Language and interaction in online asynchronous communication in university level English courses

Julie Skogs
Language and interaction in online asynchronous communication in university level English courses

Julie Skogs
Language and interaction in online asynchronous communication in university level English courses

Julie Skogs

DISSERTATION

Karlstad University Studies  |  2015:5

urn:nbn:se:kau:diva-34757

ISSN 1403-8099


© The author

Distribution:
Karlstad University
Faculty of Arts and Social Sciences
Department of Language, Literature and Intercultural Studies
SE-651 88 Karlstad, Sweden
+46 54 700 10 00

Print: Universitetstryckeriet, Karlstad 2015

WWW.KAU.SE
Language and interaction in online asynchronous communication in university level English courses
Abstract

Interaction involves people communicating and reacting to each other. This process is key to the study of discourse, but it is not easy to study systematically how interaction takes place in a specific communicative event, or how it is typically performed over a series of repeated communicative events. However, with a written record of the interaction, it becomes possible to study the process in some detail. This thesis investigates interaction through asynchronous written discussion forums in a computer-mediated learning environment.

In particular, this study investigates pragmatic aspects of the communicative event which the asynchronous online discussions comprise. The first case study examines response patterns to messages by looking at the content of initial messages and responses, in order to determine the extent to which characteristics of the messages themselves or other situational factors affect the interaction. The second study examines in what ways participants use a range of discourse devices, including formulaic politeness, humour and supportive feedback as community building strategies in the interaction. The third study investigates the role of the subject line of messages in the interaction, for example by examining how participants choose different types of subject lines for different types of messages. The fourth study examines to what extent features serving a deictic function are drawn on in the interaction and then compares the findings to both oral conversation and formal academic discourse.

The overall findings show a complex communicative situation shaped by the medium itself, type of activity, the academic discipline and topic of discussion and by the social and cultural aspects of tertiary education in an online learning environment. In addition, the findings may also provide evidence of learning.

Key words:

Discussion forums, asynchronous CMC, net-based learning, interaction, discourse
# Table of contents

Abstract 1

List of papers 3

Acknowledgements 4

1 Introduction 5
   1.1 Aims 8
   1.2 Central concepts 9
   1.3 Outline 11

2 Background and theoretical framework 12
   2.1 Spoken versus written communication 15
   2.2 Computer-mediated communication (CMC) 18
   2.3 Communicating in online learning environments 25
      2.3.1 Interaction 26
      2.3.2 Coherence 27
      2.3.3 Community building 30

3 Methodological considerations 33
   3.1 Data collection and analysis 33
      3.1.1 Material 33
      3.1.2 Participants 35
      3.1.3 Data analysis 37
      3.1.4 Questionnaire 41
   3.2 Ethical considerations 42

4 Summary of case studies 46
   4.1 Case Study 1: Getting a response to discussion thread messages 46
      in an online learning environment
   4.2 Case Study 2: Community building from a distance 49
   4.3 Case Study 3: Subject line preferences and other factors 52
      affecting coherence and interaction in student discussion forums
   4.4 Case Study 4: Features of orality, academic writing and 56
      interaction in asynchronic student discussion forums

5 Discussion 63

6 Concluding remarks and suggestions for further studies 66

References 69

Appendices 75

Papers 75
List of papers


Acknowledgements

Completing a doctoral thesis is not the work of one but the work of many. I am firstly grateful to Högskolan Dalarna for supporting my doctoral studies financially and Karlstad University for taking me on and making me feel welcome in the English department. I am also grateful to the students and teachers who gave me permission to use their interaction for this study.

I am fortunate to have a great collection of family, friends and colleagues from many parts of the world who have given support in different ways during my time as a doctoral student. These people are far too many to mention and I do not want to miss anyone out (you know who you are and what you mean to me). Fellow doctoral students at Karlstad and other universities have been so supportive and given helpful input along the way. Colleagues in Dalarna have offered and given both time and worthwhile input when needed, sometimes without their knowing just how helpful they were. I have also received invaluable help from administrative and technical support both in Karlstad and Dalarna. A special mention is needed here for Paul Flack at the Högskolan Dalarna library for his frequent and patient help with referencing software among other things.

Finally, and most importantly I would like to thank my supervisor Solveig Granath. In addition, Una Cunningham (before residing in New Zealand) and later Annelie Adel have assisted in supervision. I am grateful to all three of you for your time and energy. It is so much appreciated and has helped me grow as a researcher. Solveig’s role as my supervisor has been far beyond the call of duty. I doubt there are many doctoral students who could boast of a supervisor of the caliber of Solveig Granath. Solveig’s dedication as a supervisor is outstanding. While in a doctoral program, one of the important factors for success is the freedom to go your own way and make your own mistakes while at the same time having contact with and getting feedback from your supervisor when needed. Solveig has managed to balance these and despite a busy schedule, I always felt that she gave my needs priority and this was a decisive factor in me completing the thesis. For this I am forever grateful.
1 Introduction

According to one definition, interaction can refer to both “the activity of being with and talking to other people as well as the way that people react to each other.” It can also refer to “the process by which things affect or change each other” (“interaction,” 2009-2012). Being interactive involves “people communicating and reacting to each other” but in a computer context, a program is interactive if it “reacts to the information and instructions that you give it” (“interactive,” 2009–2012). Although these definitions may seem relatively straightforward, researchers in computer-mediated communication (CMC) are still attempting to define what interaction is and how it can be measured in CMC. One of the reasons for the difficulty in defining interaction may be that researchers from different disciplines have a different focus. How interactive something is in technology-enhanced communication may refer to user-computer interaction and not necessarily user-to-user interaction. When studies look at how users perceive interactivity there is also a problem, because users perceive different ways of communicating differently. For example, telephone conversations, which are naturally interactive in that they involve people talking and reacting to each other, may not always be perceived as being interactive by the users (Leiner & Quiring, 2008: 142). Sometimes students may feel that communicating with one another asynchronously online is not “real interaction” and they expect immediate feedback, which is only possible in synchronous communication (Teles, Gillies, & Ashton, 2002: 241). For the purpose of the present study interaction will refer primarily to written exchanges produced when students communicate with and react to each others’ contributions in online discussion forums.

Originally, the focus of the present study was online interaction in English between individuals who had no prior knowledge of each other and yet communicated online by participating in mainstream media discussion forums. However, a pilot study revealed that there was in fact very little interaction taking place in these forums. That is, the participants were not so much communicating and reacting to each other as “talking at” others, or perhaps talking into a virtual void as a way of venting feelings rather than being involved in an exchange. Instead of communicating with one another, participants tended to post a comment on a topic or on a post by another participant, and they rarely replied to the comments they received on their posts. In the rare cases when there was a response to someone else’s post, the exchange usually stopped after that first response. In most of the discussion forums that were included in the pilot study, 10% or less logged in and commented a second time. The tendencies observed in the pilot study revealed that in order to observe interaction where participants contributed more than once to the communication, a different kind of discussion forum was needed. Therefore, the material selected for the present study was taken from discussion forums in a more controlled environment, namely university level courses in English.
One of the most compelling reasons for choosing to study interaction in an online learning environment is the extent to which it is used in higher education today. Although written texts have been an important part of higher education for a long time, these written texts consisted largely of formal and structured pieces of academic work such as monographs, scientific articles and essays. As computers became more widely used and also more widely interconnected, it was possible for students to engage in written exchanges using digital tools as part of their learning. These types of exchanges, where ideas can be put forward and discussed and responses given, were previously only possible in seminars which took place in the same physical environment and in real time.

Communicating online provides both opportunities and constraints. Although there is software that provides opportunities for people to communicate in real time with audio and video, it requires that participants have access to reliable broadband and that they are able to be online at the same time. That is why asynchronous online communication is still widely used in distance education. Participating in online written asynchronous discussion means that participants in different time zones and different geographical locations are able to contribute to the discussion. Even those who have other commitments which can make it difficult for them to be available at the same time on a regular basis are able to participate in asynchronous online discussions at their convenience. Another aspect which makes asynchronous discussion forums an appealing option for participants with many commitments is that it is possible to write short messages and short responses in reply to others’ messages. The flexibility of being able to engage with fellow students and instructors at one’s convenience means that discussions and learning can continue outside any real time seminars. This also means that asynchronous discussion forums may be integrated into campus courses in order to give students the opportunity to engage in learning activities outside their scheduled lectures and seminars.

According to a report published by the Higher Education Agency in Sweden (Amnéus, Lundh, Paulsson, & Westman, 2010), the number of students involved in distance education increased by 100 percent between 2002 and 2010. At the time when this Swedish report was published, it was estimated that 20 percent of the student population were enrolled in distance courses or programs. The majority of these students were doing undergraduate studies. The report also stated that the rate of retention was lower for students enrolled in distance courses compared to students studying on campus. For the academic year of 2009/2010, 81 percent of the campus students completed the courses they were enrolled in while only 56 percent of the distance students completed theirs. The tendency to drop out of distance studies may be partly

---

1 Distance education is a term used to cover education where students and the educational institution are not necessarily at the same physical location. The term is not used uniformly by educational institutions. The amount of contact with instructors and with fellow students varies greatly and so do the learning activities. See Meyer (2009) for a brief overview.
because it is more common to study part-time and single courses by distance. Overall, there is a higher drop-out rate for students enrolled in single courses than for those enrolled in programs, which is also true for campus students. Of those studying via distance only, a total of 76 percent were registered on single courses compared to 24 percent who were enrolled in programs. This suggests that distance students are more likely to enroll in single courses rather than commit to entire programs. For those distance students who are enrolled on programs the retention rate is higher than for those enrolled on single courses (see Table 1.1).

Table 1.1. Retention rates in percent for students enrolled in distance education in Sweden 2008/2009 (Amnéus et al., 2010: 47)

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>Females</th>
<th>Males</th>
</tr>
</thead>
<tbody>
<tr>
<td>Program</td>
<td>80.3</td>
<td>84.5</td>
<td>67.1</td>
</tr>
<tr>
<td>Single courses</td>
<td>47.2</td>
<td>50.6</td>
<td>41.1</td>
</tr>
<tr>
<td>Total</td>
<td>55.9</td>
<td>60.7</td>
<td>46.1</td>
</tr>
</tbody>
</table>

One institution reported that courses that required at least some attendance on campus had a higher retention rate than those that did not. However, other institutions did not report this trend. The distribution of males and females enrolled in a course may also affect the drop-out rate (see Table 1.1). Females appear to be less likely to drop out of their studies than males. Too many variables are unaccounted for in this particular report to make it possible to draw any conclusions about what contributes to student retention or to successful distance education. For example, not much is known about students’ motivation for studying or the conditions under which they study. One of the challenges faced by higher education institutions mentioned in the report is trying to meet students’ expectations. Another challenge is creating a sense of social community and maintaining contact between students and staff when teaching asynchronously (Amnéus et al., 2010).

The figures published in the Swedish Higher Education Agency report suggest that students who are enrolled in distance education are less likely to complete their studies. This seems to be the trend at least for distance students enrolled in single courses. Some studies however show that students are more successful in distance education than their campus counterparts (Shachar & Neumann, 2003). Individuals have different learning styles and it is possible that different types of students are suited to different modes of study. Online studies may work well for those who already possess effective study skills and are less dependent on the benefits of social contact, but for those who do not possess these study skills it may be more difficult. A lack of these study skills may partly account for the higher dropout rate among distance students. Previous studies have identified certain characteristics that may predict how successful a student will be in a situation where a computer-mediated environment replaces or complements the traditional classroom environment. The most successful
online students show persistence, an ability to direct their own study and take responsibility for their achievement. These characteristics indicate that successful online students realize that they are responsible for their own learning and that successful learning does not depend on chance or anything else beyond their control (Hiltz & Goldman, 2004; Wang & Newlin, 2002).

There are, however, many factors that can affect students’ success rate and these all need investigation. One way of understanding what takes place in an online learning environment is to study communication and language use in an asynchronous discussion forum. The present study examines communication in online undergraduate English courses in literature, culture and English proficiency taught at a Swedish university. The study is an attempt to shed light on written asynchronous communication amongst undergraduate students in English who are proficient users of English according to the Common European Framework of Reference for Languages (CEFR). It is only by studying the communication that takes place in these environments that we may gain a better understanding of such environments and the interaction that occurs there. By examining discussion forums used in an educational setting and how the language is used in computer-mediated asynchronous communication, it is possible to learn more about the nature of this interaction and the functions of the language features used.

1.1 Aims

The overall aim of this study is to explore the written online interaction of tertiary level English students communicating with one another in discussion forums used as a learning activity. More specifically, the study aims to examine a number of aspects of online written communication in English. Firstly, by examining the content of messages and responses, the study attempts to determine what factors seem to affect the choices made by students and instructors regarding which messages they write a response message to. The content of the initial messages and their response messages can reveal the kind of reaction the initial message elicits. Secondly, the study attempts to determine in what ways participants use discourse devices, including formulaic politeness, humour and supportive feedback in community building by examining participants’ texts for linguistic evidence which might reflect the participants’ awareness of sense of community and what devices they may employ to build and maintain it. Thirdly, an examination of the linguistic items that participants choose to place in the subject line of their messages may help shed light on, among other things, how the different items contribute to overall coherence in the forum, what subject line preferences the participants have and if different types of subject lines have an effect on whether other participants choose to read a particular message. Finally, the study examines the frequencies of a number of features that have a deictic function and compares the frequencies with two offline registers, namely oral conversation and academic discourse. The features examined include personal pronouns as well as temporal and
spatial adverbs. It is hoped that this comparison sheds light on how participants make reference to themselves, to the discussion and to referents outside the discussion.

Overall, the present study attempts to increase our knowledge of how students communicate with one another in English using the digital tools provided and how they interact with one another in an asynchronous online learning environment. It is also hoped that furthering our knowledge of the contextual and situational constraints of asynchronous online discussions and the influence they have on communication will provide valuable knowledge which can be applied in such communicative situations. In addition, both instructors and students may benefit from an informed awareness of factors affecting communication in an online learning situation and this knowledge might be used to improve the way learning activities are designed and managed.

1.2 Central concepts

As the focus of the present study is communication in online discussion forums, it is important to discuss the environment within which the communication takes place. An *asynchronous learning network* (ALN) can be said to be a place for learning where those who participate can engage in learning activities with one another anywhere and anytime. The features that define an ALN are firstly that it is asynchronous and secondly that the learning activities that take place are cooperative or collaborative (Hiltz & Goldman, 2004: 5-6). This type of learning may not suit all students or instructors. There is an ongoing debate whether distance education requires more or less work for teachers and students than campus-based. A lot may vary depending on the type and content of the course. DiBiase (2000) argues that distance courses for a particular subject do not require more time to maintain; however, course preparation is not taken into account. Rothkopf (2003) on the other hand uses a mathematical formula showing that teaching costs for distance education may vary according to student numbers and also according to the amount of interaction. Online courses require students not only to be more active than they are when listening to a lecture but also to have good reading, writing and computer skills (Hiltz & Goldman, 2004: 13). The asynchronous discussion forums that are the focus of the four studies which make up this thesis share some features of an ALN. The forums provide a place for asynchronous cooperative learning activities but in contrast to an ALN there are time constraints on the discussion forums. The participants may engage with one another no matter where they are and contribute to the forum at any time during the time the forum is open and this time is normally only about four days. The topics are pre-determined and the deadlines must be respected.

A *learning platform* (sometimes called *learning management system*) refers to the software which allows communication in the learning environment. Most learning platforms have a variety of resources such as a chat function, a
discussion function, a quiz function as well as being a repository for pre-recorded lectures and files of various formats. *Forum* refers to the space within which written contributions are posted during the time frame designated by the instructor. The contributions within this forum deal with a topic or a choice of topics assigned by the instructor. Within the forum there can be a number of *threads*. Participants can choose to start a new thread or build on an existing one. The first message of a thread is called the *thread-initiating message*. It is possible for other participants to build on this thread-initiating message by using a function in the learning platform which visibly links their message to the original thread-initiating message, and these connected messages are called *responses*. The responses are ordered hierarchically beneath the message they are a response to. Each thread is meant to deal with a new aspect of the discussion.

Those of us that have been involved in distance education for a number of years and have used a number of different course designs, tools and learning platforms know that the kinds of activities, the type of contact and communication, the learning goals and the software available may vary. *Computer conferencing* and *discussion boards* are other terms used to refer to asynchronous discussion forums and although the software may be different, typically they allow participants to contribute to an ongoing discussion without having to be online simultaneously. Moreover, the software used normally automatically creates threaded discussions. This means that as a rule, responses are visibly linked to other messages and ordered in some way, often chronologically. Although some learning platforms order messages thematically, the forum used for the present study orders the messages chronologically.

Different types of ALNs have been in use for decades and research on them is plentiful. A meta-analysis of studies on the effectiveness of ALNs has shown that they are at least as effective as traditional learning environments (Hiltz, Zhang, & Turoff, 2010). However, a situation where students are not required to meet each other or the teacher in real time and space is not new per se. Prior to online communication, correspondence courses have been used for students living in remote locations or students who for various reasons were not able to attend campus courses. What is new is the fact that students and their instructors can now communicate through multiple channels and get responses and feedback without too much time lapsing between responses. The short bursts of communication which constitute the messages posted in the forum are not like the long academic pieces of writing that were traditionally submitted by students. The main difference is that the forum exchanges resemble seminar discussions but are performed in writing and asynchronously. In addition, the online discussions remain accessible and visible for all participants even after the exchange has ended. Data provided by the computer systems such as time and date of posting, when and how often students were logged in, and how many participants read individual messages, also provide a new dimension for research. This information is useful because, as Webb, Jones, Barker and van Schaik (2004) have shown, there may be a correlation
between students’ activity in the forum and their final grade. That is, there is a correlation between those who both spend more time in the forum and post more messages and higher grades.

The current study focuses on the written interaction that takes place in asynchronous online discussion forums in English courses in culture, literature and proficiency. Interaction refers to the content of the messages and how the participants interact with one another, as well as to the quantifiable aspects of interaction, such as how often they respond to one another. The study examines how the participants manage the interaction, the strategies used to perform different functions and the type of language used. Although it is not a comparison of communication in an online learning environment and communication in an offline learning environment as such, part of the study examines a number of linguistic features and what they might reveal about this type of communication as opposed to other offline registers. Other aspects of the interaction, such as the technical and situational factors that may affect the communication, are also taken into consideration.

1.3 Outline
This thesis is built on four separate studies of data collected from online discussion forums in undergraduate university courses in English. Each study deals with different aspects of the material even though these sometimes overlap. The background and theoretical framework is presented in chapter 2. The material and methods as well as ethical considerations are discussed in detail in chapter 3. A brief summary of the four studies and the findings are presented in chapter 4 and a discussion of the findings and suggestions for further research make up chapters 5 and 6.
2. Background and theoretical framework

The four studies are part of corpus-based research within the area of discourse analysis. Discourse can be said to be language in use and discourse analysis can be said to be the study of language in use (Wetherell, Taylor, & Yates, 2001). In practice it means gathering material which is, if possible, naturally-occurring language and analyzing it. Naturally-occurring refers to the fact that the language is produced without the interference of an observer or anything else that might cause the subjects to change the way they communicate. In answer to the question “why study discourse”, Wetherell et al. (2001: 3) argue that it is because discourse can inform us about social life. They add that “[c]ontemporary societies are mediated through discourse.”

One of the most common forms of naturally-occurring language is conversation. As defined in one dictionary, a conversation is “talk between two or more people, usually a private and informal one” (“conversation,” 2009-2012). The material examined in the present study consists of exchanges between two or more people in what might be called a semi-private setting and in that sense it can be regarded as a type of conversation. However, since the exchanges are written and not spoken it could be argued that the communication is not conversation in the strictest sense of the word. As the communication can be regarded as conversation in all other respects except for the fact that it is written, the material could be more accurately labelled ‘written conversation’. Another term to describe written conversation-like exchanges is “conversational writing” used by Jonsson (2013) to describe the exchanges in synchronous computer-mediated communication.

The material used in the present study is naturally-occurring language, in that the written exchanges take place without the interference of an observer. However, as opposed to informal conversations, participants know that what they write in the discussion forums will be read by their peers and their instructor and is part of the assessment of the course. The exchanges consist of a series of written conversations between participants in a particular setting. The analysis may not reveal anything about society in general but may reveal something about one particular society or group, namely one where a specific type of conversation-like exchange takes place asynchronously in an online academic learning environment. By identifying features of discourse and the patterns that occur, it is possible to learn both something about the relationships of the participants and to recognize patterns of usage previously observed in other communicative contexts.

In order to do an analysis, a text or a collection of texts has to be compiled. A collection of texts compiled for linguistic analysis is called a corpus. Prior to computers, the collection and analysis of texts was done entirely manually, which means that the amount of text that could be analyzed within a reasonable time was greatly limited. More recently, computer software is used in corpus
analyses, although usually some degree of manual analysis is still necessary. Large corpora consisting of texts from different registers serve the purpose of providing insight into the way language is used in general, while a more focused study using a smaller corpus, such as the present study, highlights the particular characteristics of one type of communication in one particular context. By analyzing linguistic features and patterns of usage we can both test and form theories about language in use and the functions that it performs.

As the data used in the four studies were collected from a specific academic context, Tribble’s (2002) analytical framework for analyzing corpora provides a base. Tribble (2002: 132-133) argues that using smaller corpora can provide useful information that will improve understanding of academic written discourse. This framework consists of seven contextual aspects and three linguistic ones as listed in Table 1.2.

Table 1.2. Analytical framework (adapted from Tribble, 2002: 133)

<table>
<thead>
<tr>
<th>CONTEXTUAL analysis</th>
<th>LINGUISTIC analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name</td>
<td>What is the genre?</td>
</tr>
<tr>
<td>Social context</td>
<td>What social setting? What are the constraints and obligations?</td>
</tr>
<tr>
<td>Communicative purpose</td>
<td>What is the communicative purpose of the text?</td>
</tr>
<tr>
<td>Roles</td>
<td>What roles are required?</td>
</tr>
<tr>
<td>Cultural values</td>
<td>What shared cultural values are required?</td>
</tr>
<tr>
<td>Text context</td>
<td>What knowledge of other texts may be required?</td>
</tr>
<tr>
<td>Formal text patterns</td>
<td>What shared knowledge of formal text features (conventions) is required to write effectively in this genre?</td>
</tr>
<tr>
<td>Lexico-grammatical features</td>
<td>What features are prominent and salient?</td>
</tr>
<tr>
<td>Text relations/textual patterning</td>
<td>Can patterns be identified and what are the reasons for these patterns?</td>
</tr>
<tr>
<td>Text structure</td>
<td>How is the text organized and for what reason?</td>
</tr>
</tbody>
</table>

Below, each of the aspects of context suggested by Tribble is discussed in relation to the material used in the four studies. With regard to the contextual aspects, the material analysed might be categorised as academic discourse owing to the context in which it is produced. However, in large corpora, discourse which is labelled academic discourse is normally the type of text found in peer-reviewed books or articles, as well as in university textbooks. Sometimes it also includes the spoken discourse used in lectures and seminars. The material used for the four studies does not fall clearly into one type of text. Although the material is written, it resembles conversation in that it involves exchanges between two or more participants responding to something that someone else has contributed to the exchange. In some respects, the material does share certain contextual features with academic discourse. For example, the content is of an academic nature and the exchange takes place in a very specific social setting, that is, between student peers (or between students and instructors) at tertiary level. This means that there are “constraints and obligations” which affect how the participants communicate. However, the constraints and
obligations that are found in this particular social setting may not be the same as in other academic settings. For instance, an online discussion forum is a place where less formal ways of communicating can be expected, compared to for example the writing of an academic paper.

The communicative purpose of these online discussions is to help students develop their knowledge and skills in collaboration with one another through the asynchronous exchanges. The exchanges in the forums express opinions, ask and answer questions and contribute to helping students develop course-related skills and competences. The exchanges are a way for students to develop and demonstrate their knowledge and understanding of the course material and the students’ messages contribute to their assessment. The participants have roles as either students or instructors. In a traditional classroom environment, the instructor normally dominates the communication but in this particular environment the instructor assumes a lower profile when it comes to presence and participation. In the forums where there is instructor input, it is meant to guide and help the students in their thinking and reasoning. Rather than transmitting knowledge, the communication is more concerned with building knowledge.

Determining the shared cultural values of the participants is more complex. Even though the students and instructors share some knowledge of communication in an academic environment, the individuals come from various backgrounds and have had different pedagogical experiences. This means that participants’ expectations as to how to express themselves when communicating with one another in this context may vary. A number of studies that specifically address students’ expectations of communication in an academic environment are discussed in Flowerdew (2002: 235-301) which concerns ethnographic approaches to academic discourse.

The exchanges in discussion forums take place in a computer-mediated environment among participants whose experience of communication in an online learning environment may vary. The choices made by participants and how they communicate may depend on the extent of their knowledge of the conventions of discourse in an academic environment. In addition, their linguistic choices may be influenced by their perceptions and beliefs about communication in different types of offline and online contexts. Students and instructors who are familiar with other types of online communication may feel that the same text conventions found in other registers of CMC, such as email or chat, may be used in the discussion forums for university courses. This might mean that participants would not feel that the obligations and constraints of other forms of academic writing, for example essay writing, necessarily apply to the discussion forum exchanges.

The contextual aspects provide a background for the linguistic analysis of the material which looks at linguistic features and the functions they perform. Such
features include the lexico-grammatical choices made by the participants, pragmatic choices, their choices concerning the organisation of the text as well as recurring patterns of usage. Language in general can be said to fill a number of functions. Jakobson (1999: 48-54) describes six functions of verbal communication, namely, the referential, emotive, poetic, phatic, metalingual and conative. These functions are understood in terms of the effect the act of speaking has on others. He also states that the structure of a message will depend on its primary function (Jakobson, 1999: 48-49). Halliday (1994) takes another approach and suggests that language is about potential choices made by speakers and writers and that these choices perform three different functions, an ideational one, an interpersonal one and a textual one.

Although the four studies do not use Hallidayan systemic-functional linguistics as the theoretical framework, the focus is, nevertheless, on the functions performed and the choices made. The first study of the four studies in this thesis is concerned mainly with the ideational function of the texts, but also to some extent with interpersonal functions. The second study focuses on the interpersonal function and specifically how a number of features contribute to community building amongst the participants. The third study focuses on the textual function as well as interpersonal functions. It examines how participants interact with one another and how the linguistic choices they make contribute to the organisation and overall coherence of the text. The focus of the fourth study is the frequency of the different lexico-grammatical features used to perform ideational, textual and interpersonal functions and how these frequencies compare to those of other types of texts.

Herring (2004) drew on a number of different disciplines to devise a framework for computer-mediated discourse analysis (CMDA). She referred to this as an approach rather than theory or method, as it allows “diverse theories to be entertained and tested” and it also allows for the use of multiple methods for analysis (Herring, 2004: 341). Underlying CMDA are two basic assumptions taken from linguistic discourse analysis: firstly, that there are identifiable observable patterns in discourse, and secondly, that those who produce the discourse have made choices. These choices can reflect both cognitive and social factors (Herring, 2004: 341). In addition to these two assumptions from discourse analysis, Herring adds a third assumption, which is that the technology may or may not have an effect on the language (Herring, 2004: 341). It is within this framework that the present studies are carried out.

2.1 Spoken and written communication

Before dealing with communication in an online discussion forum, it is important to discuss different ways of communicating in general. It is difficult to devise a model of communication that can be applied to all types of communication because of the varying nature of why, where and how we communicate with one another. Sperber and Wilson (1986: 2) draw a parallel
between communication and locomotion. Even though walking, flying and swimming are all forms of locomotion they take place in very different environments and require specific theories to explain the conditions. Elaborating on their example, it can be said that in order to achieve motion we use our knowledge of the physiology of our bodies and the environment and using this knowledge we can move in the desired direction and at the pace chosen within the physical constraints that exist when we want to walk and run on the ground. Moving through air or water successfully means learning about the conditions that are afforded by those particular environments and adapting to them. We may be able to use some of what we know about walking and running and holding ourselves upright in these environments, but this knowledge may not be enough to help us achieve our goals in these environments. We may need to adapt our knowledge or learn something new. One environment might allow us to move in ways that another one does not. This is a useful comparison which helps us understand the complexities of communication that takes place in different environments and for different purposes.

During the early part of the twentieth century, linguists regarded speech as the most important object of research while the written language was regarded as secondary. Outside of linguistics, however, written language was regarded as more important and this was emphasized in the school system where English literacy and not oral communication was given priority. Historically and culturally it can be argued that we speak before we write, but this may not be reason enough to argue for the spoken language being more important than the written. In addition, more recently the boundaries between spoken and written language have become less clear. However, they can be both considered important and equally worth investigating (Biber, 1988: 7-9).

In attempts to describe language and the variation in usage that occurs, many researchers have tried to categorize language into different modes. One commonly used division is that between spoken and written language. One of the widely used books comparing oral and written language was Horowitz and Samuels’ (1987) edited work Comprehending Oral and Written Language. In particular, the chapter by Chafe and Danielewicz (1987) which describes differences in frequencies across genres has proved to be a seminal piece of research. This work showed that there were apparent systematic differences between spoken and written language in such areas as the use of personal pronouns. However, the authors suggested that the differences might be the result of the purpose of the communication rather than the fact that the communication was spoken or written. The work of Biber (1988), among others, also showed variation between and within spoken and written language. Biber devised a method of comparing spoken and written texts by using different linguistic dimensions and sought to describe the linguistic characteristics that co-occurred in a number of different text types (Biber, 1988: 55). Biber's dimensions can be described as clusters of empirically observable
features that co-occur in texts. Similarities and differences between the texts could be attributed to these co-occurring features (Biber, 1988: 55). Biber’s notion of dimension had a number of characteristics, one of which was that the dimensions were on a continuum rather than separate and mutually exclusive. For example, instead of categorising texts as either formal or informal, it would be more accurate to describe them as more or less formal (Biber, 1988: 9). An important finding was that looking at one dimension is not enough to describe a text. Instead, a multi-dimensional approach is necessary in order to come to any conclusion about the function of the text (Biber, 1988: 24).

Based on the frequencies of different linguistic features, texts were previously classified according to three dimensions as a) interactive as opposed to edited, b) abstract as opposed to situated, and c) reported as opposed to immediate (Biber, 1988: 57). Based on the features examined by Biber, a more complex set of dimensions evolved. Sometimes, different genres could be found at opposite ends of the dimension but other times the frequencies of features placed different genres at a similar position on the dimension. For example, along the dimension for involved production as opposed to informational production, conversation has a higher score and academic prose a negative score. This means that conversation has frequent occurrences of the second person pronoun and contractions as well as a number of other features. Together with these there would be a lack of attributive adjectives, prepositions and less variety in the language (Biber, 1988: 129). Conversation and academic prose also contrasted along the dimension for situation-dependence. For example, academic prose had more nominalisation and fewer place and time adverbials than conversation (Biber, 1988: 142). Along the other four dimensions however, conversation and academic prose shared many frequencies. Biber’s approach, in contrast to earlier studies of spoken and written language, was based on the assumption that co-occurrence of linguistic features was an indication of the function of the text. Biber’s (1988) work has had a major impact on how texts are described. His work also brought to light the inadequacy of simply labelling texts as written or oral, formal or informal.

Another comprehensive account which compares linguistic frequencies across genres is Biber, Johansson, Leech, Conrad, and Finegan (1999). This descriptive grammar has been used in the present thesis as a background for describing the functions of a number of grammatical features. Biber et al. use the term grammatical features broadly and conclude that much of the variation in feature use is systematic. They also conclude that this systematic variation is related to contextual factors, including the purpose of the communication, whether a text is spoken or written, and the participants themselves (Biber et al., 1999: 5). The work of Biber et al., (1999) is particularly helpful for this thesis in that it provides frequencies for features appearing in a number of registers. Two of these registers were used for comparison in the fourth case study which focuses on comparing communication in an online discussion forum with oral conversation and spoken and written academic discourse.
In order to understand and analyse communication, it is important to be aware not only of the medium but also of the purpose of the communication, the participants, and the social situation. Hymes (1974), with his background in ethnography, emphasized the importance of analysing language in its social setting and not as something separate from people or contexts. It is not the language itself that should be the focus but communication and the communicative activities of the community or network of people who produced the material being analysed (Hymes, 1974: 4). Although the present studies focus on the language used, the particular social setting is drawn on to explain why language is used the way it is.

2.2 Computer-mediated communication (CMC)

Written language in itself, whether it is on paper or in a computer-mediated environment, has a degree of permanence that speech has not. The role of written language has traditionally been “to keep records of all kinds, and written documents have facilitated communication over long distances and long periods of time” (Milroy & Milroy, 1991: 55). In contrast, conversational speech was regarded as primarily performing a social function, conveying emotions and attitudes among other things (Milroy & Milroy, 1991: 55). However, in a computer-mediated environment, some of the written language appears to share features of conversational speech, filling a social function as well as facilitating communication over space and time (Herring, 1999; Yates, 1996). Email communication and text chatting as well as contributions to discussion forums are often carried out with very little planning and are comprised of relatively short stretches of language, unlike traditional offline written communication. One study shows that participants in a discussion forum think that a good message is one that is no longer than about 25 words (Berthold, Sudweeks, Newton, & Coyne, 1997).

Some researchers have suggested that CMC is a variety of language in itself. The term interactive written discourse (IWD) was introduced by Ferrara, Brunner and Whittemore (1991). They hypothesized that a variety of language was developing in computer-mediated environments which was characterized by “the written language occurring in simultaneous terminal-to-terminal typed dialogue” (Ferrara et al., 1991: 9). They claimed that this was a naturally occurring register and a hybrid variety sharing features of both spoken and written language and that it was becoming conventionalized (Ferrara et al., 1991: 10). That particular study was relatively small and confined to a group of computer professionals and their families, who themselves may be seen as forming a particular discourse community based on their shared work and interests. Therefore, what was observed may not have been the direct result of online communication but a result of this particular community of practice. It is an empirical question whether the patterns of usage observed in their communication would be found in other more diverse groups communicating online for different purposes. A group of subjects with a different background
may not have used the same features as those observed in the Ferrara et al. (1991) study. Herring (2003: 613) suggests that the value of their study lies in the fact that it brought into focus the language used in a computer-mediated context and consequently, it gained the attention of a number of other linguists (Herring, 2003: 613).

A number of different terms are used to refer to communication which takes place using electronic media such as computers and other technological devices. The most widely used term is computer-mediated communication (CMC). Research into CMC includes not only studies of various aspects of the language, but also aspects of the medium itself. Other terms such as computer-mediated discourse (CMD) and computer-mediated language (CML) are sometimes used to highlight the fact that it is the language used which is the focus of the study rather than the technical aspects of the communication. In any study of communication which takes place with the help of technology, the medium and the effect it may have on communication must be taken into consideration, even if the language is the focus. In the present study, the term CMC is used; even if the language is the main focus of this thesis, aspects of the medium and their effect on language use is also dealt with.

One of the pioneers in the field of research into CMC is Naomi Baron who reflected on the impact of technology on language in an early study (Baron, 1984). She uses the term electronically-mediated communication (EMC) which includes not only exchanges via a computer but also communication that takes place with the help of devices such as mobile phones. Baron was interested in how the medium itself affected the linguistic message and suggested a number of aspects of communication with regard to this question (Baron, 1984: 120). One aspect she suggested should be investigated was the range of vocabulary used in CMC compared to face-to-face communication. Another area which warranted investigation was the social aspect of computer conferencing (Baron, 1984: 131). Her work was seminal in bringing the attention of other linguists to the language being used in computer-mediated environments.

Another important contribution to research in CMC is Susan Herring’s edited volume from 1996. Here, Herring identifies three key issues of research within CMC that were in focus in the early 1990s. These concerned firstly the language of CMC and how it compared to traditional written and spoken language. Secondly, there was an interest in communication in a medium where extra-linguistic cues were not available as there was speculation that CMC was impersonal. Thirdly, there was also interest in social interaction that was computer-mediated and used in online communities (Herring, 1996: 3-5). The articles in Herring’s collection of empirical studies revealed a number of interesting findings. Much of the speculation and theorizing about the nature of CMC and its potential benefits and harmfulness seemed to be unfounded. The first section of the book, on linguistic aspects, showed that CMC shared the features of a number of existing oral and written genres. Other sections dealt
with the cross-cultural and social aspects of CMC and showed that CMC was not impersonal and could bring people from different cultures into contact with one another. Later, Herring (2003) identified four areas of research into CMC which were current at the time: classification of computer-mediated discourse (CMD), linguistic structure, social practice and interaction management (Yates, 1996: 614-625). Herring (2007) devised a classification scheme for analysis that went beyond traditional approaches to discourse such as oral or written as presented in Chafe and Daniewich (1987), or genre and register the way it was presented by Biber (1988) and Swales (1990). Her approach was inspired by Hymes (1974) and Baym (1995) and her own background in library science and classification. Her classification scheme took into account multiple aspects of the medium and the situation, with a view to providing a better base for comparison across and within the different types of communication. For example, communication that took place asynchronously would be expected to differ from that which took place synchronously. If they are compared, then the fact that one is asynchronous and the other synchronous must be taken into account when discussing findings. Crystal (2006: 139) argues that “technologically imposed length constraints are an important factor influencing the linguistic character of chat group messages,” meaning we would not expect to find a long stretches of writing in such a communicative situation.

Although Baron (1984) tried to bring the impact of technology on the language into focus, a number of researchers have questioned the assumption that there is such an impact. Renkema (2004: 70) suggests that we should examine this assumption more critically. One of the problems with early studies in CMC was, according to Herring (2003: 613), the tendency to over-generalise due to a failure to recognise in what ways technical and contextual factors affect the language. One media researcher, Boczkowski (2005), warns that we should not overestimate the effects of technological determinism on our communication and adds that what we observe in CMC has usually developed from “established communication practices” (Boczkowski, 2005: 3).

There are aspects of written CMC which differ from written offline communication. The most obvious characteristic is the fact that print versions of texts are linear, whereas web versions can have, among other things, hypertext links and theoretically allow for readers to go to a link and read another text in the middle of the first text. This allows readers to interact with the technology. Even if not everyone does this, the option is available (Crystal, 2006: 50). The same flexibility is not found in offline written texts. Another aspect of CMC is the opportunity for interaction with other people in the form of comments and responses in discussions. Baron (2000: 227) likens the opportunity for community discourse provided by computer-mediated communication to that provided by rural telephone exchanges where groups of people could come together, even though they lived some distance from each other. So while email exchanges replace the two-party telephone calls and
letters, discussion forums in networked computers replace other forms of multiparty communication (Baron, 2000: 227).

Some software allows multiple users to update and change the same text, resulting in collaborative texts with multiple authors. The closest offline equivalent to these multi-authored computer-written texts might be when one person writes comments on a text written by another person (Crystal, 2006: 216). In CMC it is not uncommon to find texts that are multi-authored. This means that it is not always clear who exactly created a text and when and under what circumstances (Crystal, 2006: 216). Wikipedia is perhaps one of the best examples of written texts produced by multiple authors. Resources such as Wikipedia are designed to facilitate and encourage multiple authors and collaborative writing. The degree to which users are allowed to participate in the production of discourse depends on what technical devices have been included and also how much the site owners monitor who can contribute and which contributions will be displayed. Although Wikipedia is relatively open, some users feel they are not allowed complete freedom (Forte, Larco, & Bruckman, 2009: 55).

Another aspect of electronic texts is that what is written often remains accessible long after the discussion has ended. This particular characteristic of written online communication is referred to as *persistence, persistence of transcript* and *conversational persistence*, and without the text being available on screen for us to trace back through, written computer-mediated conversations would be much less coherent (Herring, 1999). A computer-mediated discussion can be resumed after a long time lapse. This is not possible in face-to-face communication. There are a number of similarities between written communication in electronic environments and in offline environments and these should not be overlooked. These similarities and differences challenge the more traditional views of language prior to online communication. One example is a description from the 1990s which claims that speaking is seen as a social activity while “writing is solitary” (Milroy & Milroy, 1991: 54). CMC consisting of chatting, discussion forums and multi-authored online documents means that many traditional observations about language need rethinking and re-evaluating.

With regard to the written language itself, empirical evidence of the fact that there are both similarities and differences between CMC and non-CMC was found in a study by Yates (1996). Yates’ study used material from computer-conferencing systems where messages were posted, and among other features, it looked at the frequency of 1st, 2nd and 3rd person pronouns. The study showed that CMC was more similar to writing in an offline environment when it comes to the frequency of personal pronouns, whereas when looking at the proportion of 1st, 2nd and 3rd person pronouns, CMC appeared to be closer to speech. In total, the material examined contained fewer 3rd person pronouns than either traditional writing or speech (Yates, 1996: 41). The study also
concluded that CMC contained significantly more modal auxiliaries than both speech and writing, but CMC and speech were found to be similar with regard to the relative frequencies of different semantic clusters such as ability, possibility, volition etc. (Yates, 1996: 44). This demonstrates that there does not seem to be a clear division between the features that distinguish CMC and other communication that takes place outside the computer-mediated environment.

Computer-mediated environments provide the opportunity to communicate in a conversation-like way but in writing. However, as one researcher points out, it is more time-consuming to write conversation than to conduct a conversation orally (Hård af Segerstad, 2002: 152). When communicating in writing in a computer-mediated asynchronous environment such as a discussion forum, the participants do not share the same physical location and are individually situated at their own computer. In addition, online written communication often involves more participants than are normally involved in a face-to-face conversation and, depending on the software used, they may not know whether those they are communicating with are online at the same time as themselves. This means that writers must be more explicit than when communicating face-to-face because if not enough information is provided it may create communication problems. Communicating across time, space and cultural boundaries always means there is potential for misunderstanding. Herring (1999) shows that CMC is relatively coherent and that affordances of CMC, such as persistence of text, help make up for the lack of other features.

Another feature of CMC is that written texts tend to be edited to a lesser extent than traditional written texts and this is one aspect that makes them interesting. People have the power to shape language on the Internet more than in other media (Crystal, 2006: 215-216). Even if written CMC is less edited than traditional written texts, one study has shown that in some chat situations at least, there is a degree of editing. An example of this is students chatting in a foreign language they are learning, where recordings of their writing behaviour showed that they did edit to some degree (Sauro & Smith, 2010). Text chatting resembles conversation in that it usually takes place in real time or at least almost synchronously. It is also similar to conversation with respect to the unpredictability of the topic, but although it may resemble oral conversation, it is still not quite as unpredictable as face-to-face conversation (Crystal, 2006: 151). Technical limitations and connection problems may also affect communication through synchronous chat (Osman & Herring, 2007). This is particularly relevant in chat and other almost synchronous CMC, but does not affect written asynchronous communication to the same extent.

There have been attempts to create glossaries of the words used in computer-mediated communication. Crystal compared two of these, *Cyberspeak* and *Wired Style*, and found that only 25% of the head words were the same in both (Crystal, 2006: 70-71). Dictionaries are never identical but this is a low percentage compared to traditional dictionaries, where the majority of head
words are normally the same. He also found these glossaries to be prescriptive rather than descriptive (Crystal, 2006: 71). One of the reasons for this may be that they are not dictionaries in the strictest sense of the word, but guidelines for those wanting to communicate online. In CMC, those who do not conform to the conventions risk being ridiculed or even excluded from a particular group (Crystal, 2006: 75-76).

Studies of linguistic structures in CMC often focus on syntactic reduction and other strategies which are used to save time and sometimes space, in line with the observations made by Ferrara et al. (1991). As Crystal (2006) points out in his description of email language, there has been a tendency to focus on the informal features of messages, such as the use of contractions, loose sentence construction, ellipsis, colloquial abbreviations, and acronyms (Crystal, 2006: 127). However, he adds that these features are not indicative of the way language is used in CMC as a whole. Despite this, a number of labels have been given to try to describe CMC. Examples of labels which have been used are Internet language, electronic discourse, the language of the Internet and Netspeak (Crystal, 2006: 19). Crystal uses the term Netspeak but points out that there is no consensus on what constitutes Netspeak and that this may be because just as in other kinds of communication, there are many different reasons for using the Internet to communicate. Language users have different interests, different levels of linguistic and technical ability and the technology and quality of the medium change (Crystal, 2006: 98). At the same time, Crystal (2006) puts forward arguments for the existence of Netspeak, one of which is the fact that people allude to a new variety and he argues that this in itself is evidence for its existence. He gives examples of some lexico-graphological features unique to what he terms Netspeak. He proposes that these features, which are a result of the fact that the communication is computer-mediated, provide a solid base for the claim that Netspeak is a particular variety of the language (Crystal, 2006).

Examples of these features are the symbols found on standard keyboards which are used in new and creative ways to express things that are not normally possible in offline communication. These features include both conventional and unconventional abbreviations which represent such things as laughter (lol, ‘laughing out loud’; rofl, ‘rolling on the floor laughing’) as well as the sounds of words (8r, ‘later’) (Crystal, 2006: 91). This type of word play is not new nor is it unique to CMC. Examples of this in the English language can be found already in the 1800s. One example is Essay to Miss Catherine Jay where abbreviations and sounds are represented with letters and numbers “He says he loves U 2 X S, U R virtuous and Y’s,” (Bombaugh, 1860: 68).

Crystal (2006: 272) argues that the language used in CMC is neither ‘written speech’ nor ‘spoken writing’ but rather something completely new. One criticism of Crystal’s (2006) claim that Netspeak exists and of others such as Ferrara et al. (1991) who try to describe CMC as a new variety, is that that they have perhaps forgotten that written and spoken language vary greatly regardless of the medium. In addition, focus on what is different has meant that what is
similar to established communication may have been neglected. For example, the use of symbols can be said to be a feature of the language of CMC, but it is not the only feature and just examining one or two features will not provide a complete description of the language.

It has been difficult to place CMC at a specific point along the spoken-written continuum, and trying to do so may not even be particularly informative. A more complex system of dimensions such as Biber’s (1988) could perhaps shed more light on the nature of CMC (see section 2.1 above). As Yates (1996) points out, one of the main differences between speech and writing which researchers have focused upon is reference to the self and others (Yates, 1996: 34). Yates (1996) refers to conflicting research on the effect of the medium with regard to the level of involvement as indicated by reference to the self and others in a text. If the medium does have a major effect on linguistic features occurring and co-occurring in discourse, and if this effect is identifiable in all types of CMC, then it would follow that examining CMC as a register would be justifiable. However, if the effect is limited and there are other factors such as social practice and the purpose of communication which play a more important role in determining the features of a text, trying to identify CMC as a register or a variety of language might not be useful. In text-based communication, the only way to identify the field is through textual clues. Yates (1996) suggests that there is a lack of a defined semiotic field in CMC, meaning a lack of clearly defined social structures and physical location. This may be one explanation for some of the features found in this type of discourse, such as a higher frequency of modal auxiliaries that indicate ability and possibility (Yates, 1996: 45).

Another study of chat observed that paralinguistic cues and intonation were represented in writing to compensate for the absence of these when the communication was computer-mediated (Werry, 1996: 61). Examples of these are the use of multiple exclamation marks for volume or capital letters for stress and combinations of keyboard symbols to represent smiles or heads nodding to show agreement or understanding (Werry, 1996: 56-60).

Some researchers have suggested that CMC is a hybrid variety of written and spoken English (Collot & Belmore, 1996). Trying to categorise CMC as a single variety or type because of the mode can be somewhat misleading, as it implies that CMC is different from other kinds of language use (Thurlow, Lengel, & Tomic, 2004: 118). Just as there is much variation in offline communication depending on the situation and purpose of the text, so is there much variation in CMC. Hård af Segerstad (2002) also concludes that dealing with CMC as a single mode is not a useful way to study CMC (Hård af Segerstad, 2002: 238). Even in the limited corpus of emails that she examined in her study, she found that the styles ranged from formal to informal depending on the purpose of the communication (Hård af Segerstad, 2002: 242). Crystal (2006), too, admits to the difficulty in treating for example emails as one single type of communication (Crystal, 2006: 99). The relationship between the participants
and the purpose of the communication are important variables and so are the constraints and opportunities of the technology itself.

Trying to categorise emerging registers as CMC has occupied a number of researchers. One goal of a study on blogs by Herring, Scheidt, Bonus, & Wright (2005) was to help understand the impact of technological changes on forming new genres. However, their study did not convince them that blogs are a new genre or that they are uniquely digital. Their quantitative analysis revealed that self-expression was the reason the majority blogged and with regard to structure and function, personal blogs are very similar to written diaries. Even the non-personal blogs with an external focus, which connect to and comment on current events, can be likened to existing genres in non-CMC such as editorials or letters to the editor in a printed newspaper (Herring et al., 2005: 159). Herring et al. (2005: 160) suggest that rather than being a new genre, blogs are extensions and hybrids of other already existing genres. Many blogs have multiple purposes and combine features from different genres. Community blogs are similar to online asynchronous discussion forums in that they are multi-participant, public, text-based and are dynamically interactive online communication environments.

2.3 Communicating in online learning environments

Research into language in educational settings has often focused on the uniqueness of classroom discourse and how it reflects cultural practice, literacy development, second language development and learning (Temple Adger, 2001). Rather than viewing knowledge as something that exists outside of students and which can be acquired by a process of transfer, a constructivist view of learning sees it as a result of interaction (Hiltz & Goldman, 2004: 20-21). As mentioned in Section 1.2, an asynchronous learning network can be said to be a place for learning where it is possible to engage asynchronously in learning activities involving cooperation and collaboration anywhere and anytime. The most common way of cooperating and collaborating in an ALN is using asynchronous discussion forums. These are referred to in many studies as computer conferencing (Garrison, Anderson, & Archer, 2000; Yates, 1996) and are also referred to as discussion boards (Wijekumar & Spielvogel, 2006). There are different systems for carrying out online discussions and these systems allow communication to take place between people through computers connected to a network. This in turn allows participants to contribute in writing to an ongoing, threaded discussion. The systems, the environments and the participants may change but the systems provide a space for learning in which communication is the means and the manifestation of the learning process. In these spaces, the intention is that the participants, whose common goal is learning, will interact with one another in a mutually comprehensible way and that together they will form a learning “community” and contribute to one another’s sense of belonging to that community. Even though this may be the goal it might not necessarily be realized.
There are a number of factors which may affect the way in which participants communicate with one another in an academic environment. Not all participants can be expected to be familiar with how to communicate in the academic learning environments, particularly as the different academic disciplines seem to have their own preferences for usage (Hyland, 1998, 2005, 2010). In their study of academic discourse, Bourdieu, Passeron and de Saint Martin (1994) examined how students’ social and educational backgrounds are related to their ability to understand academic language. They tested students’ understanding of words which were frequently used in lectures and found that many of them displayed a poor understanding of the meanings of these words (Bourdieu et al., 1994: 37). They also found that results varied not only due to social and educational background but also due to a number of other variables, such as the type of course and how the student used resources such as a dictionary. Furthermore, they found that the results varied according to the type of task they had to perform when tested on their knowledge of academic vocabulary (Bourdieu et al., 1994: 38-39).

In the following sections, 2.3.1, 2.3.2 and 2.3.3, three aspects of communication in online learning environments which are dealt with in the present studies are discussed: interaction, coherence, and community building. Interaction is the main focus of the four studies. Coherence is also part of all four studies as it is an important part of successful communication. The linguistic choices made by writers help readers make sense of a text by such features as the organisation of the text, reference to the outside world and to the participants involved in the communication. A number of linguistic features perform social functions and can be used for community building among members of a group. The social function performed by and in language is important, particularly in an educational setting, as the sense of belonging is one of the deciding factors for students to successfully complete their studies (Heyman, 2010; Tinto, 1997).

2.3.1 Interaction

As interaction is in focus in all the four studies here, the terms interaction and interactive warrant some more discussion. As mentioned in the introduction, interaction can refer both to the communication that takes place and the activity of reacting to another person. For researchers in computer science, the term may refer to the human-computer interaction, but for those interested in language it concerns the communication that is produced and the activity of reacting to others involved in the communication. Biber (1988: 73), for example, counted the frequency of a number of features such as pronouns and direct questions to determine the degree of interactivité in a text. This means that texts written in non-electronic media may also be considered interactive if there is a great deal of reader-writer involvement in the text.
Researchers that try to measure learning sometimes use models of interaction to determine the level of cognitive activity. That is, they look for evidence of knowledge being acquired and used in written transcripts of discussions (L. Gunawardena, Lowe, & Anderson, 1997). The model proposed by Gunawardena et al. (1997) was designed specifically for examining transcripts of computer-mediated discussions. This model identified a number of phases; evidence of each phase could be found in such things as participants showing agreement, asking questions, disagreeing, clarifying concepts, relating personal experience and illustrating the application of knowledge. Using the model designed by Gunawardena et al., Kanuka and Andersson (1988: 16) suggest that knowledge is constructed “in online learning environments through social interchange.” Studies such as the one by Wilson and Stacey (2003) on interaction in computer-mediated learning environments resulted in concrete suggestions for how instructors needed multiple skills to facilitate interaction in these environments. Others, such as Sims (2003), suggest that, among other things, learner control may be an important aspect for interactivity which in turn may provide a learning environment that better meets the students’ expectations of an interactive learning environment (Sims, 2003: 101). Northrup (2001: 31) devised a framework of interaction strategies which included interaction with content, social interaction, collaboration, conversation and performance support.

Even though learning and knowledge construction is not the focus of the present thesis, communication is still an important part of the learning process. The importance of social presence for learning has interested a number of researchers, e.g. Rourke, Anderson, Garrison and Archer (1999). In one study of support in online interaction among students, Fahy (2003) concludes that participants generally used supportive strategies in interaction, while conceding that more research is needed on the correlation between the social aspects of communication and learning in computer-mediated environments.

In the present thesis, a number of issues concerning interactivity are examined and discussed, in an attempt to shed light on the way language choices made by participants affect and influence the interaction. Some examples from the present studies are a) what kinds of messages result in more responses, b) how participants formulate disagreement while maintaining the sense of community, and c) what responses result from participants clarifying concepts or relating personal experience. These features can be said to be part of the social interaction. Most of the interaction is between the students, but in some of the forums examined there is also some student-instructor interaction. Interaction is used here in a broad sense to refer to student-to-student exchanges, including frequency and type of exchange, as well as the organisation of the discourse. Features of text organisation aid coherence and consequently facilitate communication. In addition, strategies for creating and maintaining a sense of community are also important for the interaction itself. The present thesis also
examines the features of the communication that indicate the students’ interaction with the text itself, that is, involvement with the text.

2.3.2 Coherence

In order for a text to make sense, that is, to be comprehensible, it must be coherent (Schober & Brennan, 2003: 95). Herring (2003) identifies two potential obstacles to communication as a direct result of the communication being computer-mediated as opposed to face-to-face communication. One is “disrupted adjacency” and the other is “lack of simultaneous feedback” (Herring, 2003: 618). Hård af Segerstad (2002: 131) also points to these features that are lacking in CMC, and like Herring, she points out that other affordances of CMC, such as persistence of text, compensate for the absence of these features. Disrupted or interrupted turn adjacency in CMC has received more attention amongst researchers than lack of simultaneous feedback. In written asynchronous discussion forums, as well as in almost all synchronous chat involving many participants, disrupted turn adjacency means that responses may not come immediately after the message to which they refer. This is because contributions are normally displayed in the chronological order in which they are posted and this order may not correspond to the conversational order. Consequently, a message and its response might not be placed immediately after one another, and other messages and responses might come in between. Therefore it might be expected that disrupted adjacency can cause coherence problems in an asynchronous text-based environment. Örnberg Berglund (2009) maintains that Herring’s (2003) observation of interrupted turn adjacency in CMC may have been subsequently interpreted by some as being synonymous with misunderstanding or miscommunication (Örnberg Berglund, 2009: 4). As both Herring and Örnberg Berglund have pointed out, there are affordances in CMC which compensate for those features that are missing. The sheer number of people who communicate successfully in these environments suggests that we do perceive the communication as coherent (Bednarek, 2005). A number of studies of online communication show that disrupted adjacency does occur (Örnberg Berglund, 2009; Woerner, Yates, & Orlikowski, 2007). However, in her study of instant messaging (IM), Örnberg Berglund (2009) concluded that disrupted turn adjacency does not necessarily result in misunderstanding or confusion.

It could be that features that make oral conversation coherent are not the same as the ones that make written computer-mediated communication coherent. A study by Schallert et al. (1996) suggests that this might be the case. Their study looked at students’ different ideas on what they thought constituted coherent conversation in an oral and a written discussion. Whereas the oral conversation was perceived as coherent when turn-taking and immediate feedback was working, features such as connecting comments to previous ones or to previous knowledge or experience were what the participants felt made written conversation coherent (Schallert et al., 1996: 480). Another researcher (Lapadat, 2007) noted a reduction in turn sequencing in online discussions as compared
to oral discussions and also observed that the devices used for coherence in online written discussions were not always from conversation. Conventions such as salutations and signing off which are typical of letter writing were also used (Lapadat, 2007: 75).

Apart from disrupted turn adjacency and its potential to affect communication negatively, Herring (1999) also identifies some advantages of text-only CMC. She maintains that although text-only CMC does not seem to be as coherent as face-to-face conversation, this type of communication may have its own advantages. The most obvious advantage when it comes to coherence is conversational persistence, which refers to the permanence of writing as opposed to lack of such permanence in oral conversation. This means that participants can trace back through the exchange and also have more time for reflection compared to oral utterances, which disappear as soon as they are spoken (Herring, 1999).

Although coherence and cohesion tend to be seen as separate aspects of a text they are often talked about in conjunction with one another. A cohesive text, that is a text with explicit cohesive devices, is thought to be easier to process and therefore aid coherence, but lack of cohesive devices in a text does not automatically result in a text which is incoherent or incomprehensible. Örnberg Berglund concluded that the lack of explicit cohesive devices, which are ordinarily found in other types of writing, may indicate that there is no need for these devices (2009: 22). There are other devices than cohesive devices which may contribute to maintaining coherence. One example is background knowledge (Örnberg Berglund, 2009: 22). In her study of SMS, chat, email and IM, Hård af Segerstad (2002) noted that most of the communication took place between participants who knew each other well, and their shared background knowledge meant that they could communicate successfully despite lack of explicitness (Hård af Segerstad, 2002: 255). In line with Hård af Segerstad (2002), McNamara, Crossley & McCarthy (2010) suggest that “if the reader has sufficient background knowledge, the mental representation of a low-cohesion text may be coherent” (McNamara et al., 2010: 60). In their study of academic essays, they showed that there was no correlation between texts containing cohesive devices, defined as connectives and co-reference, and essays that were given high grades, even though coherence was one of the grading criteria. Those who graded the essays must have found the essays coherent despite a lack of explicit cohesive devices (McNamara et al., 2010: 74). The view that lack of cohesive devices does not result in an incoherent text is also supported by Hellspong and Ledin (1997). While cohesion refers to the linguistic signals that identify context, it is not absolutely necessary to have this cohesion for a text to be coherent. The main thing is that it is holistically meaningful for someone (Hellspong & Ledin, 1997: 35).

In an attempt to explain how sense is made in communicative situations, Langacker (2001) proposes what he terms the current discourse space (CDS)
which refers to participants’ presumed shared knowledge as well as their sense of the ongoing discourse. Aspects of this CDS, including previous and subsequent events of the discourse, are accessible and may be used to create meaning (Langacker, 2001: 145). How a message connects to previously posted messages in a dialogue is “a central aspect of coherence in asynchronous computer-mediated communication” (Severinson Eklundh & Rodriguez, 2004: 1). There are a number of ways of linking to previous messages. Examples of these can be repetition, directly addressing the commenter and also cutting and pasting in sections of someone else’s message (Tanskanen, 2006: 122). These connecting strategies have also been noted by other researchers. In a study of coherence in the online classroom, Schallert et al. (1996) noted that in the written conversation, postings were not very often direct responses to the immediately preceding comment. They also remarked that “it was rare for a comment to be posted without some sort of explicit marker showing how it was related to at least one comment that had been made previously” (Schallert et al., 1996; 480). This means that because responses do not necessarily come directly after the message they respond to, other explicit markers such as repetition or directly addressing another participant are used instead.

In their comparison of synchronous and asynchronous environments, Paulus and Phipps (2008) found that naming, greeting, closing and inviting were used somewhat more frequently in asynchronous discussions than in synchronous discussions (Paulus & Phipps, 2008: 471). Greetings and addressing another participant (addressivity) are often used in combination. If a response is introduced with a greeting directed at a particular participant, then it may help connect the response to that particular participant’s comments rather than the comment in closest proximity (both spatially and temporally). This connection can be made in the subject line of a message or in the body of the message. In a study of conflict in CMC, Graham (2007) gives examples of how participants use the subject line to reflect their position in a discussion and suggests that this is an indication of the importance of the subject line for putting messages in context. Graham (2007: 749-750) also claims that the fact that participants use this strategy portrays them as competent users. The present study examines the use of these features associated with creating coherence in CMC and compares their usage to previous findings on their function in asynchronous CMC.

### 2.3.3 Community building

Aspects of communication other than those promoting interaction and creating coherent texts may also be relevant for contributing to an environment that facilitates learning. One important aspect is creating a sense of community. Addressing another person will aid coherence by connecting messages but will also fill a social function. Both Vygotsky’s (1978) and Bakhtin’s (Matusov, 2011) work brought attention to the social nature of learning and had a major influence on socio-cognitive research traditions. The ‘communities of practice’ described by Lave and Wenger (1991) have been useful to help understand how...
learning takes place in both schools and workplaces where groups of people share a set of established practices. According to supporters of this theory of learning, participants construct knowledge collaboratively. Learning takes place when newer members of the community of practice first begin as observers of the practices and then gradually participate in them.

The focus on practices has been useful, but Gee (2004) raises a number of issues concerning the use of the concept ‘community’. He suggests the ‘space’ as a better alternative to refer to what he calls this form of social affiliation. Those who share practices in a learning environment might not have close ties and it may be more useful to focus on the space rather than the community itself. Gee (2004: 77-89) believes that practices are only performed in a particular and somewhat limited context and the ties between participants may only be temporary and quite loose.

Discourse and social practice in a computer-mediated environment are of great interest to researchers in a number of fields both inside and outside learning environments. As computers have become increasingly networked, computer-mediated social interaction has also increased. The availability of free, user-friendly software gave new opportunities even for those who have not been used to interacting with one another with the help of technology. The popularity of these online communication tools indicates that maintaining social relationships is important. Studies on the correlation between social presence and students’ satisfaction have suggested that there is a correlation between degree of social presence and a high final grade in an online learning environment (C. N. Gunawardena & Zittle, 1997; Liu, Gomez, & Yen, 2009; Richardson & Swan, 2003). One of the difficulties for researchers, however, has been in developing tools that can be used for determining social presence. Rovai (2002a, 2002b) used a set of questions which the students responded to in order to measure the sense of community that they experienced. She found that sense of community was important for students to succeed in completing their studies. Haythornthwaite, Kazmer, Robins and Shoemaker (2000) also examined perceived sense of community and concluded that experiencing a sense of community benefits the individuals. In their study, all the courses they observed began with a few intensive days where students met in the same physical environment and this was seen as an important part of creating a sense of community. Being able to meet other participants in person is not always an option for students studying online however, and therefore the course design must take into account the importance of community building and the tools and activities used must reflect this. Aspects of building and maintaining community are considered to some extent in all four studies in this thesis.

Chapter 2 has dealt with the theoretical framework for the four studies in this thesis, previous empirical studies of features of spoken and written communication and relevant aspects of computer-mediated communication. In addition, specific aspects of communication in an online learning environment
have been covered. The next chapter deals with the methodological considerations of the thesis.
3 Methodological considerations

Techniques, procedures and tools for studying online communication are constantly being developed as more and more research on online material is carried out. Even though electronically stored material is easily accessed and is a cost-effective way of gathering large amounts of data, any study requires careful planning and a great deal of thought needs to be given to design. Aspects of design that need to be considered include the gathering of data, the analyses and interpretation of results as well as any ethical considerations that need to be addressed (Hewson, 2008: 59). As the material for the present thesis is collected from a restricted access forum, a number of the methodological concerns could be more easily managed than if the material had been taken from an open access and public discussion forum. For example, the identity of the participants could be confirmed and they could be contacted and informed of the nature of the research project. Furthermore, it was possible for them to make informed choices about whether or not to consent to be a part of the study.

3.1 Data collection and analysis

This section is divided into four parts. The first two sections deal with the data and the participants. The third section deals with the analysis of the data and the fourth section contains a description of the questionnaire that was sent to the participants to gain additional information to help with the data analysis.

3.1.1 Material

The material gathered for this study comes from three undergraduate courses taught in an English department at a Swedish university. The material selected is from courses that ran during the first half of 2009. With the exception of one group of students in one course, all other students used online communication only in their studies. This means that some of the learning took place in real-time seminars where both instructors and students were present, while other learning activities were computer-mediated and asynchronous. The activity that is the object of the current study is the course participants’ asynchronous written interaction in the discussion forums for the different courses. The discussion forums are where opinions can be expressed and ideas discussed by the participants asynchronously and in writing for a limited time. These opinions and ideas are visually organised by computer software in a way that makes the contributions resemble a conversation in some respects. Such a “conversation” within a discussion forum consists of an original idea, which is the thread-initiating message, and a number of responses to that idea as well as responses to responses. Figure 1 is a visual representation of how the messages are ordered hierarchically according to when they are posted and to which messages they have been connected.
Participants choose which previously posted contribution they wish to connect their new message to or they can choose to start a new thread or “conversation”. What makes the interaction in a discussion thread resemble a conversation is that two or more participants are involved, the exchange is relatively informal and there is turn-taking among the participants. It is also relatively private in an online context in that permission to access the forum is required and is restricted to those enrolled in or teaching on the course. For the purpose of the current study, the terms discussion forums and forums refer to the enclosed online space which was open to a limited number of participants for a limited time. Here participants could both read what was written and write their own contributions or messages relevant to the topic designated by the course instructor. After the forum was closed for contributions, participants were still able to read all of the contributions that had been posted but no longer post messages.

The discussion forum material came from courses in literature (different novels for each thread), culture (different topics for each thread) and language proficiency (different sections of the text book dealing with grammatical issues). The literature course consisted of four groups, all studying online exclusively, one with one instructor and the three others with another instructor. In the culture course there were four groups, two with one instructor and two with another. They all studied online exclusively. In the language proficiency course there were three groups. Two studied online exclusively and the third studied on campus but used some online tools such as the discussion forum. All three proficiency groups had the same teacher. The culture course comprised eight forums (96,222 words), the literature course comprised fifteen forums (144,901 words) and the proficiency course had a total of nine forums (110,403 words). In total, the corpus collected for this study consists of 351,526 words. For an overview of the forum statistics, see Appendix II.
The communication which takes place in the discussion forums is asynchronous, message by message transmission. The message buffer can sometimes be a factor with regard to how long a text remains accessible. Sometimes in instant messaging there may be a restriction in the number of characters an individual message can manage. This restriction has not had an effect on the messages in the present corpus material as the character limit was relatively high. The character limit for each message was 65,532 characters and anything written after that disappears. This number of characters can be said to be the equivalent of around nineteen A4 pages and none of the messages in the present study exceeded this length or even came close to that limit. In fact, no message exceeded 7,000 characters and the average length was closer to 1,000 characters. It is also possible to put pictures, formulae, diagrams and website URL addresses in the messages. Of these options, the participants only used URLs as background reference to support an opinion they expressed. After a message has been posted it is possible for a participant to edit or remove their own message as long as nobody has commented on it. After it has been commented on, it cannot be changed or removed. Another affordance of the discussion forum software is that instructors are able to edit and remove their own and others’ messages. If an instructor edited another person’s comment and reposted it, it would appear to readers that the teacher was responsible for the entire message. That is, it is not possible to see which part was originally written by the student and which was changed by the instructor. However, this turned out not be a problem, since none of the instructors did this in the material used for the present study.

Participation in the discussion forums was mandatory for all students in all courses, as was participation in real-time seminars. The students were required to post a minimum number of contributions in the discussion forums for their respective courses and the minimum requirement was at least one thread-initiating message per forum and at least two responses to messages posted by their fellow students per forum. No upper limit was set and the length of the message was not specified. The students’ activity in the discussion forum contributed to their final grade for the course. All of the participants who studied online exclusively also met on a regular basis in their respective groups for synchronous online seminars in Marratech, a tool which allowed for seminars in real time using both audio and video. The students also had the possibility of meeting with each other in Marratech in their own time. The campus group of proficiency students attended real-time seminars on a regular basis at the university instead of online seminars.

3.1.2 Participants

The participants formed a homogeneous group of students in so far as the vast majority had Swedish as a first language and English as their second. In addition, their level of proficiency was high, that is B2 and above according to
the Common European Framework of Reference for Languages (CEFR). Two students had English as their native language. The students were all in their first or second term of university studies in English. There were altogether 98 individuals who contributed to the discussion forums for each course. Of these 98 individuals, 31 studied both literature and culture and of these 31, one individual was enrolled in the proficiency course as well. A further 14 students participated in literature only and another nine students in culture only. Of the remaining students, 25 studied proficiency online and 19 were enrolled in and participated in the campus proficiency course which also included online discussion forums. The literature and culture courses are normally studied in the first term of English while the proficiency course can only be studied by students who have previously taken another proficiency course.

The fact that the majority of the participants in the present study have English as a second or foreign language may have affected the way they communicated. According to Osman and Herring (2007), the effect a participant’s culture and language has in an online instruction environment should be taken into account. In one study of foreign students’ attitudes to asynchronous online learning, Zhao and McDougall (2008) found that some students believed that the lack of visual cues in asynchronous communication required a higher level of written competence. Others preferred asynchronous communication as they found it less stressful. This was partly because of the extra time for reading, thinking about and formulating ideas that asynchronous communication affords and partly because they found speaking and listening more difficult than reading and writing (Zhao & McDougall, 2008: 67). Nearly all of the respondents in the present study reported in the questionnaire that using the discussion forum gave them time to think about their ideas before writing and posting their contribution. Although some stated that they preferred to read what others had written before posting their opinion, the majority did not. It may be that individual preferences are just as important as the differences that can be attributed to culture or language.

The fact that the material was produced by non-native speakers of English is not the focus of the present study. One example of differences between native and non-native speakers when writing English was found in a study by Altenberg (1997) based on the Swedish component of the International Corpus of Learner English (ICLE). In a discussion of the results, Altenberg (1997) argues that Swedish students, like other foreign language learners, show a lack of awareness of register by not using the same degree of detachment that a native speaker might use when writing argumentative essays in English. On the other hand, Ådel (2008), who examined another corpus of English produced by Swedish students, argues that students producing written work under certain circumstances do not display the typical lack of register awareness that Altenberg (1997) observed. Ådel (2008) suggests that factors such as fewer time constraints and access to secondary sources affect students’ writing style. The written discussion forum interaction which constitutes the material for the
present study is produced in a situation where these two factors are present. Firstly, the students have access to secondary sources and as the communication is asynchronous, they have the opportunity to revise and edit what they write before posting. In addition, the writing is not a typically formal register that would be expected in an academic essay for instance. Online written discussions are more conversation-like with shorter stretches of language and turn-taking exchanges. Furthermore, there may be other factors to consider when communication is asynchronous and computer-mediated. One of them is computer literacy. Computer literacy not only includes students’ typing proficiency but also general knowledge of communicating via computers. The participants who answered the questionnaire all reported being regular computer users, and they reported that reading and writing emails and reading news and other items online were their two most prevalent online activities. More than half of them were involved in social networking as well and the majority also regarded themselves as “fast typists” and had also studied online previously.

Contributing to the forum is not normally possible for those who are not registered as students in the course. The only way it would be possible was if the instructor changed access rights to the forum. The default setting is that only those registered in a course room containing the discussion forums are able to contribute to the discussion forums. It is of course theoretically possible for a student to log in as another student or for non-students to log in as someone who is registered and write a contribution, but this would mean having access to that person’s log-in information. No students have reported this happening and there do not seem to be any unauthorised contributions. The students and instructors met in the same groups for seminars in real time, either online using software with audio and video or in a classroom on campus, and thus got to know each other quite well. Moreover, discussions that began in the discussion forum often continued in the classroom. This means that if someone else had contributed to the discussion forum as another student it should have possible for peers and instructors to recognise an outsider.

3.1.3 Data analysis

The forum material was saved as text files, which means that the forum-generated text such as the date, the sender as well as the labels on function buttons, such as reply, were also included in the material and counted in the total number of words. The technical affordances of the context also allow for filtering, which means that each individual’s aggregated contributions could easily be compiled into a text file for analysis. The literature course forums were the only ones with instructor input and as instructors’ messages made up a sizeable number of the contributions in these forums, they were examined separately. For the studies where frequencies were calculated, the software
WordSmith Tools\(^2\) was used. However, even for these studies a manual analysis of the results was always necessary to identify spelling and typing errors that might affect the results. It was also necessary to examine context to differentiate words such as then used as a temporal adverb (by then I realized) and as a linking adverb (if so, can we then understand), or there which can be a spatial adverb (Stevens works there) or a pronoun (There are differences). The manual analysis was also necessary to identify words and phrases that were quotes from course literature or where participants had copied and pasted others’ comments into their message when responding. In the proficiency course, for example, participants discussed examples of language use and it was therefore important to determine which words represented the participant expressing ideas and opinions and which words were the language items or examples under discussion. Cited texts and grammatical items such as personal pronouns that were the topic of discussion by participants were included in the total word count but these items were not analysed or counted as items used by participants in the analysis.

For the first study, which examined factors affecting response rates in the forum, the analysis was done in several steps. Firstly, a number of non-linguistic features that may have had an effect on the extent to which messages received responses were identified. These features included when and where in the thread the message was posted in the forum, how long the message was, who posted the message and if it was a message that started a new thread (thread-initiating) or if it was a reply to a previously posted message (response). Next, the actual messages were analysed to determine if certain types of messages or certain individuals’ messages might have elicited more responses than others. The model of analysis for type of message that was used was taken from a previous study by Mazzolini and Maddison (2007), who analysed instructor messages and their effects on responses but it was modified for the current study. The original model contained three categories: messages that contained questions, those that contained answers and messages that had a combination of both. The fourth category contained messages that had evaluative comments and the fifth group contained messages that could not be categorised into any of the previous four categories. As messages that contained direct questions and answers were rare in the material examined, the following new categories were devised: personal reactions (e.g. What an analysis you have done!), agreement (e.g. True), disagreement (e.g. I don’t think the message is that...), both disagreement and agreement combined (e.g. I agree with you both in saying that Stevens is looking forward to a new dawn, but I do not think that it means...) and finally a category for messages that did not fall into any of the first four categories.

The main focus of the second study was to determine in what ways students use discourse devices such as formulaic politeness expressions, humour, greetings and addressivity, support and agreement to build and maintain

\(^2\) Software designed by Mike Scott for finding patterns in text.
community and how this reflects their awareness of sense of community. The type and frequency of the devices were identified for each of the discussion forums. When analysing computer-mediated language, it is common to use the whole message as the unit of analysis and identify its purpose or function even though several functions may be performed by the one message. One way of solving this problem is to focus on the main function, as Collin (2005) does, for example. The disadvantage of using this method is that messages may have several equally important functions. The messages in the present study were not primarily posted to promote community but as part of learning activities. Thus using the message as the unit of analysis would not necessarily reveal anything about community building strategies. Many of the messages contained multiple devices for community building and as the aim of the study was to identify in what ways students use these devices, each device was categorised separately. A secondary aim of this study was to compare it with the results of a previous study (Lapadat, 2007) on community building devices used among tertiary students taking part in online discussions. Three coders were given the model from Lapadat's study and asked to identify and categorise the features listed. The model of analysis proved difficult to apply as the categories were not adequately defined or exemplified. Consequently, even though there was agreement that they did perform a social function, it was difficult for coders to agree on where some devices should be placed. However, some of Lapadat’s categories were not necessary for the purpose of the present study and so could be collapsed into fewer categories. For example, it might be not necessary to differentiate between social remarks, asides and disclosure, which could all be considered social remarks which personalised content and contributed to the sense of community. Therefore, in the present study these three categories were collapsed into one. Distinguishing between inviting comment and requesting help also seemed unnecessary. The difference lies in how these were formulated but the purpose of requesting or inviting comment was basically the same. Coding of formulaic politeness, greetings and addressivity, emoticons, alignment and inclusive language resulted in high agreement among coders (see Table 3 for examples of the categories used).
Table 3.1. Categories and examples of community building features from Case Study 2

<table>
<thead>
<tr>
<th>Categories</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Greetings &amp; addressivity</td>
<td>Hi everyone, Hello</td>
</tr>
<tr>
<td>Social remarks, disclosure, asides</td>
<td>Remember from school how our English teacher repeatedly brought up these questions</td>
</tr>
<tr>
<td>Request for help, invite comment</td>
<td>Can we use this kind of things in academic writing, or any written things?</td>
</tr>
<tr>
<td>Support</td>
<td>I like your example about snow :)</td>
</tr>
<tr>
<td>Humour</td>
<td>To the rescue! Haha, no not really</td>
</tr>
<tr>
<td>Inclusive language</td>
<td>We should use the future perfect</td>
</tr>
<tr>
<td>Alignment</td>
<td>I totally agree, Same here</td>
</tr>
<tr>
<td>Features from others genres</td>
<td>Dear NN, Regards,</td>
</tr>
<tr>
<td>Formulaic politeness</td>
<td>Thanks, please, sorry</td>
</tr>
<tr>
<td>Emoticons</td>
<td>=), 😊, 😆</td>
</tr>
</tbody>
</table>

The focus of the third study was on the content of the subject lines of the student messages and student preferences for subject line content as well as the function of the different types used. The subject line content was categorised according to four inductively identified categories and a fifth category was added for those subject lines which could not be placed in the four other types. The content of the first category, which was divided into three subcategories (see Table 3.2), reflected message content by referring to the assignment or question number or name of the assignment or book. Also included were subject lines that contained an original title which reflected the content of the actual message. The second category described is the type or function of the message, such as answer or reply, while the third type of subject line contained a greeting or addressed another person. The fourth type included subject lines that contained stance or evaluative feedback. There were other types that were identified but they were much less frequent and therefore these were placed in the category ‘other’. The different categories used for Case Study 3 are illustrated in Table 3.2.
Table 3.2. Categories and examples of subject line types from Case Study 3

<table>
<thead>
<tr>
<th>Type</th>
<th>Examples from case study material</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type 1a</td>
<td>Topic</td>
</tr>
<tr>
<td>Type 1b</td>
<td>Numbers</td>
</tr>
<tr>
<td>Type 1c</td>
<td>Own title</td>
</tr>
<tr>
<td>Type 2</td>
<td>Label</td>
</tr>
<tr>
<td>Type 3a</td>
<td>Greeting &amp; addressee</td>
</tr>
<tr>
<td>Type 3b</td>
<td>Address no greeting</td>
</tr>
<tr>
<td>Type 3c</td>
<td>Greeting no addressee</td>
</tr>
<tr>
<td>Type 4a</td>
<td>Stance</td>
</tr>
<tr>
<td>Type 4b</td>
<td>Evaluative comment</td>
</tr>
<tr>
<td>Type 5</td>
<td>Other</td>
</tr>
</tbody>
</table>

Of the four studies, the final one relied on quantitative results generated by using concordance software. Using this kind of software makes patterns of usage more visible in collections of electronic texts. The software WordSmith Tools was used to determine frequencies and distributions of a number of features used for referring. The features examined included those showing reference to the self, such as the use of the first person pronouns, reference to others, such as the use of the second person pronouns, as well as referencing to the surroundings outside the learning environment and within the context of the learning environment. The results of the software search were examined manually to determine that the identified items were not typing or spelling errors. In addition, quotes in the text were excluded as they were not the participants’ own words. Next, the results of the searches were calculated to give frequencies per thousand words. The tokens were then compared to the frequencies of the same items in Biber et al.’s (1999) analysis of conversational and academic English. To further inform the analysis, patterns of usage from an oral seminar from MICASE were investigated. The seminar only comprised a relatively small number of words compared to the forum material but this particular seminar was chosen because it was in a similar subject area as the forum material, that is, English proficiency. In addition, students working together and commenting on each others’ written work is a similar activity to that which takes place in the forum discussions. As Biber et al. gave frequencies per million words, their figures were divided by 1,000 in order compare with the frequencies in the present study. The frequencies in the MICASE seminar were also calculated per 1,000 words.

By using a number of different methods of analysis, it was hoped that each study would shed light on different aspects of the communication taking place among students and instructors in this online learning environment.

3.1.4 Questionnaire

Even though it is the communication itself that is the focus of the four studies, the communication takes place in a particular setting and information about the
context and participants is useful for a discussion of the results of the analysis of the empirical data as discussed in Section 3.1.3. In order to provide more information on the context, a questionnaire was sent to the participants after they had completed their studies (see Appendix I). The response rate was low (26 of 98) but still provided some useful background information. The purpose of the questionnaire was to find out more about the participants with regard to such things as their use of computers, experience of learning online and previous academic studies. The questionnaire was distributed as an electronic link on two separate occasions. With 26 respondents, the response rate was a little more than one quarter of the 98 students. The low response rate may be due to the fact that there was a two-year time lapse between when they participated in the course and when the questionnaire was distributed. Although it is not possible to generalize with so few respondents, the results of the questionnaire did provide useful background information and revealed some trends.

The informants were asked whether they felt that communicating in English instead of their native language (Swedish for the majority) made them less likely to contribute and whether they read and edited their contribution before posting. Participants were also asked to state what they believed the function of the subject line was and whether they considered it important. In addition, they were asked to describe their forum reading and writing habits. This included stating which messages they read first on opening the forum and what kinds of messages prompted them to write a response. Finally, participants were asked to state whether they found it important to be kind and friendly to their fellow students in the forum. This question related specifically to the second case study on community building.

The information and opinions provided by the respondents in the questionnaire are used as additional information for three of the four studies in the present thesis. Not all the information in the questionnaire was relevant to all of the studies but some of it was useful for some of the studies. For example, respondents’ reported behaviour on which messages they read and responded to was useful for the first study. The questionnaire information was particularly useful in the third case study concerning subject lines choices. Respondents reported that they believed the subject line was important and that they did make conscious choices. This information could not have been extracted from examining the empirical data alone.

3.2 Ethical considerations

In all studies where people are involved, questions need to be asked and answered by the researcher in order to ensure that basic human rights are not encroached upon. These concerns must be weighed together with the potential benefits that may come as a result of the study. According to the guidelines set out by the Association of Internet Researchers (AoIR), the primary ethical
obligation is to do no harm. Ethical guidelines are well established for social science researchers who collect material in offline situations, and these guidelines have had a strong influence on the guidelines that are used for online research (Fielding, Lee, & Blank, 2008: 24). Research that gathers primary material from an online environment is not by definition more likely to be harmful than research where material is collected offline, but it does present a number of challenges for the researcher, such as determining what constitutes a private act on the Internet (Fielding et al., 2008: 27). The AoIR guidelines recommend that when making ethical decisions concerning online material, researchers should be aware of the particular context. For example, if the content does not address sensitive topics, then the risk to the participant is most likely low. In addition, if linguistic form rather than the content is the focus of the study, this may also reduce the risk to the subject. The AoIR guidelines place an emphasis on this context-dependence, which means respecting people’s values or expectations in different settings. The problem is trying to assess what risk the researcher puts the participants at as it is not always clear what the consequences might be.

It may seem reasonably easy to distinguish between public and private on the Internet by categorising environments into those with restricted access and those with open access. Restricted access sites are those environments that are private and not intended for a general audience. These may include internal emails in an organisation or discussion groups where participants need permission to have access. The open access or public sites are those that anyone may access and are meant to be read by a wide audience. As the number and the types of electronically stored data have increased, as well as the number of users, this distinction between private and public is not always clear. Just because a site accessible for the general public does not automatically mean that it can be used as the object of research without permission from participants. Some would argue that users are aware of the public nature of the site and that if they participate in a public discussion forum they should be prepared for the possibility that it could be used for research purposes. However, there are examples where applying this reasoning may cause harm. One example is an online discussion forum used for self-help groups. Such groups are often open access and accordingly can be defined as public. Usually, they are easily accessible so that those who rely on them for support can easily and anonymously get help; however, for researchers to assume that they are able to use them without permission can sometimes result in negative consequences for members of these groups. Eyesenbach and Till (2001) give one example of a case where some members of a breast cancer support group withdrew from participation when they found out that they were being observed by researchers. If the presence of a researcher causes individuals to withdraw from participating and as a result the individual does not receive the help they need, then the primary obligation of a researcher to do no harm has been neglected.
Another example of the unclear division between private and public is the so-called Enron emails. As part of a federal investigation into this energy company in the United States, all correspondence between executives working for this company was made available on the Internet. According to the laws of the United States, evidence in such cases is made public. The entire body of email correspondence became accessible to anyone with Internet access. It became apparent that not only were those who were the target of the investigation exposed, but also their families and friends who had no role in the company or the investigation. Consequently, the body of correspondence was removed from public access and names and content that did not pertain to the investigation were removed (Klimt & Yang, 2004). What is clear is that careful consideration must be taken when publishing, storing and using electronic data so that the identity and privacy of individuals is protected. For the researcher it means storing data in a way that unauthorised people cannot gain access to it. Moreover, the researcher must carefully consider what quotes to use in published work, as these may be located and the participant identified using search engines, which means putting participants at risk.

Obtaining permission from the participants is not always an easy task and particularly in an open forum it can be problematic, as it is not always possible to determine the identity of the participants. If researchers only have contact with potential participants online and there is no visual contact, they only have the word of the participant to rely on. Even if it is possible to establish that participants are who they say they are, researchers must also assess a participant’s ability to give informed consent. For example, the material to be researched may be produced by participants with a diminished capacity to judge the consequences of consenting to participation. In the present study, this has not been difficult as it has been possible to determine the identity of the participants and contact information for all participants was available. Informed consent means that participants must be able to show they understand the general goals of the research as well as what they are agreeing to when they give consent. They must also understand the potential risks as well as the benefits of taking part in a study and they must also understand that they have the option of withdrawing from the study at any time and having their material removed from the study. The material for the present study consisted of interaction produced by students in a learning environment with restricted access which according to the AoIR recommendations should be considered private. Therefore it was necessary to contact the students and obtain their permission. For the present study it was possible to identify and contact the participants as they were students who had been registered for courses at the higher education

---

institution. It was possible to provide them with the necessary information about the goals of the present study before obtaining their permission. They were informed that their identity would not be revealed and that they had the right to withdraw at any time.

In the current study, the ethical issues have been addressed. In order to minimise the risk of participants being identified steps were taken to store the material in a way that unauthorised people could not access it. The material was produced in an online learning environment at tertiary level and only instructors and certain administrators employed at that higher education institution and the students enrolled in those particular courses that specific semester were able to access the material. It is also possible for these people to access the material after the course has ended. Therefore, by request of the researcher system administrators at the university restricted access to the material and it is no longer accessible to anyone except the researcher and system administrators. In addition, search engines are not able to access this material. Permission has been obtained from both the students and the instructors involved in the courses to use the material for this particular study. Permission to use the material was obtained after the course was completed, which means that at the time when the communication took place, neither the instructors nor the students were aware that the material might be used for research. This means that they could not have altered their behaviour because they felt that they were being observed.
4 Summary of case studies

The overall aim of the current study is to examine language and interaction in one type of communicative event, namely the asynchronous discussions that take place in a university learning environment. These may be considered communicative events in that the language is both significant and indispensable (Swales, 1990: 45). The discussions are also assumed to have a shared set of features and a shared purpose. Overall, the empirical investigation and observations show the role of language and that the choices made by those involved in the interaction serve a number of purposes. Case Study 1 focuses on features of messages in the discussion forums that prompt others to read the message and which can elicit responses, while Case Study 2 takes a closer look at linguistic strategies used for building and maintaining a sense of community. Case Study 3 examines one particular aspect of the messages in the asynchronous discussions, namely the subject line, which in itself performs specific functions. Finally, Case Study 4 investigates the frequency of a number of linguistic features used in the online discussion forums and compares them with the frequencies in two offline registers, namely oral conversation and academic discourse. In order to further inform Case Study 4, the frequencies are compared to a small corpus of a transcribed oral academic seminar in the same subject area and task type.

4.1 Case Study 1: Getting a response to discussion thread messages in an online learning environment

The aim of the first study was to examine factors that may affect the choices made by students and instructors regarding which messages they respond to. The content of the responses was examined to see if it is possible to determine what it is that prompts a response, that is, what reaction the previous message evokes and what kinds of thread-initiating messages receive a response. The material consisted of text files from 14 literature discussion forums comprising 815 messages and more than 140,000 words in total. Of these 815 messages, 168 were posted by participants to start a discussion thread in these forums, and 138 of them actually resulted in a thread by generating at least one response from another participant. Of these 138 messages, 63 continued into a longer thread, that is, these 63 threads each had a response message which in turn received its own response message. Figure 4.1 is a screenshot of a part of one thread that has many responses.
Despite the fact that the instructions for posting messages, the preparation for the forum discussions and the minimum requirements for the number of messages to be posted were the same for the different groups, the number of responses varied and certain messages and threads attracted more responses than others. In order to determine if it was the content of the thread-initiating message or other factors which affected the rate of response, it was necessary to look at a number of other factors. The forums with the most messages were also the ones that had the highest number of response messages per participant. At the same time, the messages tended to be shorter than the messages in the forums for the smaller groups. In the latter forums, where there were fewer students and fewer messages in total, the number of responses per participant was also lower. It is difficult to determine whether this was a general trend or whether it can be attributed to the individuals in the different groups. More research is needed on the effect of group size and group dynamics and how these might affect the quality and quantity of messages posted in discussion forums.

Participants were able to have some freedom of choice in and control over what they wrote in their messages and which messages they chose to respond to. Both student and instructor messages were included in the study to see if instructors’ messages in themselves somehow prompted more or fewer responses. Both the length of the messages and the time when the messages were posted were examined. Long messages take longer to read so with limited time, students may choose a shorter message to read and respond to. Posting a message early seems to have been the most decisive factor whether or not a message generated a response. This may have to do with the fact that messages are displayed in chronological order on the screen, and early messages would appear at the top of the screen. According to a number of respondents, this is where they began to read when they opened the forum. However, it was not
always the message that was posted first in the forum that developed into the most complex thread with the most responses. It was initial messages that contained explanations or interpretations that aroused memories or reminded others of their own personal experiences that grew into more complex threads.

Generally, it can hence be concluded that the messages that resulted in the most responses were the ones where the content seemed to arouse an emotional response. It is not possible to predict from an original message if it will generate a response or not, as we do not know the personal histories of participants so we cannot predict what messages will trigger memories and emotional reactions. It is in the response that it is possible to see a personal reaction to the original message. The prototypical response message would reflect either a positive or negative emotion, but most often positive. Such evaluative feedback was found to contain positive adjectives referring to what a previous participant had written, as well as a fair number of mental verbs expressing enjoyment and pleasure. Even when participants did not agree with what another participant had written, they normally began with positive feedback or supportive comments in their response before disagreeing. Being positive and supportive was not something the students were specifically instructed to do for these courses. Their supportive attitude could be interpreted as a signal that the students found it important to maintain good social relations with others in their group. This is supported by the questionnaire results, where more than half of the respondents answered that they felt this was important. Direct questions from both instructors and students were rare and did not necessarily generate any responses.

Another interesting issue is whether or not the messages written by the instructor might affect the number of response messages. The forum contributions of the two instructors who taught the literature course varied in length. One instructor contributed messages of about 80 words in length in the forums, while the other instructor’s average message length was 136 words. However, the instructor whose messages were shorter posted messages more frequently than the other instructor. The students who contributed to the forums with the instructor who posted longer messages tended to produce much longer thread-initiating messages and somewhat longer responses than the students whose instructor wrote shorter messages. Thread-initiating messages were longer than response messages in all forums, but the thread-initiating messages in the group where the instructor wrote shorter messages averaged between 295 and 565 words, while for the other instructor’s forums the thread-initiating messages averaged between 509 and 677 words. The trends observed in this study indicate that the students may model some of their messages on those of their instructor with regard to type and length. Hence, if the instructor posts short messages then the students tend to do so as well and if the instructor tends to post longer messages then the students tend to post longer messages too. For the forum where there were many participants, the length of messages may also have been affected by the volume of messages. In
this study, as in previous studies, it appeared that when the volume of messages in a forum becomes very large, the length of messages tends to be shorter. This trend follows what was found in previous studies, namely that one strategy participants use is that when there are many messages and a lot of text, they write shorter messages. However, this was not true for all forums in the study. In some of the forums where there were fewer participants and fewer messages, the messages were short. It is likely that other factors such as self-confidence, group dynamics and topic also affect message length.

The instructors gave similar types of feedback to the students, most often evaluative feedback. Most of the instructor messages that received responses were from the students they had given feedback to and not from other students. These threads could be likened to a two-party dialogue rather than a multiparty conversation, unlike other threads where responses were student-initiated. In this particular study it was more common for threads to develop where fellow students had given responses rather than the instructor. This means that instructor input does not appear to have had a noticeable effect on increasing student input in the discussion forums. Apart from other factors, such as when a message is posted, responses tended to be generated by messages that provoked an emotion of some sort in the reader. These emotions were overwhelmingly positive and concerned personal experiences and memories as well as gratefulness for others’ opinions and explanations, which students felt had been helpful in their understanding of the course material. What instructors might learn from this is that careful choice of topic as well as clear instructions for forum activities which require students to describe and explain in their own words as well as comment on other students’ contributions may result in participants contributing more actively to discussion forums.

4.2 Case Study 2: Community building from a distance

In the second study, the material used was from a course in English proficiency. The discussion forum material was taken from three different groups of students studying the same course with the same instructor. Two groups (P1 and P2) studied exclusively online and met online for seminars, never on campus, while the third group (P3) studied on campus but used the online discussion forum to complement their seminars. Group P1 comprised nine participants, group P2, fifteen and group P3, fourteen. The material from the three forums totaled 19,497 words and comprised 155 messages. The average message length for all three groups was similar and ranged between 115 and 146 words. The majority of students in groups P1 and P2 had Swedish as their first language and English as their second; two students in group P2 had English as their first language and there were also a few students in these two groups who had another language than Swedish as their first language. In group P3, five of the fourteen students had Swedish as their first language and the other nine were speakers of Spanish, Russian, Chinese, German or Turkish.
The main aim of this study was to investigate how students build and maintain social relationships in an asynchronous learning environment. This was done by identifying what social functions various expressions could be said to take on in the student messages (see Table 4.2 for an overview of the categories that were used). Unlike previous studies where the message was the unit of analysis and categorized according to primary function (see e.g. Hård af Segerstad, 2002; Rafaeli & Sudweeks, 1997), the present study examined each message and identified any expression that could be said to take on a social function. Expression here refers to single words, phrases, or emoticons, i.e., visual representations of facial expressions. As a result, most messages were taken to be multifunctional, as they were comprised of various numbers of such expressions.

This study used a modified version of a model of discourse devices which take on a social function, suggested by Lapadat (2007). The model had to be adapted to the present study, partly because some devices were not clearly defined in the previous study, but also because some distinctions between categories in Lapadat’s model were not useful for this study. Lapadat’s fifteen categories were thus reduced to thirteen. Two categories were added to Lapadat’s, as they were not included in the original model. They were categories 14 and 15, i.e., formulaic politeness expressions and emoticons. Both of these perform social functions and were therefore included in this study (see Table 4.1).

Table 4.1. Categories used in the study of community building devices (modified from Lapadat’s 2007 model)

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Greetings and addressivity</td>
</tr>
<tr>
<td>2</td>
<td>Outside social interaction</td>
</tr>
<tr>
<td>3</td>
<td>Colloquialisms</td>
</tr>
<tr>
<td>4</td>
<td>Social remarks</td>
</tr>
<tr>
<td>5</td>
<td>Disclosure</td>
</tr>
<tr>
<td>6</td>
<td>Asides</td>
</tr>
<tr>
<td>7</td>
<td>Help</td>
</tr>
<tr>
<td>8</td>
<td>Support</td>
</tr>
<tr>
<td>9</td>
<td>Humour</td>
</tr>
<tr>
<td>10</td>
<td>Invitation to comment</td>
</tr>
<tr>
<td>11</td>
<td>Inclusive language</td>
</tr>
<tr>
<td>12</td>
<td>Alignment</td>
</tr>
<tr>
<td>13</td>
<td>Features from others genres</td>
</tr>
<tr>
<td>14</td>
<td>Formulaic politeness</td>
</tr>
<tr>
<td>15</td>
<td>Emoticons</td>
</tr>
<tr>
<td></td>
<td>Formulaic greetings and naming</td>
</tr>
<tr>
<td></td>
<td>Comments not connected to the course content</td>
</tr>
<tr>
<td></td>
<td>In-group terms coined by participants, jargon</td>
</tr>
<tr>
<td></td>
<td>Any remarks with social intent</td>
</tr>
<tr>
<td></td>
<td>Remarks that personalize topic, relate to self</td>
</tr>
<tr>
<td></td>
<td>Anecdotes revealing personal aspect</td>
</tr>
<tr>
<td></td>
<td>Explicit requests for or offers of help</td>
</tr>
<tr>
<td></td>
<td>Praise, support of other participants</td>
</tr>
<tr>
<td></td>
<td>Any type of humour</td>
</tr>
<tr>
<td></td>
<td>Inviting others to comment on ideas</td>
</tr>
<tr>
<td></td>
<td>First person plural pronouns we (as a group)</td>
</tr>
<tr>
<td></td>
<td>Aligning or agreeing with other participants</td>
</tr>
<tr>
<td></td>
<td>e.g. Salutations, signing off (from letter writing)</td>
</tr>
<tr>
<td></td>
<td>e.g. Thanks, please, sorry</td>
</tr>
<tr>
<td></td>
<td>e.g. =), 😊😊😊</td>
</tr>
</tbody>
</table>
The results demonstrate that requests for help and invitations to comment (categories 7 and 10) were found quite often in the messages in all three discussion forums. This is not very surprising, since the task given to the students in this course was to collaborate with fellow students and to discuss language issues. Hence, the use of many devices may have been the direct result of the task given rather than of the participants trying to achieve a sense of community.

Community building devices were used differently by the three groups. For instance, groups P1 and P2, the two groups which studied online only, tended to use greetings and address one another by name to a greater extent than group P3, who met one another on campus. There was also a tendency for groups P1 and P2 to use politeness expressions more frequently than the campus group. Overall, it was observed that the two online groups, P1 and P2, used more and a greater variety of community building strategies than group P3. The differences might be ascribed to the fact that these two groups did not meet their peers physically. However, when the two online groups were compared, it turned out that group P2 employed a wider variety of devices and also used them more frequently than the other two groups, which suggests that it was not only the presence versus the absence of physical meetings which might explain the different usage of all devices. For instance, the students in group P2 had a higher proportion of disclosure (category 5 in Table 4.2) than the other two groups. When it comes to asides, which, like disclosures, is a way of personalizing the content of the message, they were used the most frequently by group P2, less often by group P1, and not at all by the students in group P3. Examples of supportive language (category 8) were only found in the communication between students in group P2. They also used alignment as well as features from other genres more often than the other two groups. Emoticons were not used frequently by any group, but in group P2, they were sometimes used in combination with other community-building devices, such as requests or offers for help.

In the forums, students were required by the course instructors to post a minimum of three messages, of which at least one should be a thread-initiating message and two should be responses to other students’ posts. Although there was individual variation, there was a tendency for group P2 as a whole to post messages and responses more often. Hardly anyone in this group posted fewer than the required number. On average, the participants in group P2 posted more than five messages per person, whereas the participants in the other two groups posted about three messages on average. This perhaps relates to one observation made in the first case study which revealed that messages which caused an emotional reaction in the reader tended to get responses. In the discussion forum of group P2 where there are more examples of personalizing content, there is also more interaction. Such personal content along with
expressions of gratefulness and supportive feedback indicate student involvement.

Overall, group P3 made less use of community building devices than groups P1 and P2. Some of the differences between group P3 and the other two groups might be attributed to the fact that they met on campus whereas the others did not. It must be remembered also that the students in group P3 had more varied language backgrounds than the students in the online groups, and this may have affected the results. The fact that there were differences between the frequency and variety of community building devices used by the two online groups as well indicates a need for more research into factors that may affect community building strategies in online asynchronous communication. The significance of the social aspects of language in the learning situation has been discussed by Wenger (1998) and Coupland (2003) among others.

Interestingly, it was the group that used a larger variety of community building devices and also more frequently than the other two groups that had the highest proportion of students who received a high grade in the proficiency course. How learning outcomes, language proficiency and the use of community building devices are related to one another requires further research.

4.3 Case Study 3: Use of subject line and other factors contributing to coherence and interaction in student discussion forums

The aim of the third study was to examine in what way the subject line content of messages in an asynchronous discussion forum contributes to interaction. More specifically, this study investigated whether different types of subject lines can be seen to affect which messages are read and also what different functions they fill in the interaction, for example, social functions or discourse organising functions. The study also examined student preferences for subject line type. Other factors which might affect whether certain messages are read more than others were also examined. These factors were group size, volume of messages and the position where a message was displayed on the computer screen. The material consisted of the subject lines from the first nine discussion forums for the literature, culture and proficiency courses. Two forums were from the literature course, four from the culture course and three from the proficiency course. The material was produced by 98 students enrolled in the three courses and comprised 435 messages in total. In addition to the analysis of the 435 subject lines, two of the questions from the student questionnaire (see chapter 3.1.4) were of particular relevance for the present study. Question number 7 asked participants about the importance of the subject line as well as what they thought its function was, and question 9 asked about which messages they chose to read first. Other questions also informed this case study to a certain degree, by dealing with aspects of communication such as linking messages and treating fellow students kindly, among other things.
The messages were first divided into two groups depending on whether they initiated a new thread or were a response to a previously posted message. After that, the subject lines for each message were categorised according to their perceived function. One of the challenges of this particular study was identifying and describing the categories, as the only previous studies found that examined subject lines were for email correspondence. The categories were arrived at inductively based on the perceived functions they performed and how these functions were performed linguistically. The ten types are illustrated in Table 3.2, repeated here for convenience as 4.2.

**Table 4.2. Subject line types and examples**

<table>
<thead>
<tr>
<th>Type</th>
<th>Examples from case study material</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type 1a</td>
<td>Topic</td>
</tr>
<tr>
<td></td>
<td><em>The Remains of the Day</em></td>
</tr>
<tr>
<td>Type 1b</td>
<td>Numbers</td>
</tr>
<tr>
<td></td>
<td><em>Assignment 2, Questions 2 &amp; 7</em></td>
</tr>
<tr>
<td>Type 1c</td>
<td>Own title</td>
</tr>
<tr>
<td></td>
<td><em>We are all butlers, A love story</em></td>
</tr>
<tr>
<td>Type 2</td>
<td>Label</td>
</tr>
<tr>
<td></td>
<td><em>My questions, comments</em></td>
</tr>
<tr>
<td>Type 3a</td>
<td>Greeting &amp; addressee</td>
</tr>
<tr>
<td></td>
<td><em>Hi NN!</em></td>
</tr>
<tr>
<td>Type 3b</td>
<td>Addressee no greeting</td>
</tr>
<tr>
<td></td>
<td><em>NN, To NN</em></td>
</tr>
<tr>
<td>Type 3c</td>
<td>Greeting no addressee</td>
</tr>
<tr>
<td></td>
<td><em>Hi</em></td>
</tr>
<tr>
<td>Type 4a</td>
<td>Stance</td>
</tr>
<tr>
<td></td>
<td><em>I agree, I think that...</em></td>
</tr>
<tr>
<td>Type 4b</td>
<td>Evaluative comment</td>
</tr>
<tr>
<td></td>
<td><em>Excellent, Good interpretation there</em></td>
</tr>
<tr>
<td>Type 5</td>
<td>Other</td>
</tr>
<tr>
<td></td>
<td><em>Sorry I’m late!</em></td>
</tr>
</tbody>
</table>

The first two types have a discourse organising function. Type 1, which has three subcategories, includes subject lines that refer to the content of the message, either by using the topic or title given, using numbers or an appropriate topic made up by the participant. Type 2 comprises subject lines that describe the function of the message by using a label such as *question*. Type 3 is made up of subject lines that greet and address another participant in the forum. Greetings and addressivity generally have a social function, but in an asynchronous online discussion forum, addressivity also aids coherence by connecting the message to the message of another participant and can thus also be said to have a discourse organising function. Type 3a, which were subject lines containing both greeting and addressivity, can be said to have both a social and a discourse organising function. Type 3b contained no greeting and therefore had more of a discourse organising function, while Type 3c, which had a greeting but no addressee had a purely social function. Occasionally, the subject line greeted and addressed the whole group as in *Hi everyone.*
Type 4a subject lines were those which revealed what the student thought about the topic under discussion (stance). These subject lines had more of a discourse organising function by suggesting what the content of the message might be. However, a subject line such as I agree shows alignment and might also be said to have a social function. Type 4b subject lines showed the opinion of the student and contained positive evaluative comments. As there were no negative evaluative comments in the material it seems that giving positive evaluation may be important for the participants in this study. Type 5 subject lines are those that could not be assigned to any of the first four types. There were very few messages in this category, and a number of the subject lines were blank. For obvious reasons, the latter could not be assigned a function. One study of email subject lines suggests that an empty subject line might arouse curiosity so that a person would be more likely to open the message with a blank subject line than certain other types (Porter & Whitcomb, 2005: 284). As there were so few cases of blank subject lines in the present study, it is not possible to say what effect they had on whether or not the message was read. The other examples categorised as Type 5 were subject lines that were questions or apologies.

The analysis of the material revealed that the function of the subject line tended to differ depending on whether the message was thread-initiating or a response. The analysis of thread-initiating messages revealed a preference for Types 1 and 2 subject lines. For example, of the 119 thread-initiating messages, 91 subject lines were of Type 1, that is, subject lines that reflected the content of the message by referring to the task topic in some way. Another 20 subject lines were of Type 2, that is, labelling message as a question or a comment; this type was only used by the proficiency groups. Of the three proficiency groups, one used only Type 2, while the other two displayed more variation. One of the literature groups used only type, Type 1c, while the other literature group varied more in their use of the subject line. There were also clear preferences for one type of subject line in most student groups, with the exception of one proficiency group where there were two preferred types. This proficiency group was also the group that in Case Study 2 used the most community building devices. Group preferences might be partly explained by the forum tasks, as only students in the proficiency course used Type 2 subject lines in their messages. However, it is noteworthy that group preferences tended to develop despite identical course content and forum tasks. The type of subject line used early in a forum tended to be the type used for the majority of messages in that particular forum.

Subject lines containing greetings and addressivity, Type 3, were not used in any of the subject lines of thread-initiating messages. In contrast, Type 3 subject lines were the ones that were used most frequently for response messages. This trend also connects this case study to Case Study 2, which focussed on community building strategies. The frequent use of a subject line that has a social function may indicate that maintaining good social relations with other
students is so important to the participants that this is done in the subject line as well as in the message. Stance was the next most frequently used subject line for response messages, closely followed by students’ own title and labelling. Overall, the choice of subject line varied more for response messages than for thread-initiating messages.

In addition to identifying the function of the subject line, this study aimed to determine whether certain types of subject lines would encourage others to open and read a message. Based on the recorded statistics for the number of students who read each message in the forum, it appeared that the subject line alone could not predict how many would open and read a message. In the survey, the students reported that they believed the subject line to be important and that they had made strategic choices both about what to write in their subject line and about whether to read a message or not based on the subject line content. The quantitative analysis, however, indicated that there were other factors that affected how many students read a message. One of these was when a message was posted in the forum. Messages were organised chronologically by the learning management system, which means that those posted first always appeared at the top of the computer screen. Messages posted early in the forum were also more likely to be read. Messages posted later risked not being visible when students logged in, and students would consciously have to scroll down the page to see the message and its subject line. This relates to Case Study 1, which revealed a similar trend. Another factor affecting whether a message was accessed and recorded as read was the number of students in the group and the volume of messages. A discussion forum with a large number of students participating is likely to produce more messages and this may also affect how many participants read the messages. Many messages or a lot of text means that participants may not have time to read them all. This trend is shown in Figure 2 in the published article for Case Study 3.

The results of Case Study 3 may not necessarily reflect trends in general. However, they do indicate that there are different ways in which subject lines can contribute to coherence. The results also show aiding coherence is not the only function of the subject line. Particularly for response messages, the most frequently used subject line types across all messages examined were those that had a social function. There were group preferences that developed almost immediately, so that the subject lines preferences of those who posted the first messages had a normative effect on those posted subsequently. The impact of this for teaching is that it is possible that instructors as well as students can influence linguistic choices by posting their preferred type early in the discussion forum. There is a clear tendency for differences to be greater between groups than within groups. More research is needed into the way group norms develop and the effect they have on linguistic choices and communication strategies in asynchronous online learning environments; however, there does seem to be a tendency for individuals in the different groups to align with regard to the types of subject lines they use. The
preference for using subject lines that have a social function indicates that, at least for these particular students, the social side of learning is important. Both Case Study 2 and 3 show a need for more research concerning the social aspects of learning online and their significance in creating a favourable learning environment.

4.4 Case Study 4: Features of orality, academic writing and interaction in asynchronic student discussion forums

While the first three studies focused mainly on what drives the interaction, the fourth study focuses specific linguistic aspects of the interaction. In particular, this study compares the frequency and usage of selected linguistic features to other spoken and written registers. Previous studies argue that synchronous or almost synchronous online written CMC is more conversation-like than asynchronic written CMC and this type of communication has been described as 'conversational writing' (Jonsson, 2013). The reason for examining linguistic features in a body of text is that systematic variation of linguistic features reflects the main function of a text and can provide information about a particular register which is related to the situation. These situational differences can reflect the purpose of the text, the circumstances under which the text is produced as well as the degree of interactivity (Biber et al., 1999: 11). The main aim of the study was to examine how written asynchronous interaction in the discussion forums relates to spoken registers on the one hand (a general oral corpus and an oral academic seminar) and written academic prose on the other hand with respect to a number of linguistic features whose frequencies are associated with interaction. Secondly, the frequencies of these selected features were examined as a reflection of how students interact with one another, with the course content, and with the interaction itself. The study was not intended to give a complete picture of the level of orality or literacy in the discussion forums but instead to investigate how the frequency and usage of these features compare to other oral and written registers both within and outside an academic environment. The tendencies observed in this study suggest that this particular asynchronous type of written CMC leans towards conversation in some respects but also has tendencies observed in academic prose.

A quantitative approach was used where the frequencies of a limited number of linguistic features were singled out. For this study all the material collected from the three courses and 32 discussion forums was used (see Appendix II for an overview). There were 1,446 messages in total, and after allowances were made for automatically generated text, the corpus totalled 337,370 words. The frequencies of the selected features in this specific register of online communication were compared to the frequencies of other registers. Firstly, a comparison was made to the frequencies of these features in two larger corpora - oral conversation and academic prose - part of the Longman Corpus of
Spoken and Written English. The oral conversation corpus consists of 6,410,300 words and the academic prose corpus, 5,331,300. In addition to these two corpora, the frequencies of these items in a transcribed English composition seminar from the MICASE were also included to further inform the study. After the transcription annotations had been removed the oral seminar consisted of 19,942 words. The MICASE material does not represent a general register like the Longman corpora but is used as an indication of how context and content may affect linguistic choices.

Three sets of features were examined in this study. They were chosen for a number of reasons. Firstly, all words examined belong to the closed-class group of words. That is, they are function words, they occur frequently and they are used in all types of text as opposed to words from the open-class groups such as verbs and nouns whose frequencies vary greatly (Biber et al., 1999: 55). Secondly, these particular features “reflect the fact that the speaker and listener typically interact with one another while reader and writer do not” (Biber, 1988: 43). As the discussion forum material is written but still contains short conversation-like exchanges, it is interesting to see if the frequencies indicate that the interaction found in these forums leans towards oral registers. For example, if the participants are interacting with one another in conversation-like exchanges, higher frequencies of I and you in the forums might be expected than in written texts which do not involve conversation-like exchanges. The third person pronouns were excluded in this study because they not reflect interaction between participants in the same way as the first and second person pronouns do and often have no personal reference. In addition to the first and second person pronouns, the other items that were chosen include demonstrative pronouns and determiners and some common context-dependent spatial and temporal adverbs which normally rely on shared context for their spatial and temporal referents. The interaction in the discussion forums took place asynchronously between participants in different locations, so the situational aspects of this register might be reflected in the frequencies. A further reason for the choice of these particular items in this case study is because with the exception of those, the frequencies of the other items also contrast somewhat between the oral conversation and academic prose registers and therefore are more useful for comparison than features whose frequencies do not vary between registers. A list of the features is provided in Table 4.3.

---

4 This is the corpus on which the Longman Grammar of Spoken and Written English (Biber et al. 1999) is based, and the frequencies used for comparison in this study were taken from the grammar.

5 Michigan Corpus of Academic Spoken English
Table 4.3. Table of features indicating interaction

<table>
<thead>
<tr>
<th>Groups of features</th>
<th>Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>First and second person pronouns</td>
<td><em>I, me, we, us, you</em></td>
</tr>
<tr>
<td>First and second person possessive determiners</td>
<td><em>my, our, your</em></td>
</tr>
<tr>
<td>Demonstrative pronouns</td>
<td><em>this, that, these, those</em></td>
</tr>
<tr>
<td>Demonstrative determiners</td>
<td><em>this, that, these, those</em></td>
</tr>
<tr>
<td>Temporal adverbs</td>
<td><em>now, then</em></td>
</tr>
<tr>
<td>Spatial adverbs</td>
<td><em>here, there</em></td>
</tr>
</tbody>
</table>

The software WordSmith Tools was used for identifying the features, followed by a manual analysis of the results. For instance, *then* could be a misspelling of *than* and vice versa. Furthermore, *then* can function as a spatial adverb as in *then he came to Flint* and also as a linking adverb as in *if it permeates his very being, then what is there to regret.* As the focus of this study was only on the spatial uses of *then*, all occurrences of the linking adverb *then* were excluded from the frequency count. The same procedure was necessary for *now*, which can be used as a temporal adverb as in *now I obviously don’t get anything,* and but also as a linking adverb or discourse marker as in *Now, is there a message.* Only temporal *now* was included in the frequency count. *There* also occurred frequently as existential *there* in the forum material and these instances, which made up 75% of the total number identified by the software, were removed from the concordance lists before frequencies were calculated.

Out of the twenty items examined, the frequencies of fifteen lay between those of oral conversation and academic prose. When compared to the frequencies in the oral seminar, the frequencies of the features in the discussion forum material lay closer overall to those of academic prose than to those of the academic seminar. On the other hand, the results revealed that the frequencies and the collocations of the first person pronoun *I* were more similar to oral conversation than to academic prose. For example, the first person singular pronoun (*I*) occurred frequently with mental verbs such as *think, believe* and *guess.* These collocations are less frequent in academic prose. The observed patterns of three-word lexical bundles containing *I* in the forum material are in line with the findings of Biber, Conrad and Cortes (2004) who found that certain lexical bundles occurred frequently in classroom discourse and conversation and others in classroom discourse and academic prose.

The contrast between the frequencies of *I* and *you* in the two oral registers, conversation and the academic seminar, and the two written registers, the discussion forum and academic prose, can be seen in Figure 4.2.
Both I and we were more frequent than the other personal pronouns and determiners and the contrast was salient across registers. The three possessive determiners, my, your and our, were used more frequently in the discussion forums than in academic prose, and the first person plural (we, us) was used more frequently in the discussion forums than in the oral seminar. This may be partially explained by the subject matter of the communication which dealt with cultural issues. Students frequently used first person plural pronouns when talking about society, culture and life experiences. The first person plural pronouns referred first of all to themselves and the discussion group, secondly, to themselves and others not in the discussion group and thirdly and most frequently to people in a more general sense. These three uses are illustrated in the examples I hope [teacher’s name] might give us an answer on Thursday, We had one bathelor [sic] sleeping over at my place once and we are all human and we all make mistakes. There was some variation between frequencies in the separate course discussion forums and this was most notable in the proficiency forums, suggesting that task type and topic affect frequencies.

This, both as a determiner and as a pronoun, occurs more frequently in academic prose than in oral conversation (see Fig. 4.3). In the present study, it is striking that the frequencies of use are so similar for the discussion forums and academic prose, especially when it comes to this as a determiner. An examination of the collocations and text in the discussion forums revealed that the referent of this could be found most often within the text, and usually in the immediately preceding text. When used as a determiner, this determined nouns that were most commonly course material such as the book or novel that was read. As a determiner that occurs less frequently than in academic prose and much less frequently than in oral conversation and the oral seminar. As a
demonstrative pronoun, *that* also occurred less frequently in the discussion forums than in oral conversation and the oral seminar and only slightly more frequently than in academic prose. Neither *these* nor *those* was very frequent in any of the four registers, yet *these* was more frequently used in academic prose followed by the discussion forums, the oral seminar, and oral conversation in that order. An examination of the collocates revealed that the usage in the discussion forums was more similar to academic prose in that a postmodifying clause or phrase typically occurred after *those*. There was some course variation for the discussion forum frequencies but on the whole, the frequencies were too low to suggest any trends.

![Figure 4.3. Frequencies of *this* and *that* across the four registers](image)

The frequencies of the four adverbs *now, then, here* and *there* in the discussion forums were more like academic prose than the two oral registers were (see Fig. 4.4). Although students who participated in the forum discussions were not in the same place nor wrote at the same time, they referred to the time and place where they composed the message using both *here* and *now*, giving their messages a sense of immediacy. In addition, both the temporal (*now, then*) and spatial (*here, there*) adverbs were used to refer to parts of the books or novels being discussed as well as to where the students themselves were in their thinking process. The spatial adverbs were also used to refer to a place in the discussion. When referring to a specific geographical location, *here* was often post-modified, such as in *here in our Nordic part of the world*, to make it clear where the referent was.
Figure 4.4. Frequencies of now, then, here and there across the four registers

The frequencies of the items examined in this study and the ways in which they were used indicate that the language use in these particular discussion forums is in some ways similar to oral conversation, but at the same time they tend towards academic prose. Thomas (2002) suggested that the technology-mediated communication of online discussion forums has an oral function like that of conversation and yet is in written form but that the interactional nature of conversation cannot be duplicated in written contexts (Thomas, 2002: 363). The trends observed in this study indicate that the language in this forum does not appear to duplicate either oral conversation or academic prose, but may share features of both. The results suggest that the participants are aware of the purpose and the context for the communication and adapt their language accordingly, addressing each other directly as in oral conversation but also providing post-modification and other explanatory text which is typical of writing where writer and reader do not share time and space.

One of the unexpected observations of this study was how students used temporal and spatial adverbs to refer to their thinking processes and understanding of the course material. This indicates an awareness of the learning process on the part of the students. The forum discussions can reveal aspects of learning that would not otherwise be visible in an academic piece of writing, which would be edited and more “polished”, or in an oral seminar, where students may not feel comfortable revealing partial or lack of understanding. As one purpose of the forum is that students develop their ideas, it can be a unique window to the reasoning processes of students. The reasonably high frequency of mental verbs such as think and believe indicate that participants are expressing opinions and thoughts that might not be expressed
in academic prose where claims are expected to be supported. The fact that modal verbs indicating tentativeness were also common collocates of first person pronoun *I* suggest uncertainty on the part of participants when expressing opinions. This tentativeness was also seen in Case Study 1 where students used question marks to show uncertainty when suggesting something rather than as a marker of a question. Discussion forum material of this kind can provide useful information about the students’ learning in a way that perception studies cannot. The communication that takes place in this type of computer-mediated learning environment provides a forum for expression of thought that neither an oral seminar nor academic prose can provide.
5. Discussion

The main aim of the four studies was to explore various aspects of the written online interaction of tertiary level English students communicating with one another in discussion forums. The studies examined the content of the messages, the subject line and other aspects of the interaction such as when messages were posted, which messages received responses and how participants referred to themselves, to each other and the course material. It is hoped that this examination will further our understanding of written online communication as well as the learning process. In addition, this thesis aims to further knowledge on some of the contextual and situational affordances (constraints and opportunities) of tertiary learning and teaching online. In addition, both instructors and students may benefit from improved awareness of factors affecting communication in an online learning situation and what the communication can tell us about the learning process. This knowledge in turn can be used to improve the way learning activities are designed and managed.

A lack of interaction, that is, very few messages in a forum, should not necessarily be interpreted as a sign of an unsuccessful learning environment. There are many factors which affect the frequency of posting in a forum, such as design factors, size of group, the time the forum is open, the topic and the clarity of instructions from the teachers and whether or not the contributions to the forum constitute part of the students’ grades or not. The results of the first case study indicate that posting early in the forum increases the chance that a message will be read and receive a response. It also seems that students choose to respond to a message when some sort of emotion or personal connection is aroused by the content. That is, they respond to messages that have in some way triggered an emotional response. These responses are often positive ones, such as gratefulness for an explanation suggested by another student or acknowledgements that students have had similar experiences. Hence, choosing topics for discussions that people feel strongly about is one possible way of promoting interaction among students. Discussion forums, although asynchronous, are conversation-like and interactive in nature and, accordingly, shorter messages are more likely to be read and get responses if interaction is the goal of the discussion forum. The mixed nature of the online discussion forums with both conversation-like and academic-like features observed in Case Study 4 suggest that the online discussion forum may provide a path to academic writing for students who have not yet mastered traditional academic discourse.

Although frequent responses from students and instructors may be seen as a sign of a well-functioning discussion forum, frequent messages alone is not always a sign of effective learning. It was observed in Case Study 1 that instructor input does not necessarily increase the amount of responses and instructor messages do not always get responses. This suggests that peer messages may be regarded to be as important as instructor messages in the
discussion forums. Although some of the literature course forums contained a lot of short messages expressing students’ opinions, some of the forums with longer but fewer messages broached theoretical aspects of the course content and tended to discuss topics more thoroughly. A mix of discussion types might prove to be the best recipe for success. Early in a course the instructor could encourage frequent shorter messages which would contribute to students building confidence and establishing rapport with other students in a group. Later in the course, students could be encouraged to apply more theory and support their opinions with secondary sources.

There is tendency for the majority of students in a group to mimic the subject lines of those who posted first in the forum. This could suggest that alignment and that this behaviour reflects a desire to conform to others in the group. It could also suggest that there may be some insecurity about how to write their contributions. There are indications that instructors, and perhaps even confident students, who post early in a forum may influence the linguistic behaviour of those who subsequently participate in the discussion forum. If the instructor gives clear instructions or examples of how to communicate and states the participation requirements, this may have an effect on the messages students produce and this, in turn, may affect the learning experience positively for all participants. The effect of a sample message by the instructor was particularly salient in studies 1 and 3, where one instructor teaching the literature course gave an example of how to write both a thread-initiating message and a response. In the first case study, for instance, it was noted that this instructor’s groups produced similar types of messages to the sample with regard to length and content. In the third case study these same groups were observed to be using a similar type of subject line as the sample given by the instructor. In other groups where instructors had not given examples, students seemed to be influenced by the messages posted early in the forum by other students. That is, their messages were of a similar length and also tended to use subject line headings that were similar to the first messages posted in that particular forum.

One feature of CMC that is considered typical but which was rarely found in the material in these studies is the emoticon. The use of symbols to communicate phatically is not necessarily restricted to CMC. For example, informal letters between friends have often contained crosses and zeroes as symbols representing hugs and kisses, and hearts representing feelings of fondness. Perhaps the reason that these features may have gained attention in CMC research has to do with the nature of the medium. It may be the persistence of text and the novelty in written online communication that has brought them to the attention of researchers and perhaps gives the illusion of CMC as unique. However, as with offline language, there is great variation depending not only on the medium but also factors such as the purpose of the communication and the relationship between participants. This can be seen in the case studies where despite the fact that the situation appears to be the same
or very similar, there is variation between the linguistic choices made by individuals and groups. Even though task and topic might be the same in the course, a discussion forum may develop in an unpredictable way, similar to conversation.

There is sometimes a tendency for comparative studies of CMC to give the medium credit for any differences that may be apparent and neglect factors that have a greater effect, such as the context and purpose of the communication and the relationship among the participants. As can be seen in Case Study 4, there are patterns of usage in these conversation-like exchanges in the discussion forum material that are similar to patterns of usage in academic discourse. This suggests that despite the fact that these online discussion forums are not the same type of academic registers that are found in traditional academic discourse such as scientific articles or textbooks, the institutional setting within which the exchanges take place does exert an influence on the linguistic choices made by the participants. At the same time, the forum communication also shares features of oral conversation, indicating that that these discussion forums might be suitable for students who have not yet mastered academic registers fully and act as an environment where students can express opinions in a more informal way than is expected in other kinds of academic discourse such as academic essays.

Nevertheless, communication produced through exchanges in an asynchronous computer-mediated environment is relatively new in comparison to other forms of academic discourse, which means that the conventions might not yet be as firmly established or agreed upon as they are in other types of academic discourse. On the other hand, even if there might not be established conventions for communicating in discussion forums in an online learning environment, there are established conventions in other forms of writing which may be used and if necessary adapted to the online environment. Even if students have used discussion forums previously, many of them may not feel confident about exactly how they are expected to communicate in a particular forum. For those who design and manage learning activities online, there are multiple aspects to consider. The most important of these is how an asynchronous discussion forum can contribute to students fulfilling the learning outcomes of a particular course. For courses where collaboration and problem-solving are goals, the asynchronous forums may work well. Students wishing to improve basic language proficiency might benefit from smaller discussion groups with shorter messages and more frequent messaging. On the other hand, those with a higher level of proficiency might benefit from more complex discussions which require the use of a more advanced vocabulary and critical analysis of course content.
6. Conclusion and suggestions for further studies

The four studies that make up this thesis investigate different aspects of interaction in online discussion forums used in higher education. The first case study focused on interaction with regard to the types of messages that resulted in one or more responses. The second case study looked at interaction in terms of the community building strategies of the participants in the discussion forums. The third case study examined the interaction from the perspective of the functionality of the subject line of messages. The final study looked at features that indicated participants’ interaction through reference to one another, the text and the context as reflected in the use of pronouns, determiners and adverbs. It also compared the frequencies and usage of these features to those in other types of discourse. Further research on the features of computer-mediated academic communication is needed. Previous studies have examined textbooks, academic prose, office conversation and classroom discourse (see Biber (2006); Biber et al., (2004); Hyland, (2004), (2008)), but neglected computer-mediated academic language.

Taken together, the four studies represent an examination of a communicative event which takes places in one particular social situation, that is, a learning environment where communication is text-based, asynchronous and online. The focus is on how participants in such an environment interact with one another, what linguistic choices they make, what strategies they use and how their preferences may or may not affect the interaction. The studies also shed some light on how linguistic choices reflect the learning process in this particular environment. Even though the studies are not grounded in pedagogical theory as such, it is hoped that the findings will contribute to the furthering of knowledge about some of the mechanisms that are at work when tertiary students communicate with one other in an online learning environment. In particular, it is hoped that these studies will provide insight for those interested learning as a communicative event and in situated learning, which regards learning as a social situation.

Although many previous studies have focused on learning outcomes in online learning environments, it has often been difficult to show how these outcomes have been achieved. Linguistic artifacts can be seen as evidence of critical thinking and other cognitive activity but there are so many variables at work that it is difficult to point to one factor, such as the technology or the task type or course design itself as the catalyst for these activities. The results and analyses from the four studies provide insight into the nature of interaction in an asynchronous learning environment and the linguistic strategies employed by participants to perform different functions as part of their course activities. Some of these functions include discourse organisation and argumentation, as well as social activities and learning support. Although not a specific aim of these studies, the features examined in the discussion forum interaction suggest that students are in the process of learning and are also aware of their own
thought processes. This was reflected in Case Study 4 where students used temporal and spatial adverbs to refer to their thoughts and their interpretations of the course material. The findings in Case Study 1, which examined responses to other students’ messages, and Case Study 2, which examined community building strategies, indicate that the social side of learning and its effect on the learning experience warrants further research.

The relationship between instructor input and student input appears to be a complex one and more research is needed on participants’ expectations and group dynamics as well as message content and purpose of the discussion forum. One study by Mazzolini and Maddison (2007) showed that the more the instructor posted, the less frequently the students posted. However, this was not necessarily the case in the different forums studied here. There seem to be other factors that influence response rate, for example the size of group and the length of the message. Discussion threads that had a lot of responses tended to contain messages that personalised the topic being discussed. The group that posted more messages per person than the other groups was also the group that used the most community building devices. Further studies on the correlation between social and personal aspects of learning may be able to shed some light on the relationship between participants’ sense of community, how it is created and maintained and the potential benefits for learning.

Results from comparative studies have previously identified perceived learning rather than actual learning. It is possible that there is a relationship between perceived and actual learning but this connection has not yet been firmly established. Although this particular aspect was not dealt with in any detail in the present case studies, it was noted in the investigation of response messages in one of the case studies (no. 1) that students who posted their messages towards the end of the forum tended to receive a different type of feedback from the instructor than those who posted early. The feedback for the later messages indicated that these students had not adequately grasped the subject matter, understood or applied theory and had not managed the task satisfactorily. This suggests that stronger students tend to post their ideas earlier in the forum while the weaker students wait until other contributions have been posted.

In an online learning environment, factors such as course design and content, student background, and instructor experience and attitude may all affect linguistic choices. For example, more detailed studies of a participant’s status in a group may provide some explanation for some of the linguistic choices made. For example, it has been suggested that the social status of participants in an online discussion forum may be based on their perceived helpfulness by other participants (Biber & Conrad, 2009: 197). Even though the work of Biber and Conrad (2009) was specifically based on help forums, the discussion forums in a learning environment function in a similar way, as students are supposed to help each other with their learning. If being helpful is associated with status in
learning environments as well, then this may partially explain the role of community building strategies and evidence of politeness which were observed in the material used for the four case studies. In the second case study, the group who used the most and the greatest variety of community building strategies was also the group who received the highest grades on completion of the course. More research is needed on pragmatic competence and subject competence to clarify the relationship between the two.

In addition, more studies which examine the correlation between linguistic artifacts and students’ experience of learning and actual learning outcomes are needed. For example, it may be useful to determine whether it is the use of community building devices that promotes students’ learning or whether highly proficient and motivated students are more likely to build rapport with their peers. It is through language that some of our thought patterns and reasoning become visible. The more we learn about interaction and how different linguistic choices function to organize discourse, create coherence, build and maintain social relations, encourage interaction and support learning, the more this knowledge can be integrated into course design and implementation.

The majority of the material in the four studies which make up this thesis is comprised mainly of student-to-student interaction; only a few forums contained written input from the instructors. It must be remembered, however, that the tasks were designed by the instructors to aid learning and help students achieve specific learning outcomes for each of the three different courses. Further studies on the ways in which instructors can have an influence on the volume and content of messages, the language used and the frequency of responses are needed.

As broadband capabilities improve, synchronous online communication is becoming more frequently used than previously and a comparison between synchronous and asynchronous online communication in the tertiary education context might reveal some interesting trends. Synchronous communication often uses multiple channels simultaneously, including audio, video, chat and a whiteboard or notepad. More information is needed on how students and teachers manage these multiple channels and also if and to what extent multimodal communication can enhance learning. In addition, the boundary between traditional campus students and net-based students is becoming blurred, as online activities are being integrated into courses for campus students. More detailed research on both spoken and written interaction in the different types of courses and activities in online and offline study can inform us of how these activities contribute to student interaction and learning.
References


Heyman, E. (2010). Overcoming student retention issues in higher education online programs. *Online journal of distance learning administration, 13*.


72


Appendix I

Questionnaire

1. What language do you consider your first language or mother tongue?
2. Mark all activities that you do daily or almost daily
   - Read news or other information online
   - Watch films/YouTube clips/listen to music online
   - Send SMS
   - Read & write emails
   - Write on your own blog
   - Use social networks (e.g. Facebook, Twitter)
   - Play multiplayer computer games
   - Use discussion forums outside the learning environment
3. Do you consider yourself a fast typist?
4. Was spring term 2009 the first time you studied online?
5. Mark all statements you think are true about the discussion forum for the course(s).
   - Writing in the forum gave me time to think before I posted my ideas.
   - It was sometimes difficult to follow which messages people were responding to.
   - It was important to be kind and friendly to other students.
   - Writing my thoughts helped me understand and learn.
   - I always read my comments thoroughly and edited them before posting them.
   - If I was more confident in English, I would have written in the forum more often.
   - It was important to link my responses to messages already posted by quoting from a previous message or using a person's name.
   - I felt others misunderstood what I wrote sometimes.
   - I preferred not to post my message until after others had posted first.
   - I did not read all messages in the forums.
   - I read the discussion forums for other groups as well as my own.
   - I learnt a lot from what other students wrote.
   - Sometimes the discussion threads had too many messages and this made it hard to follow the discussion.
   - The teacher's responses were the most important.
   - I liked being able to go back and read messages more than once.
   - The responses from my peers helped me learn.
6. What kind of messages made you write a response?
7. Was it important to write in the subject line and what do you think is its function?
8. Did you prefer to read and respond to other students before writing your own ideas? Why/Why not?
9. When you opened the discussion thread, which messages did you read first?
### Appendix II

Overview of the forum statistics

<table>
<thead>
<tr>
<th>Course discussion forums</th>
<th>Number of participants</th>
<th>Number of messages per forum</th>
<th>Number of words per forum</th>
<th>Average number of words/message</th>
</tr>
</thead>
<tbody>
<tr>
<td>culture forum topic 1 gr 1</td>
<td>11</td>
<td>30</td>
<td>9082</td>
<td>303</td>
</tr>
<tr>
<td>culture forum topic 1 gr 3</td>
<td>9</td>
<td>22</td>
<td>8349</td>
<td>380</td>
</tr>
<tr>
<td>culture forum topic 2 gr 1 &amp; 3</td>
<td>18</td>
<td>53</td>
<td>14921</td>
<td>282</td>
</tr>
<tr>
<td>culture forum topic 3 gr 1 &amp; 3</td>
<td>17</td>
<td>61</td>
<td>19489</td>
<td>319</td>
</tr>
<tr>
<td>culture forum topic 1 gr 2</td>
<td>6</td>
<td>15</td>
<td>6049</td>
<td>403</td>
</tr>
<tr>
<td>culture forum topic 1 gr 4</td>
<td>12</td>
<td>38</td>
<td>9281</td>
<td>244</td>
</tr>
<tr>
<td>culture forum topic 2 gr 2 &amp; 4</td>
<td>13</td>
<td>46</td>
<td>11249</td>
<td>245</td>
</tr>
<tr>
<td>culture forum topic 3 gr 2 &amp; 4</td>
<td>14</td>
<td>50</td>
<td>17802</td>
<td>356</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>literature forum 1 gr 2</td>
<td>15</td>
<td>57</td>
<td>13011</td>
<td>228</td>
</tr>
<tr>
<td>literature forum 2 gr 2</td>
<td>16</td>
<td>69</td>
<td>15455</td>
<td>224</td>
</tr>
<tr>
<td>literature forum 3 gr 2</td>
<td>16</td>
<td>58</td>
<td>19804</td>
<td>341</td>
</tr>
<tr>
<td>literature forum 4 gr 2</td>
<td>13</td>
<td>44</td>
<td>14422</td>
<td>328</td>
</tr>
<tr>
<td>literature forum 5 gr 2</td>
<td>4</td>
<td>5</td>
<td>2144</td>
<td>429</td>
</tr>
<tr>
<td>literature forum 1 gr 1, 4, 5</td>
<td>27</td>
<td>185</td>
<td>15138</td>
<td>82</td>
</tr>
<tr>
<td>literature forum 2 gr 1</td>
<td>7</td>
<td>36</td>
<td>7826</td>
<td>217</td>
</tr>
<tr>
<td>literature forum 3 gr 1</td>
<td>7</td>
<td>31</td>
<td>7523</td>
<td>243</td>
</tr>
<tr>
<td>literature forum 4 gr 1</td>
<td>4</td>
<td>10</td>
<td>2643</td>
<td>264</td>
</tr>
<tr>
<td>literature forum 2 gr 4</td>
<td>16</td>
<td>103</td>
<td>14786</td>
<td>144</td>
</tr>
<tr>
<td>literature forum 3 gr 4</td>
<td>13</td>
<td>76</td>
<td>12649</td>
<td>166</td>
</tr>
<tr>
<td>literature forum 4 gr 4</td>
<td>13</td>
<td>56</td>
<td>8441</td>
<td>151</td>
</tr>
<tr>
<td>literature forum 2 gr 5</td>
<td>5</td>
<td>20</td>
<td>3993</td>
<td>200</td>
</tr>
<tr>
<td>literature forum 3 gr 5</td>
<td>4</td>
<td>17</td>
<td>3152</td>
<td>185</td>
</tr>
<tr>
<td>literature forum 4 gr 5</td>
<td>5</td>
<td>18</td>
<td>3914</td>
<td>217</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>proficiency forum 1 gr 1</td>
<td>9</td>
<td>30</td>
<td>4388</td>
<td>146</td>
</tr>
<tr>
<td>proficiency forum 1 gr 2</td>
<td>15</td>
<td>81</td>
<td>10152</td>
<td>125</td>
</tr>
<tr>
<td>proficiency forum 1 gr 3</td>
<td>14</td>
<td>44</td>
<td>4957</td>
<td>113</td>
</tr>
<tr>
<td>proficiency CA gr 1 &amp; 2</td>
<td>22</td>
<td>43</td>
<td>19826</td>
<td>461</td>
</tr>
<tr>
<td>proficiency forum 2 gr 1</td>
<td>17</td>
<td>49</td>
<td>17993</td>
<td>367</td>
</tr>
<tr>
<td>proficiency forum 2 gr 2</td>
<td>16</td>
<td>42</td>
<td>20914</td>
<td>498</td>
</tr>
<tr>
<td>proficiency CA gr 3</td>
<td>17</td>
<td>17</td>
<td>9930</td>
<td>584</td>
</tr>
<tr>
<td>proficiency revised essay gr 3</td>
<td>18</td>
<td>18</td>
<td>10778</td>
<td>599</td>
</tr>
<tr>
<td>proficiency rough draft AE gr 3</td>
<td>18</td>
<td>22</td>
<td>11465</td>
<td>521</td>
</tr>
</tbody>
</table>

|                                            | 346                    | 110403                      |                            |                                  |
Language and interaction in online asynchronous communication in university level English courses

The four case studies presented in Language and interaction in online asynchronous communication in university level English courses investigate written discussion forum interaction in a computer-mediated learning environment. These studies deal with different, yet related, aspects of discussion forum communication. Aspects included are the labeling and response patterns of messages, community-building strategies among participants and features of informal conversation and formal academic writing in the messages. Building on discourse analysis combined with content analysis and corpus method, the work systematically examines the linguistic patterns of communication in the discussion forums. The findings show that there are multiple factors at work simultaneously that affect the linguistic choices by the discourse participants. The constraints and opportunities of the communication are not only connected to the fact that it is computer-mediated, but also to the fact that it is written and in a particular academic environment. Knowledge of the choices available and of what factors potentially affect them is useful for anybody involved in research on net-based teaching and learning.