxiliaries (F63) do not show any significant difference across the temporalities of CMC texts.

Three of the features, contractions (F59), *THAT* deletions (F60), and split infinitives (F62), showed significantly higher frequencies in CS than in CA texts, while only stranded prepositions (F61) occurred significantly more often in CA than in CS texts. The findings on the features of reduced forms and structures show that, generally speaking, most of these features are associated with the synchronous mode of the CMC medium. This is not surprising as, in the synchronous temporality, it is easier for the "speakers" to use the reduced forms and structures, as well as the contractions. The significantly higher frequency of stranded prepositions (F61) in CA than in CS texts is out of expectation and is thought to be

related to the low frequency of *THAT* relative clauses on object position (F30), as discussed earlier in Section 5.9.3.

5.16 Coordination

This section describes the findings from the features in the category of coordination, which covers phrasal coordination and non-phrasal coordination. Coordination is when words like *and*, *but*, *or*, and *so* are used as coordination devices. However, most of these words have other functions and cannot be counted precisely by the computer as coordinators (Biber, 1988a). In the present study, the operational definition of the feature coordination is limited to the use of *and*. The coordinator *and* can serve either the function of phrasal coordination (F64), connecting adjectives, adverbs, verbs, or nouns together in a phrase, or that of non-phrasal coordination (F65), connecting independent clauses. Between the two, non-phrasal coordination has been found to have significant differences in the present study between the synchronous and asynchronous CMC texts.

5.16.1 Findings

Phrasal coordination (F64) and non-phrasal coordination (F65) are the two features that constitute coordination. In Table 5.19, as well as Figure 5.47, the statistical findings of the coordination device *and* serving as both phrasal and non-phrasal coordinators are depicted to show how the use of *and* may differ between the two temporalities of CMC texts.

Phrasal coordination (F64):
I did find Kombucha, a discussion about using ***and*** *consuming the Kombucha beverage.* (ntcho-002)
I'd like to get to know the differences between British ***and*** *American in grammar* ***and*** *vocabulary. (tesln-002)*
JoelE smiles ***and*** *nods GregS.* (rmoo-0402)

Non-Phrasal coordination (F65):
TESOL FRANCE Annual Colloquium Grammar, how it is learnt ***and*** *how it is taught 25 - 26 October 1996 Paris, France.* (ntcho-009)
I did, ***and*** *nothing happened.* (hmoo-0417)
The long-time tenured here were given it gratis ***and*** *then they established the hoops.* (rmoo-0402)

From both the table and the figure below, it is clear that *and* used as a non-phrasal coordinator (F65) reached a significantly higher frequency in CA than in CS texts (1.55 vs.

1.03). The other feature, phrasal coordination (F64), though also occurring more highly in CA than in CS texts, did not reach the significant level of difference (α =0.005).

In Table 5.19, we can see that significant difference between asynchronous and synchronous CMC texts occurs on non-phrasal coordination (F65), but not on phrasal coordination (F64). A closer look at Figure 5.47 gives a clearer idea of how the non-phrasal coordination occurred more often in CA than in CS texts.

It would seem, then, that, in the asynchronous more than in the synchronous temporality of CMC communication, *and* is used to connect clauses together. This indicates that, in the asynchronous CMC, like email or discussion lists, clauses are often expanded with the coordinator *and.*

Table 5.19
Comparison Between Sync. and Asyn. CMC Texts on *Coordination*

FEATURES	CA > CS	CS > CA	P Value
(O) COORDINATION			
F64 Ph. coord.	4.54 > 2.38		0.044
F65 Non-ph. coord.	1.55 > 1.03		0.001*
Total cases	1 case significant	0 case significant	

Remarks: CA: CMC Asynchronous; CS: CMC Synchronous.
*: Significant difference reaches α = 0.005.

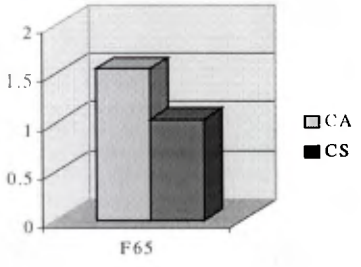

Figure 5.47 Comparison of Feature from F65
F65 Non-Phrasal Coordination (*I did, **and** nothing happened.* -hmoo-0417)

5.16.2 Discussion of coordination

Subordination and coordination are two features often used to distinguish the styles of writing and speech. Kress (1994, p. 28) says, "speech is structured by sequences or chains of clauses, which are, generally speaking, co-ordinated."

Our finding here is that, in the asynchronous more often than in the synchronous temporality of CMC communication, the coordinator *and* is used to connect clauses together. This indicates that, in the asynchronous CMC, like email or discussion lists, clauses are often expanded with the coordinator *and.* This finding is quite interesting, compared with the argument of Kress (1994), who claims that the clausal structure is "both evidence of immediate thinking, in that the speaker does not have time to assemble complex structures, and evidence of the needs of the hearer as the recipient of information" (p. 28). Kress observes that the traditional speech tends to be coordinated with a chain of clauses. In this study, however, it seems that CA writers somehow adopt strategies typically used by speakers in the spoken context to which Kress refers.

The relatively less often use of *and* as a non-phrasal coordinator (F65) in the CS texts suggests that, in the simultaneous CS context, the chaining of clauses occurs relatively less often. It may be that the physical behaviour of keyboard typing has somehow hindered such a long sequence of clause chains and has resulted in a certain impact on the style of CMC texts.

We can also compare the frequencies of the coordination device *and* as used in CMC and in non-CMC genres. In Figures 5.48 and 5.49 below, we can see that the rankings of CA and CS are different when phrasal coordinator and non-phrasal coordinator are concerned. Figure 5.48 shows that CA texts stand much higher than CS texts in the ranking of the frequency of the phrasal coordination device *and.* Besides, we can see that all the speech genres stand at the bottom of the ranking in the use of *and* as a phrasal coordinator. This quite clearly indicates that writing genres use the phrasal coordinator *and* (F64) relatively more often than speech genres.

In Figure 5.49, we can see that both CA and CS texts stand very low in the ranking of the frequency of the non-phrasal coordination device *and.* Besides, we can see that all the speech genres use *and* as a non-phrasal coordinator more often than writing genres.

The observation from Figures 5.48 and 5.49 above further verifies the earlier observation by Kress (1994) that speech is characterised by sequences of clauses coordinated by the use of *and.* On the other hand, we find in the present study that CMC users tend to use the coordinator *and* more often in the asynchronous than in the synchronous temporality.

The relatively lower use of the non-phrasal coordinator *and* in synchronous CMC (CS) texts is most likely to be a result of the time constraints experienced by users in the CS mode. This has been observed in Section 5.9, where we discussed the use of subordinators. There

we found that the users of synchronous CMC (CS) tend to use less subordination constructs than those of asynchronous CMC (CA) users.

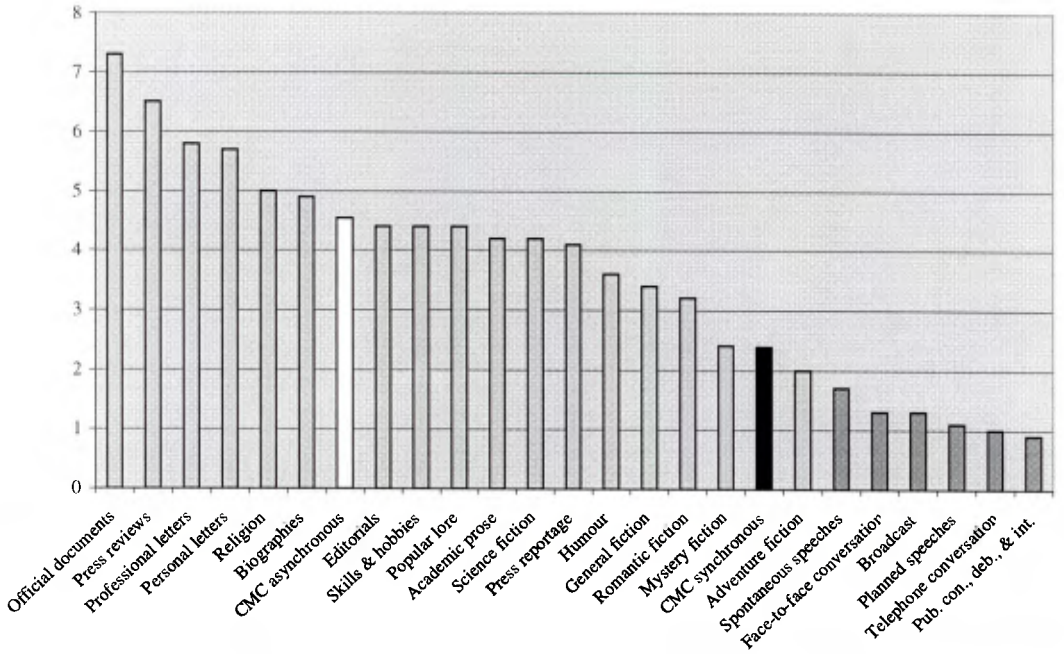

Figure 5.48 Use of *Phrasal Coordinator* Between CMC and Non-CMC Genres
F64 Phrasal Coordination (*JoelE smiles **and** nods GregS.* -rmoo-0402)

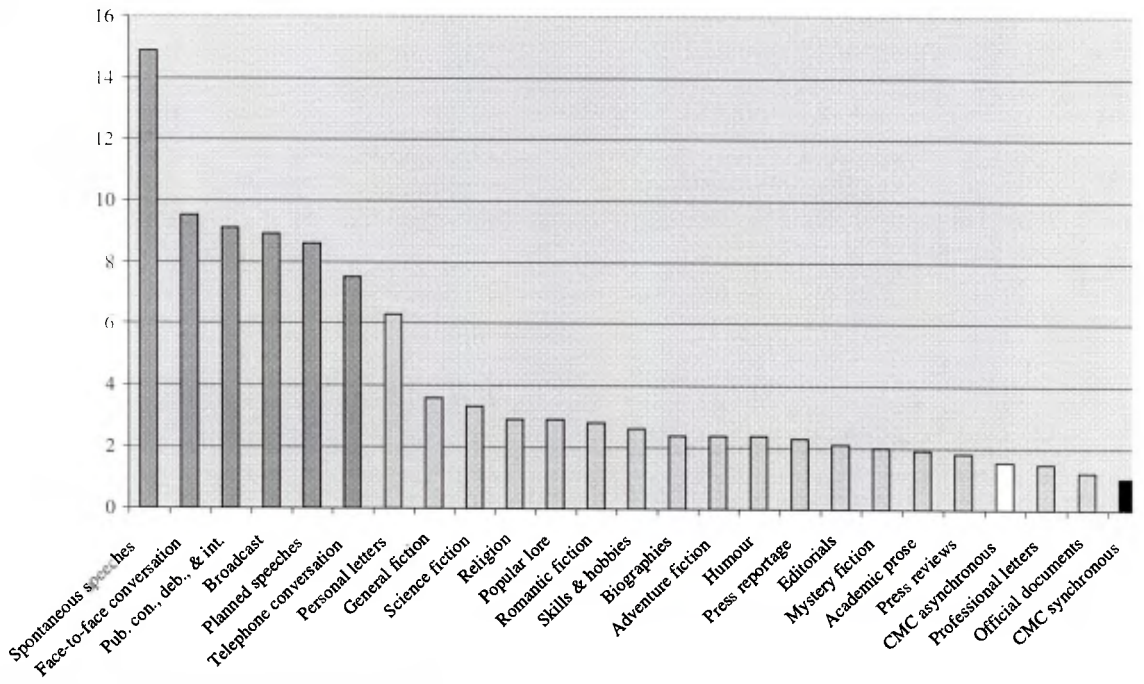

Figure 5.49 Use of *Non-Phrasal Coordinator* Between CMC and Non-CMC Genres
F65 Non-Phrasal Coordination (*I did, **and** nothing happened.* -hmoo-0417)

It is believed that the strict time constraints in synchronous CMC force the communicators to choose simpler and shorter sentences, rather than producing long and complex sentences with the help of subordination devices. While people are found to use the coordinator *and* to link sequences of clauses in speech genres, they are constrained by time and therefore use simpler sentences when communicating synchronously in CMC.

We must understand that CS participants have to communicate synchronously but with a slow means of message production, i.e. typing. The speed of typing is definitely much slower than that of vocal speech. If they still adopt the habit of using a string of coordinated clauses and lengthy sentences, as they would in speech, the synchronicity of communication would be hindered. To facilitate a frequent exchange of messages and give the turn-taking right to the counterpart, participants must use shorter expressions and simpler sentences.

5.16.3 Summary of coordination

The observation of the two features in the category of coordination reveals that the coordinator *and* is used more often as a non-phrasal coordinator in the asynchronous than in the synchronous CMC texts. As non-phrasal coordination, *and* is used to mark a series of coordinated clauses, its high frequency indicates a larger number of clauses, and consequently a lower information density than is typically found in spoken language (Halliday, 1989). The finding in this study is interesting, but quite in line with similar findings on the use of subordinators in Section 5.9. The reason for the significantly lower use of coordinators and subordinators in synchronous CMC is believed to be related to the time constraints on synchronous CMC users. Those users in the asynchronous temporality may feel more at ease to prolong their texts with coordination than synchronous CMC users, who are busy responding to their counterparts.

5.17 Negation

There are two features covered in the category of negation: the synthetic negation (F66) and analytic negation (F67). Synthetic negation is when the word *no* is placed before a noun, such as *no computers*, *no attention,* or before a noun with an adjective, *no further hesitation, no qualified captain.* It does not extend, however, to when *no* is used as a response, such as *No, I didn't*, in replying to *Did you read it?* Synthetic negation also includes *neither...nor*,

where neither of the entities referred to in a statement is accepted. On the other hand, the analytic negation is the *not* used as an adverb in a non-affirmative construction, such as *This is not the correct way.* In the present study, only the synthetic negation had been found to exhibit significant difference between the two temporalities of CMC texts.

5.17.1 Findings

Among the two features covered in the category of negation, only the difference of synthetic negation (F66) reached the significant level (α=0.005), though both appeared to have higher frequency in CS than in CA texts. The relevant findings are illustrated in Table 5.20, as well as Figure 5.50 below. Some examples of these two features are listed here to give an idea of how they were used in the CMC texts:

Synthetic negation (F66):
I have ***no*** *computers and* ***no*** *access so what can I do?* (ntcho-089)
I seem to remember this seeing this as a discussion topic that I paid ***no*** *attention to some time ago, therefore, apologize for raising it again.* (tesln-003)
I don't think we gain much by comparing them, really, ***nor*** *by trying to accomplish the same ends. (rmoo-1029)*
No. They're not and ***neither*** *am I.* (hmoo-0509)
Nor *what their level of MOO experience would be.* (hmoo-0417)

Analytic negation (F67):
Also, please feel free to pop in early and contribute to the list if you can't attend the cafe. (ntcho-001)
Answers to me please, ***not*** *the list.* (tesln-009)
Why don't you try it too? (hmoo-0417)
He must ***not*** *have been here, Marcy.* (rmoo-0402)

From the table and the figure, we can see that both the synthetic negation (F66) and analytic negation (F67) occurred more often in the CS texts. The synthetic negation is higher in CS than in CA texts (0.72 vs. 0.67), and so is the analytic negation (7.86 vs. 7.50). However, only the difference in the former appeared to be significant.

Table 5.20
Comparison Between Sync. and Asyn. CMC Texts on *Negation*

FEATURES	CA > CS	CS > CA	P Value
(P) NEGATION			
F66 Synthetic neg.		0.72 > 0.67	0.000*
F67 Analytic neg.		7.86 > 7.50	0.234
Total cases	17 cases significant	12 cases significant	

Remarks: CA: CMC Asynchronous; CS: CMC Synchronous.
*: Significant difference reaches α = 0.005.

5.17.2 Discussion of synthetic negation

The two negation structures, synthetic negation (F66) and analytic negation (F67), are examined in this study. Synthetic negation is the use of *neither*, *nor*, and *no* (not as a response) to effect negation. Analytic negation is the use of *not* in either the full or the contracted form. Earlier studies found that the structure of negation is used much more frequently in speech than in writing, suggesting a higher frequency of repetitions, denials, rejections, questions, and mental verbs in speech (Tottie, 1981, 1982). Tottie (1983) also distinguished synthetic negation from analytic negation, the former being more literary and integrated, the latter being more colloquial and fragmented.

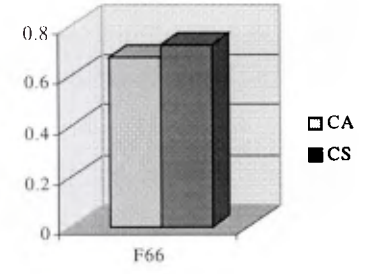

Figure 5.50 Comparison of Feature from F66
F66 Synthetic Negation (*I have **no** computers and **no** access so what can I do?* -ntcho-089)

In Biber's study (1988a), the feature synthetic negation was found to be an underlying factor of narration. It was thought to be related to its stronger emphasis than analytic negation, e.g. "*he said nothing*" versus "*he did not say anything*" (p. 109). With this in mind, it is rather surprising to find that in CMC texts, synthetic negation was used significantly more often than analytic negation. It would be interesting to study why synchronous CMC participants use a more emphatic type of negation than those in the asynchronous CMC.

If we compare the synthetic negation used in CMC texts with that used in non-CMC texts, it is also interesting to see that, though the two differ significantly from each other, both are very low in frequency when they stand among the 23 non-CMC genres. In Figure 5.51, it can be seen that non-CMC genres with a high frequency of synthetic negation are mostly written. This is in line with Tottie's argument (1983) that synthetic negation was more literary. In contrast, the use of synthetic negation in CMC texts should not be considered a characteristic.

5.17.3 Discussion of analytic negation

While synthetic negation showed significant differences in the CS texts, the analytic negation also had a higher frequency in the CS texts, though it did not reach the significant level.

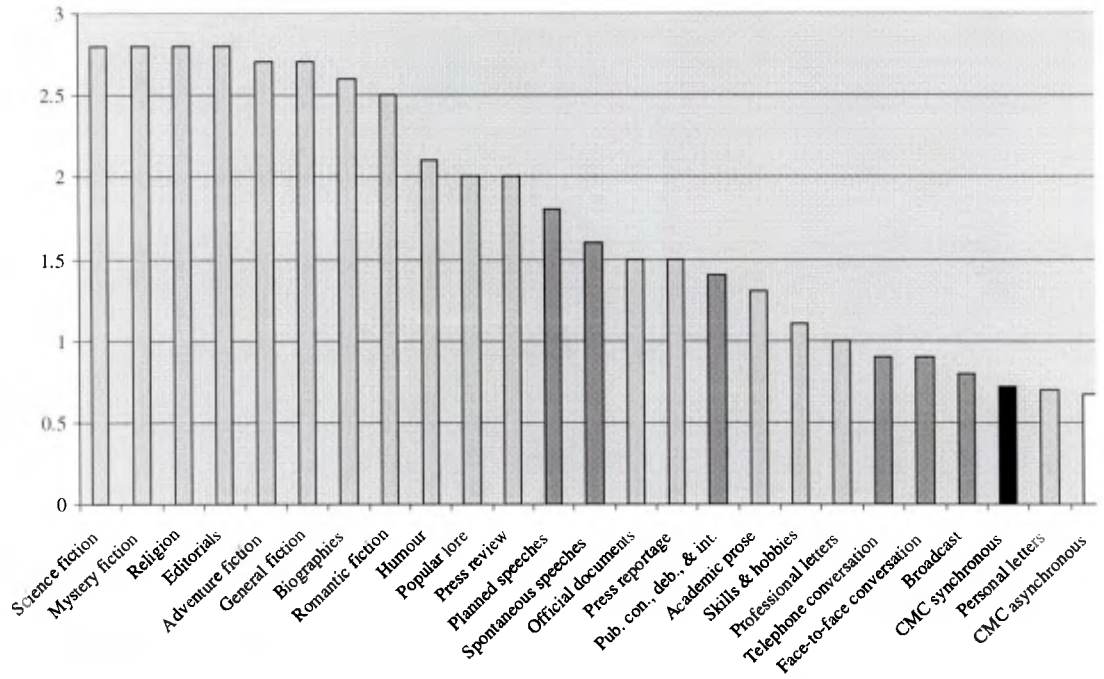

Figure 5.51 Use of *Synthetic Negation* Between CMC and non-CMC Genres
F66 Synthetic Negation (*I have **no** computers and **no** access so what can I do?* -ntcho-089)

The analytic negation has a characteristic of interaction that is similar to many other interactive features in this study, such as second and third person pronouns, *DO as* pro-verbs (F07, F08 and F12, all discussed in Section 5.4), *WH*-questions (F13, Section 5.5), and hedges, emphatics and discourse particles (F47, F49 and F50, Section 5.12). All of these features occur more often in synchronous than in asynchronous CMC texts. The analytic negation falls within this group of features. In other words, the synchronous CMC is supposed to be more interactive than the asynchronous CMC. However, the difference here did not reach the significant level 0.005.

A comparison of the two CMC genres with the 23 non-CMC genres on the analytic negation (F67, see Figure 5.52), shows that the two CMC genres stand somewhere in the middle of all the genres being compared. Most synchronous CMC texts, with the exception of broadcast, have a higher frequency of analytic negation, a finding quite in line with Tottie's argument.

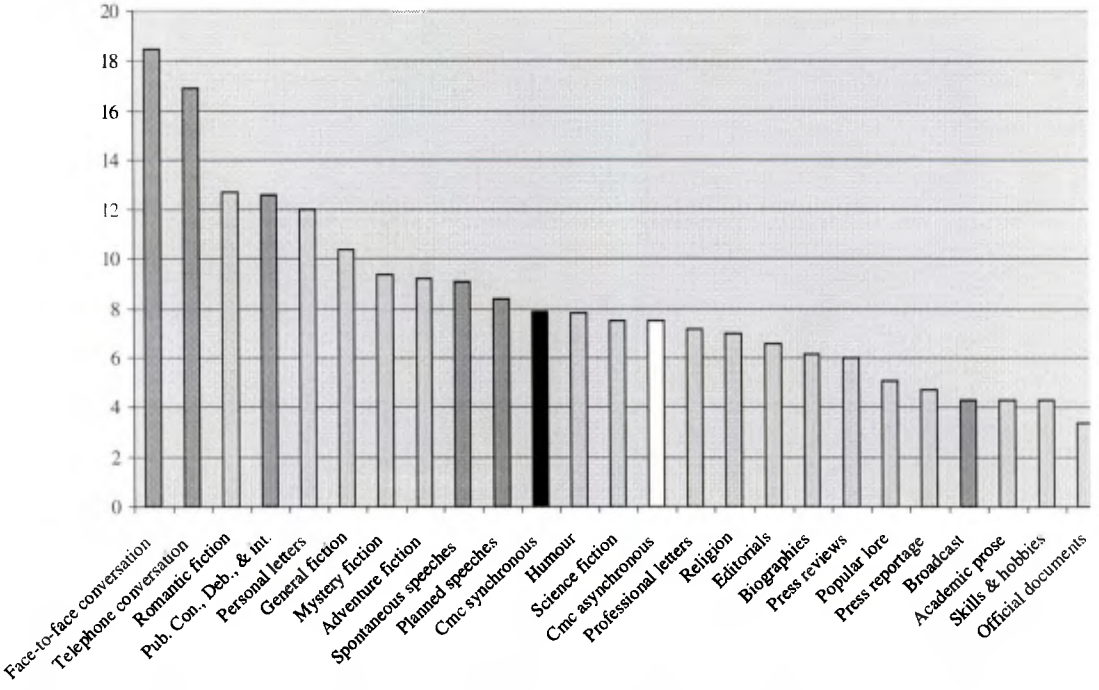

Figure 5.52 Use of *Analytic Negation* Between CMC and non-CMC Genres
F67 Analytic Negation (*He must **not** have been here, Marcy.* -rmoo-0402)

5.17.4 Summary of negation

This section deals with findings related to the category of negation. There are two types of negation examined: synthetic negation (F66), the use of *neither*, *nor*, and non-response *no,* and analytic negation (F67), the use of *not.* The synthetic negation (F66) showed a significantly higher frequency in the CS texts than in CA. However, F66 was actually very low in frequency in both CA and CS texts when compared with non-CMC genres. It is interesting to find that synthetic negation is used much less often in CMC than in most non-CMC genres. It confirms that CMC texts, whether synchronous or asynchronous, do not exhibit the characteristic of narration. The finding of this section also confirms that the analytic negation exhibits a colloquial and interactive nature, as most speech genres exhibit a higher frequency than written.

5.18 Summary of Findings and Discussion

5.18.1 Introduction

This chapter presented the findings on the nature of CMC texts in the synchronous and

asynchronous temporalities. A total of 67 linguistic features under 16 categories were examined for their frequencies of occurrence in the CMC texts, and the results of the statistical procedures of the Non-Parametric Statistical Test was reported and discussed. For some features, a comparison was also made on the use of the feature between CMC and non-CMC genres. This did not involve the results of any strict referential statistics, but still generated some very interesting ideas.

With the level of significance set at α=0.005, significant difference was found in 29 of the 67 features, as shown in Table 5.1 of Section 5.1.2 (pp. 132-135). Among the 67 features, 17 were found to be relatively higher in frequency in asynchronous CMC (CA) texts, and 12 features were found to be relatively higher in frequency in synchronous CMC (CS) texts. There are 38 features that do not show any significant differences in frequency between CA and CS texts. The 17 features on which CA texts have a significantly higher frequency are nominalisations (F14), *BY* passives (F18), *THAT* clause as adjective complements (F22), *WH* clauses (F23), past participial clauses (F26), past participial WHIZ deletions (F27), *THAT* clauses on subject position (F29), *THAT* clauses on object position (F30), *WH* clauses on object position (F32), sentence relatives (F34), causative adverbial subordinators: *because* (F35), concessive adverbial subordinators: *although, though* (F36), other adverbial subordinators (F38), attributive adjectives (F40), *SEEM/APPEAR* (F58), final prepositions (F61), and non-phrasal coordination (F65) (see pp. 132-135).

The 12 features in which the CS texts exhibit a significantly higher frequency than CA texts include: present tense verbs (F03), time adverbials (F05), *DO* as a pro-verb (F12), *WH* questions (F13), present participial clauses (F25), general hedges (F47), discourse particles (F50), public verbs (F55), contractions (F59), *THAT* deletions (F60), split infinitives (F62), and synthetic negation (F66).

5.18.2 Findings in the different categories

The 29 features that showed a significant difference across the asynchronous and synchronous CMC texts have been grouped together on the basis of their nature and are discussed under 16 categories in the previous sections from Section 5.2 to Section 5.17. These findings will be discussed in more depth in Section 7.1. A brief review of each category is summarised below.

5.18.2.1 Tense and aspect markers

Among the three tense and aspect markers, present tense verbs (F03) showed a significantly higher frequency in the CS than in the CA texts, making CS significantly more of the involved style than CA. On the other hand, past tense (F01) and the perfect aspect (F02) are quite seldom used when compared with the cases of non-CMC genres. As they are taken as markers of the discourse of narration, CMC seems a medium not suitable for such discourse. It may be that this is connected with the nature of the short length of texts and the efficiency of communication that is emphasised in the CMC environment.

5.18.2.2 Place and time adverbials

The feature of time adverbials (F05) showed a significantly higher frequency in the CS texts than in the CA. This finding suggests the relative importance of the time referents in the synchronous temporality of CMC. On the other hand, it is noteworthy that such relative difference is not found for place adverbials (F04). This suggests that the conception of the space relationship is not much different between the synchronous and asynchronous CMC users.

5.18.2.3 Pronouns and pro-verbs

The pro-verb *DO* (F12) showed a significantly higher frequency in synchronous CMC texts. This is thought to be a technique of synchronous CMC participants, used to shorten their utterances under the time constraints. However, it is also found that both types of CMC use pro-verbs quite rarely in comparison with non-CMC genres. This may be the result of the limitation of time and text length that forces the CMC users to use more specific verbs.

5.18.2.4 *WH*-questions

Participants in synchronous CMC, more often than in asynchronous CMC, use *WH-questions* (F13) as a strategy to enhance interpersonal relationships, as well as to get more information. It may be the presence of an immediate addressee that helps foster the context in which direct questions are often raised. It is also interesting to see that people, when under the limitation of time and text length in either synchronous or asynchronous CMC settings, use the direct question forms more often than they do in non-CMC settings.

5.18.2.5 Nominal forms

Among the three features of nominal forms, nominalisations (F14) showed a significantly higher frequency in the CA than in the CS texts. Another interesting finding in this category is that CMC users use nominal forms quite often, suggesting that they tend to use words of higher lexical density to offset the pressure of time constraints.

5.18.2.6 Passives

The two passive features examined in this study both show a tendency to occur more frequently in asynchronous CMC (CA) than in synchronous CMC (CS) texts. The difference reached the level of significance (α=0.005) in *BY*-passives (F18), though not in agentless passives (F17). The difference is believed to be connected with the more formal register of passive devices, which the participants applied in the asynchronous mode of CMC.

5.18.2.7 Stative forms

Neither of the two features in this category of stative forms, *BE* as a main verb (F19) and the existential *THERE* (F20), showed significant differences between the two temporalities. These forms suggest a fragmented presentation of information and low information density, and occur mostly in speech genres. The fact that CMC genres generally have low frequencies seems to suggest that CMC users actually try to increase the information density under the limitation of time constraints and have to refrain from using the fragmented discourse patterns.

5.18.2.8 Subordination

Among the 18 features examined, 12 demonstrate significant differences between CA and CS texts, and 11 out of the 12 occur significantly more often in the CA texts. Only one feature, the present participial clause (F25), appears significantly more often in CS than in CA texts.

The relatively lower frequency of most subordination devices used in synchronous CMC is thought to be due to the tense time constraints, which force the participants to choose their words more carefully, and use shorter, simpler sentences. Due to the limitation on the speed of typing, the general practice of using subordinate and coordinate sentences in speech is not practical in synchronous CMC. To avoid the synchronicity of communication being hindered, CMC participants tend to use less devices of subordination in the synchronous mode than they do in the asynchronous mode.

5.18.2.9 Adjectives and adverbs

Attributive adjectives (F40) is the only feature in this category that showed a significantly higher frequency in CA than in CS texts. This is thought to be related to the association of attributive adjectives with nominal forms. Besides, as it is found that adjectives and adverbs are used much less in CMC than in non-CMC texts, we can conclude that CMC users do not tend to use these features to achieve integration and high information density in their texts. The features more typically used in CMC for text integration purposes are nominalisation (F14), higher type/token ratio (F43), and longer words (F44).

5.18.2.10 Lexical specificity

Both type/token ratio (F43) and word length (F44) serve to mark the density of information. The general belief is that these two features are more characteristic of rich content and higher lexical density. They would be expected to occur more often in asynchronous than in synchronous CMC texts. This, however, was not confirmed in the findings.

It is interesting to find that CMC texts tend to have a higher type/token ratio and higher word length than non-CMC speech genres, i.e. face-to-face conversation and telephone conversation. This shows how CMC users try to incorporate more dense information within the limitation of time and space. This is believed to be associated with the specific topic being discussed and the background of participants in the sampled CMC texts, and also to the time constraints and the physical workload of typing in the CMC setting.

5.18.2.11 Lexical classes

Two features examined in this category exhibited significant differences between the two temporalities. They are general hedges (F47) and discourse particles (F50), and both of them occur more often in the synchronous than in the asynchronous CMC texts. The general hedges and discourse particles are a popular practice in synchronous CMC as a means to mark uncertainty or facilitate coherence in a colloquial way, just as in speech.

5.18.2.12 Modals

The category of modals includes three features, namely possibility modals (F52), necessity modals (F53), and predictive modals (F54). In the present study, none of the three features was found to demonstrate significant differences between synchronous and

asynchronous CMC texts, though they generally occur more often in the asynchronous CMC temporality.

5.18.2.13 Specialised verb class

Two features in this category, i.e. public verbs (F55) and *SEEM/APPEAR* (F58), showed significant differences between the asynchronous and synchronous CMC texts.

The finding that public verbs occur much more highly in synchronous CMC than in asynchronous texts is suspected to be due to the large number of computer-generated line headings "[So and so] says..." The large number of public verbs thus found in the CMC data does not represent the choice of the addressers themselves. The more significant finding is that *SEEM/APPEAR* are associated with the asynchronous CMC texts. As this feature is one of those marking the academic hedging, it suggests that users in asynchronous CMC try to sound more modest and careful than those in the synchronous condition when discussing topics of their interest.

We also found that susaive verbs are generally used more often in CMC than in non-CMC texts. This is suspected to be an effect of the limited online time of CMC users, who tend to use suasive verbs more often to overtly persuade others.

5.18.2.14 Reduced forms and structures

Among the five types of reduced forms and structures examined in this category, three of them, contractions (F59), *THAT* deletions (F60), and split infinitives (F62), showed a significantly higher frequency in CS than in CA texts, while stranded prepositions (F61) occurred significantly more often in CA than in CS texts.

The finding that most of these features are associated with the synchronous mode of the CMC medium is not surprising as, in the synchronous temporality, it is easier for the "speakers" to use the reduced forms and structures, as well as the contractions. The significantly higher frequency of stranded prepositions (F61) in CA than in CS texts is out of expectation, and is thought to be related to the low frequency of *THAT* relative clauses on object position (F30).

5.18.2.15 Coordination

The coordinator *and* is found to be used more often as a non-phrasal coordinator (F65) in the asynchronous than in the synchronous CMC texts, to link clauses together. Its high

frequency indicates a larger number of clauses, and consequently a lower information density. The finding in this study that the non-phrasal coordinator has a relatively higher frequency in the asynchronous CMC is interesting. It suggests less use of coordinators in synchronous CMC under the time constraints, so that synchronous CMC users, unlike speakers, have to cut their sentences short.

5.18.2.16 Negation

Between the two types of negation examined, synthetic negation (F66) and analytic negation (F67), the former showed a significantly higher frequency in CS than in CA texts. However, synthetic negation was actually very low in frequency in both CA and CS when compared with non-CMC genres. As it is supposed to be a marker of narration, the lower frequency of synthetic negation in CMC than in most non-CMC genres confirms that CMC texts, whether synchronous or asynchronous, do not exhibit the characteristic of narration. The finding in this section also confirms that the analytic negation exhibits a colloquial and interactive nature, as most speech genres exhibit a higher frequency of it than written genres.

5.18.3 General discussion

Several observations can be made from the above findings. Firstly, of the 67 features studied, CMC texts vary between the two temporalities in 29 cases, which is less than half. In each of the 16 categories of features, significant differences can be found in only some. Therefore, it is hard to decide which category of features can serve as a marker to distinguish the temporalities of CMC texts.

Secondly, among the features on which significant difference was noted, the asynchronous CMC (CA) texts usually use relatively more subordination and coordination features than synchronous CMC (CS) texts. This suggests longer and more complex sentence patterns in CA. The more frequent use of *BY*-passives may stand for a more abstract characteristic of the CA texts, while the more frequent use of *SEEM/APPEAR* in CA marks a more formal style of hedging.

Thirdly, the more frequent occurrence of certain features in synchronous CMC (CS) texts, such as general hedges (F47), discourse particles (F50), contractions (F59) and *THAT* deletions (F60) suggests a more colloquial style of utterance in these texts. The frequent use of present tense verbs (F03) marks the CS texts as a symbol of involved discourse.

The above comparison suggests, on the features where asynchronous CMC (CA) and synchronous CMC (CS) texts show significant differences, a tendency for the CA to adopt more complex structures and a somewhat more formal tone, while CS communication tends to be more colloquial and simple in structure.

If we also compare the features used in CMC texts with those used in non-CMC texts, there are also several interesting findings:

(1) CMC texts rarely use *DO* pro-verbs (F12), as do non-CMC texts. It seems the CMC users prefer more precise verbs and do not have time to recycle the same topic for too long. The fact that CMC users use more suasive verbs (F57), but rarely use stative forms, like *BE* as a main verb (F19) and the existential *THERE* (F20), also indicates the efforts of CMC users to increase the information density.

(2) CMC users raise *WH*-questions (F13) more often than non-CMC speakers/writers, which suggests a more direct manner of discourse.

(3) CMC users use subordination and coordination features much less than non-CMC speakers to facilitate a faster pace of communication.

(4) CMC users exhibit a higher information density than speakers in non-CMC settings. This can be found in the higher type/token ratio (F43) and the longer words (F44) used in CMC texts over in non-CMC genres.

The above findings all suggest that CMC users utilise certain measures to increase the information density in the texts, especially in the synchronous mode, when the time constraints are high. The speakers in non-CMC settings are also under time constraints, but the type of constraints seems very different. In non-CMC settings, the time constraints urge the speaker to speak more and ignore a precise choice of words and expressions. In CMC settings, the users have to communicate to each other by typing on the keyboard, and there are few non-verbal cues like facial expression or prosodic features. The use of measures to keep the texts short and increase information density seems a widely used strategy by CMC users to give all of the participants adequate chance of turn-taking and to keep communication continuing smoothly.

Though CMC users try to achieve higher information density, they do so mainly by the measures described above. The use of attributive adjectives and adverbs is rare in comparison with the cases of non-CMC texts.

5.18.4 Conclusion

The findings reached so far on the characteristics of CMC texts suggest that they are a product of personal characteristics of the user, the social relations of the communication participants, and also the conditions of using the medium. The use of CMC has proliferated in the past few years and may even become a major means of communication. We can expect the CMC genres to develop further and to affect the styles of the language we use. It is natural, then, to ask what kind of impact this change will bring to language education. In the next chapter, I will discuss some of the pedagogical implications of these findings.

Chapter 6: Pedagogical Implications

6.1 Introduction

Language changes with time and the development of media, and different varieties of language evolve through the different contexts of communication. A very important lesson for language learners is to learn to say the right thing in the right way and to understand the meaning conveyed by others, despite the various forms of language used. This is what language is for: communication. A responsible language teacher should watch for the varieties of language that may come into existence as a result of changing contexts. The teacher needs to consider the value of the variety for their students. Is it a variety that is worth teaching? If yes, what is the best way to help the learner develop a sense of awareness for this variety?

The latest technological development of the computer network has fostered the new medium of computer-mediated communication (CMC), which is virtually revolutionising communication behaviour in many ways. Moreover, it is rapidly turning into a major means of communication. One of the impacts of this new medium is the new language varieties emerging in the new communication contexts. For the billions of English speakers and learners, these varieties will frequently appear around them and can never be neglected.

In the present study, I have provided empirical evidence for the new language variety and also some specific features associated with it. Aside from the possible reference for linguistic studies, I hope the findings presented here could also be of some pedagogical use.

In this chapter, I first describe the new CMC variety identified in the study and discuss the value of developing language awareness for this variety. I also introduce two approaches to developing awareness: a syllabus for teaching the features involved, and the teaching of the technique of concordancing.

6.2 Awareness of the Newly Emerged CMC Variety

Language awareness is a concept advocated to promote greater understanding of the ways in which language functions in society (Hawkins, 1987). Donmall (1985) defines language awareness as "a person's sensitivity to and conscious awareness of the nature of

language and its role in human life." Donmall (1985) and Wolfram (1998) state that a very important objective of language teaching is to foster the linguistic awareness of the varieties of language. All language learners will encounter different varieties of the language in daily communication. It is agreed that, instead of being given prescriptive rules to follow, learners should be trained to observe and analyse the varieties of language as they change in different contexts. To apply this concept in education, Hawkins (1987) proposes a curriculum in which students practice skills, such as listening, for particular language features and language patterns.

As has been reported in Chapters Four and Five, this study has verified that a new language variety has emerged in CMC that is very different from any non-CMC genres. Several features that are characteristic of this new variety have also been identified. We first present a brief description of how this new form looks, followed by a discussion of the pedagogical value of these findings.

6.2.1 The new language variety of CMC and its characteristics

As reported in Section 4.6.2 of Chapter Four, the factor analysis was implemented on the sampled CMC texts using the same 67 features that Biber (1988a) used on his non-CMC texts. While Biber extracted a total of seven factors, 26 factors (see Appendix 8), were extracted from the same 67 features in the CMC sampled texts. Besides, it was found that the features in each factor are not related in any way to those features gathered from Biber's seven factors from the non-CMC texts.

The finding suggests the extraordinary diversity of the features used in CMC texts. The 67 features, when occurring in CMC texts, do not gather into clusters as they do in non-CMC texts. This means that CMC texts are much more diversified in language use and cannot be easily represented as several types.

The finding here provides a preliminary piece of evidence that suggest that CMC texts are newly emerging variaties, which differ in nature from non-CMC texts in so much as they are quite diversified.

Further comparisons in Chapter Five revealed that, in terms of the frequencies of language features involved, certain differences exist between CMC and non-CMC genres, and also between the synchronous and asynchronous temporalities of the CMC texts.

In the medium of CMC, involvement and explicitness are the two major characteristics. Compared with the non-CMC texts, CMC participants more often use *WH*-questions to show interactional, involved discourse and to establish interpersonal relationships. CMC participants also convey dense information by adopting devices that

can express explicitness, such as the more frequent use of nominalisations, passives, suasive verbs, higher type/token ratio and longer words.

The phenomenon of this high density information is further confirmed by the less frequent use of certain features in the CMC texts, such as the pro-verb *DO* and the existential *THERE*, and *BE* as the main verb. This shows CMC participants try to avoid fragment meanings, so they rarely use these two features.

Another interesting phenomenon observed is the less frequent use of most subordination structures and the non-phrasal coordinator *AND*. This suggests that CMC participants are mostly restricted by time when interacting with each other. They do not have sufficient time to carefully embed their sentence structure, nor to freely prolong their sentences.

Another factor involved here is the means of language production. As CMC users produce their texts mainly through typing rather than though speech, they are hindered from elaborating upon their ideas, and must resort to simple, short sentences. Generally speaking, then, we observed more lexical density and less structural complexity in CMC than in non-CMC genres.

Aside from the comparison between CMC and non-CMC texts, we also observed that, within the CMC, the asynchronous text (CA) also showed differences from the synchronous texts (CS). The CA employs the frequent use of several features, such as passives, *THAT* clause, *THAT* clause on the subject position, past participial clauses, sentence relatives, other adverbial subordinators, suasive verbs, as well as the phrasal coordinator *AND*. These features reveal that participants in asynchronous CMC tend to use explicit expressions to convey more information, as they experience less time constraints compared to when they communicate in synchronous CMC contexts.

On the contrary, the CS texts show a higher usage of present tense verbs and hedges, as well as discourse particles. All of these present a picture illustrating that interpersonal relationships are established by these features. This is because present tense verbs provide involvement as participants engage in simultaneous communication. The hedges provide a style of informality, and the discourse particles enhance closeness for interpersonal interaction.

The above findings are in general agreement with the belief that synchronous discourse uses more interactive features, such as those used in speaking, and asynchronous discourse uses more complex structures and more formal, reserved styles.

In short, the finding that a new variety of language has developed in CMC is in itself a major achievement, which should arouse the attention of linguists and language

teachers. Through the process of the study, I discovered that a new variety of discourse was emerging out of the use of the new medium of the computer network. The new variety that has emerged is no less complex and sophisticated than that of any established language varieties.

6.2.2 Teaching the new language variety of CMC

In an age in which communication is increasingly mediated by electronic medium, we need to be familiar with the language forms used in CMC. Moreover, as the major language used in CMC is English (Kehoe, Pitkow, Sutton, Aggarwal, & Rogers, 1999), ELT teachers need to prepare their learners to master this international form in the medium. The findings gathered in this research hence serve to help teachers and learners in this task.

As previously discussed, the development of communicative competence has been taken as the main objective in English language teaching. The sense of awareness of the various varieties of the language in different social contexts is an important construct in communicative competence. On a large scale, English has been a major language for international communication for many decades (Kehoe et al., 1999). However, problems occur when non-fluent, non-native speakers communicate with each other. This issue has been raised by Jenkins (1995, 1998, 2000) and Crystal (1997), and they both appeal for *international intelligibility.* While Crystal calls for World Standard Spoken English (WSSE), Jenkins stresses the importance of recognising, as well as performing, various phonological forms of utterances in an international context.

While Jenkins and Crystal draw attention to the spoken form in international conversation, I would plead the same for the CMC context. I believe that intelligibility has the same high degree of importance in CMC, as more and more people communicate via a computer network worldwide. I also believe that mutual intelligibility in CMC would mainly, if not exclusively, lie on the understanding of the nature and characteristics of the language variety used in this new medium.

Linguistic awareness is an important goal in language education. The patterns found in this study have revealed that a new variety of language is emerging in CMC. This study has disclosed that certain language features appear significantly more frequently in some text groups, but less so in others. The indication of this propensity to high and low frequency is due to the effect of the specific media and temporalities. As reported earlier, involvement and explicitness are the two major characteristics in CMC

texts. Therefore, the CMC texts are also characterised by the high lexical density and low structural complexity.

I hope that the findings on the propensity of feature can help English teachers and students to develop a better awareness of this new variety, as well as of the different language features appearing in CMC and non-CMC texts.

Other than as a set of reference data in the study of language, the findings of this investigation might serve as a basis for syllabus design for English classes. As argued by Willis (1990, p. 124), we should include "commonest words and phrases" in English and their meanings in the course book. The findings in this study can be provided to ELT material developers so that they can consider including adequate language features in their teaching materials. The materials thus compiled will then represent more adequate language features used in different contexts of media and temporalities.

Using CMC has become a major practice in English language teaching. The present study should serve to arouse adequate attention to the new language variety of CMC and to provide concrete information on the features associated with the variety.

In the next sections, I introduce two approaches to developing an awareness for this language variety in CMC: a syllabus for teaching the features involved, and the teaching of the technique of concordancing.

6.3 Introducing the New CMC Variety to Learners

In this section, I plan to provide some suggestions on how to introduce the special CMC variety to language learners. There are two stages involved. The first is to introduce the learners to the asynchronous and synchronous contexts. The second is to draw their attention to the special language features involved.

6.3.1 Introducing learners to asynchronous and synchronous CMC

From the findings discussed in Chapter Five, we have noticed that the CMC texts are new varieties currently emerging from the CMC medium. For the time being, CMC is playing an important role in developing new language varieties. It is also a good medium through which learners can learn English by using authentic language for communication. I would advise teachers to select several pieces of CMC texts to demonstrate their specific features to learners by explaining how they are different from the traditional written and spoken texts. These CMC texts can be taken from asynchronous texts, such as postings

from discussion lists, and synchronous texts such as transcripts of MOO meetings.

Discussion lists, such as those hosted by LISTSERV, ListProc, Majordomo or JISCmail usually allow public access to their archives, or allow their members to retrieve particular postings. For instance, the archives of various discussion lists hosted by JISCmail in the UK can be publicly accessed by anyone. All the user has to do is click on the list of names in an alphabetical order at the web site <http://www.jiscmail.ac.uk/>. Further clicking on the archive directory and individual postings can retrieve the messages required.

For lists hosted by other network communication software, teachers obtain the right to access the archives only after subscribing to the lists. My advice to teachers is to subscribe to the discussion lists with a theme of language teaching, so that they can at the same time share and exchange their knowledge with other teachers member. However, teachers might want to subscribe to lists dealing with other themes. All of these can be searched on the LISTSERV site at <http://www.lsoft.com/lists/list_q.html>. For example, typing in "teachers of English as a second language" will generate two entries, i.e. TESL-L and TESLK-12. The former is a discussion list for international English teachers, while the latter is mainly for English teachers in primary schools. Users can then subscribe by further clicking the entry to send the message, *SUBSCRIBE TESL-L.*

After becoming an official member of a list, a teacher is then allowed to retrieve their archives. It is a simple process, as the teacher only needs to email a message, *get TESL-L <logyymmw> f=mail* to the address *listserv@cunyvm.cuny.edu.* The message, *get TESL-L <logyymmw> f=mail* retrieves the TESL-L daily discussion archives dated in whatever year, month and week from the mail archives. For instance, if I want to retrieve the file of the first week of discussion in February 2001, I simply send a message of, *get TESL-L log0102a f=mail* to *listserv@cunyvm.cuny.edu.* Within a few minutes, the archives arrive in my email box.

The synchronous CMC texts can be downloaded from web sites. For instance, Tuesday Café, part of its texts used in this research, can be accessed from the Netoric Project site at <http://bsuvc.bsu.edu/~gsiering/netoric/logs/>. The other synchronous CMC texts, NETEACH-MOO – the texts of which are also used in this study, can be downloaded directly from <http://www.dyvic.com/~greg/netmoo/moologs.html>.

6.3.2 Introducing learners to the features of CMC

English teachers are encouraged to enrich their students' language learning experience by introducing them to the new varieties in CMC. Of course, careful planning

and preparation are still needed, and the teaching activities may be conducted in several stages.

First, as mentioned in the previous section, Section 6.3.1, it is a good idea to show learners what the CMC context is like. Materials similar to those presented in Chapter Three, (Figure 3.1 *A Linear Structure of Message Organisation*, and Figure 3.2 *A Hierarchical Structure of Message Organisation*) are useful for familiarising students with the format of the asynchronous CMC setting. Likewise, Figure 3.3 *A MOO Conversation Sample*, and Figure 3.4 *The Discourse Structure of the MOO Conversation* are good examples to start with in the synchronous CMC context.

The next step is to introduce the students to some asynchronous discussion lists, for example, TESL-L. The teacher can subscribe to the list and show the students what the daily messages look like. Synchronous CMC usually imposes very tight time constraints and is therefore better arranged after the students are used to the asynchronous CMC contexts.

When showing students the asynchronous texts from the discussion list, the teacher can draw the learners' attention to the features that exemplify the new variety of language, for instance those characterising involvement and explicitness. Specifically, the teacher should draw the learners' attention to the prone features found in this study, such as the *WH*-questions, nominalisations, passives, suasive verbs, higher type/token ratio and longer words. At a later stage, perhaps, the teacher can remind the students of the less frequent use of other features, such as the pro-verb *DO*, the existential *THERE*, *BE* as a main verb, subordination and coordination structures.

Further to familiarising students with the characteristics of asynchronous texts, the teacher can encourage students to subscribe to a list on their own. As mentioned in Section 6.3.1, discussion lists can be searched and joined either from LISTSERV or JISCmail. Moreover, a discussion list that is specially prepared for students is Student-L, which is available at <http://www.latrobe.edu.au/www/education/sl/sl.html>. By encouraging the students to observe and eventually participate in the discussion, the teacher is providing the learners with an authentic reading and writing environment.

Once comfortable with the environment of the asynchronous discussion lists, the third step is to introduce students to synchronous CMC. A warm-up activity is necessary for this task. This includes steps like: (1) again showing students how the CMC conversation flows (i.e. Figures 3.3 and 3.4 in Chapter Three), (2) showing them more materials, i.e. the texts downloaded from sites like Tuesday Café and Neteach-MOO, as mentioned in Section 6.3.1, (3) teaching them the basic commands with which to talk in

the synchronous environment, (4) reminding them of the frequently occurring features, and (5) encouraging students to lurk long enough before they can comfortably converse with others in real time.

For Step (3), a list of basic commands for synchronous communication can be found at <http://www.hunter.cuny.edu/ieli/commands.htm>. The commands here are good for MOO. Moo is a multi-user, object-oriented site, similar to a conference call. A MOO that is created specially for English learners is SchMOOze University, which can be found at <http://schmooze.hunter.cuny.edu/test.html>.

For Step (4), the teacher needs to remind the students of the features which are assumed to be popular or rare in MOO. The teacher can also encourage the students to verify if these assumptions are true, based on their experience of using the MOO.

Step (5) is necessary as synchronous communication allows only a limited time to respond to questions raised by others, or topics which everyone is discussing. Once students are familiar with the linguistic features used here and the time constraints, they can go on to join the discussion. The teacher can also organise a real time meeting with classes in other schools worldwide to encourage their students to use English with an authentic audience. This also fulfils the main trend of the communicative approach in ELT.

In summary, the more exposure that learners have to the asynchronous and synchronous CMC texts, such as the discussion lists and the MOOs, the more chance they have to develop an awareness of the new variety of the English language. It is good if the learners can find patterns of features that meet the assumptions proposed in this study. It would be even more interesting if, in the many contexts in which new language forms may emerge in CMC, the learners can identify more representative features that have not been located in this study.

6.4 Concordancing and the Sense of Language Awareness

In the previous sections, we discussed the emergence of the new language varieties of CMC, the language features characteristic of the new variety, and also a possible syllabus in which to present to learners the associated language features.

In this study, it is hoped that not only the findings, but also the process itself can be of pedagogical inspiration to language teachers. As this study used a quantitative corpus linguistics, I can demonstrate the usefulness of this approach in analysing texts. For

English teachers, the use of corpus linguistics can help learners to make discoveries for themselves, which is a very effective method for self-learning. Nowadays, the role of the teacher, as Willis (1990, p. 130) argues, is "not so much to teach as to manage learning" but "to create an environment in which learners can operate effectively." By taking this further, the job of the teacher is "to help the learners manage their own learning." This marks a shift from teachers as "knowers" to learners as "discoverers." The process of my text analysis demonstrates the use of the concordancer, and will hopefully inspire English teachers to use this tool.

Concordancing is a technique that enables users to automatically find a key word, a key phrase, a sentence, or a certain structure in many possible contexts in a large number of texts. Many linguists frequently use computer concordancing devices to find language patterns and collocation of a specific structure, which they then use to form an assumption for analysis (Gledhill, 1996; Lewandowska-Tomaszczyk, 1998; Mandala, 1998; Scott, 2000). Some teachers also use such devices to find authentic evidence of certain language usage as supplementary materials to help learners understand how language is actually used in real life (Coniam, 1996; Flowerdew, 1998; Johns, 2000; Kettemann, 1996; Tribble, 1991, 1997, 1998). All of these reflect the importance of using the technique of corpus linguistics for theory building and language learning.

To further explain how a corpus linguistic technique can aid language learning, I would like to demonstrate an example with COMBO, a small concordancing software within CLAN, available at <http://childes.psy.cmu.edu/html/clan.html>. It is free for academic research, and I used it for this study, as mentioned in Section 4.4.1. I will examine the pronoun *it* from a collection of texts.

First, some texts have been collected. To make it simple, we assume that we only have five texts. They are numbered from *n001* to *n005*. Each text is saved as a text type under a directory of named *data* in the *c* drive of a computer. It is necessary to save the COMBO programme in the *c* drive, too. A notation is designed in order to retrieve the *it* pronoun from the five text samples. Below is the notation for this task.

combo +y +m +sit c:\data*.* +ff09

In this notation, *combo* is the command to activate the COMBO programme. The *+y* is an option to enable COMBO to accept text samples as text type. The *+s* is another option to inform COMBO to search for the word following *+s*. In this case, it is the pronoun *it*. The *c:\data*.** is the command to tell COMBO to look for sample texts under the *data* directory in the *c* drive. The *+f* is an option to inform COMBO to save the retrieval result with *f09* as its file extension name.

In sum, this command notation is to search for the pronoun *it* construction from the sample texts. In the result files generated by COMBO, shown in Figure 6.1 below, we learn that the pronoun *it* construction only occurs in Lines 5, 10 and 11 of the Text No. 002 and Lines 8, 23, 25 and 40 of the Text No. 004.

```
COMBO.EXE +y +m +sit c:\data\N002 +ff09
**** line 5; file c:\data\NO0028 ****
We have been experimenting with cool talk, but are having a hard
time coming up with meaningful ways of using it in the ESL classroom.
                                                    1
**** line 10; file c:\data\N002 ***
It is too easy to interrupt the other users.
1
**** line 11; file c:\data\N002 ****
However, it seems as though it could be interesting integrating the
         1                  2
paint feature, the messaging feature and the www.
Strings matched 4 times


COMBO.EXE +y +m +sit c:\data\N004 +ff09
**** line 8; file c:\data\N004 ****
I find it is a good way to simply develop and reinforce all of those
       1
fundamental skills necessary for the beginning language learner.
**** line 23; file c:\data\N004 ****
It teaches keyboarding using real English in words, phrases, and
sentences.
1
**** line 25; file c:\data\N004 ****
It doesn't have the color and flash, but it does the job.
1                                    2
**** line 40; file c:\data\N004 ****
I don't think it will make anyone feel very good.
             1
Strings matched 5 times
```

Figure 6.1 Two Excerpts of COMBO Output for F09 Pronoun *IT*

In the above figure, COMBO marked numbers below all the words of *it.* At the end, COMBO also reported "*Strings matched 4 times*" or "*Strings matched 5 times.*" This means that it found four or five occurrences of *it* construction in these two texts. If we have more texts, i.e. more than the five texts we assume for this experiment, COMBO will continue to search for the same construction through other texts, and reports back how many occurrences it finds in every text. With the occurrences marked in the output, it is easy to locate the *it* construction, and its collocation environment, i.e. the words co-occur before and after *it*, at the same time.

The English teacher can use this COMBO result for a language learning exercise. For example, he or she can ask students what the first *it* construction refers to in the Text No. 002 (the occurrence in Line 5), and in the Text No. 004 (the occurrence in Line 8). For the referents of the other *it* constructions, the teacher needs to go back to the original

sample text, or to enlarge the text line by giving another option during COMBO searching, i.e. the *–w3* gives three lines before the key word occurrence, and the *+w2* gives two lines after the construction. In doing this, we can locate the text that *it* has referred to.

The use of the concordancing technique is not limited to linguistic researchers only. Language teachers and students can also benefit from imaginative implementation of the concordancing techniques. For teachers, concordancing can be a tool to find concrete evidence of language features in all kinds of text genres. For students, it can also be a tool for learning English, either from good model texts or in examination of students' own inter-language writing, as developed in different stages of language learning.

In summary, language features associated with a particular variety can be introduced to the language learners either by a syllabus, or by a process of discovery learning through the use of concordancing programmes. In either case, I hope to demonstrate the importance of developing the sense of language awareness in the new varieties of English in CMC.

6.5 Summary

The main findings in the study are the emergence of a new language variety, among others, in the CMC contexts and the several language features that are closely associated with the CMC varieties in its different temporalities. In this chapter, I attempted to establish the pedagogical values of the findings. I also proposed two approaches to developing learners' language awareness of the new CMC texts. One is through involving the learners in the asynchronous and synchronous CMC contexts; the other is through the teaching of the techniques of concordancing.

Both the instrument and the genres involved in this study are based on the computer. The coming of the electronic age has certainly brought to us new ways to use language and new ways to study language. It is my hope that English teachers and learners do not miss this important aspect in the learning and teaching of the English language.

Chapter 7: Conclusion

7.1 Summary

The many varieties of the English language, their backgrounds, their special characteristics and functions, their connections to human social relations and their status in language teaching have been an important topic of study in sociolinguistics. The rapid development of the new technology of computer-mediated communication (CMC), like the advancement of other media, has resulted in the emergence of new language genres. This study has examined the impact of the media on the language used in the CMC context. This chapter summarises the whole study from various perspectives.

7.1.1 Purpose of the study

Language in past studies has usually been depicted as being more oral or written in style, with oral communication characterised by immediacy and writing marked for its temporal or geographical distance (Baron, 1998) and its permanence (Olson, 1996). However, new communicative modalities often blend the presuppositions of spoken and written language. The telephone (Baron, 1998), radio broadcasts (Oksenholt, 1977; Scanlan, 1980) and tape recording (Frommer & Weitz, 1981; Long & Povey, 1982; Wherritt, 1979), for instance, have broken the barriers of space and time, so that the content of speaking can be transmitted far and remain permanent. Written messages sent by fax now virtually in real-time and also exhibit new styles of written communication (Baron, 1998; Basch, 1995; Coyle & Spitzer, 1992; Gerrey, Brabyn, & Crandall, 1990; Strasser, 1995). Voice mail offers the vocal cues of speech without the opportunity for feedback from the interlocutor, thus truncating the expected parameters of spoken language (Baron, 1998). It seems, therefore, that the development of a new medium is very likely to result in new styles of language use (McLuhan & Nevitt, 1974).

The complexity of the communication contexts developed in CMC, however, may be even more sophisticated than that fostered by any other media, and the new language varieties developed can be an immense sphere for investigation. CMC can be either synchronous or asynchronous. It can take the form of a closed exchange of information between intimate friends, or open discussion forums in which hundreds of thousands of interested parties participate. The versatility of the CMC contexts is unprecedented. Moreover, the wide application of CMC in almost all aspects of human society has made

it such an important means of communication, and the language varieties developed may become major genres. Language researchers should not neglect the importance of the new language varieties of CMC, and language teachers should also try to seek a better understanding of these new varieties so as to foster an adequate awareness of them in their students.

The development of language variety is a continuing question. CMC, as the latest technology, is bringing about enormous changes to communication contexts, which in turn may result in new varieties of language. Can we confirm that such new varieties have been developed? What are the language features associated with these new varieties? These questions can best be answered by empirical studies. This is the purpose of the present work: to conduct a comprehensive study on a large set of language features found in synchronous and asynchronous CMC texts to determine if the CMC texts do form new varieties of language use, which are different from the texts found in non-CMC settings.

7.1.2 Methods used and procedures followed

The methods adopted and procedures undertaken in this study were described in detail in Chapter Four. Here is just a short review.

7.1.2.1 Approach and objects of the study

In the study, I adopted the approach of corpus linguistic studies and that of quantitative comparison. The large quantity of texts stored in the electronic version forms becomes an ideal data pool, which can be accessed and directly examined with the help of the computer. I hope the quantitative data gathered from the study can provide sufficient objective evidence of the nature of the new variety of the language.

For the objects of investigation, I decided to adopt the list of 67 linguistic features used in earlier studies by Biber (1988a), which I believe to be comprehensive enough for my research purpose. The 67 features are grouped in 16 categories: (A) tense and aspect markers, (B) place and time adverbials, (C) pronouns and pro-verbs, (D) questions, (E) nominal forms, (F) passives, (G) stative forms, (H) subordination, (I) prepositions, adjectives and adverbs, (J) lexical specificity, (K) lexical classes, (L) modals, (M) specialised verb classes, (N) reduced forms and structures, (O) coordination, and (P) negation.

As most previous studies on the characteristics of CMC texts focused on some limited sets of language features, I hoped that by adopting this rather complete list of

features, I could have a comprehensive perspective in my study.

The sample texts of analysis were chosen from archives of two Internet asynchronous discussion lists, NETEACH-L and TESL-L, as well as two synchronous MOO forums, NETEACH-MOO and Netoric-MOO. The theme of these texts is teaching English as a second or foreign language, and almost all contributors are English teachers. The first selected text sample in the asynchronous CMC group (abbreviated as CA) had a total of 100,194 words in 503 email texts. However, it was adjusted to 171,689 words in 220 texts in the final analysis – the Non-Parametric Statistical Test. The synchronous CMC (CS) text sample had a total of 104,162 words in 17 MOO session logs.

For cross-reference purposes, I also intended to collect a set of comparable non-CMC texts. To avoid the technical difficulty of converting a large amount of non-CMC texts into a computer readable form, however, I decided to adopt the quantitative data from already-existent non-CMC text samples of similar genres. The data were collected from the work by Biber (1988a). The corpus of Biber was built by using 15 written genres from the LOB Corpus (Johansson, 1982; Johansson, Leech, & Goodluck, 1978), 6 spoken genres from the London-Lund Corpus (Svartvik & Quirk, 1980), and two types of unpublished letters (Biber, 1988a), all of which totalled 960,000 words in 481 texts, distributed in 23 genres, 17 of which were written and 6 spoken.

7.1.2.2 Data preparation and tools used

To handle the large quantities of CMC text data, I selected the following tools: (1) the concordancing programme CLAN (MacWhinney, 1995, 1996a, 1996b), developed at Carnegie Mellon University, and (2) the part-of-speech tagging programme TAGGER (Mason, 1996) developed at the University of Birmingham.

All the CMC data downloaded from the web sites were manually checked to remove all irrelevant lines, such as the email headers and lines generated by the computer system. The texts were then sent by e-mail to TAGGER for the part-of-speech tagging. The tagged files, together with their original untagged version, are the objects for analysis using the CLAN concordancing programme later.

To enable the concordancing programme to find the patterns and structures of the linguistic features, the 67 linguistic features had to be depicted in notations. The CMC texts were then processed using the COMBO command in CLAN. Although the tasks of tagging and concordancing were mainly processed by the computer programmes, it was still necessary to carry out manual proofreading and correction.

In order to obtain the most adequate model of analysis, I considered and attempted several methods, including a 2 x 2 Factorial Experiment between CMC and non-CMC texts, a Factor Analysis on CMC texts, and a Non-Parametric Statistical Test of synchronous versus asynchronous CMC texts. Eventually, the Non-Parametric Statistical Test was decided upon as the major statistical measure for analysis.

7.1.3 Major findings

Biber (1988a) took the approach of factor analysis to extract from a large number of linguistic features a limited set of factors that account for variation between speech and writing. Biber's approach took the oral/written distinction as on a continuous scale of variation. Based on the 67 linguistic features examined in a corpus of 960,000 words of 23 written and spoken genres (Biber, 1988a, p. 67), the procedure of factor analysis extracted a total of seven factors, which may serve as dimensions on which texts may be measured as being more speaking-like or writing-like. These include: informational versus involved production, narrative versus non-narrative concerns, explicit versus situation-dependent reference etc., as depicted in Sections 4.6.2 of Chapter Four. After their publication, these seven dimensions have been quoted by many (e.g. Allomong, 1996; Collot & Belmore, 1996; Ljung, 1991; Tribble, 1997) in the study of language styles and in the design of language teaching syllabus and materials.

7.1.3.1 Different sets of factors generated

In order to see if a similar set of factors can also be extracted from CMC texts, the procedure of factor analysis was implemented with the same 67 linguistic features on the CMC sample texts collected in this study. If the analysis resulted in a similar set of factors in the non-CMC texts, i.e. if the same dimensions are shown to exist in CMC texts, then we could easily plot where each of the CMC data sets fitted along these dimensions.

The results, as presented in Appendix 8 and discussed in Section 4.6.2, showed that among the CMC texts, 26 factors were extracted from 67 features! It marks a complete difference from the conditions in non-CMC texts. Besides, it was found that the features in each factor are not related to those gathered in Biber's seven factors for the non-CMC texts.

Usually, it is hoped that Factor Analysis will derive a relatively small set of underlying variables, called "factors," from large sets of variables. Biber extracted seven factors (Biber. 1988a). Previous studies on language features (Bainbridge, 1986; Des

Brisay, Duquette, & Dirir, 1993; Dunning, 1995; McEnery & Wilson, 1996; Warschauer, 1996d) extracted between two and five factors.

The fact that we found 26 factors out of the 67 features in the CMC texts indicates the extraordinary diversity of CMC texts. The 67 features, when occurring in CMC texts, do not gather into clusters as they do in non-CMC texts. This suggests that CMC texts are much more diversified in language use and cannot be easily represented as being of several types.

As Factor Analysis can be used to uncover patterned variation, it would be interesting to investigate further why the CMC texts are so diversified in nature. However, the finding here has provided a preliminary piece of evidence that CMC texts are a new emerging variation that is quite diversified in nature. It is very different in nature from the non-CMC texts.

7.1.3.2 Differences in the frequencies of linguistic features

To compare the relative frequencies of linguistic features as they occur in CMC and non-CMC texts, the procedure of a Non-Parametric Statistical Test was implemented. The analysis showed that, at the significance level of α=0.005, there are 17 features that occur significantly more often in asynchronous CMC (CA) than in synchronous CMC (CS) texts. They include the features of *nominalisations* (F14), *BY passives* (F18), *THAT clauses as adj. complements* (F22), *WH clauses* (F23), *past participial clauses* (F26), *past participial WHIZ deletions* (F27), *THAT clauses on subject position* (F29), *THAT clauses on object position* (F30), *WH clauses on object position* (F32), *sentence relatives* (F34), *causative adverbial subordinators: because* (F35), *concessive adverbial subordinators: although, though* (F36), *other adverbial subordinators* (F38), *attributive adjectives* (F40), *SEEM/APPEAR* (F58), *final prepositions* (F61), and *non-phrasal coordination* (F65).

On the other hand, there are 12 features which occur significantly more often in synchronous CMC (CS) than in asynchronous CMC (CA) texts. They include *present tense verbs* (F03), *time adverbials* (F05), *DO as a pro-verb* (F12), *WH questions* (F13), *present participial clauses* (F25), *general hedges* (F47), *discourse particles* (F50), *public verbs* (F55), *contractions* (F59), *THAT deletions* (F60), *split infinitives* (F62), and *synthetic negation* (F66).

There are 38 features that do not show significantly different frequencies between the synchronous and the asynchronous CMC texts.

Among the 17 cases in which the features occur significantly higher in asynchronous CMC (CA) than in synchronous CMC (CS), we can see that most belong to

the category of subordination. This includes those in the sub-categories of complementation, participial forms, relatives and adverbial clauses. This generally meets the expectation that the complex structures of subordination are more associated with the written genres and are therefore adopted in the asynchronous CMC texts. Other than the subordination features, the features of *nominalisations* (F14) and *attributive adjectives* (F40) also have the function of reducing text length and increasing lexical density. The feature *final prepositions* (F61) is also linked with the complexity of sentences. The higher frequency of *BY passives* (F18) and *SEEM/APPEAR* (F58) are believed to be associated with the reserved style often found in written genres.

The above analysis generally shows that the asynchronous CMC texts usually exhibit features typically found in written genres.

Among the 12 features that occur significantly more often in synchronous CMC, most, e.g. *present tense verbs* (F03), *time adverbials* (F05), *DO as pro-verb* (F12) and *WH questions* (F13), show the styles of an involved, situation-dependent, overt-persuasion type of discourse, which is usually associated with speech. The significantly higher frequency of the features *discourse particles* (F50), *contractions* (F59), *THAT deletions* (F60), and *split infinitives* (F62) also suggests a typical colloquial style.

While the above findings generally indicate that asynchronous CMC texts and synchronous CMC texts match the styles of written and spoken discourses, we must still note that this is not a clear distinction.

The most important point to note is that 38, over half, of the 67 features do not show any significant difference. That means the synchronous and the asynchronous CMC texts do not vary on those features. As most of the 67 features were found to belong to one or more of the factors determining the speech versus writing orientation of non-CMC texts (Biber, 1988a), it is interesting that more than half of the features do not have such a function of distinction in CMC texts. This means that CMC users use these features more freely in both synchronous and asynchronous contexts. We can use several specific cases for further discussion. Take for example the category of time and place adverbs. Though *time adverbials* (F05) showed significantly higher frequency in the CS text than in the CA text, *place adverbials* (F04) do not show any significant difference between the two temporalities. The finding seems to suggest a different way of conceiving of the space relationship when CMC participants interact in cyberspace.

In the category of pronouns and pro-verbs, only the pro-verb *DO* shows a significantly higher frequency in synchronous CMC (CS) texts. The pronouns do not show a significant difference in frequency between CMC texts, although first and second

person pronouns are believed to indicate an interpersonal focus and a generally involved style (Chafe, 1982, 1985). Third person pronouns are found to co-occur frequently with past tense and perfect aspect forms to serve as a marker of narrative, reported styles (Biber, 1988a). The different patterns of the use of pronouns between CMC and non-CMC texts show that users of CMC do not necessarily follow the discourse styles in the use of pronouns as they do when interacting in non-CMC settings.

7.1.3.3 Rankings of frequencies between CMC versus non-CMC genres

Aside from the procedure of Non-Parametric Statistical Test conducted to measure the difference between asynchronous and synchronous CMC texts, the rankings of the frequencies of the features were also compared between CMC and non-CMC genres. There were some interesting findings.

Taking *WH*-questions (F13) as an example, when we compared the rankings of CMC and non-CMC genres for the frequencies of *WH*-questions, as shown in Figure 5.8 of Chapter Five, we found that synchronous CMC (CS) had the highest frequency of *WH*-questions (1.27). Speech genres such as telephone conversation (1.1) and spontaneous speeches (1.0) also stood high in the ranking. Asynchronous CMC (CA) also stood among the highest in the frequency of *WH*-questions (0.73). That both CS and CA texts have higher rankings on the use of *WH*-questions than most non-CMC genres indicates that CMC users tend to raise questions or make requests directly, instead of resorting to euphemism. This may be due to the time constraints and also to the less controlled, anonymous contexts of CMC.

Another example is *nominalisation*. The use of this feature is usually taken as a means of increasing the lexical density and reducing the sentence length. In Figure 5.10 of Chapter Five, we can see that asynchronous CMC (CA) (22.1) ranks sixth among all of the 25 genres, and synchronous CMC (CS) (13.01) ranks 15th, higher than six writing genres and three speech genres. Both of the CMC genres examined here are relatively higher than non-CMC genres in the use of nominalisation. This suggests that nominalisation is a measure to which CMC users resort to reduce the text length.

Subordination is believed by some scholars to mark greater elaboration and be characteristic of informational discourse typical in writing genres (O'Donnell 1974, Key 1977). On the other hand, Halliday (1979) believes that subordination is associated with the production constraints characteristic of speech.

Among the many subordination features examined in this study, it is interesting to find that synchronous CMC tends to always stand at the bottom of the ranking in the use

of subordination features. This seems to contradict Halliday's argument for the higher frequency of subordination devices used in the interactive discourse of speech. According to Halliday, the time constraints in the communication contexts are the cause of the higher frequency of subordination devices used in speech. We also suspected that the high frequency of the use of the feature nominalisation in CMC texts was due to time constraints. Why have the time constraints not caused a higher frequency of subordination features? This is a question we will attempt to discuss in the next section. For the time being, we can say that CMC users tend to pack information together by using words of higher density, rather than resorting too more complex sentence structures.

A related finding is in the category of lexical specificity. It was found that CMC texts tend to have higher type/token ratio and higher word length than synchronous non-CMC texts, i.e. face-to-face and telephone conversations. Again, this suggests that CMC users tend to increase lexical density as a means to shorten the time spent communicating.

Another important observation we can make is on the use of coordination devices, especially the non-phrasal coordinator *and* (F65), which is meant to connect clauses. In Figure 5.49 of Chapter Five, we can see that both CA and CS texts stand very low in the ranking of the frequency of the use of the non-phrasal coordinator *and*. Besides, we can see that all the speech genres use *and* as a non-phrasal coordinator more often than writing genres.

The observation on Figure 5.49 verifies the earlier observation by Kress (1994) that speech is characterised by sequences of clauses coordinated by the use of *and*. On the other hand, we find in the present study that CMC users tend to use the coordinator *and* less than in non-CMC contexts. The reasons for this phenomenon will be discussed in the next section.

The general findings we have reached in this section are that CMC texts, in comparison with non-CMC genres, are characterised by the extensive use of nominalisation, higher type/token ratio and longer words. On the other hand, devices to increase the complexity of the sentences, including the use of subordination and coordination devices, are rarely used in CMC contexts. In summary, we can sense a lexical density but structural simplicity in the CMC texts.

7.1.3.4 Reasons accounting for the unique CMC characteristics

In the comparison between CMC and non-CMC genres discussed in the previous section, a general observation is that CMC texts tend to achieve higher lexical density than non-CMC genres by using more nominalisation, higher type/token ratio and longer

words. The higher lexical density reduces text length, which is important under the tense time constraints of CMC contexts

Secondly, we also noticed that CMC texts use subordination devices much less than in the case of non-CMC texts. This seemed quite unusual, because the use of subordination devices marks the increased complexity of sentence structures, also an effective measure of reducing text length to offset the time constraints. As a matter of fact, according to Halliday (1979), the time constraints are the cause of the higher frequency of subordination devices used in speech. Now that synchronous CMC also proceeds under very tense time constraints, why do CMC participants use fewer subordination devices than they do in non-CMC speech contexts?

The observation on the use of the coordination device *and* also raises a similar question. Kress (1994) found that speech is characterised by sequences of clauses coordinated by the use of *and.* In the present study, however, we found that *and* is rarely used as a coordinator to connect clauses in CMC contexts. What causes such difference?

The explanation I proposed is an interaction of the elements of time constraints and the means of production.

In a synchronous non-CMC setting such as face-to-face or telephone conversations, the speaker does feel more pressure to express his/herself fully within certain time constraints than he/she would in writing a letter in an asynchronous non-CMC setting. However, the speaker in this setting has an easy means of production, the speech organs. He/she can easily produce dozens of words in a matter of seconds. He/she also can use as many subordinators or coordinators as needed to expand his sentences in order to put his/her meaning across. That is to say, speakers in traditional non-CMC contexts have the advantage of an easy means of production that compensates for the time constraints. They feel free to expand the sentences in length and complexity, and need not worry about low lexical density.

For the writer in the asynchronous non-CMC setting, the situation is different. There may be ample time for him or her to ponder on the language to be used, but the means of production, either by handwriting or typing, is much less convenient than the speaking. The limitation is not of time, but of the length of text that can be possibly produced. The writer's duty is to express his/her ideas in as compact a form as possible. The length of the text within which the writer must strive to convey the meaning with higher lexical density and structural complexity is limited.

Users of CMC are faced with different conditions of constraints of time and means of production. For synchronous CMC users, the time constraints can be stronger than in

the case of non-CMC speaking contexts. He or she must maintain synchronous interaction with the interlocutors by using a not-so-convenient means of production, key-pressing. There is no time for him/her to expand the clauses infinitely, either by subordination or by coordination. With the little time lag of synchronous CMC, interlocutors usually pick up a point and continue the discussion before the other party finishes the line. Here, the CMC users are like the writer who must achieve high lexical or grammatical density within short sentences. However, the time constraints imposed on CMC users force them to appeal to lexical density, perhaps at the expense of structural complexity.

7.2 Reflections on the Study and Suggestions for Further Research

As a language learner and teacher, I have a strong interest in the many varieties of the English language, and I believe it is a major responsibility of language teachers to help their students to learn to understand and use the different language varieties adequately in different contexts. Some of the thoughts I have experienced towards learning a language form throughout this study are reflected below. At the same time, I will also suggests some studies for researchers who might be interested in investigating further the language form used in the CMC.

7.2.1 Achievement of the present study

In the past decade, CMC has developed so rapidly as to serve as a major context of communication. Many people can sense the emergence of new varieties in this medium. However, most studies on the language used in CMC were limited to too small a scope of features to verify the existence of the new variety. This particular work is meant to be a comprehensive study on the linguistic features found in CMC texts to determine if we can justifiably claim CMC as new varieties of language and to understand the special features associated with this particular variety. Basically, as reported in the previous section, the objectives of the study have been reached.

The factor analysis using the same set of 67 features generated completely different sets of factors in CMC and non-CMC texts. The result showed that the functions of the linguistic features when used in CMC are much more diversified than when used in non-CMC texts. In terms of the unique uses of some features in CMC texts, it was found that, among other features, CMC texts are characterised by lexical density, while non-CMC

texts are characterised by structural complexity. I also proposed that the interaction of the two factors of time constraints and the means of production interact to determine the characteristics of synchronous and asynchronous CMC texts and their non-CMC counterparts.

I believe the above findings and the explanations form an important part of our understanding of the new language variety of CMC, and the understanding is also an important part of the understanding of language varieties in general.

In summary, I have provided a descriptive basis of one of the language varieties in CMC for the two modalities, and I also considered some of the implications for education

7.2.2 Limitations of the present study

Like most other academic research projects, this study also suffered certain limitations.

The most obvious limitation is the scope of CMC texts examined. Due to the limited time and resources available, I could only retrieve the text files from the archives of two asynchronous discussion forums and two synchronous MOO forums. While the contexts covered both the synchronous and asynchronous temporalities, the themes of discussion in those forums were limited to issues on English teaching. The participants in those forums are mostly English teachers and their purpose of participation is to exchange academic and professional information. This is certainly a specialised group of participants with a specialised interest. What the study has found can be taken as one variety of the English language in a particular type of CMC context. There are most likely to be many more varieties developed when research is carried out on other CMC contexts.

Another limitation of the study was the sampled texts. Ideally, the comparison between the CMC and non-CMC texts should be carried out on two sets of comparable original texts. However, due to the technical difficulty of locating a large quantity of non-CMC texts and converting them into machine-readable format, I had to adopt the figures reported in previous research and make comparisons between the CMC and non-CMC texts based on the figures. This has limited my ability to conduct a more sophisticated statistical analysis.

Last but not least, I must admit that the project of a comprehensive study of 67 language features on a corpus of more than 200,000 words was too ambitious for an individual student. Perhaps I over-estimated the structure recognition capability of the computer concordancing programme and thought the figures reported could be ready for

use once I gave the programme the correct instruction. I later found that certain features could not be depicted adequately by notations. For those features, I needed to examine the output files and carry out manual proofreading line by line. For most other features, there was always the chance of ambiguity in the structures, and the programme would count some cases which did not really fit the requirements. Eventually, I found I had to proofread almost all of the output files for each and every feature examined. This has taken many months of tedious work and has cost me too much time. A study on a scale such as this would be better undertaken by a research team rather than by an individual researcher.

7.2.3 Suggestions for further research

For further studies, I can suggest two directions: the research that can compensate for the incompleteness of the present study, and the study that can be carried out on the basis of the findings of this present study. The former is mainly research for descriptive linguistics, the latter is mainly pedagogy for educational practice.

7.2.3.1 Linguistic research

As stated above, the study was carried out on a certain CMC context, with a limited group of participants and a specialised theme of discussion. Though the study was comprehensive in terms of the language features examined, the scope of the subjects was limited. I did establish a substantial base for a descriptive language form used in the CMC. However, the establishment of this full base of CMC language is ideal only when it is used by more researchers. It would not only enrich our knowledge, but would also form the basis of important pedagogical decisions. It would help us to address questions of how far students should go to access CMC for their learning of English, or whether they should be protected from this variety to avoid being internalised by the language form criticised by Lundstrom (1995).

To do this, more studies are needed to cover more text groups from different interests of CMC, i.e. on thems of hobbies, leisure, entertainment, community development, chat rooms, various MOOs, etc. With diverse texts gathered, there are several statistical analyses that can be implemented. For instance, we can run a Factor Analysis to see if these diverse texts also present the same pattern of 26 factors generated in this study.

Using the above diverse text data, another possibility for further studies is to compare these CMC texts to some parallel non-CMC texts in order to examine the

similarity and differences of the linguistic features involved. In doing so, I suggest the collection and preparation of original texts for both the CMC and non-CMC groups, so that sophisticated statistical procedures can be implemented for some detailed analysis.

7.2.3.2 Pedagogical studies

On the pedagogical side, there have been quite a number of studies on students' learning using CMC, as has been discussed in Section 3.4 of Chapter Three. The value of CMC in providing a collaborative learning context has been emphasised by many researchers. With the findings proposed in the present study, I believe there can be more studies on the students' language performance using CMC.

For instance, we can develop a syllabus to teach students using the common features found in this study, and guide them to practice these features in CMC. Later, we can evaluate how much they learn by examining their language performance in the CMC medium.

Students should be encouraged to develop an awareness of the new varieties of the English language in CMC. More than being told of the special features of the CMC varieties and asked to imitate, the students need more learning activities to develop their awareness of the varieties. I believe we need more pedagogical studies on how to foster the sense of language awareness in the students so that they truly master the characteristics of the new language varieties.

In conclusion, among the many studies on CMC, the present study may be the first to examine a large number of linguistic features in an attempt to depict the characteristics of CMC. It is a study on CMC from the point of view of language variety. I have tried to establish a substantial base for a descriptive language form used in CMC. I hope to arouse people's attention to the CMC varieties used in these two modalities, i.e. synchronous and asynchronous. There are still limitations to the scope and procedure of the study, but I believe I have taken an important step in the study of the new language variety, for both academic and pedagogical purposes.

Appendices

Appendix 1

The list of tags used by TAGGER hosted in University of Birmingham:

??? -- no tag assigned
CC -- coordinating conjunction
CD -- cardinal number
DT -- determiner
EX -- existential 'there'
FW -- foreign word
IN -- preposition
or subordinating conjunction
JJ -- adjective
JJR -- adjective, comparative
JJS -- adjective, superlative
LS -- list item marker
MD -- modal
NN -- noun, singular or mass
NNS -- noun, plural
NP -- proper noun, singular
NPS -- proper noun, plural
PDT -- predeterminer
POS -- possessive ending
PP -- personal pronoun
PP$ -- possessive pronoun
RB -- adverb
RBR -- adverb, comparative
RBS -- adverb, superlative
RP -- particle
SYM -- symbol
TO -- infinitive marker 'to'
UH -- interjection
VB -- verb, base form
VBD -- verb, past tense
VBG -- verb, gerund or present participle
VBN -- verb, past participle
VBP -- verb, non-3rd person singular present
VBZ -- verb, 3rd person singular present
WDT -- wh-determiner
WP -- wh-pronoun
WP$ -- possessive wh-pronoun
WRB -- wh-adverb
" -- simple double quote
$ -- dollar sign
# -- pound sign
‘ -- left single quote
’ -- right single quote
‘‘ -- left double quote
” -- right double quote
(-- left parenthesis(round, square, curly or
angle)
) -- right parenthesis(round, square, curly
or angle)
, -- comma
. -- sentence-final punctuation
: -- mid-sentence punctuation

Appendix 2

An Original CMC Asynchronous Text with the Computer-Generated Mail Header.[+]

>From wmagno@amazon.com.br Fri Oct 11 11:41:10 1996
Received: from sucuri.amazon.com.br (root@sucuri.amazon.com.br [200.241.240.1]) by thecity.sfsu.edu (8.7.3/8.7.3) with SMTP id LAA28751 for <neteach-l@thecity.sfsu.edu>; Fri, 11 Oct 1996 11:41:09 -0700 (PDT)
Received: from wmagno (uxi.amazon.com.br [200.241.240.13]) by sucuri.amazon.com.br (8.6.12/8.6.12) with SMTP id PAA07753 for <neteach-l@thecity.sfsu.edu>; Fri, 11 Oct 1996 15:40:24 -0300
Date: Fri, 11 Oct 1996 15:40:24 -0300
Message-Id: <199610111840.PAA07753@sucuri.amazon.com.br>
X-Sender: wmagno@popmail.amazon.com.br
X-Mailer: Windows Eudora Version 1.4.4
Mime-Version: 1.0
Content-Type: text/plain; charset="us-ascii"
To: neteach-l@thecity.sfsu.edu
From: wmagno@amazon.com.br (Walkyria Magno e Silva)
Subject: ESP for reading computer sciences texts

I don't know if this is the correct address to do this, but could you post this message in your discussion list?

Are the difficulties students of computer sciences have when reading materials on their specific area centered on vocabulary, structure, or background knowledge? What do you think?
What type of materials would they need to read the most?

Thanks,
Wal

[+] with special permission from the writer to include this message here.

Appendix 3

A CMC Asynchronous Text with the Mail Header Removed.[+]

no0046
I don't know if this is the correct address to do this,but could you post this message in your discussion list?

Are the difficulties students of computer sciences have when reading materials on their specific area centered on vocabulary, structure, or background knowledge? What do you think?
What type of materials would they need to read the most?

Thanks,
Wal

[+] with special permission from the writer to include this message here.

Appendix 4

Notations Used in This Study for the 67 Features

F01 past tense verbs
e.g. They ***arrived*** a few minutes ago.
combo +y +m -w1 +w1 +s"!(@have+@be+@01get)^*^(*ed+@vbd_irr)" +ff01

F02 perfect aspect verbs
e.g. I ***have found*** him not doing his homework since last week.
1. HAVE + (ADV) + (ADV) + VBN
2. HAVE + N/PRO + VBN (question)
combo +y +m -w1 +w1 +s"@have^*^(*ed+@vbn_irr)" +ff02

F03 present tense verbs
all present tense verbs
combo +y +m -w1 +w1 +s"!to/TO^(*/VBP+*/VBZ)" +ff03

F04 place adverbials
e.g. Studying ***abroad*** is not an easy job.
Covering: aboard, above, abroad, across, ahead, ... etc.
combo +y +m -w1 +w1 +s@place_ad +ff04

F05 time adverbials
e.g. Never do that ***again***!
Covering: afterwards, again, earlier, early, ... etc.
combo +y +m -w1 +w1 +s@time_ad +ff05

F06 first person pronouns
e.g. ***I*** believe that she'd noticed this too.
Covering: I, me, we, us, my, our, myself, ... etc.
combo +y +m -w1 +w1 +s@ppfirst +ff06

F07 second person pronouns
e.g. ***You*** are my only friend.
Covering: you, your, yourself, yourselves, ... etc.
combo +y +m -w1 +w1 +s@ppsecond +ff07

F08 third person pronoun
e.g. I've reported this to ***his*** teacher.
Covering: she, he, they, her, him, them, ... etc.
combo +y +m -w1 +w1 +s@ppthird +ff08

F09 pronoun IT
e.g. But, ***it*** is also a challenge.
combo +y +m -w1 +w1 +s@pro_it +ff09

F10 demonstrative pronouns
e.g. ***This*** is ridiculous.
1. that/this/these/those + V/AUX/CLP-P/T#/WHP/and
2. that's
3. T# + that
combo +y +m -w2 +w2
+s"(((these/*+those/*+this/*)^!@noun)+(that/*^'s/*)+(!@noun^that/*^!@noun))" +ff10

F11 indefinite pronouns
e.g. That is a most ridiculous story, which ***nobody*** would believe.
Covering: anybody, anyone, anything, everybody, ... etc.
combo +y +m -w1 +w1 +s@ind_pro +ff11

Appendix 4 (Continued)

F12 DO as pro-verb
e.g. "Who broke the window?" "The cat ***did*** it!"
1. not DO+(ADV)+V DO as auxiliary
2. not ALL-P/T#/WHP+DO DO as question
combo +y +m -w2 +w2 +s"!@12to^*^@12do^*^!@12verb" +ff12

F13 WH questions
e.g. "***Who is he?***"
CL-P/T#+WHO+AUX
CL-P has clause punctuation ". ! ? : ; -";
combo +y +m -w1 +w1 +s"@13who^*^@auxi^*^*?" +ff13

F14 nominalisations
e.g. Your ***recommendation*** is indeed a great honour to me.
Covering: -tion, -ment, -ness, -ity, ... etc.
combo +y +m -w1 +w1 +s@14nominl +ff14

F15 gerunds
e.g. He still comes to the ***meeting*** though he is very sick today.
combo +y +m -w1 +w1 +s@15gerund +ff15

F16 nouns
e.g. Some ***books*** were put on the ***bookshelf*** earlier.
combo +y +m -w1 +w1 +s@noun +ff16

F17 agentless passives
e.g. I ***have been told*** that you are going to America.
1. BE+(ADV)+(ADV)+VBN
2. BE+N/PRO+VBN
combo +y +m -w1 +w1 +s"@verb_be^*^(*ed+@vbn_irr)^*^!by" +ff17

F18 BY passives
e.g. The window ***is opened by*** Jenifer.
1. BE+(ADV)+(ADV)+VBN+by
2. BE+N/PRO+VBN+by
combo +y +m -w1 +w1 +s"@verb_be^*^(*ed+@vbn_irr)^*^by" +ff18

F19 BE as main verb
e.g. "No, I ***am*** not."
combo +y +m -w1 +w1 +s"@19be^*^!(*ing+*ed+@vbn_irr)" +ff19

F20 existential THERE
e.g. ***There*** are several explanations.
1. there+(xxx)+BE
2. There's
combo +y +m -w1 +w1 +s"((@20there^*^@be)+(@be^@20there^*^*?))" +ff20

F21 THAT verb complements
e.g. I said ***that he went to the office***.
1. and/nor/but/or/also/ALL-P + that + DET/PRO/there/plural noun/proper noun/TITLE
2. PUB/PRV/SUA/seem/appear/seemed/appeared + that + xxx
where xxx is not V/AUX/CL-P/T#/and
3. PUB/PRV/SUA + PREP + xxx + N + that
where xxx is not N
4. T# + that
ALL-P has clause punctuation ". ! ? : ; - ,";
combo +y +m -w2 +w2 +s"((@21conect^that/*^@21det)+(@21prsv^that/*^
!(@21v+@21aux+and/*))+(@21prsv^@21prep^*^@noun^that/*))" +ff21

Appendix 4 (Continued)

F22 THAT adj. complements
e.g. I am glad ***that you like it***.
ADJ + T# + that
combo +y +m -w2 +w2 +s"@adj^that/*" +ff22

F23 WH clauses
e.g. I know ***what you mean.***
PUB/PRV/SUA + WHP/WHO + xxx
where xxx is not AUX (to exclude WH questions)
PUB has: acknowledge, add, admit, affirm, agree, ... etc.
PRV has: accept, anticipate, ascertain, assume, ... etc.
SUA has: agree, allow, arrange, ask, beg, command, ... etc.
combo +y +m -w1 +w1 +s"@prsverb^@23wh^!@auxi" +ff23

F24 infinitives
e.g. She told me what ***to do***.
to + (ADV) + VB
combo +y +m -w1 +w1 +s"((to/TO^@24vb)+(to/TO^@adv^@24vb))" +ff24

F25 present participial clauses
e.g. ***Stuffing his mouth with cookies***, Joe ran out the door.
T#/ALL-P + VBG + PREP/DET/WHP/WHO/PRO/ADV
combo +y +m -w1 +w1 +s"*ing^@25a" +f25a
combo +y +m -w2 +w2 +s"*ing/*^@25b" +f25b

F26 past participial clauses
e.g. ***Built in a single week,*** the house would stand for fifty years.
T#/ALL-P + VBN + PREP/ADV
combo +y +m -w1 +w1 +s"@allp^(*ed+@vbn_irr)" +ppunct +ff26

F27 past participial WHIZ deletions
e.g. The ***solution produced*** by this process is good for this problem.
N/QUANPRO + VBN + PREP/BE/ADV
combo +y +m -w2 +w2 +s"@27nounq^@17vbn^@27prepbe" +ff27

F28 present participial WHIZ deletions
e.g. The ***event causing*** this decline was examined by the investigators.
N + VBG
combo +y +m -w2 +w2 +s"@noun^*/VBG" +ff28

F29 THAT relative clauses on subject position
e.g. The dog ***that bit me*** has run away.
N + (T#) + that + (ADV) + AUX/V
combo +y +m -w2 +w2
+s"((@noun^that/*^@29auxv)+(@noun^that/*^@adv^@29auxv))" +ff29

F30 THAT relative clauses on object position
e.g. The dog ***that I saw*** was playing with Sarah.
N + (T#) + that + DET/SUBJPRO/POSSPRO/it/ADJ/plural noun/proper noun/
possessive noun/TITLE
combo +y +m -w2 +w2 +s"@noun^that/*^@30all" +ff30

F31 WH relative clauses on subject position
e.g. The man ***who likes popcorn*** is my uncle.
xxx + yyy + N + WHP + (ADV) + AUX/V
where xxx is not "ask, tell"
combo +y +m -w2 +w2 +s"((!@31no^*^@noun^@whp^@29auxv)
+(!@31no^*^@noun^@whp^@adv^@29auxv))" +ff31

Appendix 4 (Continued)

F32 WH relative clauses on object position
e.g. The man ***who Sally likes*** sits over there.
xxx + yyy + N + WHP + zzz
where xxx is not "ask, tell";
where zzz is not ADV, AUX, V;
combo +y +m -w2 +w2 +s"((!@31no^*^@noun^@whp^!@32no)" +ff32

F33 WH relative clauses pied pipes
e.g. The manner ***in which*** he was told is really bad.
PREP + WHP
combo +y +m -w1 +w1 +s@prep^@33whp +ff33

F34 sentence relatives
e.g. Bob likes fried mangoes, ***which is the most disgusting thing I've ever heard of.***
T#, + which
combo +y +m -w1 +w1 +s@comma^which +p34punct +ff34

F35 causative adverbial subordinators
e.g. He didn't come today ***because*** he has been ill since last week.
because, cos
combo +y +m -w1 +w1 +s(because+cuz+coz)^!of +ff35

F36 concessive adverbial subordinators
e.g. He still comes to the meeting ***though*** he is very sick today.
although, though, thou, tho
combo +y +m -w1 +w1 +s"although+though" +ff36

F37 conditional adverbial subordinators
e.g. ***If*** you pour oil on water, it floats.
if, unless
combo +y +m -w1 +w1 +s"if+unless" +ff37

F38 other adverbial subordinators
e.g. Be ready to start ***as soon as*** I give the signal.
Covering: as long as, as soon as, ..., so that xxx, such that xxx.
where xxx is not N/ADJ... etc.
combo +y +m -w1 +w1 +s(@38a+@38b+!@38no) +ff38

F39 prepositions
e.g. He does have the best car ***among*** his friends.
Covering: against, amid, amidst, among, ... etc.
combo +y +m -w1 +w1 +s@prep +ff39

F40 attributive adjectives
e.g. ***The big horse*** is running.
ADJ + ADJ/N
combo +y +m -w2 +w2 +s@adj^(@adj+@noun) +ff40

F41 predicative adjectives
e.g. The horse ***is big***.
1. BE + ADJ + xxx
where xxx is not ADJ, ADV, N;
2. BE + ADJ + ADV + xxx
where xxx is not ADJ, N.
combo +y +m -w2 +w2
+s"((@41be^@adj^!@41noa)+(@41be^@adj^@adv^!@41nob))" +ff41

Appendix 4 (Continued)

F42 adverbs
e.g. She dances ***beautifully***.
combo +y +m -w1 +w1 +s@adv +ff42

F43 type/token ratio
freq +y +m +z1W-400W +d4 +f43

F44 word length
mean length of the words in a text.
WordPerfect - File - Document Info - Average Word Length, then record one by one.

F45 conjuncts
e.g. I didn't go to the movie. ***Instead***, I went on a picnic.
Covering: alternatively, altogether, consequently, ... etc.
combo +y +m -w1 +w1
+s(@45conj1+(@allp^@45conj2^@allp1)+(@allp^rather^(@allp1+!@45no)))
+ppunct +ff45

F46 downtoners
e.g. I ***almost*** dropped the cake.
Covering: almost, barely, hardly, merely, ... etc.
combo +y +m -w1 +w1 +s@downton +ff46

F47 hedges
e.g. We don't count pennies when it comes to ***something like*** this.
Covering: something like, more or less, ... xxx sort of, xxx kind of;
where xxx is not DET/ADJ/POSSPRO/WHO.
combo +y +m -w2 +w2 +s"(@47hedg1+(!@47no^@47hedg2))" +ff47

F48 amplifiers
e.g. It is difficult to cross the desert by car, but not ***absolutely*** impossible.
Covering: absolutely, altogether, completely, ... etc.
combo +y +m -w1 +w1 +s@amplif +ff48

F49 emphatics
e.g. I ***do*** believe you.
for sure/a lot/such a/real + ADJ/so + ADJ/DO + V/just/really/most/more
combo +y +m -w2 +w2
+s"(@49emph1)+(real/*^@adj)+(so/*^@adj)+(@49do^*/VB)+(@49emph2)" +ff49

F50 discourse particles
e.g. ***Well***, it is time to leave now.
CL-P/T# + well/now/anyway/anyhow/anyways
combo +y +m -w1 +w1 +s@50clp^@particle +ppunct +ff50

F51 demonstratives
e.g. ***That*** beautiful dog belongs to Mr. White.
that/this/these/those
combo +y +m -w1 +w1
+s@51dem^(@noun+(@adj^@noun)+(@adv^@adj^@noun)) +ff51

F52 possibility modals
e.g. I ***might*** do it if I have time.
Covering: can, may, might, could, ... etc.
combo +y +m -w1 +w1 +s@possmod +ff52

Appendix 4 (Continued)

F53 necessity modals
e.g. You ***should*** brush your teeth before going to bed.
Covering: ought, should, must, ... etc.
combo +y +m -w1 +w1 +s@necemod +ff53

F54 predictive modals
e.g. I ***shall*** celebrate this event with you.
Covering: will, would, shall, ... etc.
combo +y +m -w1 +w1 +s@predmod +ff54

F55 public verbs
e.g. He ***admitted*** that was his fault.
Covering: acknowledge, add, admit, affirm, agree, ... etc.
combo +y +m -w1 +w1 +s@pubv +ff55

F56 private verbs
e.g. I ***believe*** that you have finished your homework by now.
Covering: accept, anticipate, ascertain, assume, believe, ... etc.
combo +y +m -w1 +w1 +s@priv +ff56

F57 suasive verbs
e.g. I ***agree*** with your view.
Covering: agree, allow, arrange, ask, beg, command, concede, ... etc.
combo +y +m -w1 +w1 +s@suasv +ff57

F58 SEEM/APPEAR
e.g. There ***seem*** to be every hope that business will get better.
combo +y +m -w1 +w1 +s@seemap +ff58

F59 contractions
e.g. ***He's*** here now.
combo +y +m -w1 +w1 +s@59cont +ff59

F60 THAT deletion
e.g. I ***think he went to*** school this morning.
1. PUB/PRV/SUA + (T#) + demonstrative pro/SUBJPRO
2. PUB/PRV/SUA + PRO/N + AUX/V
3. PUB/PRV/SUA + ADJ/ADV/DET/POSSPRO + (ADJ) + N + AUX/V
combo +y +m -w2 +w2 +s"@60prsv^!that/*" +ff60

F61 stranded prepositions
e.g. He was the candidate that I was ***thinking of***.
PREP + ALL-P/T#
combo +y +m -w1 +w1 +s@prep^@50clp +ppunct +ff61

F62 split infinitives
e.g. He tried ***to quietly leave*** the room.
to + adv + (ADV) + VB
combo +y +m -w2 +w2 +s"to/TO^@adv^*^@62vb" +ff62

F63 split auxiliaries
e.g. Her hair ***has been nicely cut*** and styled by the hair stylist.
AUX + ADV + (ADV) + VB
combo +y +m -w2 +w2 +s"@63aux^@adv^*^@63vb" +ff63

Appendix 4 (Continued)

F64 phrasal coordination
e.g. She can sign ***and*** dance.
1. ADV and ADV
2. ADJ and ADJ
3. V and V
4. N and N
combo +y +m -w2 +w2
+s"(@a^and/*^@a)+(@b^and/*^@b)+(@c^and/*^@c)+(@d^and/*^@d)" +ff64

F65 non-phrasal coordination
e.g. I went to his apartment, ***and*** I found he had moved out two months before.
1. T#/, + and + it/so/then/you/there + BE/demonstrative pronoun/SUBJPRO
2. CL-P + and
3. and + WHP/WHO/adverbial subordinator/discourse particle/conjunct
combo +y +m -w1 +w1 +s"@clp^and" +ppunct +f65a
combo +y +m -w2 +w2 +s"(and/*^@65it^@65be)+(and/*^@65whp)" +f65b

F66 synthetic negation
e.g. I have ***no*** any idea whether he would come or not.
1. no + QUANT/ADJ/N
2. Neither, nor
combo +y +m -w2 +w2 +s"(no/*^@66quant)+(neither/*+nor/*)" +ff66

F67 analytic negation
e.g. I have being told you that he is ***not*** here.
Not
combo +y +m -w1 +w1 +s@67neg +ff67

Appendix 5 Mean, Standard Deviation and Number of Texts from Four Sample Text Groups

	CA			CS			NA			NS			Features
	M	SD	N	M	SD	N	M	SD	N	M	SD	N	
F01	9.51	14.10	503	11.55	8.57	17	20.59	20.36	90	33.03	19.47	93	past tense verbs
F02	5.08	6.76	503	4.03	1.30	17	5.50	4.71	90	9.70	4.32	93	perfect aspect verbs
F03	42.95	25.85	503	102.62	10.56	17	67.14	24.51	90	126.96	26.81	93	present tense verbs
F04	2.55	5.49	503	2.02	0.73	17	2.31	3.19	90	1.82	2.11	93	place adverbials
F05	2.43	4.50	503	4.00	1.25	17	2.71	2.07	90	5.86	3.09	93	time adverbials
F06	32.44	22.72	503	26.00	5.11	17	9.61	13.98	90	59.87	17.42	93	first pers. pronouns
F07	10.31	13.99	503	11.10	3.32	17	1.87	6.74	90	29.14	13.31	93	second pers. pronouns
F08	9.79	12.25	503	13.85	5.66	17	11.19	10.48	90	25.37	14.74	93	third pers. pronouns
F09	8.19	10.35	503	8.17	2.48	17	6.03	4.21	90	18.72	9.09	93	pronoun IT
F10	4.15	6.01	503	4.47	1.12	17	2.49	2.07	90	11.59	5.19	93	demons. pronouns
F11	3.33	5.83	503	2.40	0.83	17	0.30	1.43	90	3.55	2.61	93	ind. pronouns
F12	0.34	1.74	503	0.65	0.34	17	0.91	1.82	90	7.49	4.19	93	DO as pro-verb
F13	0.93	3.29	503	1.27	0.52	17	0.00	0.00	90	0.77	1.03	93	WH questions
F14	21.59	16.23	503	13.01	4.30	17	36.73	13.42	90	10.46	7.34	93	nominalizations
F15	5.85	7.79	503	4.15	2.49	17	8.86	4.59	90	4.81	3.31	93	gerunds
F16	229.50	100.10	503	192.10	18.31	17	186.38	24.36	90	142.20	25.29	93	nouns
F17	6.03	8.17	503	3.34	0.91	17	15.92	7.75	90	4.87	3.16	93	agentless passives
F18	0.86	3.06	503	0.34	0.22	17	1.84	1.68	90	0.12	0.39	93	BY passives
F19	13.39	11.95	503	14.96	2.16	17	24.16	7.24	90	39.90	8.41	93	BE as main verb
F20	1.89	4.33	503	1.01	0.36	17	1.68	1.70	90	3.06	2.19	93	existential THERE
F21	1.91	4.06	503	1.23	0.60	17	3.32	2.70	90	4.44	3.10	93	THAT verb compl.
F22	0.23	1.01	503	0.10	0.14	17	0.41	0.76	90	0.23	0.50	93	THAT adj. compl.
F23	0.15	1.04	503	0.17	0.24	17	0.38	0.99	90	1.08	1.11	93	WH clauses
F24	11.72	11.00	503	12.89	1.96	17	14.06	6.97	90	14.27	4.75	93	infinitives
F25	0.09	0.72	503	0.35	0.34	17	1.19	1.47	90	0.00	0.17	93	pres. part. clauses
F26	0.19	1.35	503	0.08	0.11	17	0.38	0.68	90	0.00	0.00	93	past part. clauses

Appendix 5 Mean, Standard Deviation and Number of Texts from Four Sample Text Groups (Continued)

	CA			CS			NA			NS			Features
	M	SD	N	M	SD	N	M	SD	N	M	SD	N	
F27	0.80	2.38	503	0.37	0.35	17	5.12	3.85	90	0.15	0.44	93	past prt. WHIZ del.
F28	1.25	3.58	503	1.27	0.70	17	2.50	2.36	90	0.45	0.78	93	pres. prt. WHIZ del.
F29	1.08	3.38	503	0.50	0.29	17	0.26	0.62	90	0.71	1.23	93	THAT rel.: subj. pos.
F30	0.37	1.56	503	0.15	0.16	17	0.83	1.20	90	1.06	1.29	93	THAT rel.: obj. pos.
F31	1.14	2.96	503	0.68	0.46	17	2.69	2.52	90	1.23	1.65	93	WH rel.: subj. pos.
F32	0.15	0.85	503	0.08	0.11	17	2.10	1.98	90	0.80	1.26	93	WH rel.: obj. pos.
F33	0.15	1.21	503	0.00	0.00	17	1.26	1.20	90	0.36	0.87	93	WH rel.: pied pipes
F34	0.24	1.18	503	0.15	0.17	17	0.00	0.00	90	0.54	0.82	93	sentence relatives
F35	0.60	2.18	503	0.41	0.31	17	0.59	1.65	90	3.05	1.83	93	causative adv. sub.
F36	0.71	2.46	503	0.60	0.37	17	0.50	0.79	90	0.32	0.68	93	concessive adv. sub.
F37	4.76	6.81	503	2.71	1.15	17	2.09	2.12	90	4.03	2.64	93	conditional adv. sub.
F38	1.20	2.92	503	0.76	0.46	17	1.76	1.53	90	0.65	0.84	93	other adv. sub.
F39	72.50	29.16	503	76.61	8.01	17	137.21	18.51	90	86.61	19.19	93	prepositions
F40	27.80	16.35	503	21.15	2.76	17	76.86	16.21	90	43.68	11.54	93	attributive adj.
F41	3.20	5.38	503	2.51	0.89	17	5.27	2.76	90	4.98	2.75	93	predicative adj.
F42	33.41	20.19	503	38.10	4.38	17	51.58	10.60	90	83.37	14.20	93	adverbs
F43	55.44	6.02	43	58.22	4.12	17	50.87	5.15	90	46.56	3.53	93	type/token ratio
F44	4.97	0.88	503	4.82	0.40	17	4.80	0.20	90	4.14	0.22	93	word length
F45	2.20	3.93	503	1.16	0.61	17	2.94	2.24	90	0.52	0.91	93	conjuncts
F46	1.32	3.52	503	1.10	0.44	17	2.40	1.86	90	1.60	1.47	93	downtoners
F47	0.51	2.40	503	1.07	0.50	17	0.21	0.53	90	1.87	2.10	93	hedges
F48	2.29	4.19	503	1.04	0.66	17	1.46	1.67	90	6.02	2.77	93	amplifiers
F49	6.32	6.97	503	7.85	1.68	17	4.07	3.29	90	11.35	4.21	93	emphatics
F50	0.03	0.31	503	1.15	0.63	17	0.02	0.28	90	4.47	2.97	93	discourse particles
F51	4.17	6.05	503	3.37	1.01	17	11.47	4.67	90	10.29	4.14	93	demonstratives
F52	8.01	10.73	503	6.75	2.96	17	5.83	3.53	90	8.11	3.60	93	possibility modals

Appendix 5 Mean, Standard Deviation and Number of Texts from Four Sample Text Groups (Continued)

	CA			CS			NA			NS			Features
	M	SD	N	M	SD	N	M	SD	N	M	SD	N	
F53	1.63	4.06	503	1.46	0.74	17	2.20	2.30	90	2.29	1.72	93	necessity modals
F54	8.10	9.64	503	5.46	1.86	17	4.61	5.32	90	6.32	3.69	93	predictive modals
F55	6.46	8.11	503	44.79	64.98	17	6.09	5.30	90	7.97	3.93	93	public verbs
F56	14.72	12.02	503	17.11	2.60	17	13.01	6.03	90	32.69	10.09	93	private verbs
F57	5.72	8.40	503	3.85	1.00	17	3.97	5.62	90	1.81	1.86	93	suasive verbs
F58	1.11	2.81	503	0.80	0.41	17	1.00	1.21	90	0.56	0.90	93	SEEM/APPEAR
F59	11.45	11.99	503	17.51	2.45	17	0.61	3.11	90	43.66	15.82	93	contractions
F60	1.98	4.31	503	3.54	1.16	17	0.57	1.21	90	8.46	4.55	93	THAT deletions
F61	0.82	3.17	503	0.53	0.31	17	0.99	2.75	90	4.62	2.86	93	stranded prepositions
F62	0.08	0.59	503	0.14	0.17	17	0.00	0.00	90	0.00	0.00	93	split infinitives
F63	3.27	4.72	503	1.92	0.64	17	5.82	2.55	90	4.79	2.02	93	split auxiliaries
F64	3.72	6.02	503	2.38	0.89	17	4.38	2.73	90	1.12	1.12	93	phrasal coordination
F65	1.07	3.32	503	1.03	0.79	17	1.86	1.43	90	8.82	4.06	93	non-phr. coordination
F66	0.65	2.28	503	0.72	0.41	17	1.27	1.30	90	1.02	1.32	93	synthetic negation
F67	7.00	8.79	503	7.86	2.48	17	4.62	3.31	90	16.64	6.03	93	analytic negation

Appendix 6

Findings from 2 x 2 Factorial Experiment at $f \leqq 0.05$ by the Two Variables: Media (CMC vs. Non-CMC) and Temporalities (Asynchronous vs. Synchronous)

FEATURES	Interactive Effect								Main Effect			
	(C, N) x (A, S)											
	(C, N) for A		(C, N) for S		(A, S) for C		(A, S) for N		(C, N)		(A, S)	
	1	2	3	4	5	6	7	8	9	10	11	12
	CA > NA	NA > CA	CS > NS	NS > CS	CS > CA	CA > CS	NS > NA	NA > NS	C > N	N > C	S > A	A > S
(A) TENSE AND ASPECT MARKERS												
F01 past tense V.		V		V			V					
F02 perf. asp. V.				V			V					
F03 pres. tense V.										V	V	
(B) PLACE AND TIME ADVERBIALS												
F04 place adv.												
F05 time adv.											V	
(C) PRONOUNS AND PRO-VERBS												
(C1) PERSONAL PRONOUNS												
F06 1st per. pron.	V			V			V					
F07 2nd per. pron.	V			V			V					
F08 3rd per. pron.				V			V					
(C2) IMPERSONAL PRONOUNS												
F09 pron. IT				V			V					
F10 demonst. pron				V			V					
F11 indefinite pron.	V						V					
(C3) PRO-VERBS												
F12 DO as pro-V.				V			V					
(D) QUESTIONS												
F13 WH question												
(E) NOMINAL FORMS												
F14 nominal.		V				V		V				
F15 gerund												V
F16 noun									V			V
(F) PASSIVES												
F17 agentless pass.		V						V				
F18 BY-passive												V
(G) STATIVE FORMS												
F19 BE main V.		V		V			V					
F20 exist. THERE				V								

Appendix 6 (Continued)

Findings from 2 x 2 Factorial Experiment at $f \leqq 0.05$ by the Two Variables: Media (CMC vs. Non-CMC) and Temporalities (Asynchronous vs. Synchronous)

	Interactive Effect								Main Effect			
	(C, N) x (A, S)											
SIGNIFICANT/	(C, N) for A		(C, N) for S		(A, S) for C		(A, S) for N		(C, N)		(A, S)	
NON-SIGNIFICANT FEATURES	1	2	3	4	5	6	7	8	9	10	11	12
	CA > NA	NA > CA	CS > NS	NS > CS	CS > CA	CA > CS	NS > NA	NA > NS	C > N	N > C	S > A	A > S
(H) SUBORDINATION												
(H1) COMPLEMENTATION												
F21 THAT cl. as V. compl.										V		
F22 THAT cl. as adj.compl.												
F23 WH clause				V			V					
F24 infinitive												
(H2) PARTICIPIAL FORMS												
F25 pres parti. cl.		V	V					V				
F26 past parti. cl.												
F27 past parti. WHIZ del.		V						V				
F28 pres. part. WHIZ del.								V				
(H3) RELATIVES												
F29 THAT cl. on subj.												
F30 THAT cl. on obj.										V		
F31 WH cl. on subj.										V		V
F32 WH cl. on obj.		V		V				V				
F33 pied pip. const.		V						V				
F34 sentence rel.							V					
(H4) ADVERBIAL CLAUSES												
F35 causative subo.				V			V					
F36 concessive subo.												
F37 condit. subo.	V											
F38 other adv. subo.												V
(I) ADJECTIVES AND ADVERBS												
(I1)PREPOSITIONAL PHRASES												
F39 preposition		V						V				
(I2) ADJECTIVES AND ADVERBS												
F40 attrib. adj.		V		V		V		V				
F41 pred. adj.										V		
F42 adverb		V		V			V					

Appendix 6 (Continued)

Findings from 2 x 2 Factorial Experiment at $f \leqq 0.05$ by the Two Variables at: Media (CMC vs. Non-CMC) and Temporalities (Asynchronous vs. Synchronous)

SIGNIFICANT/ NON-SIGNIFICANT FEATURES	Interactive Effect								Main Effect			
	(C, N) x (A, S)								(C, N)		(A, S)	
	(C, N) for A		(C, N) for S		(A, S) for C		(A, S) for N					
	1	2	3	4	5	6	7	8	9	10	11	12
	CA > NA	NA > CA	CS > NS	NS > CS	CS > CA	CA > CS	NS > NA	NA > NS	C > N	N > C	S > A	A > S
(J) LEXICAL SPECIFICITY												
F43 type/token ratio	V		V		V			V				
F44 word length			V					V				
(K) LEXICAL CLASSES												
F45 conjunct												V
F46 downtoner												
F47 gen. hedge											V	
F48 amplifier				V			V					
F49 gen. emphatic				V			V					
F50 discourse part.				V	V		V					
F51 demonstrative										V		
(L) MODALS												
F52 possib. modal												
F53 necess. modal												
F54 predic. Modal												
(M) SPECIALIZED VERB CLASSES												
F55 public verb			V		V							
F56 private verb				V			V					
F57 suasive verb												
F58 SEEM, APPEAR												
(N) REDUCED FORMS AND STRUCTURES												
F59 contraction	V			V	V		V					
F60 THAT del.				V			V					
F61 final prep.				V			V					
F62 split inf.												
F63 split auxiliary										V		V
(O) COORDINATION												
F64 ph. coord.												V
F65 non-ph. coord.				V			V					

Appendix 6 (Continued)

Findings from 2 x 2 Factorial Experiment at $f \leqq 0.05$ by the Two Variables: Media (CMC vs. Non-CMC) and Temporalities (Asynchronous vs. Synchronous)

	Interactive Effect								Main Effect			
	(C, N) x (A, S)											
SIGNIFICANT/	(C, N) for A		(C, N) for S		(A, S) for C		(A, S) for N		(C, N)		(A, S)	
NON-SIGNIFICANT FEATURES	1	2	3	4	5	6	7	8	9	10	11	12
	CA > NA	NA > CA	CS > NS	NS > CS	CS > CA	CA > CS	NS > NA	NA > NS	C > N	N > C	S > A	A > S
(P) NEGATION												
F66 synthetic neg.												
F67 analytic neg.				V			V					
SUBTOTAL	6	11	4	24	4	2	23	11	1	7	3	8
OVERALL TOTAL	17		28		6		34		8		11	

Remarks: C stands for the medium CMC; N stands for the medium Non-CMC; A stands for the Asynchronous temporality; S stands for the Synchronous temporality;
CA: CMC Asynchronous; CS: CMC Synchronous; NA: Non-CMC Asynchronous; NS: Non-CMC Synchronous.
"V" means where the difference is statistically significant.

Appendix 7

Summary of the Factorial Structure: 7 Factors in Biber's Study (Biber, 1988a, p. 89-91)

Factor 1 (informational vs. involved production)

private verbs	.96
THAT deletion	.91
contractions	.90
present tense verbs	.86
2nd person pronouns	.86
DO as pro-verb	.82
analytic negation	.78
demonstrative pronouns	.76
general emphatics	.74
1st person pronouns	.74
pronoun IT	.71
BE as main verb	.71
causative subordination	.66
discourse particle	.66
indefinite pronouns	.62
general hedges	.58
amplifiers	.56
sentence relatives	.55
WH questions	.52
possibility modals	.50
non-phrasal coordination	.48
WH clause	.47
final prepositions	.43
(adverbs	.42)
(conditional subordination	.32)
nouns	-.80
word length	-.58
prepositions	-.54
type/token ration	-.54
attributive adjs.	-.47
(place adverbials	-.42)
(agentless passive	-.39)
(past participial WHIZ deletions	-.38)
(present participial WHIZ deletions	-.32)

Factor 2 (narrative versus non-narrative concerns)

past tense verbs	.90
third person pronouns	.73
perfect aspect verbs	.48
public verbs	.43
synthetic negation	.40
present participial clauses	.39
(present tense verbs	-.47)
(attributive adjs.	-.41)
(past participial WHIZ deletions	-.34)
(word length	-.31)

Factor 3 (explicit versus situation-dependent reference)

WH relative clauses on obj. positions	.63
pied piping constructions	.61
WH relative clauses on subj. Positions	.45
phrasal coordination	.36
nominalizations	.36
time adverbials	-.60
place adverbials	-.49
adverbs	-.46

Appendix 7 (Continued)

Factor 4 (overt expression of persuasion)

infinitives	.76
prediction modals	.54
suasive verbs	.49
conditional subordination	.47
necessity modals	.46
split auxiliaries	.44
(possibility modals	.37)

-- no negative features --

Factor 5 (abstract versus non-abstract information)

conjuncts	.48
agentless passives	.43
past participial clauses	.42
BY-passives	.41
past participial WHIZ deletions	.40
other adverbial subordinators	.39
(predicative adjs.)	.31
(type/token ratio	-.31)

Factor 6 (online informational elaboration)

THAT clauses as verb complements	.56
demonstratives	.55
THAT relative clauses on obj. Positions	.46
THAT clauses as adj. Complements	.36
(final prepositions	.34)
(existential THERE	.32)
(demonstrative pronoun	.31)
(WH relative clauses on obj. positions	.30)
(phrasal coordination	-.32)

Factor 7 (academic hedging)

SEEM/APPEAR	.35
(downtoners	.33)
(adverbs	.31)
(concessive subordination	.30)
(attributive adjs.	.30)

-- no negative features --

Appendix 8

Summary of the Factorial Structure: 26 Factors of CMC Texts in This Study

	Factor 1	
F42	Total adverbs	.745
F49	Emphatics	.584
F02	Perfect aspect	.432
F08	3rd person pron.	.410
F63	Split auxiliaries	.320
F15	Gerunds	.265
F16	Total other nouns	-.559
F44	Word length	-.497
	Factor 2	
F54	Predictive modals	.715
F59	Contractions	.595
F06	1st person pron.	.465
F24	Infinitives	.437
	-- no negative features --	
	Factor 3	
F36	ALTHOUGH, THOUGH: Concessive adverbial sbordinators	.689
F19	BE as main verb	.614
F09	Pron. IT	.376
	-- no negative features --	
	Factor 4	
F60	THAT deletion	.640
F56	Private verbs	.605
F51	Demonstratives	.534
F32	WH rel. cl. on obj. position	.273
	-- no negative features --	
	Factor 5	
F29	THAT rel. cl. on subj. position	.659
F11	Indefinite pron.	.633
	-- no negative features --	
	Factor 6	
F31	WH rel. cl. on subj. position	.747
F23	WH clauses	.551
	-- no negative features --	
	Factor 7	
F18	BY-passives	.741
F27	Past part. WHIZ deletion	.499
F04	Place adverbials	.468
	-- no negative features --	

Appendix 8 (Continued)

	Factor 8	
F61	Stranded prepositions	.736
F52	Possibility modals	.534
F67	Not: Analytic negation	.504
	-- no negative features --	
	Factor 9	
F46	Downtoners	.629
F07	2nd person pron.	-.503
F37	IF, UNLESS: Conditional adverbial subordinators	-.354
	Factor 10	
F12	Pro-verb DO	.788
F58	SEEM/APPEAR	.609
F21	THAT as verb complements	.432
	-- no negative features --	
	Factor 11	
F50	Discourse particles	.765
F55	Public verbs	.695
F03	Present tense	.410
	-- no negative features --	
	Factor 12	
F30	THAT rel. cl. on obj. position	.736
F20	Existential THERE	.512
F45	Conjuncts	.307
	-- no negative features --	
	Factor 13	
	-- no positive features --	
F13	Direct WH questions	-.775
	Factor 14	
F28	Present participial WHIZ deletion	.770
F41	Predicative adjectives	.567
F22	THAT as adj. Complements	.375
	-- no negative features --	
	Factor 15	
F35	Because: Causative adverbial subordinators	.743
F38	Other adverbial subordinators	.355
	-- no negative features --	

Appendix 8 (Continued)

	Factor 16	
F01	Past tense	.716
F05	Time adverbials	.530
	-- no negative features --	
	Factor 17	
F47	Hedges	.812
F51	Demonstrative pron.	.543
	-- no negative features --	
	Factor 18	
F14	Nominalizations	.586
F17	Agentless passives	.529
	-- no negative features --	
	Factor 19	
F25	Present participial clauses	.760
F33	Pied-piping relative clauses	.709
	-- no negative features --	
	Factor 20	
F57	Suasive verbs	.812
F39	Total prepositional phrases	.408
	-- no negative features --	
	Factor 21	
F65	Independent cl. coordination	.814
F40	Attributive adj.	.486
	-- no negative features --	
	Factor 22	
F62	Split infinitives	.524
	-- no negative features --	
	Factor 23	
F34	Sentence relatives	.684
F53	Necessity modals	.376
	-- no negative features --	
	Factor 24	
F26	Past participial clauses	.739
	-- no negative features --	

Appendix 8 (Continued)

	Factor 25	
F64	Phrasal coordination	.665
F48	Amplifiers	.586
	-- no negative features --	
	Factor 26	
F66	Synthetic negation	.806
	-- no negative features --	

Appendix 9

An Excerpt of Synchronous CMC Text from NETEACH-MOO

SESSION 6-A (MOO: What to DOO)

Gregor says, "So, let's start by sharing anything we have already tried with students here..."
Gregor nudges Frizzy.
Millie [Guest] says, "I am a newbie...HELP!!""
Yoshi [to Millie [Guest]]: OK, what can I do?
Chrissy [Guest] is a newbie, too.
Millie [Guest] says, "Just be patient...or maybe I should and just take it all in!""
Yoshi [to Gregor]: Show them the basic commands OHP sheet, will ya?
Frizzy says, "I've used schMOOze to hold conferences with students. I logged them (with permission, of course), then sent the log to the students via e-mail."
Gregor says, "I have briefly experimented with students for collaborative brainstorming... Iwant to expand my use of that."
Puzzled [to Gregor]: "you could expand on what you mean.
Gregor [to Yoshi]: I'm going to skip the basic commands for newbies today... they know how to talk, and that's enough for today! ;o)
Yoshi says, "I offered an online class at Frizzy's FUN, but it didn't turn out very well."
Millie [Guest] says, "How about cloosing off the systyem just for your purposes so that no one else can access it...Is that possible?""
Yoshi nods to Gregor.
Frizzy says, "This summer, I had a group of students do role play at schMOOze. They had a political debate (one was Bob Dole, another was Clinton, and one student logged on as God!)."
Yoshi says, "Page me if you have trouble: page yoshi "
Gregor [to Puzzled]: I mean having two (or more) students in the same virtual room brainstorming a topic and logging what they say. They can then look at their conversationlater and pice together what they need to make a next draft.
Millie [Guest] says, "Frizzy, that sounds like fun!""
Frizzy [to Millie [Guest]]: "Do you mean you'd like to have a "private" MOO session?"
George [Guest] Yoshi, what kind of problems did you encounter?
Frizzy nods to Millie. "it was really fun. The students had a blast."
Millie [Guest] says, "yes, that's correct, Friizzy!""
Smoody is going to try to join you.
Smoody has arrived.
Frizzy says, "You can lock the door of a private room here."
Yoshi says, "Hi Smoody."
Millie [Guest] says, "Hi Smoody...who are you really?"
Smoody [to all]: "Hi!
Dasha says, "Since access to the MOO is just beginning and since I've got it set up only in my office and in a lab where my class is a guest, which isn't our regular lab, I have started students with just basic conversing, mostly from the entrance gates. We haven't gotten beyond the novelty stage yet. Everyone who has tried it likes it, but access spots as well as lab times are extrememly limited"
Gregor greets Smoody. "Introduce yourself, Smoo!"
Smoody says, "I'm Suzan Moody at the University of Kansas."
Gregor [to Yoshi]: What activities did you have your students do?
Yoshi says, "I tried to start out my class with a class mailing list but there was some techie trouble on the list and didn't work out right...."
Yoshi says, "Then I tried to gather my students here at the same time but couldn't get over the time difference. That's a very important factor we have to keep in mind."
Millie [Guest] says, "I just wanted to see if I can properly input the page command...thanks- hope I didn't waste your time!""
Gregor says, "Anyone else have activities they've already tried? If not, we can start brainstorming some ideas."
Gregor drops a pin and goes on.

Bibliography

Abrams, M. H. (Gen. Ed.). (1987). *The Norton anthology of English literature, 5th Edition: The Major Authors*. London, UK: W.W. Norton & Co.

Adkins, M., & Brashers, D. E. (1995). The power of language in computer-mediated groups. *Management Communication Quarterly, 8*, 289-322.

Agre, P. E. & Schuler, D. (Eds.). (1997). *Reinventing technology, rediscovering community: Critical explorations of computing as a social practice.* Greenwich, CT: Ablex.

Ahern, T. C. (1994). The effect of interface on the structure of interaction in computer-mediated small-group discussion. *Journal of Educational Computing Research, 11*(3), 235-250.

Akinnaso, F. Niyi. (1982). On the differences between spoken and written language. *Lanaugage and Speech, 25*, 97-125.

Allerton, D. J.; Carney, Edward; & Holdcroft, David (Eds.). (1979). *Function and context in linguistic analysis*. Cambridge, UK: Cambridge University Press.

Allomong, S. (1996). Teaching the grammar of the "involved" factor: A resource model. *Journal of English Grammar on the Web, 1.* Retrieved September 19, 2000, from the World Wide Web: http://www.gsu.edu/~wwwesl/issue1/steph.htm.

Altenberg, Bengt. (1986). Constrastive linking in spoken and written English. In Gunnel Tottie & Ingegerd Backlund (Eds.), *English in speech and writing: A symposium* (pp. 13-40). Stockholm, Sweden: Almqvist and Wiksell.

Amoroso, H. C.; MacDonald, T. B.; Shedletsky, L. J.; & Travers, T. (1993). *Instructing K-12 teachers in computer networking and K-12 instructional practices in computer networking: Linking teachers and students to the global village.* Proceedings of the Annual Conference of the Eastern Communication Association. New Haven, CT: University of Maine. ERIC Document No. ED364212.

Anderson, Carl. (1998). *Word lore.* Retrieved January 28, 2001, from the World Wide Web: http://hea-www.harvard.edu/ECT/Words/.

Andersson, E. (1974). *Sentence and clause: A Taxonomy in a generative framework.* ERIC document ED154602.

Andreev, S. (1997). Classification of verbal characteristics based on their implication force by means of cluster and factor analysis. *Journal of Quantitative Linguistics, 4*(1-3), 23-29.

Asteroff, Janet F. (1987). *Paralanguage in electronic mail: A case study.* Unpublished educational doctoral dissertation, Columbia University, New York, USA.

Backlund, Ingegerd. (1986). Beat until stiff: Conjunction-headed abbreviated clauses in spoken and written English. In Gunnel Tottie & Ingegerd Backlund (Eds.), *English in speech and writing: A symposium* (pp. 13-40). Stockholm, Sweden: Almqvist and Wiksell.

Baecker, Ronald M. (Ed.). (1993). *Readings in groupware and computer-supported cooperative work: Assisting human-human collaboration.* San Mateo, CA: Morgan Kaufmann.

Bainbridge, William Sims. (1986). The structure of science fiction. In William Sims Bainbridge (Ed.), *Dimensions of science fiction* (pp. 28-50). San Francisco, USA: Harvard University Press.

Bainbridge, William Sims (Ed.). (1986). *Dimensions of science fiction.* San Francisco, USA: Harvard University Press.

Barlow, Michael, & Neumann, Steven. (1995). *MonoConc for Windows.* Houston, TX: Athelstan.

Barnitz, John G.., & Speaker, Richard B., Jr. (1999). Electronic and linguistic connections in one diverse 21st-century classroom. *Reading Teacher, 52*(8), 874-877.

Baron, Naomi S. (1984). Computer mediated communication as a force in language change. *Visible Language, 18*(2), 118-141.

Baron, Naomi S. (1998). Letters by phone or speech by other means: The linguistics of email, *Language and Communication, 18*(2), 133-170.

Barrett, E., & Lally, V. (1999). Gender differences in an on-line learning environment. *Journal of Computer Assisted Learning, 15*(1), 48-60.

Barson, J.; Frommer, J.; & Schwartz, M. (1993). Foreign language learning using e-mail in a task oriented perspective: Interuniversity experiments in communication and collaboration. *Journal of Science Education and Technology, 4*(2), 565-584.

Basch, R. (1995). The marriage of fax and online. *Online, 19*(1), 12-14, 16, 18-21.

Beaman, Karen. (1984). Coordination and subordination revisited: Syntactic complexity in spoken and written narrative disourse. In Deborah Tannen (Ed.), *Coherence in spoken and written discourse* (pp. 45-80). Norwood, NJ: Ablex.

Beauvois, M. H. (1992). Computer-assisted classroom discussion in the foreign language classroom: Conversation in slow motion. *Foreign Language Annals, 25*(5), 455-464.

Bernstein, Basil. (1970). *Class, codes, and control. Vol. 1: Theoretical studies towards a sociology of language.* London, UK: Routledge & Kegan Paul.

Bernstein, Herbert J.; Bourne, P.; & Bernstein, F. C. (1997). *Miniconcord.* Bellport, NY: Bernstein & Sons.

Berquist, B. J. (1999). Quote of the month in TAPPED IN monthly newsletter. *On the TAPIS August/September, 22.* Retrieved January 2, 2001, from the World Wide Web: http://www.tappedin.org/HyperNews/get/forums/tapis/31.html.

Berry, Wendell. (1992). *Sex, economy, freedom & community.* New York, NY: Pantheon Books.

Besnier, Niko. (1986). Register as a sociolinguistic unit: Defining formality. In J. Connor-Linton; C. J. Hall; & M. McGinnis (Eds.), *Social and cognitive perspectives on languages* (pp. 25-63). Los Angeles, CA: University of Southern California.

Bhatia, V. K. (1993). *Analyzing genre: Language use in professional settings.* London, UK: Longman.

Biber, Douglas. (1986). Spoken and written textual dimensions in English: Resolving the contradictory findings. *Language, 62*, 384-414.

Biber, Douglas. (1987). A textual comparison of British and American writing. *American Speech, 62*, 99-119.

Biber, Douglas. (1988a). *Variation across speech and writing.* Cambridge, UK: Cambridge University Press.

Biber, Douglas. (1988b). Adverbial stance types in English. *Discourse Processes, 11*(1), 1-34.

Biber, Douglas. (1989). A typology of English texts. *Linguistics, 27*, 3-43.

Biber, Douglas. (1993). The multi-dimensional approach to linguistic analyses of genre variation: An overview of methodology and findings. *Computers and the Humanities, 26,* 331-345.

Biber, Douglas. (1995). *Dimensions of register variation – A cross-linguistic comparison.* Cambridge, UK: Cambridge University Press.

Biber, Douglas; Conrad, Susan; & Reppen, Randi. (1994). Corpus-based approaches to issues in applied linguistics. *Applied Linguistics, 15*(2), 169-189.

Biber, Douglas; Conrad, Susan; & Reppen, Randi. (1998). *Corpus Linguistics: Investigating language structure and use.* Cambridge, UK: Cambridge University Press.

Bikson, Tora K.; Eveland, J. D.; & Gutek, Barbara A. (1989). Flexible interactive technologies for multi-person tasks: Current problems and future prospects. In Margrethe H. Olson (Ed.), *Technological support for work group collaboration* (pp. 89-112). Hillsdale, NJ: Lawrence Erlbaum Associates.

Björklund, Martina, & Virtanen, Tuija. (1991). Variation in narrative structure: A simple text vs. an innovative work of art. *Poetics, 20*(4), 391-403.

Blankenship, Jane. (1962). A linguistic analysis of oral and written style. *Quarterly Journal of Speech, 48*, 419-422.

Blankenship, Jane. (1974). The influence of mode, submode, and speaker predilection on style. *Speech Monographs, 41*, 85-118.

Blankenship, Jane. (1978, April). *In the presence of the word: Rhetoric and responsibility.* Paper presented at the Annual Meeting of the Southern Speech Communication Association, Atlanta, USA.

Blass, Thomas, & Siegman, Aron W. (1975). A psycholinguistic comparison of speech, dictation, and writing. *Language and Speech, 18*, 20-33.

Blom, Jan-Petter, & Gumperz, John J. (1968). Some social determinants of verbal behavior. In John J. Gumperz & Dell H. Hymes (Eds.), *Directions in sociolinguistics* (pp. 407-434). Texas, TX: Holt, Rinehart, and Winston College Language Publishers.

Blount, Ben G., & Sanches, Mary (Eds.). (1977). *Sociocultural dimensions of language change.* New York, NY: Academic Press.

Boardman, Phillip C. (Ed.). (1987). *The legacy of language.* Amsterdam, The Netherlands: John Benjamins.

Bonham, L. A.; Cifuentes, L.; & Murphy, K. (1995, March). Constructing cultures in distance education. Paper presented in the 6th International Conference of the Society of Information Technology and Teacher Education, Texas, USA.

Botta, R. (1993, August). *Does shorter mean easier to understand?-- A study of comprehension of USA Today information stories.* Paper presented at the annual meeting of the Association for Education in Journalism and Mass Communication, Kansas City, USA.

Bowers, R. (1995). A computer-mediated scientific writing program. *TESL-EJ 1*(3), 1-7.

Brainerd, Barron. (1972). An exploratory study of pronouns and articles as indices of genre in English. *Language and Style, 5*, 239-259.

Bright, William. (1960). Social dialect and language history. *Current Anthropology, 6*(5-6), 424-425.

Bright, William. (1966). The dimensions of sociolinguistics. In William Bright (Ed.), *Sociolinguistics* (pp. 11-15). The Hague, The Netherlands: Mouton.

Bright, William (Ed.). (1966). *Sociolinguistics.* The Hague, The Netherlands: Mouton.

Broota, K. D. (1989). *Experimental design in behavioral research.* New York, NY: John Wiley & Sons.

Brown, H. Douglas. (1994). *Principles of language learning and teaching.* Englewood Cliffs, NJ: Prentice Hall, Inc.

Brown, H. Douglas; Yorio, C. A.; & Crymes, R. H. (Eds.). (1977). *On TESOL '77: Teaching and learning English as a second language: Trends in research and practice.* Washington, DC: TESOL.

Brown, Penelope, & Fraser, Colin. (1979). Speech as a marker of situation. In Klaus R. Scherer, & Howard Giles (Eds.), *Social markers in speech* (pp. 33-62). Cambridge, UK: Cambridge University Press.

Brown, Gillian, & Yule, George. (1983). *Discourse analysis.* Cambridge, UK: Cambridge University Press.

Bruffee, Kenneth A. (1984). Collaborative learning and the "conversation of mankind." *College English, 46*(7), 635-652.

Bruffee, Kenneth A. (1986). Social construction, language and the authority of knowledge: A bibliographical essay. *College English, 48*(8), 773-790.

Buckland, M., & Dye, C. (1991). *The development of electronic distance education delivery systems in the United States: Recurring and emerging themes in history and philosophy of education.* ERIC Document No. ED 345 713.

Burling, Robbins. (1970). *Man's many voices: Language in its cultural context.* New York, NY: Holt, Rinehard and Winton.

Burnard, Lou. (1995). *Reference Guide to the SARA Windows client.* Retrieved April 12, 1999, from the World Wide Web: http://info.ox.ac.uk/bnc/getting/chap4.htm.

Burton, P. F. (1994). Electronic mail as an academic discussion forum. *Journal of Documentation, 50*, 99-110.

Cain, Gordon. (2000). *Re: Corpora: minimum size of corpus?* Retrieved February 10, 2000, from the World Wide Web: http://www.hd.uib.no/corpora/2000-1/0136.html.

Caldas, Celia Pereira. (1998, July 29). *Corpora: Corpora and software.* Retrieved January 22, 1999, from the World Wide Web: http://www.hd.uib.no/corpora/archive.html.

Canale, Michael. (1983). From communicative competence to communicative language pedagogy. In J. C. Richards & R. Schmidt (Eds.), *Language and communication* (pp. 2-27). London, UK: Longman.

Canale, Michael, & Swain, Merrill. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics, 1*, 1-47.

Carey, J. (1980). Paralanguage in computer mediated communication. In N. K. Sondheimer (Ed.), *The 18th annual meeting of the association for computational linguistics and parasession on topics in interactive discourse: Proceedings of the conference* (pp. 67-69). Philadelphia, PA: University of Pennsylvania.

Cedergren, H., & Sankoff, D. (Eds.). (1981). *Variation omnibus.* Edmonton, Canada: Linguistic Research.

Chafe, Wallace L. (1982). Integration and involvement in speaking, writing, and oral literature. In Deborah Tannen (Ed.), *Spoken and written language: Exploring orality and literacy* (pp. 35-54). Norwood, NJ: Ablex.

Chafe, Wallace L. (1984). Speaking, writing, and prescriptivism. In Deborah Schiffrin (Ed.), *Meaning, form, and use in context: Linguistic applications* (pp. 95-103). Washington, DC: Georgetown University Press.

Chafe, Wallace L. (1985). Linguistic differences produced by differences between speaking and writing. In David R. Olson; N. Torrance; & A. Hildyard (Eds.), *Literature, language, and learning: the nature and consequences of reading and writing* (pp. 105-123). Cambridge, UK: Cambridge University Press.

Chafe, Wallace L., & Danielewicz, Jane. (1987). *Properties of spoken and written language.* Berkeley, CA: Center for the Study of Writing.

Chapelle, C. (1990). The discourse of computer-assisted language learning: Toward a context for descriptive research. *TESOL Quarterly, 24*(2), 199-225.

Chenault, Brittney G. (1998). Developing personal and emotional relationships via computer-mediated communication. *Computer-Mediated Communication Magazine, May, 1998.* Retrieved June 16, 1999, from the World Wide Web: http://www.december.com /cmc/mag/1998/may/chenault.html.

Chesebro, J. W. (1985). Computer-mediated interpersonal communication. In B. D. Ruben (Ed.), *Information and behavior, Vol. 1* (pp. 202-222). New Brunswick, NJ: Transaction.

Chesebro, J. W., & Bonsall, D. G. (1989). *Computer-mediated communication: Human relationship in a computerized world.* Tuscaloosa, AL: The University of Alabama Press.

Child, J., & R. Loveridge. (1990). *Information technology in European services-Towards a microelectronic future.* Oxford, UK: Blackwell.

Chou, C. T. (1985a). Television and computer and their integration in teaching English to speaker of other language. *Studies in English Literature and Linguistics, 11*, 173-184.

Chou, C. T. (1985b). Some considerations on the design of programmed-instruction type of CAI courseware. *Teaching and Research, 7*, 213-226. Taipei: National Taiwan normal University, College of Arts.

Chou, C. T. (1996). *Intensity devices used in Chinese BBS texts as compared with those in newspaper forums.* Paper presented in the international conference of Social Linguistic Problems in Various Regions of the World, Moscow, Russia.

Chou, C. T. (1998). *A corpus study on the English intensifying devices in computer-mediated communication.* Taipei, Taiwan: Crane Publishing Corp.

Chun, Dorothy M. (1994). Using computer networking to facilitate the acquisition of interactive competence. *System, 22*, 17-31.

Close, E. O. (1994). Recollect orality. *Visible Language, 28*(2), 100-109.

Cole, George. (1998). *New wave learning.* TES Online editorial dated January 16, 1998. Times Educational Supplement, London, UK.

Cole, Peter (Ed.). (1978). *Syntax and Semantics, Vol. 9: Pragmatics.* New York, NY: Academic Press.

Cole, Peter (Ed.). (1981). *Radical pragmatics.* New York, NY: Academic Press.

Collot, Milena. (1991). Electronic language: A pilot study of a new variety of English. *Computer Corpora des Englischen, 5*(1-2), 13-31.

Collot, Milena, & Belmore, Nancy. (1996). Electronic Language: A new variety of English. In Susan C. Herring (Ed.), *Computer-mediated communication: Linguistic, social, and cross-cultural perspectives* (pp. 13-28). Amsterdam, The Netherlands: John Benjamins Publishing Company.

Coniam, David. (1996, August). *Using corpus word frequency data in the automatic generation of English language cloze tests.* Paper presented at the Teaching and Language Corpora Conference, Lancaster, UK.

Connor, Ulla, & Kaplan, Robert B. (Eds.). (1987). *Writing across languages: Analysis of L2 text.* Reading, MA: Addison-Wesley.

Connor-Linton, J.; Hall, C. J.; & McGinnis, M. (Eds.). (1986). *Social and cognitive perspectives on languages.* Los Angeles, CA: University of Southern California.

Conrad, K. B. (1996). CALL-non-English L2 instruction. *Annual Review of Applied Linguistics, 16,* 158-181.

Cook, Guy. (1989). *Discourse.* Oxford, UK: Oxford University Press.

Cook, Guy. (1997). Language play, language learning. *ELT Journal, 51*(3), 224-231.

Corio, Ron. (1996). [Review of the book *E-mail for English teachers: Bringing the Internet and computer learning networks into the language classroom* by Mark Warschauer]. *TESL-EJ, 2*(1), R-5, 1-4. Retrieved August 20, 1999, from the World Wide Web: http: //www.zait.uni-bremen.de/wwwgast/tesl_ej/.

Costello, R. B. (Ed.). (1993). *The American heritage college dictionary, 3rd ed.* Boston, MA: Houghton Mifflin Company.

Coulthard, M. (Ed.). (1986). *Talking about text.* Birmingham, UK: Birmingham Instant Print Ltd.

Coyle, Larry, & Spitzer, Kathleen L. (1992). Technology news; distance education project: Extending extension programming via telecommunications technology; and fax for library services. *Educational Technology, 32*(8), 44, 57-60.

Crenn, Triphaine. (1996). *Register and register labeling in dictionaries.* Unpublish PhD thesis, School of Translation and Interpretation, University of Ottawa. Retrieved January 21, 2001, from the World Wide We: http://www.dico.uottawa.ca/theses/crenn/crenn.htm.

Crystal, David. (1969). *What is linguistics?* London, UK: Edward Arnold.

Crystal, David. (1987). *The Cambridge encyclopaedia of language.* Cambridge, UK: Cambridge University Press.

Crystal, David. (1997). *English as a global language.* Cambridge, UK: Cambridge University Press.

Crystal, David. (1999). *The Penguin dictionary of language.* London, UK: Penguin Books.

Crystal, David, & Davy, D. (1969). *Investigating English Style.* London, UK: Longman.

Daft, R. L., & Lengel, R. H. (1984). Information richness: A new approach to managerial behavior and organization design. In B. M. Staw & L. L. Cummings (Eds.), *Research in organisational behavior* (pp. 191-233). Greenwich, CT: JAI.

Daft, R. L., & Lengel, R. H. (1986). Organizational information requirements, media richness and structural design. *Management Science, 32*(5), 554-571.

Dahl, Osten. (1976). What is new information? In Nils Erik Enkvist & Viljo Kohonen (Eds.), *Report on text linguistics: Approaches to word order* (pp. 37-50). Abo, Finland: Abo Akademi Foundation.

Davie, L. (1988). Facilitating adult learning through computer mediated distance education. *The American Journal of Distance Education, 3*(2), 55-69.

Davis, B. H., & Brewer, J. P. (1997). *Electronic discourse: Linguistic individuals in virtual space.* Albany, NE: SUNY Press.

Davis, Randall S. (1999). ESL syber listening. *American Language Review, 3*(5), 86.

Davis, Randall S. (2000). Digital voice journals: Beyond pencil and paper. *American Language Review, 4*(2), 86.

Derewianka, B. (1996). *Exploring the writing of genres.* Herts, UK: United Kingdom Readding Association.

Des Brisay, Margaret; Duquette, Lise; & Dirir, Mohamed. (1993). *Acquisition versus learning as variables in the trait structure of L2 test data sets.* Paper presented at the Language Testing Research Colloquium Conference in Cambridge, UK.

DeVito, Joseph A. (1966). Psychogrammatical factors in oral and written discourse by skilled communicators. *Speech Monographs, 33,* 73-76.

DeVito, Joseph A. (1967). Levels of abstraction in spoken and written language. *Journal of Communicaiton, 17,* 354-361.

Dewaele, Jean-Mare. (1993). Variation in the morphosyntactic and lexical systems of French-based interlanguages. In Bernhard. Ketteman & W. Wieden (Eds.), *Current issues in European second language acquisition research* (pp. 130-140). Tubingen, Germany: Narr.

Dillard, Joey Lee. (1972). *Black English: Its history and usage in the United States*. New York, NY: Random House.

Donmall, G. (1985). *Language awareness, NCLE Papers and Reports 6*. London, UK: Centre for Information on Language Teaching and Research.

Dorak, M. Tevfik. (2000). *Common Concepts in statistics*. Retrieved October 1, 2000, from the World Wide Web: http://dorakmt.tripod.com/mtd/glosstat.html.

Drieman, G. H. J. (1962). Differences between written and spoken language. *Acta Psychologica, 20*, 36-57, 78-100.

D'Souza, P. V. (1992). Electronic mail in academic settings: A multipurpose communications tool. *Educational Technology, 22*(3), 22-25.

Dunning, Ted. (1995). *Factor analysis*. Retrieved August 8, 1995, from the World Wide Web: http://www.hd.uib.no/corpora/.

Easton, Valerie J., & McColl, John H. (2000). *Statistics Glossary*. Retrieved October 1, 2000, from the World Wide Web: http://www.cas.lancs.ac.uk/glossary_v1.1/main.html.

Ebo, B. (Ed.). (1998). *Cyberghetto or cybertopia: Race, class and gender on the Internet*. Westport, CT: Prager.

Egan-Robertson, A., & Bloome D. (Eds.). (1998). *Students as researchers of culture and language in their own communities*. Cresskill, NJ: Hampton.

Egbert, Joy, & Hanson-Smith, Elizabeth (Eds.). (1998). *CALL environments: Research, practice, and critical issues*. Alexandria, VA: TESOL Publications.

Ellis, R. (1985). Understanding second language acquisition. Oxford: Oxford University Press.

Elsness, Johan. (1983). The preterite and the perfect in BBC news bulletins: The case for a text linguistic approach. In Hakan Ringbom, & Matti Rissanen (Eds.), *Proceedings from the Nordic Conference for English Studies, 2nd*. Hanaholmen, Finland: Hanasaari.

Emig, J. (1977). Writing as a mode of learning. *College Composition and Communication, 28*(2), 122-128.

Erickson, Thomas. (1999). Persistant conversation: An introduction. *Journal of Computer-Mediated Communication, 4*(4). Retrieved November 28, 2000, from the World Wide Web: http://www.ascusc.org/jcmc/vol4/issue4/ericksonintro.html

Ervin-Tripp, Susan. (1972). On sociolinguistic rules: Alternation and co-occurrence. In John J. Gumperz & Dell H. Hymes (Eds.), *Directions in sociolinguistics* (pp. 213-250). New York, NY: Holt, Rinehart and Winston.

eTForecasts. (2001). *Internet users will surpass 1 billion in 2005*. Retrieved February 11, 2001, from the World Wide Web: http://www.etforecasts.com/pr/pr201.htm.

Eveland, J. D., & Bikson, T. K. (1993). Work group structures and computer support: A field experiment. In Ronald M. Baecker (Ed.), *Readings in groupware and computer-supported cooperative work: Assisting human-human collaboration* (pp. 443-456). San Mateo, CA: Morgan Kaufmann.

Fairclough, Norman. (1989). *Language and power*. London, UK: Longman.

Fairclough, Norman. (1992). *Discourse and social change*. Cambridge, UK: Polity Press.

Fairclough, Norman. (1995). *Media discourse*. London, UK: Edward Arnold.

Falk, Jim. (1995, April). *The meaning of the web*. Paper presented at the AusWeb95 Conference, South Wales, Australia.

Falsetti, Julie. (1995). *What the heck is a MOO? and what's the story with all those cows?* ERIC document ED389222.

Feenberg, A. (1989). The written word: On the theory and practice of computer conferencing. In Robin Mason & Anthony Kaye (Eds.), *Mindweave: communication, computers and distance education* (pp. 22-39). Oxford, UK: Pergamon Press.

Feigenbaum, Irwin. (1978). *The use of the perfect in an academic setting: A study of types and frequencies*. Unpublished PhD dissertation. Milwaukee, WI: University of Wisconsin-Milwaukee, USA.

Ferguson, Charles A. (1977). Baby talk as a simplified register. In C. E. Snow & Charles A. Ferguson (Eds.), *Talking to children: Language input and acquisition* (pp. 219-236). Cambridge, UK: Cambridge University Press.

Ferguson, Charles A. (1959). Diglossia. *Word, 15*, 325-340.

Ferrara, K.; Brunner, H.; & Whittemore, G. (1991). Interactive written discourse as an emergent register. *Written Communication, 8*, 8-34.

Finegan, Edward. (1980). *Attitudes toward English usage: The history of a war of words.* New York, NY: Teachers College Press.

Finegan, Edward. (1987). On the linguistic forms of prestige. In Phillip C. Boardman (Ed.), *The legacy of language* (pp. 391-398). Amsterdam, The Netherlands: John Benjamins.

Finegan, Edward, & Biber, Douglas. (1986a). Toward a unified model of sociolinguistic prestige. In David Sankoff (Ed.), *Diversity and diachrony* (pp. 391-398). Amsterdam, The Netherlands: John Benjamins.

Finegan, Edward, & Biber, Douglas. (1986b). Two dimensions of linguistic complexity in English. In J. Connor-Linton, C. J. Halls, & M. McGinnis (Eds.), *Social and cognitive perspectives on language* (pp. 1-24). Los Angeles, LA: University of Southern California.

Finegan, R. (1988). *Literacy and orality: Studies in the technology of communication.* Oxford, UK: Basil Blackwell.

Finholt, T., & Sproull, L. (1990). Electronic groups at work. *Organization Science, 1*, 41-64.

Firth, J. R. (1935). The technique of semantics. Reprinted in J. R. Firth (1957), *Papers in linguistics 1934-1951* (pp. 7-33). London, UK: Oxford University Press.

Firth, J. R. (1957). *Papers in linguistics 1934-1951.* London, UK: Oxford University Press.

Fischer, John L. (1964). Social influences in the choice of a linguistic variant. In Dell H. Hymes (Ed.), *Language in culture and society: A reader in linguistics and anthropology* (pp. 483-488). New York, NY: Harper & Row.

Fishman, J. (Ed.). (1962). *Readings in the Sociology of Language.* The Hague, The Netherlands: Mouton.

Flowerdew, Lynne. (1998, August). *A corpus-based analysis of referential and pragmatic errors in student writing.* Paper presented at the Teaching and Language Corpora Conference, Oxford, UK.

Fotos, S. (Ed.). (1996). *Multimedia language teaching.* Tokyo, Japan: Logos International.

Fowler, Roger; Hodge, Bob; Kress, Gunther; & Trew, Tony (Eds.). (1979). *Language and control.* London, UK: Routledge.

Frawley, William (Ed.). (1982). *Linguistics and literacy.* New York, NY: Plenum Press.

Freeman, David E.; Fennacy, Jean W.; & Freeman, Y. (1993). California's reading revolution: A review and analysis. *The New Advocate, 6*(1), 41-60.

Frizler, Karla L. (1995). *EFL composition through intercultural understanding.* Online workshop. Retrieved December 20, 2000, from the World Wide Web: http://thecity.sfsu.edu /~funweb/.

Frommer, Judith G., & Weitz, Margaret Collins. (1981). *Sound* learning: Using taped interviews in foreign language courses. *French Review, 55*(2), 233-240.

Fulcher, G. (Ed.). (1997). *Writing in the English language classroom.* Gallen, Switzerland: Prentice Hall Europe.

Fulk, Janet; Schmitz, J. A.; & Schwarz, D. (1992). The dynamics of context-behaviour interactions in computer-mediated communication. In Martin Lea (Ed.), *Contexts of computer-mediated communication* (pp. 7-29). London, UK: Harvester Wheatsheaf.

Fuller, R. (1994). Human-computer-human interaction: How computers affect interpersonal communication. *The Electronic Journal on Virtual Culture, 2*(2). Retrieved November 30, 1996, from the World Wide Web: http://www.uni-koeln.de/themen/cmc/text/fuller.94.txt.

Furmanovsky, Michael. (1996, November). *Content video in the EFL classroom.* Paper presented at the 22nd Annual Meeting of the Japanese Association of Language Teachers, Japan.

Gaies, S. J. (1977). The nature of linguistic input in formal second language learning: Linguistic and communicative strategies in ESL teachers' classroom language. In H. Douglas Brown, C. A. Yorio, & R. H. Crymes (Eds.), *On TESOL '77: Teaching and learning English as a*

second language: Trends in research and practice (pp. 204-212). Washington, DC: TESOL.

Geddes, Marion, & White, Ron. (1978). The use of semi-scripted simulated authentic speech in listening comprehension. *Audio-visual Language Journal, 16*(3), 137-145.

Gerrey, W.; Brabyn, J.; & Crandall, W. (1990). The use of fax technology to address the reading needs of blind and visually impaired persons. *Journal of Visual Impairment & Blindness, 84*(10), 509-513.

Giddens, Anthony. (1979). *Central problems in social theory.* London, UK: Macmillan.

Givon, Tamly (Ed.). (1979). *Discourse and syntax.* New York, NY: Academic Press.

Gledhill, Chris. (1996, August). *The ideology of science as a collocation: How corpus linguistics can expand the boundaries of genre analysis.* Paper presented at the Teaching and Language Corpora Conference, Lancaster, UK.

Goody, J. (1986). *The logic of writing and the organisation of society.* Cambridge: Cambridge University Press.

Grabe, William. (1987). Constrastive rhetoric and text-type research. In Ulla Connor & Robert B. Kaplan (Eds.), *Writing across languages: Analysis of L2 text* (pp. 115-137). Reading, MA: Addison-Wesley.

Grabowski, B.; Pusch, S.; & Pusch, W. (1990). Social intellectual value of CMC in a graduate community. *Educational & Training Technology International, 27*(3), 276-283.

Graham, Philip Bruce. (1980). *Teaching oral communication skills via authentic material: A description and analysis of a learning experience in Japan.* Unpublished Master thesis, School of International Training, Vermont, USA.

Granger, Sylviane (Ed.). (1998). *Learner English on computer.* Essex, UK: Addison Wesley Longman Limited.

Granger, Sylviane, & Tribble, Chris. (1998). Learner corpus data in the foreign language classroom: Form-focused instruction and data-driven leaning. In Sylviane Granger (Ed.), *Learner English on computer* (pp. 199-209). Essex, UK: Addison Wesley Longman Limited.

Grice, H. P. (1981). Presupposition and conversational implicature. In Peter Cole (Ed.), *Radical pragmatics* (pp. 183-198). New York, NY: Academic Press.

Grune, D. (1997). *Klingon is an artificial language.* ftp://ftp.cs.vu.nl/pub/dick/publications /MutsunKlingonComparison.txt.

Gumperz, John J. (1982). *Language and social identity.* Cambridge, UK: Cambrdige University Press.

Gumperz, John J., & Hymes, Dell H. (Eds.). (1968). *Directions in sociolinguistics.* Texas, TX: Holt, Rinehart, and Winston College Language Publishers.

Gumperz, John J., & Hymes, Dell H. (Eds.). (1972). *Directions in sociolinguistics: The ethnography of communication.* New York, NY: Holt, Rinehart and Winston.

Gumperz, John J.; Kaltman, Hannah; & O'Connor, Mary Catherine. (1984). Cohesion in spoken and written discourse. In Deborah Tannen (Ed.), *Coherence in spoken and written discourse* (pp. 3-20). Norwood, NJ: Ablex.

Gunawardena, C. N. (1994). *Social presence theory and implications for building online communities.* Paper presented at the Third International Symposium on Telecommunications in Education (TEL-ED '94), New Mexico, USA.

Hadley, Alice Omaggio. (1993). *Teaching language in context.* Boston, MA: Heinle & Heinle Publishers.

Hall, Cathy. (1998). *"Constructing" language at MundoHispano.* Unpublished manuscript, George Mason University, Fairfax, Virginia, USA.

Halliday, M. A. K. (1973). *Explorations in the functions of language.* London, UK: Edward Arnold.

Halliday, M. A. K. (1974). *Language and social man.* London, UK: Longman Group Ltd.

Halliday, M. A. K. (1975). *Learning how to mean: Explorations in the development of language.* London, UK: Edward Arnold.

Halliday, M. A. K. (1976a). The form of a functional grammar. In Gunther Kress (Ed.), *Halliday: System and function in language* (pp. 7-25). Oxford, UK: Oxford University Press.

Halliday, M. A. K. (1976b). Functions and universals of language. In Gunther Kress (Ed.), *Halliday: System and function in language* (pp. 26-31). Oxford, UK: Oxford University Press.

Halliday, M. A. K. (1976c). English system networks. In Gunther Kress (Ed.), *Halliday: System and function in language* (pp. 99-135). Oxford, UK: Oxford University Press.

Halliday, M. A. K. (1976d). Modality and modulation in English. In Gunther Kress (Ed.), *Halliday: System and function in language* (pp. 189-213). Oxford, UK: Oxford University Press.

Halliday, M. A. K. (1978). *Language as social semiotic: The social interpretation of language and meaning*. London, UK: Edward Arnold Ltd.

Halliday, M. A. K. (1979). Differences between spoken and written language: Some implications for literacy teaching. In Glenda Page; John Elkins; & Barrie O'Connor (Eds.), *Communication through reading: Proceedings of the 4th Australian Reading Conference, Vol. 2* (pp. 37-52). Adelaide, Australia: Australian Reading Association.

Halliday, M. A. K. (1985). *An introduction to functional grammar*. London, UK: Edward Arnold.

Halliday, M. A. K. (1989). *Spoken and written language*. Oxford, UK: Oxford University Press.

Halliday, M. A. K. (1990). New ways of meaning: A challenge to applied linguistics. *Journal of Applied Linguistics, 6*, 7-36.

Halliday, M. A. K., & Hasan, Ruqaiya. (1976). *Cohesion in English.* London, UK: Longman.

Halliday, M. A. K. & Hassan, Ruqaiya. (1989). *Language, context, and text: Aspects of language in a social-semiotic perspective*. Oxford: Oxford University Press.

Halliday, M. A. K.; McIntosh, Angus; & Strevans, Peter. (1964). *The linguistic sciences and language teaching*. London: Longman Group Ltd.

Halliday, M. A. K.; McIntosh, Angus; & Strevans, Peter. (1964). *The linguistic sciences and language teaching*. London: Longman Group Ltd.

Harasim, L. M. (1987, June). *Computer-mediated cooperation in education: Group learning networks*. Paper presented at the Guelph Symposium on Computer Conferencing, Guelph, Ontario, Canada.

Harasim, L. M. (1989). On-line education: A new domain. In R. Mason, & A. Kaye (Eds.), *Mindweave* (pp. 50-62). Oxford, UK: Pergamon Press.

Harasim, L. M. (Ed.). (1990). *Online education: Perspectives on a new environment*. New York, NY: Praeger.

Hardy, V.; Hodgson, V.; & McConnell, David. (1994). Computer conferencing: A new medium for investigating issues in gender and learning. *Higher Education, 28*, 403-418.

Hart, R. (1996). *Interview with Clifford Stoll: E-mail connection - Love it or leave it?* Retrieved January 18, 1998, from the World Wide Web: http://zeppo.cnet.com/Content/Features /Dlife/Email/stoll.html.

Havelock, E. A. (1986). Orality, literacy, and star wars. *Written Communication, 3*(9), 411-20.

Hawkins, E. (1987). *Awareness of language: An introduction*. New York, NY: Cambridge University Press.

Haynes, Cynthia, & Holmevik, Jan Rune (Eds.). (1998). *High wired: On the design, use, and theory of educational MOOs*. Ann Arbor, MI: The University of Michigan Press.

Henzl, Vera M. (1973). Linguistic register of foreign language instruction. *Language Learning, 23*,207-22.

Hermans, Judith. (1999). *Klingon and its users: A sociolinguistic profile*. Unpublished MA thesis, Language and Culture Studies, Tilburg University. Retrieved January 16, 2001, from the World Wide Web: http://www.geocities.com/Athens/Atrium/3593/scripdef.htm.

Herring, Susan C. (1992, January). *Gender and participation in computer-mediated linguistic discourse*. Paper presented at the Annual Meeting of the Linguistic Society of America, Philadelphia, USA.

Herring, Susan C. (1994, July). *Gender differences in computer-mediated communication: Bringing familiar baggage to the new frontier*. Keynote paper presented at the American Library Association Annual Convention, Miami, USA. Retrieved October 22, 1996, from the World Wide Web: http://cpsr.org/cpsr/gender/herring.txt.

Herring, Susan C. (Ed.). (1996). *Computer-mediated communication: Linguistic, social, and cross-cultural perspectives.* Amsterdam, The Netherlands: John Benjamins Publishing Company.

Herring, Susan C. (2000). Gender differences in CMC: Findings and implications. *The Computer Professionals for Social Responsibility Newsletter, 18*(1). Retrieved January 14, 2001, from the World Wide Web: http://www.cpsr.org/publications/newsletters/issues /2000/Winter2000/herring.html.

Herring, Susan C., & Lombard, Robin. (1995). *Negotiating gendered faces: Requests and disagreements among computer professionals on the Internet.* Paper presented at Computer-Mediated Discourse Analysis Conference at Georgetown, USA. Retrieved January 13, 2001, from the World Wide Web: http://trill.berkeley.edu/users/sutton /Abstracts.html.

Herrmann, T. (Ed.). (1987). *Show & Tell. Proceedings of the Ontario Universities' Conference.* Guelph, Canada: Ontario University.

Hesser, L. A.; Hogan, R. P.; & Mizell, A. P. (1992, June). *The sum of the parts is greater than the whole in online graduate education.* Paper presented at the ICCAL '92, Nova Scotia, Canada.

Higgins, John J. & Johns, Tim. (1984). *Computers in Language Learning.* London, UK: Collins.

Hiltz, Starr Roxanne. (1984). *Online communities: A case study of the office of the future.* Norwood, NJ: Ablex.

Hiltz, Starr Roxanne. (1986). The virtual classroom: Using computer-mediated communication for university teaching. *The Journal of Communication, 36*, 162-175.

Hiltz, Starr Roxanne. (1995, October). *Impacts of college-level courses via asynchronous learning networks: Focus on students.* Paper presented at Sloan Conference on Asynchronous Learning Networks, Philadelphia, USA.

Hiltz, Starr Roxanne, & Turoff. M. (1978). *The network nation: Human communication via computer.* Reading, MA: Addison-Wesley.

Hiltz, Starr Roxanne; Johnson, K.; & Turoff, M. (1986). Experiments in group decision making: Communication process and outcome in face-to-face versus computerized conferences. *Human Communication Research, 13*(2), 225-252.

Hodge, Robert & Kress, Gunther. (1988). *Social semiotics.* Cambridge, UK: Polity Press.

Hogg, M. A., & Abrams, D. (1988). *Social identifications: A social of intergroup relations and group processes.* London, UK: Routledge.

Holden, M., & Mitchell, W. (1993). The future of computer-mediated communication in higher education. *EDUCOM Review, 28*(2), 31-37.

Holden, M., & Wedman, J. F. (1993). Future issues of computer-mediated communication: The result of a delphi study. *Educational Technology Research and Development, 41*(4), 5-25.

Holmes, Janet. (1984). Hedging your bets and sitting on the fence: Some evidence for hedges as support structures. *Te Reo, 27*, 47-62.

Holmes, Martin. (1996). Marking Student Work on the Computer. *The Internet TESL Journal, 2*(9). Retrieved May 14, 1998, from the World Wide Web: http://www.aitech.ac.jp/~iteslj /Articles/Holmes-ComputerMarking/index.html.

Howarth, P. (1995). *A computer-assisted study of collocations in academic prose, with special reference to grammatical structure and stylistic value.* Unpublished PhD thesis, University of Leeds.

Howell-Richardson, Christina, & Bish, D. (1997). Electronic communications and the teaching of writing in ESOL. In G. Fulcher (Ed.), *Writing in the English language classroom* (pp. 41-60). Gallen, Switzerland: Prentice Hall Europe.

Howell-Richardson, Christina, & Mellar, Harvey. (1996). A methodology for the analysis of patterns of participation within computer mediated communication courses. *Instructional Science, 24*, 47 - 69.

Hudspeth, D., & Maxey, D. (1995, March). *Electronic bulletin board systems (BBSs) and student teacher education.* Paper presented at the 6th International Conference of the Society of Information Technology and Teacher Education, Texas, USA.

Hunter, C. S. (1987). Computer conferencing systems: Educational communications facilitators. In T. Herrmann (Ed.), *Show & Tell. Proceedings of the Ontario Universities' Conference* (p 7-75). Guelph, Canada: Ontario University.

Hymes, Dell H. (1962). The ethnography of speaking. In J. Fishman (Ed.), *Readings in the Sociology of Language* (pp.99-138). The Hague, The Netherlands: Mouton.

Hymes, Dell H. (1964a). Introduction: Toward ethnographies of communication. *American Anthropologist, 66*(2), 2:1-34.

Hymes, Dell H. (Ed.). (1964b). *Language in culture and society: A reader in linguistics and anthropology*. New York, NY: Harper & Row.

Hymes, Dell H. (1972). On communicative competence. In J. B. Pride & Janet Holmes (Eds.), *Sociolinguistics* (pp. 269-293). Baltimore, NY: Penguin Books.

Hymes, Dell H. (1974). *Foundations in sociolinguistics.* Philadelphia, PA: University of Pennsylvania Press.

Jaffe, J. Michael; Lee, Young-Eum; Huang, LiNing; & Oshagan, Hayg. (1995). *Gender, pseudonyms, and CMC: Masking identities and baring souls*. Retrieved February 10, 2001 from the World Wide Web: http://members.kr.inter.net/yesunny/genderps.html.

Janda, Richard D. (1985). Note-taking in English as a simplified register. *Discourse processes, 8*, 437-454.

Jenkins, Jennifer. (1995). *Variation in phonological error in interlanguage talk.* Unpublished PhD thesis, University of London Institute of Education.

Jenkins, Jennifer. (1998). Which pronunciation norms and models for English as an international language? *ELT Journal, 52*(2), 119-126.

Jenkins, Jennifer. (2000). *The phonology of English as an international language.* Oxford, UK: Oxford University Press.

Jespersen, Otto. (1922). *A modern English grammar.* Heidelberg, Germany: Carl Winter's Universitatsbuchhandlung.

Johanson, Terri L. (1996). *The virtual community of an online classroom.* Retrieved January 9, 2001, from the World Wide Web: http://www.lucyonline.com/pub.html.

Johansson, Stig (Ed.). (1982). *Computer corpora in English language research.* Bergen, Norway: Norwegian Computing Centre for the Humanities.

Johansson, Stig; Leech, Geoffrey N.; & Goodluck, Helen. (1978). *Manual of information to accompany the Lancaster-Oslo/Bergen Corpus of British English for use with digital computers*. Department of English, University of Oslo.

Johns, Ann. (1993, April). *Issues in ESP for the 90's.* Paper presented at the Annual Meeting of the Southeast Asian Ministers of Education Organization Regional Language Center Seminar, Singapore.

Johns, Tim. (1990). From printout to handout: Grammar and vocabulary teaching in the context of data-driven learning. *CALL Austria, 10*, 14-34.

Johns, Tim. (2000, July). *Data-driven learning: The perpetual challenge.* Paper presented at the Teaching and Language Corpora Conference, Graz, Austria.

Johnston, V. L. (1992). *Towards a global classroom: Using computer-mediated communications at UAA*. University of Alaska, Anchorage. ERIC Document No. ED356759.

Jones, Steven G. (1995). Understanding community in the information age. In Steven G. Jones (Ed.), *CyberSociety: Computer-mediated communication and community* (pp. 10-35). Thousand Oaks, CA: Sage Publications.

Jones, Steven G. (Ed.). (1995). *CyberSociety: Computer-mediated communication and community*. Thousand Oaks, CA: Sage Publications.

Jones, Barry (Ed.). (1984). *Using authentic resources in teaching French, Teaching Languages, 4*. London, UK: Centre for Information on Language Teaching and Research. (DOE: NA. IE: NA)

Juola, Patrick. (1998, July 29). *Re: Corpora: Corpus linguistics user needs.* Retrieved January 13, 1999, from the World Wide Web: http://www.hd.uib.no/corpora/archive.html.

Kachru, Braj B. (1992). Teaching world Englishes. In Braj B. Kachru (Ed.), *The other tongue*, 2nd edition (pp. 355-365). Chicago, USA: University of Illinois Press.

Kachru, Braj B. (Ed.). (1992). *The other tongue*, 2nd edition. Chicago, USA: University of Illinois Press.

Kaiser, H. F. (1958). The Varimax criterion for analytic rotation in factor analysis. *Psychometrika, 23*, 187-200.

Kapoli, Ireneus Joseph. (1992). *The effects of interaction on the writing of English composition: An exploratory study in secondary schools in Tanzania.* Unpublished PhD Thesis, University of London Institute of Education, London, UK

Kaptelinin, V. (1996). Computer-mediated activity: Functional organs in social and developmental contexts. In Bonnie A. Nardi (Ed.), *Context and consciousness: activity theory and human-computer interaction* (pp. 45-68). Cambridge, MA: MIT Press.

Karlsson, Voutilainen, Heikkila, & Antilla. (1995).

Kay, Paul. (1977). Language evolution and speech style. In Ben G. Blount & Mary Sanches (Eds.), *Sociocultural dimensions of language change* (pp. 21-33). New York, NY: Academic Press.

Kayany, J. M. (1998). Contexts of uninhibited online behavior: Flaming in social newsgroups on Usenet. *Journal of the American Society for Information Science, 49*(12), 1135-1141.

Kaye A. (1989). Computer-mediated communication and distance education. In R. Mason, & A. Kaye (Eds.), *Mindweave* (pp.3-21). Oxford, UK: Pergamon Press.

Kaye, A. (1992). Computer conferencing and mass distance education. In M. D. Waggoner (Ed.), *Empowering networking and computer conferencing in education* (pp. 69- 100). Englewood Cliffs, NJ: Educational Technology Publications.

Keenan, Elinor O., & Bennett, Tina (Eds.). (1977). *Discourse across time and space (SCOPIL no. 5).* Los Angeles, LA: University of Southern California.

Kehoe, Colleen; Pitkow, Jim; Sutton, Kate; Aggarwal, Gaurav; & Rogers, Juan D. (1999). *Graphic, visualization, & usability center's (GVU) 10th WWW user survey.* Atlanta, GA: Georgia Institute of Technology. Retrieved February 11, 2001, from the World Wide Web: http://www.gvu.gatech.edu/user_surveys/survey-1998-10/.

Kettemann, Bernhard. (1996, August). *Concordancing in English language teaching.* Paper presented at the Teaching and Language Corpora Conference, Lancaster, UK.

Ketteman, Bernhard, & Wieden, W. (Eds.). (1993). *Current issues in European second language acquisition research.* Tubingen, Germany: Narr.

Kiesler, Sara. (1984). Computer mediation of conversation. *American Psychologist, 39*, 1123-1134.

Kiesler, Sara; Siegel, Jane; & McGuire, Timothy W. (1984). Social psychological aspects of computer-mediated communication. *American Psychologist, 39*(10), 1123-1134.

Kim, Jongbok, & Sag, Ivan A. (1995). *English and French negation: A lexicalist perspective.* Stanford, CA: Stanford University Press.

Kim, Seongcheol. (1998). *Cultural imperialism on the Internet.* Retrieved October 13, 2000, from the World Wide Web: http://kumo.swcp.com/biz/theedge/kim.htm.

Kitao, S. Kathleen. (1998). Interaction and on-line synchronous communication in English language learning. *CALL-EJ, 3*(1). Retrieved August 12, 1999, from the World Wide Web: http://www.lerc.ritsumei.ac.jp/callej/index.html.

Klein-Andreau, Flora (Ed.). (1983). *Discourse perspectives on syntax.* New York, NY: Academic Press.

Kling, Rob, & Scacchi, W. (1982). The web of computing: Computer technology as social organization. *Advances in Computers, 21*, 1-90.

Kling, Rob. (1980). The social analyses of computing. *ACM Computing Surveys, 12*(1), 61-110.

Kling, Rob. (1994). Reading "all about" computerization: How genre conventions shape non-fiction social analysis. *The Information Society, 10*, 147-172.

Kling, Rob. (1995). *Computerization and controversy: Value conflicts and social choices.* San Diego, CA: Academic Press.

Kling, Rob. (1996). Social relationships in electronic forums: Hangouts, salons, workplaces and communities. *Computer-Mediated Communication Magazine, July, 1996.* Retrieved June 16, 1999, from the World Wide Web: http://www.december.com/cmc/mag/1996/jul/kling.html.

Klotz, Michael. (2000). *Re: Corpora: minimum size of corpus?* Retrieved February 10, 2000, from the World Wide Web: http://www.hd.uib.no/corpora/2000-1/0137.html.

Kochman, Thomas. (1981). *Black and white styles in conflict.* Chicago, IL: University of Chicago Press.

Koiranen, Matti. (1995). *North-European metaphors of "entrepreneurship" and "an entrepreneur."* Babson Park, MA: Arthur M. Blank Center for Entrepreneurship.

Kolko, B.; Nakamura, L.; & Rodman, G. (Eds.). (2000). *Race in Cyberspace.* New York, NY: Routledge.

Kollock, P., & Smith, M. (1994). *Managing the virtual commons: Cooperation and conflict in computer communities.* Retrieved December 12, 1996, from the World Wide Web: http://www.uni-koeln.de/themen/cmc/text/kollo94a.html.

Kress, Gunther (Ed.). (1976). *Halliday: System and function in language.* Oxford, UK: Oxford University Press.

Kress, Gunther. (1994). *Learning to write.* London, UK: Routledge & Kegan Paul Ltd.

Kress, Gunther. (1995). *Writing the future: English and the making of a culture of innovation.* London, UK: National Association for the Teaching of English.

Kroch, Anthony S., & Hindle, Donald M. (1982). *A quantitative study of the syntax of speech and writing.* Final report to the National Institute of Education. New York, NY: National Institute of Education.

Kroll, Barbara. (1977). Ways communicators encode propositions in spoken and written English: A look at subordination and coordination. In Elinor O. Keenan & Tina Bennett (Eds.), *Discourse across time and space (SCOPIL no. 5)* (pp. 69-108). Los Angeles, LA: University of Southern California.

Kruskal, Joseph B. (1978). Factor analysis and principal components: Bilinear methods. In William H. Kruskal & Judith M. Tanur (Eds.), *International encyclopedia of statistics* (pp. 307-330). New York, NY: Macmillan and the Free Press.

Kruskal, Joseph B., & Shepard, Roger N. (1974). A nonmetric variety of linear factor analysis. *Psychometrika, 39*(2), 123-157.

Kruskal, William H., & Tanur, Judith M. (Eds.). (1978). *International encyclopedia of statistics.* New York, NY: Macmillan and the Free Press.

Kuo, Chih-Hua. (1992). Problematic Issues in EST materials development. *English for Specific Purposes, 12*(2), 171-181.

Labov, William. (1972). *Sociolinguistic patterns.* Philadelphia, PA: University of Pennsylvania Press.

Lado, Robert. (1988). *Teaching English across cultures.* New York, NY: McGraw-Hill, Inc.

Lai, K. (1996). Computer-mediated communication: A new learning context. In K. Lai (Ed.), *Words have wings: Teaching and learning with computer networks* (pp. 1-17). Dunedin, New Zealand: University of Otago Press.

Lai, K. (Ed.). (1996). *Words have wings: Teaching and learning with computer networks.* Dunedin, New Zealand: University of Otago Press.

Laina, H. (1992). American teen books easier than British ones. *Journal of Reading, 35*(4), 324-327.

Lancaster University. (1999). *CLAWS part-of-speech tagger.* Retrieved February 18, 1999, from the World Wide Web: http://www.comp.lancs.ac.uk/ucrel/claws/.

Lapakko, D. (1997). Three cheers for language: A closer examination of a widely cited study of nonverbal communication. *Communication Education, 46*, 63-67.

Lawley, E. L. (1992). *Discourse and distortion in computer-mediated communication.* Retrieved November 28, 1996, from the World Wide Web: http://www.itcs.com/elawley /discourse.html.

Lea, Martin (Ed.). (1992). *Contexts of computer-mediated communication.* London, UK: Harvester Wheatsheaf.

Lea, Martin, & Spears, R. (1991). Computer-mediated communication, de-individuation and group decision-making. *International Journal of Man-Machine Studies, 34*, 283-301.

Lea, Martin, & Spears, R. (1992). Paralanguage and social perception in computer-mediated communication. *Journal of Organizational Computing, 2*(3), 321-341.

Lea, Martin; O'Shea, T.; Fung, P.; & Spears, R. (1992). *Flaming* in computer-mediated communication: Observations, explanations, implications. In Martin Lea (Ed.), *Contexts of computer mediated communication* (pp. 89-112). New York, NY: Harvester Wheatsheaf.

Lee, J. Y. (1996). Charting the codes of cyberspace: A rhetoric of electronic mail. In L. Strate, R. Jacobson, & S. B. Gibson (Eds.), *Communication and cyberspace: Social interaction in an electronic environment* (pp. 275-296). Cresskill, NJ: Hampton Press.

Leech, Geoffrey N. (1966). *English in advertising: A linguistic study of advertising in Great Britain.* London, UK: Longman.

Leitner, Gerhard. (Ed.). (1986). *The English reference grammar: Language and linguistics, writers and readers.* Tubingen, Germany: Niemeyer.

Lever, Judy C. (1992). Meeting increasing demand using distance education. *Leadership Abstracts, 5,* 2.

Levin, J. A.; Kim, H.; & Riel, M. M. (1990). Analyzing instructional interactions on electronic message networks. In L. M. Harasim (Ed.), *Online education: Perspectives on a new environment* (pp. 187-213). New York, NY: Praeger.

Levinson, Paul. (1990). Computer conferencing in the context of the evolution of media. In Linda M. Harasim (Ed.), *Online Education: Perspectives on a New Environment* (pp. 3-10). New York, NY: Praeger.

Lewandowska-Tomaszczyk, Barbara. (1998, August). *Grammar and meaning contrasts in language corpora.* Paper presented at the Teaching and Language Corpora Conference, Oxford, UK.

Lewandowska-Tomaszczyk, Barbara, & Melia, Patrick James (Eds.). (2000). *PALC'99: Practical applications in language corpora, papers from the international conference at the University of Lodz, 15-18 April 1999.* Frankfurt am Main, Germany: Peter Lang.

Lin, C. S. (1992). *Statistics for psychology and education.* [In Chinese]. Taipei, Taiwan: Tong-Hwa Books.

Little, G. D., & Sanders, S. L. (1989). Classroom community: A prerequisite for communication. *Foreign Language Annals, 22*(3), 277-281.

Littlefair, A. B. (1991). *Reading all types of writing: The importance of genre and register for reading development.* Milton Keynes, UK: Open University Press.

Liou, Hsien-Chin. (1997). The impact of WWW texts on EFL learning. *Computer Assisted Language Learning, 10*(5), 455-478.

Liu, C.; Peek, J.; Jones, R.; Buus, B.; & Nye, A. (1994). *Managing Internet information services.* Sebastopol, CA: O'Reilly & Associates.

Ljung, Magnus. (1991). *Swedish TEFL meets reality.* Eric Document ED343133.

Long, N. R., & Povey, T. A. (1982). Audio-tapes in distance teaching: A New Zealand experience. *Teaching at a Distance, 21,* 60-65.

Lopes, Gabriel Pereira. (2000). *Re: Corpora: minimum size of corpus?* Retrieved February 14, 2000, from the World Wide Web: http://www.hd.uib.no/corpora/2000-1/0154.html.

Luk, Robert. (2000). *Re: Corpora: minimum size of corpus?* Retrieved February 10, 2000, from the World Wide Web: http://www.hd.uib.no/corpora/2000-1/0145.html.

Lundstrom, Phoenix. (1995). *Synchronous computer-mediated communication: Will internet talkers improve the communicative competence of ESL/EFL students?* Unpublished Master's Thesis. The University of Hawaii, Manoa, USA. Retrieved October 20, 1998, from the World Wide Web: http://elicos.qut.edu.au/truna/lan618/phoenix.txt.

Lyons, John. (1977). *Semantics.* Cambridge, UK: Cambridge University Press.

Mabrito, M. (1989). *Writing apprehension and computer-mediated peer-response groups: A case study of four high- and four low-apprehensive writers communicating face-to-face versus electronic mail.* Unpublished doctoral dissertation, Illinois State University, Normal, IL.

MacWhinney, Brian. (1995). *The CHILDES project: Computational tools for analysing talk.* Hillsdale, N J: Lawrence Erlbaum.

MacWhinney, Brian. (1996a). *CD-ROM: The CHILDES database.* Pittsburgh, PA: Carnegie Mellon University.

MacWhinney, Brian. (1996b). *Readme manual for CD-ROM: The CHILDES database.* Pittsburgh, PA: Carnegie Mellon University.

Maingay, Susan M. (1980). Selection and grading of authentic material for the reading class. *ELT Journal, 34*(3), 217-221.

Malinowski, Bronislaw. (1923). The problem of meaning in primitive languages.

Supplementary essay in C. K. Ogden & I. A. Richards, *The meaning of meaning* (pp. 296-336). New York, NY: Harcourt, Brace.

Mandala, Susan. (1998, August). *Searching corpora for speech acts.* Paper presented at the Teaching and Language Corpora Conference, Oxford, UK.

Mantovani, Giuseppe. (1996). Computer-mediated communication. In Giuseppe Mantovani (Ed.), *New communication environments from everyday to virtual* (pp. 98-106). London, UK: Taylor and Francis Ltd.

Mantovani, Giuseppe (Ed.). (1996). *New communication environments from everyday to virtual.* London, UK: Taylor and Francis Ltd.

Marckworth Mary L., & Baker, William J. (1974). A discriminant function analysis of co-variation of a number of syntactic devices in five prose genres. *American Journal of Computational Linguistics, Microfiche 11.*

Mason, Oliver. (1996). *TAGGER: Experimental email tagging on-line.* <mail to tagger@clg.bham.ac.uk>. Birmingham, UK: University of Birmingham.

Mason, Oliver. (1998, July 29). *Re: Corpora: Corpus linguistics user needs.* Retrieved July 31, 1998, from the World Wide Web: http://www.hd.uib.no/corpora/archive.html.

Mason, Robin, & Kaye, Anthony (Eds.). (1989). *Mindweave.* Oxford, UK: Pergamon Press.

Matheson, K., & Zanna, M. P. (1990). Computer-mediated communications: The focus is on me. *Social Science Computer Review, 8*(1), 1-12.

Matrix Information & Directory Services (MIDS). (1995). *Third MIDS Internet demographic survey.* Retrieved December 12, 1999, from the World Wide Web: http://www.mids.org /ids3/ids3sum.603.

Maynor, Natalie. (1994). The language of electronic mail: Written speech? In G. D. Little & M. Montgomery (Eds.), *Centennial usage studies.* Tuscaloosa, AL: Alabama UP.

McCarthy, Michael, & Carter, Ronald. (1994). *Language as discourse: Perspectives for language teaching.* London, UK: Longman.

McCawley, James D. (1979). Presupposition & discourse structure. In Choon-Kyu Oh & David A. Dinneen (Eds.), *Syntax & Semantics, Vol. 2: Presupposition* (pp. 371-388). New York, NY: Academic Press.

McCormick, N. B., & McCormick, J. W. (1992). Computer friends and foes: Content of undergraduates' electronic mail. *Computers in Human Behavior, 8*, 379-405.

McEnery, Tony, & Wilson, Andrew. (1996). *Corpus linguistics.* Edinburgh, UK: Edinburgh University Press.

McKnight, Alex. (1995, May). *Pragmatics and TESOL: Using authentic language data.* Paper presented at the Biennial Meeting of the MELTA International Conference on innovations in Approaches to the Teaching and Learning of English, Kuala Lumpur, Malaysia.

McLuhan, Marshall, & Nevitt, Barrington. (1974). Medium meaning message. *Communication, 1*(1), 27-33.

Medley, Frank W. (1984). *Motivating the language learner through real communication, ACTFL Master Lecture Series.* Monterey, CA: Defense Language Institute, Foreign Language Center.

Mehan, H. (1978). Structuring school structure. *Harvard Educational Review, 48*, 32-64.

Mehrabian, A. (1981). Silent messages: Implicit communication of emotions and attitudes. Belmont, CA: Wadsworth.

Mejrkova, S., & Frantidek, P. (Eds.). (1994). *The syntax of sentence and text: A festschrift for Frantidek Daned.* Philadelphia, PA: John Benjamins.

Meloni, Christine. (1998). The Internet in the classroom: A valuable tool and resource for ESL/EFL teachers. *The ESL Magazine, January.* Retrieved October 2, 2000, from the World Wide Web: http://www.eslmag.com/Article.htm.

Melvin, Bernice S., & Stout, D. F. (1987). Motivating language learners through authentic materials. In Wilga M. Rivers (Ed.), *Interactive language teaching* (pp. 44-56). New York, NY: Cambridge University Press.

Metz, Jason Michel. (1994). Computer-mediated communication: Literature review of a new context. *Interpersonal Computing and Technology, 2* (2), 31-49. Retrieved January 7, 2001, from the World Wide Web: http://www.helsinki.fi/science/optek/1994/n2/metz.txt.

Meunier, Lydie E. (1997). Affective factors and cyberteaching: Implications for a postmodern pedagogy. *CALICO Monograph Series, 4*, 122-132.

Micro-OCP. (1988). *The Oxford concordance Program.* Oxford, UK: Oxford University.

Mindt, Dieter. (1986). Corpus, grammar, and teaching English as a foreign Language. In Gerhard Leitner (Ed.), *The English reference grammar: Language and linguistics, writers and readers* (pp. 125-139). Tubingen, Germany: Niemeyer.

Mizrach, Steve. (2000). *The electronic discourse of the computer underground.* Retrieved January 6, 2001, from the World Wide Web: http://www.fiu.edu/~mizrachs/cudisc.html.

Moffat, Dave. (1998, July 31). *Re: Corpora: Corpus linguistics user needs.* Retrieved August 5, 1998, from the World Wide Web: http://www.hd.uib.no/corpora/archive.html.

Moody, Suzan. (2000, March). *After the PCI: Beginner's guide to navigating the Internet.* Workshop presented at the TESOL 2000 Conference, Vancouver, Canada.

Morgan, Deborah. (1999). *ESOL case study in UK WELL Project.* Retrieved July 23, 2000, from the World Wide Web: http://www.well.ac.uk/casestud/morgan.htm.

Morgan, Janice C. (1997). *Community ties and a community network: Cupertino's computer-mediated Citynet.* Unpublished Master's thesis, San Jose State University, USA.

Morton, Robert. (1999). Abstracts as authentic material for EAP classes. *ELT Journal, 53*(3), 177-182.

Murray, Denise E. (1988). The context of oral and written language: A framework for mode and medium switching. *Language in Society, 17*, 351-373.

Murray, Denise E. (1995). *Knowledge machines: Language and information in a technological society.* London, UK: Longman.

Murray, P. J. (1996). Nurses' computer-mediated communications on NURSENET - A case study. *Computers in Nursing, 14*(4), 227-234.

Myers, D. (1987). Anonymity is part of the magic: Individual manipulation of computer-mediated communication contexts. *Qualitative Sociology, 10*(3), 251-166.

Nardi, Bonnie A. (Ed.). (1996). *Context and consciousness: activity theory and human-computer interaction.* Cambridge, MA: MIT Press.

NETEACH-L. (1996). *Daily logfiles.* <mailto listserv@thecity.sfsu.edu>.

Netoric-MOO. (1996). *Logs of MOO sessions.* Retrieved December 5, 1996, from the World Wide Web: http://bsuvc.bsu.edu/~00gjsiering/netoric/request/.

Novak, Stanley M. (2000). *American shibboleth: Ebonics.* Retrieved January 16, 2001, from the World Wide Web: http://www.csa.com/hottopics/ebonics/oview.html

Nunamaker, J. F.; Dennis, A. R.; Valacich, J. S.; Vogel, D. R.; & George, J. F. (1993). Electronic meeting systems to support group work. In Ronald M. Baecker (Ed.), *Readings in groupware and computer-supported cooperative work: Assisting human-human collaboration* (pp. 718-739). San Mateo, CA: Morgan Kaufmann.

Ochs, Elinor. (1979). Planned and unplanned discourse. In Tamly Givon (Ed.), *Discourse and syntax* (pp. 51-80). New York, NY: Academic Press.

O'Donnell, Roy C. (1974). Syntactic differences between speech and writing. *American Speech, 49,* 102-110.

O'Donnell, Roy C.; Griffin, W. J.; & Norris, R. C. (1967). A transformational analysis of oral and written grammatical structures in the language of children in grades three, five, and seven. *Journal of Educational Research, 61*, 36-39.

Oh, Choon-Kyu, & Dinneen, David A. (Eds.). (1979). *Syntax & Semantics, Vol. 2: Presupposition.* New York, NY: Academic Press.

Oksenholt, Svein. (1977). Put a short-wave radio in your foreign language classroom. *Audiovisual Instruction, 22*(5), 19-22.

Olson, David R. (1977). From utterance to text: The bias of language in speech and writing. *Harvard Educational Review 47*, 257-281.

Olson, David R. (1996). *The world on paper: The conceptual and cognitive implications of reading and writing.* Cambridge: Cambridge University Press.

Olson, David R.; Torrance, N.; & Hildyard, A. (Eds.). (1985). *Literature, language, and learning: the nature and consequences of reading and writing.* Cambridge, UK: Cambridge University Press.

Olson, Margrethe H. (Ed.). (1989). *Technological support for work group collaboration.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Ong, Walter J. (1982). *Orality and literacy: The technologizing of the word.* New York, NY: Methuen.

Ord, J. (1989). Who's joking? The information system at play. *Interacting with Computers, 1,* 118-128.

Osgood, Charles E. (1960). Some effects of motivation on style of encoding. In Thomas A. Sebeok (Ed.), *Style in language* (pp. 293-306). Cambridge, MA: MIT Press.

Page, Glenda; Elkins, John; & O'Connor, Barrie (Eds.). (1979). *Communication through reading: Proceedings of the 4th Australian Reading Conference, Vol. 2.* Adelaide, Australia: Australian Reading Association.

Patterson, Holly. (1996). *Computer-mediated groups: A study of a culture in Usenet.* Unpublished doctoral dissertation, Texas A&M University, Texas, USA. Retrieved January 10, 2001 from the World Wide Web: http://www.sci.tamucc.edu/~hollyp/pubs/dis/dissert.html.

Perren, G. E., & Trim, J. L. M. (Eds.). (1971). *Applications of linguistics.* Cambridge, UK: Cambridge University Press.

Peretz, Arna S. (1988, April). *An interactional, learner-centered reading course.* Paper presented at the 22nd Annual Meeting of the International Association of Teachers of English as a Foreign Language, Edinburgh, Scotland.

Pica, T.; Lincoln-Porter, F.; Paninos, D.; & Linnel, J. (1996). Language learners' interaction: How does it address the input, output, and feedback needs of L2 learners? *TESOL Quarterly, 30* (1), 59-84.

Pickering, Andrew (Ed.). (1992). *Science as practice and culture.* Chicago, IL: University of Chicago Press.

Pincas, Anita. (1999). Reference in online discourse. *TESL-EJ, 4*(1), 1-24. Retrieved August 20, 1999, from the World Wide Web: http://www-writing.berkeley.edu/TESL-EJ/ej13/a1.html.

Pincas, Anita. (2000). Features of online discourse for education. *Learning Technology, 2*(1). Retrieved December 19, 2000, from the World Wide Web: http://lttf.ieee.org/learn_tech/issues/january2000/#pincas.

Poon, Simpson. (1993). Internet relay chat (IRC): A real-time multi-user computer collaborative learning medium. *Proceedings of the Australian and South Pacific External Studies Association Conference* (pp. 63-72), Adelaide, Australia.

Porter, Don, & Roberts, Jon. (1981). Authentic listening activities. *ELT Journal, 36*(1), 37-47.

Postmes, T.; Spears, R.; & Lea, Martin. (1998). Breaching or building social boundaries? SIDE-effects of computer-mediated communication. *Communication Research, 25*(6), 689-715.

Pride, J. B., & Holmes, Janet (Eds.). (1972). *Sociolinguistics.* Baltimore, NY: Penguin Books.

Purgason, Katherine Barnhouse. (1991). *Planning lessons and units.* In Marianne Celce-Murcia (Ed.), *Teaching English as a Second or Foreign Language* (pp. 419-431). New York, NY: Newbury House

Quirk, Randolph; Greenbaum, Sidney; Leech, Geoffrey; & Svartvik, Jan. (1972). *A grammar of contemporary English.* London, UK: Longman.

Quirk, Randolph; Greenbaum, Sidney; Leech, Geoffrey; & Svartvik, Jan. (1985). *A comprehensive grammar of the English language.* London, UK: Longman.

R. R. A. Associates (Ed.). (1984). *The new media: Communication yearbook, Vol. 6.* Beverly Hills, CA: Sage Publications.

Reetz, Henning. (1998, July 29). *Re: Corpora: Corpus linguistics user needs.* Retrieved August 5, 1998, from the World Wide Web: http://www.hd.uib.no/corpora/archive.html.

Reid, Elizabeth. (1991). *Electropolis: Communication and community on Internet Relay Chat.* Unpublished honors thesis, Department of History, University of Melbourne, Melbourne, Australia.

Reid, Elizabeth. (1994). *Cultural formations in text-based virtual realities.* Unpublished Master's thesis, University of Melbourne, Australia.

Rheingold, Howard. (1994). *The virtual community: Homesteading on the electronic frontier.* New York, NY: Harper Perennial.

Rice, R. E. (1984). Mediated group communication. In R. R. A. Associates (Ed.), *The new media: Communication yearbook, Vol. 6* (pp. 129-154). Beverly Hills, CA: Sage Publications.p communication. In R. R. A. Associates (Ed.), *The new media: Communication yearbook, Vol. 6* (pp. 129-154). Beverly Hills, CA: Sage Publications.

Rice, R. E., & Love, G. (1987). Electronic emotion: Socioemotional content in a computer-mediated communication network. *Communication Research, 14*(1), 85-108.

Richards, J. C., & Schmidt, R. (Eds.). (1983). *Language and communication.* London, UK: Longman.

Ringbom, Hakan, & Rissanen, Matti (Eds.). (1983). *Proceedings from the Nordic Conference for English Studies, 2nd.* Hanaholmen, Finland: Hanasaari.

Rivers, Wilga M. (1987). Interaction as the key to teaching language for communication. In Wilga M. Rivers (Ed.), *Interactive language teaching* (pp. 3-16). New York, NY: Cambridge University Press.

Rivers, Wilga M. (Ed.). (1987). *Interactive language teaching.* New York, NY: Cambridge University Press.

Rivers, Wilga M., & Temperley, Mary S. (1978). *A practical guide to the teaching of English as a second or foreign language.* New York, NY: Oxford University Press.

Roget, Peter M. (1982). *Roget's Thesaurus of English words & phrases.* London, UK: Longmans.

Romaine, S. (1988). *Pidgin and creole languages.* New York, NY: Longman.

Rome, Ute. (2000). *Re: Corpora: minimum size of corpus?* Retrieved February 09, 2000, from the World Wide Web: http://www.hd.uib.no/corpora/2000-1/0134.html.

Romiszowski, A. J., & deHaas, J. A. (1989). Computer mediated communication for instruction: Using e-mail as a seminar. *Educational Technology, 29*(10), 7-14.

Rosen, Harold, & Burgess, Tony. (1980). *Languages and dialects of London school children.* London, UK: Ward Lock Educational.

Rosenbaum, H., & Snyder, H. (1991). An investigation of emerging norms in computer mediated communication: An empirical study of computer conferencing. *Proceedings of the ASIS Annual Meeting, 28*, 15-23.

Royar, R. (1994). New horizons, clouded vistas. *Computers and Composition, 11*, 93-105.

Ruben, B. E. (Ed.). (1985). *Information and behavior, Vol. 1.* New Brunswick, NJ: Transaction.

Rubin, Andee. (1980). A theoretical taxonomy of the differences between oral and written language. In Rand J. Spiro, Bertram C. Bruce, & William F. Brewer (Eds.), *Theoretical issues in reading comprehension: Perspectives from cognitive psychology, linguistics, artificial intelligence, and education* (pp. 411-438). Hillsdale, NJ: Lawrence Erlbaum Associates.

Sacks, Harvey. (1972). On the analyzability of stories by children. In John J. Gumperz, & Dell H. Hymes (Eds.), *Directions in sociolinguistics: The ethnography of communication* (pp. 325-345). New York, NY: Rinehart & Winston.

Sadock, Jerrold M. (1978). On testing for conversational implicature. In Peter Cole (Ed.), *Syntax and Semantics, Vol. 9: Pragmatics* (pp. 281-297). New York, NY: Academic Press.

Sampson, Geoffrey. (1998, July 29). *Re: Corpora: Corpus linguistics user needs.* Retrieved August 4, 1998, from the World Wide Web: http://www.hd.uib.no/corpora/archive.html.

Sankoff, David (Ed.). (1986). *Diversity and diachrony.* Amsterdam, The Netherlands: John Benjamins.

Sapir, Edward. (1915). Southern Paiute and Nahuatl, a study in Uto-Aztekan, part II. *American Anthropologist, 17*, 98-120, 306-28.

Sapir, Edward. (1929). The status of linguistics as a science. *Language, 5*, 207-214.

Saunders, Carol S.; Robey, Daniel; & Vaverek, Kelly A. (1994). The persistence of status differentials in computer cnferencing. *Human Communication Research, 20*(4), 443-472.

Savicki, Victor; Lingenfelter, Dawn; & Kelley, Merle. (1996). Gender language style and group composition in Internet discussion groups. *Journal of Computer-Mediated Communication, 5*(1). Retrieved November 30, 1998, from the World Wide Web: http://www.ascusc.org /jcmc/vol2/issue3/

Scanlan, Timothy M. (1980). The use of Radio Canada and Radio France recordings in the teaching of language and culture. *French Review, 53*(3), 402-410.

Scardamalia, M., & Bereiter, C. (1993). Higher levels of agency for children in knowledge building: A challenge for the design of new knowledge media. *Journal of the Learning Sciences, 1*(1), 37-68.

Schaeffner, Christina. (2000). *German and English case study in UK WELL Project.* Retrieved July 23, 2000, from the World Wide Web: http://www.well.ac.uk/casestud/schaeffner2.htm.

Scherer, Klaus R., & Giles, Howard (Eds.). (1979). *Social markers in speech.* Cambridge, UK: Cambridge University Press.

Schiffrin, Deborah. (1981). Tense variation in narrative. *Language, 57*, 45-62.

Schiffrin, Deborah (Ed.). (1984). *Meaning, form, and use in context: Linguistic applications.* Washington, DC: Georgetown University Press.

Schmidhuber, Juergen, & Prelinger, Daniel. (1993). Discovering predictable classifications. *Neural Computation, 5*(4), 625-635.

Schneider, S. M. (1996). Creating a democratic public sphere through political discussion - A case study of abortion conversation on the Internet. *Social Science Computer Review, 14*(4), 373-393.

Scott, C. R., & Easton, A. C. (1996). Examining equality of influence in group decision support system interaction. *Small Group Research, 27*(3), 360-382.

Scott, Mike. (1996). *WordSmith tools.* Oxford, UK: Oxford University Press.

Scott, Mike. (1998, July 29). *Re: Corpora: Corpus linguistics user needs.* Retrieved August 5, 1998, from the World Wide Web: http://www.hd.uib.no/corpora/archive.html.

Scott, Mike. (2000, July). *Picturing the key words of a very large corpus and their lexical upshots or getting at the Guardian's view of the world.* Paper presented at the Teaching and Language Corpora Conference, Graz, Austria.

Sebeok, Tomas A. (Ed.). (1960). *Style in language.* Cambridge, UK: Cambridge University Press.

Selfe, C. L., & Meyer, P. R. (1991). Testing claims for online conferences. *Written Communication, 8*(2), 163-192.

Sherblom, John. (1988). Direction, function, and signature in electronic mail. *The Journal of Business Communications, 25*(4), 39-54.

Short, Deborah J. (Ed.). (1999). *New ways in teaching English at the secondary level, New ways in TESOL series II: Innovative classroom techniques.* Waldorf, MD: TESOL Publications.

Short, John, Williams, Ederyn, & Christie, Bruce. (1976). *The social psychology of telecommunications.* London, UK: John Wiley & Sons.

Siegel, J.; Dubrovsky, V.; Kiesler, S.; & McGuire, T. (1986). Group processes in computer-mediated communication. *Ofganizational Behaviour and Human Decision, 37*, 157-187.

Smeltzer, D. K. (1992). Computer-mediated communication: An analysis of the relationship of message structure and message intent. *Educational Technology, June* 1992.

Smutek, Malinda M. (1986). *English as a second language: In search of an acquisition rich environment.* Boston, MA: University of Massachusetts.

Snow, C. E., & Ferguson, Charles A. (Eds.). (1977). *Talking to children: Language input and acquisition.* Cambridge, UK: Cambridge University Press.

Sondheimer, N. K. (Ed.). (1980). *The 18th annual meeting of the association for computational linguistics and parasession on topics in interactive discourse: Proceedings of the conference.* Philadelphia, PA: University of Pennsylvania.

Spears, R., & Lea, Martin. (1992). Social influence and the influence of the "social" in computer-mediated communication. In Martin Lea (Ed.), *Contexts of computer-mediated communication* (pp. 30-65). London, UK: Harvester Wheatsheaf.

Spiro, Rand. J,; Bruce, Bertram C.; & Brewer, William F. (Eds.). (1980). *Theoretical issues in reading comprehension: Perspectives from cognitive psychology, linguistics, artificial intelligence, and education.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Spitzer, M. (1986). Writing style in computer conferences. *IEEE Transactions on Professional Communications, 29*(1), 19-22.

Sproull, Lee, & Kiesler, Sara. (1986). Reducing context cues: Electronic mail in organizational communication. *Management Science, 32*, 1492-1512.

Sproull, Lee, & Kiesler, Sara. (1991). *Connections: New ways of working in the networked organization.* Cambridge, MA: MIT Press.

Staw, B. M., & Cummings, L. L. (Eds.). (1984). *Research in organisational behavior.* Greenwich, CT: JAI.

Stoll, Clifford. (1996). *Silicon snake oil: Second thoughts on the information highway.* New York, NY: Anchor Books.

Strate, L.; Jacobson, R.; & Gibson, S. B. (Eds.). (1996). *Communication and cyberspace: Social interactin in an electronic environment.* Cresskill, NJ: Hampton Press.

Straumann, H. (1935). *Newspaper headlines: A study in linguistic method.* London, UK: Allen & Unwin.

Strasser, Gerhard F. (1995). Fax-technology for essay exchanges with German students as an enhancement of "c"-culture. *Die Unterrichtspraxis: Teaching German, 28*(2), 159-164.

Straus, S. G., & McGrath, J. E. (1994). Does the medium matter? The interaction of task type and technology on group performance and member reactions. *Journal of Applied Psychology, 79*(1), 87-97.

Streiff, Virginia. (1983, March). *Correlations and distinctions among ESL skills of university students.* Paper presented at the Fifth Annual Language Testing Colloquium, Ottawa, Canada.

Stubbs, Michael. (1980). *Language and literacy: The sociolinguistics of reading and writing.* London, UK: Routledge and Kegan Paul.

Stubbs, Michael. (1986). Lexical density: A technique and some findings. In M. Coulthard (Ed.), *Talking about text* (pp. 27-42). Birmingham, UK: Birmingham Instant Print Ltd.

Svartvik, Jan, & Quirk, Randolph (Eds.). (1980). *A corpus of English conversation.* Lund, Sweden: CWK Gleerup.

Swan, Michael. (1985). A critical look at the communicative approach (2). *ELT Journal, 39*(2), 76-87.

Sykes, J. B. (Ed). (1987). *Concise Oxford dictionary of current English.* Oxford, UK: Oxford University Press

TACT. (1990). *TACT text analysis computing tools.* Charlottesville, VA: University of Virginia.

Tajfel, H., & Turner, J. C. (1986). The social identity theory of intergroup behaviour. In S. Worchel & W. G. Austin (Eds.), *The psychology of intergroup relations* (pp. 7-24). Chicago, IL: Nelson-Hall.

Tannen, Deborah (Ed.). (1982). *Spoken and written language: Exploring orality and literacy.* Norwood, NJ: Ablex.

Tannen, Deborah (Ed.). (1984). *Coherence in spoken and written discourse.* Norwood, NJ: Ablex.

Tannen, Deborah. (1985). Relative focus on involvement in oral and written discourse. In David R. Olson, N. Torrance, & A. Hildyard (Eds.), *Literacy, language, and learning: The nature and consequences of reading and writing* (pp. 124-147). Cambridge, UK: Cambridge University Press.

Tannen, Deborah. (1989). *Talking voices: Repetition, dialogue, and imagery in conversational discourse.* New York, NY: Cambridge University Press.

Tella, Seppo. (1999). Media education for our times. *Kasvatus,* 30(3), 20-221.

TESL-L. (1996). *Weekly logfiles.* Mailto <listserv@cunyvm.cuny.edu> with message "get TESL-L <logyymmw> f=mail."

Thomas, B. J. (1979). The use of authentic material in English language teaching. *Zielsprache Englisch, 8*(1), 8-10.

Thompson, William Irwin. (1992). The American replacement of nature: The everyday acts and outrageous evolution of economic life. New York, NY: Doubleday.

Thompson, Sandra A. (1983). Grammar and discourse: The English detached participial clause. In Flora Klein-Andreau (Ed.), *Discourse perspectives on syntax* (pp. 43-65). New York, NY: Academic Press.

Thompson, Sandra A. (1984). "Subordination" in formal and informal discourse. In Deborah Schiffrin (Ed.), *Meaning, form, and use in context: Linguistic applications* (pp. 85-94). Washington, DC: Georgetown University Press.

Thompson, Sandra A. (1985). Grammar and written discourse: Initial vs. final purpose clauses in English. *Text 5*, 55-84.

Tillyer, Anthea. (1997, April). *Does sex count? – Gender differences in Internet discourse.* Paper presented at the Southeastern Conference on Linguistics, Charlotte, USA.

Tillyer, Anthea. (1998). *Mailing lists for teachers of English as a second language – A BBC interview.* Retrieved September 11, 1999, from the World Wide Web: http: //www.bbc.co.uk/worldservice/learningenglish/network/interv10.shtml

Tonnies, F. (1957). *Community and society.* East Lansing, MI: Michigan State University Press

Tottie, Gunnel. (1981). Negation and discourse strategy in spoken and written English. In H. Cedergren & D. Sankoff (Eds.), *Variation omnibus* (pp. 271-184). Edmonton, Canada: Linguistic Research.

Tottie, Gunnel. (1982). Where do negative sentences come from? *Studia Linguistica, 36*, 88-105.

Tottie, Gunnel. (1983). *Much about 'not' and 'nothing': A study of the variation between analytic and synthetic negation in contemporary Aerican English.* Lund, Sweden: CWK Gleerup.

Tottie, Gunnel. (1986). The importance of being adverbial: Focusing and contingency adverbials in spoken and written English. In Gunnel Tottie & Ingegerd Backlund (Eds.), *English in speech and writing: A symposium* (pp. 93-118). Stockholm, Sweden: Almqvist and Wiksell.

Tottie, Gunnel, & Backlund, Ingegerd (Eds.). (1986). *English in speech and writing: A symposium.* Stockholm, Sweden: Almqvist and Wiksell.

Tottie, Gunnel; Altenberg, Bengt; & Hermeren, Lars. (1983). *English in speech and writing* (ETOS Report 1). Lund, Sweden: Engelska Institutionen.

Traweek, S. (1992). Border crossings: Narrative strategies in science studies and among physicists in Tsukuba Science City, Japan. In Andrew Pickering (Ed.), *Science as practice and culture* (pp. 429-465). Chicago, IL: University of Chicago Press.

Trevino, L. K.; Lengel, R. H.; & Daft, R. L. (1987). Media symbolism, media richness, and media choice in organizations: A symbolic interactionist perspective. *Communication Research, 14*(5), 553-74.

Tribble, Chris. (1991). Concordancing and an EAP writing program. *CAELL Journal, 1*(2), 10-15.

Tribble, Chris. (1997, April). *Improvising corpora for ELT: Quick-and-dirty ways of developing corpora for language teaching.* Paper presented at the Practical Applications in Language Corpora, Lodz, Poland.

Tribble, Chris. (1998, July). *Genres, keywords, teaching: Towards a pedagogic account of the language of project proposals.* Paper presented at the Teaching and Language Corpora Conference, Oxford, UK.

Tsai, Su-hsun. (1997a, March). *A comparative analysis of computer-mediated communication (CMC) versus non-CMC text along the dimension of abstracts vs. non-abstract information.* Paper presented at the Symposium on Text Analysis Conference, Taipei, Taiwan.

Tsai, Su-hsun. (1997b, April). *Abstractness and non-abstractness: The differences between computer-mediated communication (CMC) and non-CMC texts.* Paper presented at the Register and Corpus Dynamics Conference in Aston University, Birmingham, UK.

Tsai, Su-hsun. (1997c, November). *The styles of English computer-mediated communication (CMC) texts in comparison with those in non-CMC genres.* Paper presented at the Major Varieties of English (MAVEN), Vaxjo, Sweden.

Tsai, Su-hsun. (2000). Syntactic variations across texts of computer-mediated communication (CMC) and non-CMC media. In Barbara Lewandowska-Tomaszczyk & Patrick James Melia (Eds.), *PALC'99: Practical applications in language corpora, papers from the international conference at the University of Lodz, 15-18 April 1999* (pp. 129-158). Frankfurt am Main, Germany: Peter Lang.

Turbee, Lonnie. (1995). What can we do in a MOO?: Suggestions for language teachers. In Mark Warschauer, (Ed.), *Virtual Connections* (pp. 235-238). Manoa, HI: University of Hawaii at Manoa.

Turbee, Lonnie. (1996). *MOO and IRC: What's the big difference to the language learner?* Retrieved July 16, 1998, from the World Wide Web: http://web.syr.edu/~lmturbee /mooyirc.html.

Turbee, Lonnie. (1997). *Educational MOO: Text-based virtual reality for learning in community.* ERIC Digest, ERIC Clearinghouse on Information & Technology, Syracuse, NY.

Turbee, Lonnie. (1998). MOO, WOO and more: Language learning in virtual environments. In Joy Egbert, & Elizabeth. Hanson-Smith (Eds.), *CALL environments: Research, practice, and critical issues* (pp. 346-361). Alexandria, VA: TESOL Publications.

Turner G. J., & Pickvance, R. E. (1971). Social class differences in the expression of uncertainty in five-year-old children. *Language and Speech, 14,* 303-325.

Uhlioova, L. (1994). E-mail as a new subvariety of medium and its effects upon the message. In S. Mejrkova & P. Frantidek (Eds.), *The syntax of sentence and text: A festschrift for Frantidek Daned* (pp. 273-282). Philadelphia, PA: John Benjamins.

Underwood, John. (1984). *Linguistics, computers, and the language teacher.* Rowley, MA: Newbury House.

University of Essex. (1999). *Corpus linguistics – Software.* Retrieved August 25, 1999, from the World Wide Web: http://clwww.essex.ac.uk/w3c/corpus_ling/content/software.html.

University of Leeds. (1997). *Automatic mapping among lexico-grammatical annotation models (AMALGAM).* Retrieved June 23, 1998, from the World Wide Web: http: //www.scs.leeds.ac.uk /amalgam/amalgam/amalgsoft.html.

Ure, Jean. (1971). Lexical density and register differentiation. In G. E. Perren & J. L. M. Trim (Eds.), *Applications of linguistics* (pp. 443-452). Cambridge, UK: Cambridge University Press.

Vachek, Josef. (1979). Some remarks on the stylistics of written language. In D. J. Allerton; Edward Carney; & David Holdcroft (Eds.), *Function and context in linguistic analysis* (pp. 206-215). Cambridge, UK: Cambridge University Press.

van Dijk, Teun Adrianus. (1977). *Text and context: Explorations in the semantics and pragmatics of discourse.* London, UK: Longman.

van Hout-Wolters, B., & Schnotz, W. (Eds.). (1992). *Text comprehension and learning from text.* Amsterdam, The Netherlands: Swets & Zeitlinger.

Vande-Kopple, W. J. (1995, March). *Some explorations of the synoptic and dynamic styles.* Paper presented at the 46th annual meeting of the Conference on College Composition and Communication, Washington, DC.

Vilmi, Ruth M. (1994, October). *Global communication by email: An ongoing experiment.* Paper presented at TESOL France, Paris, France.

Vilmi, Ruth. M. (1996). An overview of HUT e-mail writing project. *IATEFL Newsletter No. 5.* UK: IATEFL English for Specific Purposes Special Interest Group.

Vygotsky, Lev Semenovich. (1962). *Thought and language* (E. Hanfmann & G. Vakar, Trans.). Cambridge, MA: The MIT Press.

Waddell, Ellen. (2000). *The virtual job interview: Recruitment online.* Retrieved January 6, 2001, from the World Wide Web: http://sjsu.sjweb.net/2/.

Waggoner, M. D. (Ed.). (1992). *Empowering networking and computer conferencing in education.* Englewood Cliffs, NJ: Educational Technology Publications.

Walther, Joseph B. (1990). *Relational communication in computer-mediated interaction.* Unpublished doctoral dissertation, University of Arizona, Tucson.

Walther, Joseph B. (1992). Interpersonal effects in computer-mediated interaction: A relational perspective. *Communication Research, 19,* 52-90.

Walther, Joseph B. (1994). Anticipated ongoing interaction versus channel effects on relational communication in computer-mediated interaction. *Human Communication Research, 20*(4), 473-501.

Walther, Joseph B. (1996). Computer-mediated communication: Impersonal, interpersonal, and hyperpersonal interaction. *Communication Research, 23*(1), 3-43.

Walther, Joseph, & Burgoon, J. (1992). Relational communication in computer-mediated interaction. *Human Communication Research, 19*(1), 50-88.

Warschauer, Mark (Ed.). (1995). *Virtual Connections.* Manoa, HI: University of Hawaii at Manoa.

Warschauer, Mark. (1996a). *Computer-mediated collaborative learning: Theory and practice.* Honolulu, HI: University of Hawaii.

Warschauer, Mark. (1996b). Comparing face-to-face and electronic discussion in the second language classroom. *CALICO Journal, 13*(2/3), 7-26.

Warschauer, Mark. (1996c). Motivational aspects of using computers for writing and communication. In Mark Warschauer (Ed.), *Telecollaboration in foreign language learning* (pp. 29-46). Honolulu, HI: University of Hawaii.

Warschauer, Mark (Ed.). (1996d). *Telecollaboration in foreign language learning*. Honolulu, HI: University of Hawaii.

Warschauer, Mark. (1996e). Computer-assisted language learning: An introduction. In S. Fotos (Ed.), *Multimedia language teaching* (pp.3-20). Tokyo, Japan: Logos International.

Warschauer, Mark. (1998). Online learning in sociocultural context. *Anthropology & Education Quarterly, 29*(1), 68-88.

Warschauer, Mark. (2000). Language, identity, and the Internet. In B. Kolko, L. Nakamura, & G. Rodman (Eds.), *Race in Cyberspace* (pp. 151-170). New York, NY: Routledge.

Weber, Elizabeth G. (1985). *From feelings to knowledge: Verbs of cognition.* Paper presented at NWAVE XIV Conference. Georgetown, USA.

Weisband, S., & Kiesler, S. (1996, April). *Self disclosure on computer forms: Meta-analysis and implications.* Paper presented in the CHI 96 Conference on Human-Computer Interaction, Vancouver, Canada.

Wellman, B., & Gulia, M. (1995). *Net surfers don't ride alone: Virtual communities as communities.* Retrieved February 28, 1996, from the World Wide Web: http://www.sscnet.ucla.edu/soc/csoc/cinc/wellman.htm.

Wells, Rulon. (1960). Nominal and verbal style. In Thomas A. Sebeok (Ed.), *Style in language* (pp. 213-220). Cambridge, UK: Cambridge University Press.

White, Ronald V. (1975). The language, the learner and the syllabus. *RELC Journal, 6*(1), 31-52.

Wherritt, Irene. (1979). Developing listening comprehension in beginners. *Journal of the National Association of Learning Laboratory Directors, 13*(3-4), 7-10.

Widdowson, Henry G. (1990). *Aspects of language teaching*. Oxford, UK: Oxford University Press.

Widdowson, Henry G. (1994). The ownership of English. *TESOL Quarterly 28*(2), 377-389.

Williams, H. L., & Merideth, E. (1995, March). *Internet outposts: Communication patterns in computer-mediated classrooms.* Paper presented at the 6th International Conference of the Society of Information Technology and Teacher Education (SITE '95), San Antonio, Texas.

Wilkins, H. (1991). Computer talk: Long-distance conversations by computer. *Written communication, 8*(1), 56-78.

Wilson, T., & Whitelock, D. (1997). Monitoring a CMC environment created for distance learning. *Journal of Computer Assisted Learning, 13*(4), 253-260.

Willis, Dave. (1990). *The lexical syllabus – A new approach to language teaching*. London, UK: Collins Cobuild.

WinBrill. (1997). *WinBrill.* Retrieved November 2, 1999, from the World Wide Web: http://jupiter.inalf.cnrs.fr/Winbrill/.

Winter, Eugene. (1982). *Towards a contextual grammar of English: The clause and its place in the definition of sentence.* London, UK: George Allen and Unwin.

Wolf, Alecia. (1998). Exposing the great equalizer: Demythologizing Internet equity. In B. Ebo (Ed.), *Cyberghetto or cybertopia: Race, class and gender on the Internet* (pp. 15-32). Westport, CT: Prager.

Wolfram, W. (1998). Dialect awareness and the study of language. In A. Egan-Robertson & D. Bloome (Eds.), *Students as researchers of culture and language in their own communities.* Cresskill, NJ: Hampton.

Wolski, Stacy. (1999). *Introduction to communication technology: CMC theory lecture.* Retrieved January 7, 2001, from the World Wide Web: http://www.u.arizona.edu/ic/polis/courses991/COMM_209/CMCTheory.html.

Wolters, Maria. (1998, July 29). *Re: Corpora: Corpus linguistics user needs.* Retrieved August 6, 1998, from the World Wide Web: http://www.hd.uib.no/corpora/archive.html.

Wootton, A. (1975). *Dilemmas of discourse.* London, UK: Allen & Unwin.

Worchel, S., & Austin, W. G. (Eds.). (1986). *The psychology of intergroup relations*. Chicago, IL: Nelson-Hall.

Younger, Greg. (1996). *Logs of NETEACH-MOO sessions*. Retrieved December 5, 1996, from the World Wide Web: http://spot.colorado.edu/~youngerg/moologs.html.

Younger, Greg. (1997). NETEACH-L: An international teachers' forum on your desktop. *CoTESOL Newsletter, 20*(1), 11.

Yuen, Anna, & Foley, Claire. (1998). *CALLing on communicative competence*. Retrieved December 27, 2000, from the World Wide Web: http://rkenner.concordia.ca/t430_98 /Anna/FinalAnnaJol.html.

Zipf, G. K. (1949). *Human behaviour and the principle of least effort.* Cambridge, MA: Addison-Wesley.