Making Sense Digitally

By Signe Rejdvik

Coherence, broadly defined, is that which in a discourse connects utterances with utterances, utterances with people, and people with other people. It is, in short, the 'glue' of text and conversation.

Erickson et al. (2002: 2)

Making Sense Digitally
Conversational Coherence in
Online and Mixed-Mode Contexts

Therese Örnberg Berglund

Abstract

Successful interaction makes sense to its participants – it is, in other words, coherent. As different resources are employed to indicate mutual orientation by showing which actions are linked and where attention is paid, coherent conversation can be said to be achieved multimodally. This thesis builds on an interactional and ecological theoretical basis, and investigates how strategies for sense-making vary depending on the context of interaction.

Specifically, this is a study of interaction in different types of multimodal and multiplex communication situations. The characteristics of these technology-rich communication situations are mapped and, primarily through ethnographic methods and interaction analysis, it is investigated how conversational coherence is maintained. Furthermore, it is suggested how the findings can be applied to inform interaction design.

The analysis presented is based on results from four case studies, dealing with different aspects of coherence creation in different English speaking contexts. Study 1 investigates coherence in intertwined threads in dyad instant messaging (IM) interaction, and Study 2 focuses on coherence through conversational feedback strategies in interaction in a multimodal desktop video conferencing system. Study 3 has a somewhat broader focus, as the multiple communication channels in which one participant is involved are all taken into consideration in the investigation of sequential coherence. Finally, in Study 4, the physical and digital interaction in a partly shared studio space is investigated, with an emphasis on how mutual orientation is established during conversation initiation.

The findings show that when addressing coherence it is important to acknowledge both dynamic aspects of context, such as activity and participants, and more static aspects, such as communicative affordances of environments and tools. In online and mixed-mode interaction, the notion of context becomes particularly complex, as participants are simultaneously part of both the individual context, which may include digital tools for communication, and the shared context of interaction (*polycontextuality*). In this thesis, it is further shown that this has consequences for coherent conversation initiation and possibilities to engage in multiple semi-simultaneous conversations (*polyfocality*). Additionally, the results emphasize the importance of explicit linguistic strategies in computer-mediated interaction.

Keywords:
Conversational coherence, context, communicative affordances, conversation analysis, multimodality, cohesion, sequential structure, attention, polyfocality, polycontextuality

Table of Contents

List of papers

I.   Örnberg Berglund, T. Forthcoming 2009. Disrupted turn adjacency and coherence main-tenance in Instant Messaging conversations. To appear in *Language@Internet* 6(2).

II.   Örnberg Berglund, T. 2009. Multimodal student interaction online: An ecological perspec-tive. *ReCALL* 21(2): 186-205.

III. Örnberg Berglund, T. 2007. Multiplex conversations afforded by technology. *Proceedings of the 40th Annual Hawaii International Conference on System Sciences*, Hawaii.

IV.   Örnberg Berglund, T. Submitted. Conversation initiation in multiplex communicative ecologies.

Acknowledgements

Many people have contributed towards the completion of this thesis. First of all, I would like to thank all the participants in my studies, who so kindly have allowed me to observe and record their every conversational interaction. This study would not have been possible without you!

My main supervisor Patrik Svensson has been a superb guide throughout the whole journey. As the supervisor for my C- and D-essays in English linguistics, he was the one who got me excited about online communication as an area for research. Patrik's encouragement and good advice has been invaluable, and he deserves my most sincere thanks. My auxiliary supervisor Mathias Broth, has been an excellent reader and discussant during the last intense year of completing this thesis, and has given numerous important suggestions for improvements on my texts. We have not always agreed, but our discussions have always been very constructive. I want to thank you both for your strong commitment and your quick replies to my many queries over email – please excuse my many interruptions!

Over the years, I have had the opportunity to be part of many different inspiring contexts. As a doctoral student in the research profile "The humanities and information technology", I got my position through HUMlab at Umeå University. This is an innovative meeting place for people interested in how information and communication technology might be explored in the humanities – as a tool, a platform and a study object. In HUMlab, I have had the opportunity to experiment with technology, and with expertise at hand when needed. To my friends in HUMlab, among others, Emma Olofsson, Magnus Olofsson, Magnus Nordström, Jon Svensson, Fredrik Palm, Johan Lindskog, Ellen Jacobsson, Cecilia Lindhé, Stefan Gelfgren and Erika Lokander, thanks for technical support, stimulating discussions, social coffee breaks and fun Friday lunches. My doctoral colleagues in the lab, Jim Barrett, Stefan Blomberg and Van Leavenworth, have also provided valuable input and a positive social atmosphere. A special thanks goes out to my doctoral colleague, good friend and "sister in arms" Stephanie Hendrick, for many stimulating conversations about work and private matters, as well as for thorough last-minute, late-night proof-reading! Many thanks also to Erica Robles, postdoctoral fellow in HUMlab, who gave valuable input on the design related parts of a late draft of my text.

At the Department of Language Studies (previously Modern Languages), I have also had the pleasure of engaging with inspiring people. For instance, my roomies from the doctoral student office during my first years provided many appreciated chats and helpful comments. Thanks to Anette Svensson and Martin Shaw! After moving offices to the pavilion I had the opportunity to engage in stimulating lunch conversations with colleagues there. Thanks

Prologue

As digital technology is becoming a common component in social, educational and workplace contexts, we are beginning to see how this has an impact on interactional patterns. It is these effects which are the topic of the book you are currently reading.

For example, in a workplace today you will often have access to a number of different tools for communication, and you may choose whether to interact with your colleagues face-to-face or digitally. In addition, it is very likely that you will be focusing on several different tasks, both work-related and social, semi-simultaneously.

The current study addresses some of the practical problems that arise when we rapidly switch between different tools and different conversations. For instance, how do we know which actions are linked and where attention is paid during conversational interaction? How can we find appropriate times to initiate conversations? And how can we manage conversational multitasking and still be able to work productively?

One example of a very practical problem can be found in a report from the American research and advisory firm Basex from 2005. Here it was shown that in the US, 28% of the knowledge worker's day was going to waste because of unnecessary interruptions. Their calculations further showed that this costs US companies $588 billion each year (Spira & Feintuch 2005). Whereas one might argue that it is impossible to measure the exact cost of interruptions, the conclusion that such interruptions are many and costly is nonetheless significant.

The research presented in the current study can contribute to a better understanding of these problems through a detailed investigation of how mutual orientation is established and maintained in some different types of multiplex and multimodal communication situations. Potentially, this research may also contribute towards a partial remedy, through the introduction of the notion of *balanced awareness* as a design strategy to ensure coherent interaction.

1. Introduction

Simply defined, coherence is connectedness and meaningfulness in conversation. A coherent conversation seems well structured and sensible to the participants. Coherence is normally taken for granted, yet the production of coherence is complex and not altogether understood.

Littlejohn (2002: 81)

Communication is sense-making. When we communicate we continuously need to make our actions accountable and negotiate both the process and the product of our interaction (Garfinkel 1967, Clark & Brennan 1991, Clark 1996). At our disposal in the sense-making practice, we have the unfolding discourse and the context in which it is situated. This implies that sense is something that is *made* collaboratively during interaction. From such a pragmatic viewpoint, coherence can be seen as "[...] a kind of organization which actors accomplish or construct communicatively during their interactions" (Korolija 1998: 26). This organization serves to establish and maintain mutual orientation and links between related actions, and coherence thus concerns "[...] the flow of attention during joint activities and the continuity of actors' shared apprehension of the situation" (Korolija 1998: 115).

Traditionally, studies of coherence within linguistics have been concerned with lexical and grammatical cohesion.[1] With such an approach, the focus is on lexical and grammatical means of realizing connections in longer stretches of spoken or written discourse, and as such it has been described as an attempt at defining systematic grammar on the discourse level (Paltridge 2006). In the current study, it is argued that in order to comprehend fully how mutual orientation is established and maintained in conversation, the scope has to be broadened and alongside strategies of lexical and grammatical cohesion, other aspects of coherence creation need also to be taken into account.[2]

In face-to-face interaction, one important clue as to which actions are to be understood as linked can be found in the sequential structures of a conversation (Heritage 1984, Schegloff 1990). When we communicate, we work together to ensure that adjacent actions make sense, even if utterances lack formal cohesive links. This is one of the skills that competent communicators possess (Coates 1995). Apart from the fact that participants in interaction strive to make sense of each other's actions (cf. also Garfinkel 1967), other explanations to how utterances lacking structural cohesion are made

---

[1] The perhaps most prominent example is Halliday and Hasan's (1976) influential work reported in *Cohesion in English*.

[2] Examples of analyses of coherence from a broader perspective are available in, for instance, Craig & Tracy (1983), Neubauer (1983), Gernsbacher & Givón (1995) and Korolija (1998).

coherent can be found in ways of indicating continued shared attention (Clark & Brennan 1991). For example, when ensuring that adjacent actions are linked in face-to-face interaction, shared attention is a prerequisite, and gaze an important clue. Further clues can be found in the broader context of interaction, including activity type and participants. As situated communicative action is intrinsically linked with context (cf. e.g. Linell 1978), conversational sense-making is context-dependent and multimodal, drawing upon different temporal, visual and auditory resources.

Context can be defined in a variety of ways, and can encompass notions such as *cotext*, referring to the text surrounding the utterance in question, *social situation* and *material context*.[3] Moreover, context is often seen as something which is configured during interaction and which is made relevant by the participants themselves (Goodwin & Duranti 1992, Linell 1998, Goodwin 2000). In the current study, different dimensions of context are explored, but the main focus lies on the ways in which the material environment not only is a resource for coherence maintenance, but also presents an ecology for action, an environment with different *affordances* relative to the living organism (Gibson 1977, 1979). These affordances influence communicative patterns (Hutchby 2001) and thus also strategies for sense-making and awareness of attention levels.

As digital tools for interaction are becoming common components of our communicative ecologies, additional practical problems arise in relation to conversational coherence. For instance, when having access to a number of different modes, it is possible to engage in several different activities semi-simultaneously. When involved in such *polyfocal* interaction (cf. Scollon et al. 1999, Jones 2004) and when conversing in intertwined threads we may need to find alternative ways of showing which actions are linked and where attention is paid. These *polycontextual* situations (cf. Engeström et al. 1995, Leander 2002, Saarenkunnas 2004) also call for alternative ways of achieving coherent conversation initiations, as we no longer can depend on a shared context when finding appropriate times to initiate interaction.

1.1 Aims and outline
In the current study, I take a pragmatic approach to coherence, viewing sense-making as a co-constructed, multimodal and context-dependent achievement. Building on an interactional and ecological framework, this study explores how coherence is co-created and maintained in multiplex[4]

---

[3] For examples and discussions of contextual models, see for instance Hymes (1974), Halliday & Hasan (1985), Goodwin & Duranti (1992), Korolija (1998) and Jones (2004).

[4] The term *multiplex* might be misleading, as for the participants in interaction, these situations often do not appear to be very complex. The term is used here to indicate that the participants need to pay attention to several things at once.

and multimodal communication situations. It is argued that the findings relate to different dimensions of context, and not least the communicative affordances of different modes, media and environments.

The main aim of this study is to identify practices of coherence establishment and maintenance in technology-rich communication situations based on four case studies. More specifically, this investigation deals with the situational, communicative, theoretical and applied questions presented in Table 1.

Situational:	What characterizes some technology-rich communication situations and the interaction taking place there?
Communicative:	What characterizes the strategies for coherence establishment and maintenance in these situations?
Theoretical:	What is the relation between strategies for sense-making and context of interaction in these situations?
Applied:	How might the results be applied in the design of tools for communication?

Table 1: Questions explored

One step in the process of achieving the aims of this study is to refine an analytical model describing the relationship between coherence and context. This is done by merging interactional views of context with the ecological notion of *affordances* (Gibson 1977, 1979), thus emphasizing material context. Furthermore, it is shown how the resulting model can be employed to counter some previous tendencies in research on computer-mediated communication (CMC), namely the tendency to view face-to-face interaction as superior, the tendency to ascribe either all or nothing to technology, and the tendency to perform de-contextualized analyses.

One of the contributions of the current study is thus that it shows the importance of including different contextual dimensions in the analysis of conversational coherence in technology-rich situations. By separating the material context from other contextual dimensions, it is also possible to contribute to the area of interaction design. In addition, this research adds to the field of interactional linguistics by applying a common linguistic analytical focus to "new" types of texts, as coherence is analyzed within and across modes and digitally-supported conversations.

Considering the multitude of possible communicative ecologies arising as a result of new tools for interaction being introduced, it is by no means possible to give an exhaustive account of coherence in digitally-mediated interaction. Rather, four case studies in situations showing different degrees

of multiplexity have been conducted and different aspects of coherence creation are addressed.

In all case studies, it is investigated how communicative affordances influence conversational structures. More specifically the analysis concerns the influence of different aspects of conversational context on the kind of sense-making strategies identified in the unfolding discourse. Here the qualities of the modes employed and the qualities of the interfaces and environments which combine them are taken into account. Other contextual factors such as activity type, participants and background knowledge are addressed with different foci for each of the four studies. The analysis mainly concerns interactional patterns, and based on these findings, conclusions about the affordances of the environments and tools are drawn.

Study 1 and 2 deal with communicative affordances on the tool level. Study 1 provides a close analysis of instant messaging (IM) interaction, and a special focus is on how coherence is maintained during disrupted turn adjacency, whereby the system causes logically adjacent turns to appear typographically separated. This is complemented by an investigation of additional signs of problematic coherence maintenance. Study 2 analyzes student interaction in a multimodal desktop video conferencing environment. Here, tool and task design are taken into consideration when analyzing multimodal strategies of conversational feedback and establishment of shared attention.

Study 3 deals with the communicative affordances of the complete communicative landscape available to one participant at her computer during conversational multitasking. In this study, sequential coherence, where, for example, replies are delivered upon request, and switches between different online conversations are at the core of the analysis. In Study 4, the scope is further widened, as the communicative affordances of a shared environment with technological tools are investigated. Here, the focus is on in-group interaction in a partly shared physical environment and more specifically on how conversation initiation is made coherent in this physical-digital communicative ecology.

Before we move on to a presentation of the results from the different studies, some central concepts need to be clarified, and the theoretical basis for the analysis needs to be further explicated. In order to position my research in relation to previous tendencies in research of CMC, in Section 2, I survey previous models aiming to explain the influence of mediation on communicative patterns. In Section 3, I describe the interactional basis of the present study, and the main concepts addressed, *coherence*, *context* and *affordances*, are further developed. This section also includes reviews of previous research on coherence, context and affordances in CMC. In Section 4, I present how the concepts are used in the current study, and introduce the analytical model employed.

In Section 5, I give an account of the methodology used in the different studies as well as discuss ethical concerns. Section 6 consists of a summary of the included papers, and the results are further discussed in Section 7. In Section 8, I make some concluding remarks. The complete papers corresponding to the four case studies can be found after the references section.

1.2 Defining conversation

Even though, or perhaps because, *conversation* is a commonly used term, it has to be defined in the context of this study. When I refer to *conversation*, I do not restrict myself to face-to-face interaction; rather this term is also used to refer to digitally-supported interaction. This is a conscious choice made in order to emphasize that, for example, text-based interactions depend on some of the same principles as face-to-face interactions, and can be analyzed in a similar manner. For the sake of analysis, it is still important to differentiate between the different modes of interaction, and thus these are specified when deemed necessary (for instance, face-to-face conversations vs. IM conversations). Furthermore, in the current study the term conversation is used about different types of talk-in-interaction and type-in-interaction[5], and is not reserved for a specific *activity type* (Levinson 1992). This can be compared with other uses of the term to differentiate the activity type of conversation as an informal verbal encounter, as opposed to other activity types, such as a discussion or a debate (Schegloff 1987).

---

[5] The term *talk-in-interaction* was introduced by Schegloff (1987) in order to illustrate that not only informal conversations, but also institutional encounters are part of the field of interest. I expand this notion by introducing the concept of *type-in-interaction* to refer to the exchanges taking place in the written mode.

2. Previous models for analyzing computer-mediated communication

The study of computer-mediated communication (CMC)[6] is by default an interdisciplinary endeavour, and has been approached from a variety of different disciplines and through diverse theoretical and methodological frameworks. For example, it has been studied in computer science, cognitive science, sociology, pedagogy and linguistics. Associated cross-disciplinary fields include, for instance, *computer-supported collaborative work* (CSCW), *computer-supported collaborative learning* (CSCL) and *workplace studies*.[7] The common denominator for these different approaches has been the object of study: human interaction via computer systems. This multitude of research, as well as disciplinary and theoretical approaches, makes it difficult to present a complete overview. The following critical account is structured with regard to some observed tendencies that will be assessed in light of the current study.[8] Further examples from previous research will be presented in Section 3.

2.1 Tendency to view face-to-face interaction as superior

Much early research on CMC tended to view face-to-face interaction as superior to technologically-mediated interaction. As Whittaker (2003) illustrates, many of the early studies on CMC were based on the *bandwidth hypothesis*, which he formulates as follows:

> *The closer the modes supported by a technology correspond to those of face to face communication, the more efficient the communication with that technology. Specifically, adding visual information to speech should improve the efficiency of communication.*
>
> *Whittaker (2003: 9)*

These theories assume a one-to-one correspondence between cues and functions, and if there are no extra-linguistic cues this would mean that CMC would have to be impersonal (Walther & Parks 2002). Among the earliest examples illustrating this approach is *social presence theory* (Short et al. 1976). Under this view, it is stipulated that social interaction at a distance benefits from the cues that are available in face-to-face interaction. More

---

[6] It should be noted that I use a broad definition of CMC, including, for example, both interactions mediated by personal computers and interactions mediated by mobile devices, such as cell phones.

[7] For introductions to and examples of these three areas of research, see e.g. Grudin 1994 (CSCW), Stahl et al. 2006 (CSCL) and Button 1992 (workplace studies).

[8] For previous, inclusive reviews of different approaches to CMC, see e.g. Walther & Parks (2002), Whittaker (2003), Nardi (2005) and Sassenberg & Jonas (2007).

specifically, it is hypothesized that in order to achieve efficient communication, the tools employed should support non-verbal interaction, such as gaze, facial expressions and prosody. Similar ideas were proposed by Rutter (1987, Rutter et al. 1987) in his *cuelessness model*, and subsequently by Daft & Lengel (1986) in their *media richness theory*. Media richness theory stipulates that media can be matched with activity types based on the types of cues supported. More precisely, it is argued that rich media is needed when expressing emotional, ambiguous and uncertain content.

As Whittaker (2003) points out, the bandwidth hypothesis has been disproven in much subsequent research. Alternative models have instead shown how mediated interaction might become more personal than face-to-face interaction. For example, Walther (1996) refers to how text-based communication is often *hyperpersonal*, as participants in text interaction might reveal more personal information online than they would in a face-to-face situation.[9]

Nevertheless, more recent research at least partially supports the notion that face-to-face interaction is superior. For instance, while acknowledging that it may well be possible to use lean media in rich ways and that it is misleading to focus too much on media bandwidth, Nardi (2005) maintains that:

> [a]t the same time, the impact of face to face experience remains unique. Popular discourse retains a clear distinction between 'RL' (real life) and what is experienced online, even among younger people who have been raised on a steady diet of technology. The efficiency of face to face communication in supplying high impact bodily experiences that promote affinity, commitment, and attention is undeniable. While face to face communication is most certainly information rich in valuable ways, the relational aspects of communication are also efficiently served with face to face communication.
>
> *Nardi (2005: 124)*

Nardi (2005: 125) goes on to state that she is "[...] generally in sympathy with arguments confirming the uniqueness of face to face communication", while reminding the reader that this is not a straightforward issue.

In the present work, face-to-face interaction is seen as a basic mode of interaction which can be referred to for comparison. However, it is not argued that face-to-face interaction is necessarily and always more efficient or personal than digitally-mediated interaction. Rather, the participants studied here are shown to exploit different affordances to deal with different tasks.

---

[9] In the *reduced social cues approach*, first advocated by Kiesler et al (1984), these ideas are taken to an extreme. Here it is hypothesized that CMC should lead to reduced self-awareness, which in turn would lead to behaviour breaking the social norms of face-to-face interaction (cf. Sassenberg & Jonas 2007: 276).

2.2 Tendency to ascribe either all or nothing to technology

The bandwidth theories discussed in the previous section exemplify an approach to CMC which puts high emphasis on the influence of the tools employed. In that respect, they might be seen as adhering to the idea of technological determinism. Contrarily, proponents of other approaches often tend to avoid ascribing importance to the inherent qualities of technological tools, but instead they emphasize the social construction of technology.[10]

Hutchby (2001) points to the divide between technological determinist and social constructivist approaches, and claims that social constructivist theories are problematic since they often fail to acknowledge the influence of technology completely. As an example of strong social constructivism, he refers to Grint & Woolgar (1997: 37), who argue that many theories within their own school demonstrate "a residual technicism" as they do not completely move away from the deterministic approach. Instead, Grint & Woolgar's alternative, according to Hutchby, is to see technology as a text, whereby the reader has the possibility to read, interpret and use technology differently than intended. Hutchby argues that in the technology as text approach:

> [i]t is the interpretation which makes of the technology what it is, rather than there being elements of the technology which constrain the possible range of interpretative moves that can be made [...].
>
> Hutchby (2001: 29)

Linell (2009) also points to this divide, and refers to how *monologists* tend not to take the users into account, whereas extreme *interactionists* disregard the properties of the technology and only focus on how it is used. Similarly, Nardi & O'Day (1999) review and discuss common metaphors for technology applications. They contrast the metaphors of *technology as tool* and *technology as text*, and introduce the more complex notion of *technology as system*, which incorporates ideas from both previously mentioned metaphorical systems.

Furthermore, Hutchby (2001) claims that:

> [t]echnologies do not impose themselves on society, mechanistically altering the pattern of human relations and social structure. Neither does human agency encounter technologies as blank slates, as infinitely malleable forms.
>
> Hutchby (2001: 206)

---

[10] See e.g. the introduction to Button (1992: 10-28) on the social construction of technology.

A fruitful middle road, he argues, is rather to give attention to the *affordances* that the artefacts offer (cf. also Linell 2009). In relation to interpersonal interaction, Hutchby introduces the notion of *communicative affordances*, to be reviewed in detail in Section 3.2.3. As will become apparent, I follow the view that a focus on communicative affordances presents an alternative to strong constructivist and strong deterministic approaches.

2.3 Tendency to perform de-contextualized analyses

One of the criticisms which Nardi & O'Day (1999) present regarding the technology as a tool metaphor is that the tool is of main concern and the context is thus left out. In fact, much previous research has been decoupled from context, as the empirical data often consisted of log files only without taking the situational context into consideration (cf. Jones 2004).

Log files are often important sources of data, and in many respects they provide the analyst with the same information that is available to the participants in distributed interaction. However, by only focusing on log files, there is a risk of missing out on potentially important information regarding, for example, exact timing of messages and co-occurring activities. Of course, the drawbacks and benefits of using log files only also depend on the aim of the analysis.

Jones (2004) criticizes analyses of CMC for seldom moving beyond the screen, thus reinforcing the false distinction between virtual and actual reality. Similarly, Walther & Parks (2002: 550) suggest that we should not limit our investigations to one mode of interaction, but that we need to consider what they refer to as *mixed-mode relationships*. Some CSCW research that focuses on workplaces has been moving in a direction where the complete, often complex, situation is taken into account (see e.g. Suchman 1987, 1992; Luff & Heath 1992; Goodwin & Goodwin 1996). Other examples of approaches taking the local and multimodal context into account can be found in multimodal discourse analysis (see e.g. Levine & Scollon 2004).

Not least the relatively recent possibilities to capture online interaction easily, for instance through key-stroke logging and screen capturing devices, have led to an increase in studies that include data other than log files only (cf. Erickson 2004). The question then becomes exactly how much and what type of data is necessary, as some researchers include all information available in their transcripts (see e.g. Markman forthcoming). From a purely interactional perspective, it could be argued that detailed accounts of, for example, revisions not visible to the other participant are irrelevant. However, if one seeks to explain the interrelation between different contextual dimensions, information concerning the activities of the individual is also of relevance in the analysis. In addition, when collecting rich situational data, ethical concerns arise which the researcher needs to resolve, and which

might restrict what types of data can be collected (see Section 5.1 for further discussion of this).

As we shall see, in the current study, it is argued that interaction is best analyzed in non-experimental settings, and that different contextual dimensions need to be taken into account when making sense digitally.

3. Theoretical framework

The work presented in this study mainly builds on an interactional framework, incorporating ideas from ecological psychology. In this section, some of the basic assumptions of the interactional approach are presented, and central concepts are discussed.

3.1 The interactional approach

Broadly speaking, interactional theories of language fall within the scope of empirical pragmatics, which can be defined as "[...] the study of the making or construction of meaning among people using language in naturally occurring interaction" (Korolija 1998: 31). The primary interest is thus to come to terms with how actions are carried out through communication. To employ an interactional approach entails an emphasis on language in use and a view of language use as something which is conditioned by the close interactional context. In other words, interest does not lie in hypothetical examples and well structured imagined utterances, as has often been the concern in traditional structural linguistics, rather in analyzing empirical observations and data (cf. Hutchby & Wooffitt 1998).

One of the most influential scholars in the interactional school is Erving Goffman. He introduced an interactional approach to sociological analysis, where the aim was to describe the *interaction order* (Goffman 1983). Among the first to point to the relevance of talk as data for analysis, he used different concepts, such as *frames* and *footing* to account for everyday activities, and the ways in which we adapt our behaviour depending on context (Goffman 1974, 1981).[11] Goffman (1963) also distinguished between different types of co-presence, and between focused and unfocused interaction, which he sees as expressing different levels of involvement that are relevant when analyzing interaction.

Other important contributions to the interactional perspective on sociology originate in the work of Harold Garfinkel. Garfinkel coined the term *ethnomethodology* to refer to "[...] the study of members' methods for producing recognizable social orders" (Rawls 2003: 546). This is an approach to the analysis of social interaction that does not depend on *a priori* theoretical constructs, but where the patterns identified in the interaction are used to draw conclusions about the orderliness of human interaction. One of the

---

[11] The notion of frames has been employed also by linguist scholars. See, for instance, Fillmore (1976) on *frame semantics* and Tannen (1993) on *framing in discourse*.

principles which Garfinkel has brought to the front is that all actions are *accountable*[12]. By this he means that:

> *[...] any social setting [should] be viewed as self-organizing with respect to the intelligible character of its own appearances as either representations of or as evidences-of-a-social-order. Any setting organizes its activities to make its properties as an organized environment of practical activities detectable, countable, recordable, reportable, tell-a-story-aboutable, analyzable – in short, accountable.*
>
> *Garfinkel (1967: 33)*

This implies both that participants in interaction make their contributions accountable (intelligible, describable) for the other participants, and that as researchers we can depend on this accountability when analyzing unfolding interaction.

Another basic tenet of the ethnomethodological approach is the *emic* perspective, which implies that the researcher should try to take the perspective of the participants in interaction (ten Have 2007). From a strong interactional point of view, this means that the researcher should focus on the actions of the participants as these are displayed to the others taking part in interaction, since what can be observed and analyzed is how participants in interaction react to each others' actions. However, from a less strong interactional viewpoint it can be claimed that also other ethnographic approaches, such as interviews, can be employed to gain an emic perspective. In the current study, this broader ethnographic view is advocated, something which is especially clear in Case study 4.

A central interactional approach, originating in ethnomethodology, is *conversation analysis* (CA).[13] This approach shares the ethnomethodological interest in taking an emic perspective and in avoiding a priori theories or categories. Instead, the material should be analyzed in detail in order to find emerging and recurrent patterns. This leads us to a CA hypothesis, namely that it may be possible to identify structural patterns at all levels of interaction. From the possibility that there is "order at all points" (Sacks 1984: 22), it follows that no level of interaction is too mundane to analyze. Even the most everyday of activities, such as conversing, will exhibit orderliness and this can be an important clue in the project of understanding human interaction.

Some structural components of interaction can be found in the turn-taking and sequencing systems. In face-to-face interaction participants take

---

[12] As will become apparent, this definition of accountability is different from the one proposed by e.g. Erickson & Kellogg (2000) as part of their *social translucence* concept, which also is relevant in relation to the current thesis (see Section 3.2.2 and Section 7.4).

[13] For introductions to conversation analysis, see for instance Levinson (1983), Hutchby & Wooffitt (1998), Markee (2000) and ten Have (2007).

turns at talking and reciprocally contribute to the on-going emergent project of conversation. Sacks et al. (1974) refer to *turn-constructional units* (TCUs), and show that when a TCU has reached its completion a transition relevance place (TRP) appears. Sacks et al. (1974: 704) identified three rules which apply when a TRP has been reached, and which participants orient towards when taking turns in interaction. In short, these rules state that if a next speaker has been selected, that person has to take the turn, and if no next speaker has been selected, either another person may self-select or the current speaker may continue.

Another relevant conversation analytic concept is *sequencing*. Conversations are not built up of random turns, but rather, turns are sequentially organized and are built for the sequential slot in which they occur. This is most clearly evidenced in the notion of *adjacency pairs*, where certain communicative actions require certain types of responses (Schegloff 1968). For example, once a question has been asked, a reply is normatively expected. Other adjacency pair types include greeting – greeting, request – acceptance/refusal, offer – acceptance/refusal, assessment – agreement/ disagreement and blame – denial/admission (Levinson 1983: 303-304, 336). It should be further noted that certain types of second pair parts are said to be *preferred*, and for a tactful participant, in most situations, it takes more effort to give a negative *dispreferred* reply than a positive preferred one (cf. e.g. Levinson 1983: 336).

The CA approach has received some criticism. For instance, Duranti (1997: 266) identifies three major problems with CA from the perspective of linguistic anthropology and ethnography, namely:

(i)   *a repeated disinterest in the 'larger context,' for instance, where and when the exchanges being analyzed took place, and a disregard for non-verbal or gestural aspects of face-to-face communication;*

(ii)   *a rudimentary notion of what constitutes speech (as demonstrated by a transcription system that does not take full account of the prosodic features of spoken language);*

(iii)   *a disregard for the interpretations that the participants themselves might provide of their own behaviour.*

Duranti (1997: 266)

Criticism has also been delivered in relation to the claim that analysts can take an emic perspective and completely leave out theoretical presuppositions, where it has been argued that it is impossible for a researcher to gain the same perspective as the participants (Hammersley 2003).

Other interactional approaches have succeeded in meeting some of this criticism. For instance, several scholars have acknowledged the importance of bringing in a larger context in the analysis of interaction. *Linguistic anthropology* (Duranti 1997), *interactional sociolinguistics* (Gumperz 1982)

and the *ethnography of speaking* (Hymes 1974) are examples of such approaches where contextual factors are prominent in the analysis. Similarly, many proponents of CA have come to include different semiotic modes alongside the verbal in their analyses (see e.g. Heath 1982, Goodwin 2000, Stivers & Sidnell 2005, Broth 2009, Mondada 2009).

In the current study, a multimodal approach is advocated, and it is argued that in order to increase the understanding of the situation analyzed, also background facts concerning, for example, participants and activity type should be taken into account. Also interview data, with the participants' own views on the topic, can be a valuable resource. When combined with a detailed analysis of unfolding action, this gives us a variety of important clues as to how mutual orientation is established and maintained.

3.2 Central concepts

3.2.1 Coherence

Coherence has long been a central concern in pragmatics. Formal pragmatics-based approaches to the concept are concerned with theoretical models explaining how *presuppositions*, or implicit assumptions, are linked through language (see e.g. Habermas 1983). Another direction has been to investigate coherence empirically, in actual conversational data (see e.g. Craig & Tracy 1983, Gernsbacher & Givón 1995, Korolija 1998, Tanskanen 2006).

Previous research on conversational coherence typically defines coherence on two levels: *local coherence* and *global coherence*. Local coherence concerns the influence of the surrounding text, the cotext, on coherence strategies, whereas global coherence is concerned with the bigger picture, relating to situational context and background knowledge (cf. Korolija 1998, Littlejohn 2002).

On the most local level, coherence research deals with the closely related concept *cohesion*. Several scholars have explained the difference between the two concepts (see e.g. Coates 1995, Tanskanen 2006), and there seems to be a consensus that cohesion relates to the links which can be found in the verbal content of the text, whereas coherence, while including cohesion, also involves links with additional contextual factors.

Halliday & Hasan (1976) present five strategies for textual cohesion. These are *reference*, where for instance pronouns are used to point back to previously mentioned antecedents, *substitution,* where instead another lexeme is used in its place, *lexical cohesion*, where lexical repetition is used as a strategy to convey links, *ellipsis*, where certain words are excluded but understood from the context and thus linking to the preceding text, and *conjunction*, where linking words are used to connect utterances.

In conversation, cohesive devices can help create coherence, but also other strategies are employed. For example, the structural patterns of conversation can be an important aid (Schegloff 1990). Because conversation is structured so that there are logical links between adjacent utterances, we assume by default that each new contribution is relevant in relation to the previous one, as long as no clear signs are presented which indicate that this is not the case.[14]

Highly relevant in relation to sequencing and coherence is the role of attention. Attention does not have to be seen as a purely cognitive phenomenon, but it can be analyzed as a *transactional resource* (Jones 2004: 28), or as part of how experiences are expressed in line with the *phenomenological concept of the mind* (Norris 2004: 4). Korolija (1998: 13) notes that "[...] coherence has to do with shared attention and coordinated sense-making [...]", and so she points to the fact that one prerequisite for coherent conversation in face-to-face interaction is shared attention. The importance of shared attention was emphasized by Goffman (1963) as he concluded that during encounters we move between *unfocused* and *focused engagements*, between *primary* and *secondary involvements* and between different *attentional tracks* (cf. Jones 2004).

Norris (2004) also discusses the importance of detecting attention levels in multimodal interaction, and claims that participants in interaction are able to switch focus rapidly between simultaneous actions. These can be mapped onto a continuum from fore-grounded via mid-grounded to back-grounded action. At their disposal in discerning which actions are fore-grounded, participants can rely on the ways in which *modal density* is created. According to Norris, it is either obvious that the actions of one particular mode are fore-grounded, implying that modal density is created through intensity, or, modal density is achieved by combining many different modes, through complexity.

This relates to the reasoning of *common ground theory* (Clark & Brennan 1991). Here, it is argued that participants in interaction continuously need to confirm that common ground has been reached, both regarding the process and the content of interaction. Clark & Brennan argue that it is not enough to seek for negative evidence, as apparent in, for example, repair requests or communicative breakdown, but that we also always seek positive evidence that current contributions have been understood by the other participants in interaction. They identify three types of positive evidence: *acknowledgements*, *relevant next turns* and *continued attention*. Furthermore, Clark &

---

[14] This is in line with Grice's (1975) conversational maxim of relation, stating that contributions should be relevant.

Brennan (1991) claim that grounding techniques differ depending on both medium and purpose of interaction.

Common ground theory has been criticized, for instance, by Koschmann & LeBaron (2003). In their analysis of interaction between competent practitioners and an intern in an operational room, they conclude that learning did take place, even though the participants did not reach agreement and common ground (cf. Stahl 2006). It should further be noted that common ground theory with its emphasis on mental activity, might be seen to conflict with strong interactional approaches where interest lies only in the actions of the participants as they are observable in conversational data. Whereas Clark and Brennan (1991) situate common ground theory in the conversation analytical tradition, Schegloff (1991) is critical of approaches suggesting that participants strive towards a common knowledge. He argues that such theories deal with the transfer of beliefs between participants in interaction, whereas conversation analysis prefers an explanatory model building on the ideas of intersubjective sense-making. In the current study, grounding is used to refer to processes by which participants in interaction indicate that conversations are on track. I believe that the types of positive evidence presented here are relevant in relation to display of attention, and as such they can be approached also from an interactional perspective.

Also when initiating conversation, we are apt at discerning the level of attention of the person we want to speak with (cf. Norris 2004, Baron 2008). This is a key to ensuring coherent conversation initiation. Mondada (2009) describes the detailed sequential work involved in moving from unfocused interaction to focused interaction in face-to-face encounters. She draws on Goffman's (1961, 1963) notions and illustrates how *pre-beginning sequences* (Schegloff 1979), where social contact is established, also display a logical structure.

As previously mentioned, and as is indicated in the following quote from Goodwin (1995), coherence can also be created on a more global level:

> *Coherence encompasses not only relationships between linguistic elements within the stream of speech, but also the fit between the action and content of an utterance, and the social situation within which it is embedded.*
>
> *Goodwin (1995: 118)*

Alongside *verbal devices* (in this case what Goodwin (1995: 117) terms *prospective indexicals* framing the current utterance and pointing forward to receive their meaning) and *sequential structures*, Goodwin discusses how coherence is maintained through *recipient design*. This implies that participants adapt their contributions to the recipient and their assumed background knowledge. A key in Goodwin's argument here is that participants have an *adaptive flexibility* (1995: 118), meaning that coherence can

be renegotiated as contextual conditions are altered, for example, as new participants enter the conversation.

Also Scollon & Scollon (2001: 53), analyze a combination of strategies, and in so doing they focus on *cohesive devices*, *schemata*, *prosodic patterning* and *conversational inference.* The cohesive devices they discuss are *reference*, *verb forms* and *conjunctions.* By schemata, Scollon & Scollon (2001: 57) mean an "expected sequence of activities", and in this category they include both world knowledge and adjacency sequences. Prosodic patterning is an additional way in which links between utterances in oral discourse can be indicated, both through intonation and timing. Finally, conversational inference refers to the ways in which participants in conversation continuously need to make inferences in real-time based on interpretations of preceding messages, something which is part of what has been referred to as *interactive intelligence.*[15]

A strong global perspective on coherence can be found in the inferential-strategic approach to coherence, introduced by Jacobs & Jackson (1983). Here, a game metaphor is used to explain how coherence is maintained during interaction. It is claimed that, just as when playing a game, participants in interaction have different objectives and need to follow the relevant rules applying to the situation at hand in order to accomplish their goals in a coherent manner.

Important to remember is that both local and global coherence is something which is presupposed by participants in interaction. When conversing, we take for granted that the contributions of other participants should be designed to make sense in the local context, and so we will strive to create coherence between even the most seemingly unrelated utterances (cf. e.g. Garfinkel 1967, Grice 1975, Coates 1995).

Previous studies of coherence in CMC have sometimes been concerned with what has become one of the most central issues in CMC research – the categorization of CMC as primarily written or primarily spoken interaction.[16] For instance, Tanskanen (2006) compares cohesive strategies in face-to-face interaction, mailing-list interaction, academic writing and prepared speeches, and concludes that the written dialogue in the mailing list shares attributes with both writing and speech.

Other approaches to the analysis of coherence in CMC have centered on the structural components of CMC and the ways in which coherence is maintained despite disrupted turn adjacency and a lack of simultaneous feedback (Herring 1999, Simpson 2005, Lam & Mackiewicz 2007, Lapadat 2007,

---

[15] Further examples of inclusive approaches to coherence in conversation can be found in the collections by Craig & Tracy (1983) and Gernsbacher & Givón (1995).

[16] Examples of studies dealing with the speech and writing-like qualities of CMC include Crystal (2001) and Hård af Segerstad (2002).

Markman forthcoming). Often these investigations have dealt with multi-party interaction, and the successful elements identified include *embedded repeats*, *adjacency pairs*, *sentence structure* and *speaker selection* (Markman forthcoming). Research has further identified additional obstacles to coherence maintenance, such as *multitasking* and *authority* (Woerner et al. 2007).

Nardi et al. (2000) investigate conversation initiation in relation to IM (instant messaging), showing that acknowledgement that a message has been received is often missing. They claim that this can either be because the other person is not available, or because they exploit the "plausible deniability of presence" which the IM medium affords (Nardi et al. 2000: 6). In face-to-face or telephone conversations, the initiator of a conversation has the highest degree of control, due to *participant asymmetry*. In IM, however, a higher degree of control is given to the receiver of the message, concerning when to read and reply. Baron (2008: 32) refers this to as *volume control* – different media provide participants with different options concerning the level of involvement and intrusiveness of incoming messages.

In the present study, strategies relating to cohesive links, conversational structures and display of attention are all taken into consideration. Furthermore, certain aspects relating to coherence on a more global level are included in the analysis.

3.2.2 Context

As previously stated, coherence is closely related to context. Goodwin & Duranti (1992) claim that it is impossible to provide a single definition of context. Instead, what we can do is to discern different ways of viewing the phenomenon. When reviewing the literature on context, it is possible to identify three main approaches to the subject. A first approach has been to describe context in relation to specific categories (see e.g. Hymes 1974, Halliday & Hasan 1985, Scollon & Scollon 2001). A second approach has been to view it as something which is actively construed during interaction (see e.g. Gumperz 1982, Goodwin 2000). Finally, a third approach has been to ignore the concept of context completely, and instead include different aspects of what would traditionally have been referred to as context in the description of multimodal action (see e.g. Scollon & Scollon 2003, Norris 2004).

As was seen above, one way to describe context is to specify what should be understood as belonging to a particular context by means of different categories. For example, Hymes (1974), introduces a complex set of contextual categories within his theory of the Ethnography of speaking, the initials of which form the acronym SPEAKING: *setting* and *scene*, *participants*, *ends* (purposes, goals, outcomes), *act sequence* (order of the event), *key* (tone, manner, spirit), *instrumentalities* (speaking style), *norms*, and *genre*.

A similar set of categories can be found in Scollon & Scollon (2001) where the categories included are *scene*, *key*, *participants*, *message form*, *sequence*, *co-occurrence patterns* and *manifestation*. A less complex categorization is available in Halliday and Hasan's (1985) model, where they distinguish between *field* (the topic of the interaction), *tenor* (participants) and *mode* (the mediational means).

Another tendency has been to view context as something which participants in interaction mutually construct. Gumperz (1982) shows how context is negotiated during interaction, by introducing the notion of *contextualization cues*:

> *[...] the means by which speakers signal and listeners interpret what the activity is, how semantic content is to be understood and how each sentence relates to what precedes or follows.*
>
> *Gumperz (1982: 131)*

Similarly, Goodwin (2000) introduces the concept of *contextual configuration*, whereby he emphasizes how participants in interaction draw on different semiotic resources when co-constructing their moment-by-moment context. Furthermore, he introduces the notion of *reflexive awareness* to account for public displays of attention. This in turn affects *participation frameworks*, the continuously adapted constellations of relevant participants in interaction.

The concept of contextual configuration has parallels in other approaches describing multimodal interaction (see e.g. Scollon & Scollon 2003, Norris 2004), in that dynamic aspects of context are emphasized. However, rather than referring to different semiotic resources as context, in approaches such as *mediated discourse analysis*, context is often seen as incorporated in the multimodal action, and thus figure is not distinguished from ground.

The notion of context becomes particularly interesting in CMC, which can be seen as *polycontextual*, as participants are simultaneously part of different contexts (Engeström et al. 1995, Leander 2002, Saarenkunnas 2004). Context awareness is an often discussed topic in CSCW, and several attempts have been made to develop tools enhancing awareness concerning attention levels and availability for interaction, ensuring, for instance, coherent conversation initiation (cf. e.g. Nardi 2005).

Early research on awareness was conducted at Palo Alto Research Center (PARC) with the implementation of the Portholes system for distributed collaboration (see e.g. Dourish & Bly 1992). Rather than promoting focused synchronous video interaction, this system allowed participants to passively gain awareness concerning the activities of their co-workers at any point in time, primarily through video images. The findings show that important functions of the system were to give information concerning availability for

interaction and to help create a sense of community, despite physical distance. Similarly, as part of their work on social computing, Erickson & Kellogg (2000) introduce the concept of *social translucence* to illustrate how interaction can benefit from a higher awareness of the concurrent activities in which distant participants engage.

Dourish (2004) discusses how to view context in relation to CMC, with a special interest in context awareness. He argues that designers should pay greater attention to the sociological conception of context as emerging during interaction. This is in line with the reasoning of Jones (2004), who in his most insightful contribution on the relationship between discourse and technology claims that we need to reconceptualise context in relation to CMC. He argues that much research on context depends on false dichotomies, such as virtual – material, figure – ground, sender – receiver and text – context. Referring to Scollon et al. (1999), Jones draws attention to the fact that interaction is increasingly *polyfocal*, and illustrates how it can be difficult to separate text from context; which action is to be seen as context to which when continuously switching in between? Instead of investigating figure – ground, he suggests that we should focus on *involvement screens* (cf. Goffman 1963), *mutual monitoring possibilities* and *attentional choreography*.

Similarly, studies of workplace practices have often dealt with polyfocal interaction. Some very relevant examples in relation to the current study can be found in analyses of interaction in flight operations rooms (Suchman 1992, Goodwin & Goodwin 1996), where it is described how mutual orientation is multimodally achieved.

In IM interaction research, one of the problems identified in connection with the lack of awareness cues concerns the way in which the notifications of new messages may interrupt current activities. Previous research shows somewhat contradictory results concerning the interruptive power of IM. IM communication has generally been seen as interruptive (cf. Rennecker & Godwin 2003); however, an interview study conducted by Garrett & Danziger (2007) showed that those participants who used IM reported on being interrupted less frequently than those who did not. This indicates that interruptions through IM might be seen as less disruptive than those face-to-face or over the phone.

In their research on context-aware mobile interactions, Danninger et al. (2008) make a further distinction between research on *interruptability* and research on *awareness*. Interruptability research deals with how to design technical solutions that can automatically help filter interruptions and awareness research is concerned with how to design tools which allow the users themselves to share contextual information. They present solutions for mobile telephony integrating both approaches, and show how user configurability can support users to successfully carry out conversations in a

coherent and context-sensitive manner. For instance, they present and test a tool which allows the receiver of a phone call to use pre-programmed touch-talk features. These suggestions will be further discussed in Section 7.4 dealing with implications for design.

As we shall see, I agree with Jones' (2004) propositions, presented above, in many respects, but I argue that it is fruitful to retain a figure – ground perspective for the sake of analysis. This means that in the present study, the current communicative actions are seen as the text (figure) and other factors, including just prior actions, are considered part of the context (ground). A more detailed description of how the interrelation between figure and ground is viewed in the current study is available in Section 4.

3.2.3 Affordances

Despite an increased interest in notions such as space and place as semiotic resources and their interrelation with discourse (cf. Goodwin 2000, Scollon & Scollon 2003, McIlvenny et al. 2009), relatively few studies stress the role of the material context in constraining and affording communicative action.

One way of emphasizing the importance of the material context is by incorporating ideas from ecological psychology. Gibson (1977, 1979) describes the complementary relationship between the animal (humans included) and its environment through the concept of *affordance*:

> *The affordances of the environment are what it offers the animal, what it provides or furnishes, either for good or ill. The verb to afford is found in the dictionary, but the noun affordance is not. I have made it up. I mean by it something that refers to both the environment and the animal in a way that no existing term does.*
>
> *Gibson (1979: 127)*

Building on theories of direct perception, Gibson (1977) argues that most affordances are directly perceivable, and so, extensive learning is not needed. Further, he claims that by focusing on the relationship between the animal and the environment, it is possible to find an alternative to subjective – objective dichotomies (Gibson 1977: 70).

In Gibson's view, affordances are facts of the environment, but only relative to the individual trying to act in the environment. Different species will perceive different affordances. In the concept, Gibson also includes man-made objects, and in reference to modified environments he claims that no distinction is needed between cultural and natural environments, as the former build on the latter and do not comprise "[...] a world of mental products distinct from the world of material products" (Gibson 1977: 70).

As has been pointed out by several scholars, the concept of affordance has been applied in different contexts and with different definitions.[17] It should be noted that I do not see it as my task to set straight the several and varying definitions of affordance. Instead, I would like to show what assumptions my own definition and use of the term build upon.

In the field of *human-computer interaction* (HCI), the concept has been introduced as a tool in interaction design, notably by Norman (1988, 1999) and Gaver (1991, 1992, 1996). Norman (1988) applies the affordances framework in his discussion of the design of everyday things, and mainly tries to find explanations for why some things are so difficult to use. He claims that the study of affordances is "[...] the study of the psychology of materials and of things" (Norman 1988: 9), and that affordances are but one of many factors influencing the usability of a designed object. Other factors that must be taken into consideration are cognitive models and constraints.

The earlier work of Norman has often been accused of creating confusion concerning the definition of affordances, as here he argues that affordances are not part of the objects, rather they only would exist if perceived. In later work, Norman (1999) tries to resolve this by distinguishing between real and perceived affordances. He concludes that what is most relevant for interface design concerns perceived affordances, cognitive models and constraints. Real affordances are, in his view, of more interest when dealing with physical objects.[18]

Gaver (1991) has stayed close to Gibson in his use of the concept, in that he also sees affordances as a result of direct perception. For example, he claims that:

> [t]he concept of affordances points to a rather special configuration of properties. It implies that the physical attributes of the thing to be acted upon are compatible with those of the actor, that information about those attributes is available in a form compatible with a perceptual system, and (implicitly) that these attributes and the action they make possible are relevant to a culture and a perceiver. [...] The actual perception of affordances will of course be determined in part by the observer's culture, social setting, experience and intentions. Like Gibson I do not consider these factors integral to the notion, but instead consider culture, experience, and so forth as highlighting certain affordances.
>
> Gaver (1991: 81)

According to Gibson, the most important affordances for an animal are provided by animals of the same species, and other humans are thus a central part of the environment (Gibson 1977: 75). In his discussion of

---

[17] For a comparison between and critique of different approaches to affordances, see for instance McGrenere & Ho (2000), Oliver (2005) and Rambusch & Susi (2008).

[18] McGrenere & Ho (2000:6) question Norman's claim that real affordances only are relevant in relation to physical objects, and claim that "application software also provides possible actions".

affordances for interaction and sociality, Gaver (1996) changes the emphasis slightly in that he chooses not to deal with the affordances that people provide each other (*social affordances*), but rather how the physical environment influences possibilities for interaction (*affordances for sociality*):

> Research on 'social affordances' [...] focuses on the possibilities for action that people offer one another and on the role of other people in pointing out new affordances (e.g., to babies). What I am concerned with here is the possibilities offered by the physical environment for social interaction. These are not social affordances, as defined above, but affordances for sociality. I believe they offer new opportunities for basic research and a powerful tool for design.
>
> *Gaver (1996 :3)*

The view of affordances proposed in the current study is closely related to Gaver's notion of *affordances for sociality*, in that the main interest lies in the interrelation between the environment, the tools employed and conversational patterns. However, other participants are included in the larger contextual model provided (see Figure 3, Section 4). These areas are separated, isolating the influence of the design of the material environment and the digital tools employed, while still emphasizing the role of other participants in highlighting affordances and influencing the norms that develop among users.

In CMC research, the importance of the material context and its affordances has been discussed by Hutchby (2001), who introduced the concept of *communicative affordances* to explain "[...] the relationship between forms of technology and structures of social interaction" (Hutchby 2001: 2). Building on the work of Gibson (1977, 1979), Hutchby claims that this concept makes it possible to find a middle road between constructivist and deterministic approaches. He criticizes both these models, and claims that a focus on communicative affordances acknowledges the influence of the materiality of technology while not arguing for technological determinism. Although technology provides affordances and constraints regulating how we can interact with and through it, we are often also able to adapt the technology to fit our needs.

Hutchby further couples the concept of affordances with insights from CA, and claims that communicative affordances are reflected in conversational interaction. From a communication perspective, it could be argued that affordances relate to the options for action that are relevant when we communicate with others. The options chosen are reflected in conversational structures and conventions, and Hutchby shows that by looking at specific instances of conversation, it is possible to investigate what affordances are actually perceived and utilized in a specific communicative situation.

Hutchby exemplifies research within different areas of technologically-mediated interaction and shows how communicative affordances might influence the norms for conversation. He begins by applying his theoretical framework to the analysis of telephone conversations, and in so doing he argues that we should focus on the ways in which humans have dealt with the communicative affordances of the telephone by adopting certain, sometimes innovative, forms of talk-in-interaction. Among other things, his findings show how telephone conversations afford "intimacy at a distance" which has resulted in, for instance, new conventions for opening sequences (Hutchby 2001: 81; cf. Arminen & Weilenmann 2009). Video-mediated interaction is also briefly touched upon by Hutchby (2001: 124-130). Here he argues that the main obstacle to successful interaction in this context is our expectations on how interaction should unfold, based on our ideas of proto-typical face-to-face interaction. We fail to acknowledge that the medium has certain affordances and constraints, relating for example to the quality of the video images and to the limited opportunities to employ gaze as a semiotic resource.

Hutchby's notion of communicative affordances has caused some debate. For instance, Rappert (2003) argues that Hutchby's description of the constructivist approaches is unjust, and that this middle road between determinism and constructivism has been suggested by others before him. Another type of criticism is presented by Nardi (2005), who takes research dealing with affordances as an example of the channel metaphor, thus grouping it with approaches such as social presence and media richness.

In the following section, the proposed analytical model will be presented. Here, I will also return to and comment on some of the criticism presented above.

4. Towards an analytical model

As will have become apparent, it is possible to find a number of studies concerned with coherence and context in CMC. What my contribution aims to add is a multimodal analysis of coherence where different types of strategies are taken into account. Furthermore, the current research provides an empirical investigation across different types of contexts where rich situational data have been collected. In addition, I focus on the role of technology, but strive to strike a balance between constructivist and deterministic approaches. My research also provides a model combining a number of previously suggested concepts, such as coherence, context and affordance, allowing for an analysis of their interrelations.

In the current study, coherence is used as an umbrella term to refer to different strategies employed to establish connections between related actions. I am mainly interested in local strategies of maintaining coherent interaction in light of the communicative affordances in the communicative ecologies in which the interaction takes place. I acknowledge that people will try to make sense of utterances despite lack of cohesive links, and that conventions develop which also influence coherence strategies. I am also aware that other contextual variables influence the strategies employed. Not least activity types (Levinson 1992) are of importance, and considerations such as whether the activity is, for example, private or urgent, and whether it involves a set agenda need to be addressed.

While I acknowledge that context is dynamic and emergently co-constructed during the course of interaction in that relevant contextual resources are acted upon as interaction unfolds (cf. Goodwin 2000)[19], I also emphasize the importance of the material context in interaction, not only as a semiotic resource, but as an environment providing affordances and constraints in relation to the interaction taking place (cf. Gibson 1977, 1979; Gaver 1991, 1992, 1996). I propose to do this by giving special attention to the communicative affordances of different tools and environments and the ways in which these are acted upon during interaction. This also means that while I acknowledge that in practice different components of context are intertwined, in the model of analysis proposed here, these components are separated for analytical purposes.

---

[19] It can be noted that, in fact, parts of what has been incorporated in the category of coherence in the current thesis (such as display of focal attention and mutual orientation) could also be seen as contextual configurations (Goodwin 2000). Shifts in focal attention, activities and participant frameworks lead to shifts concerning which parts of the context are to be seen as relevant at a particular point in time, and the ways in which these shifts are made explicit ensures coherence. This emphasizes the blurry line between text and context.

In this study, communicative affordances are defined as those properties of the material context relative to the acting participants that affect how they perceive their own communicative options and others' communicative actions. It should be noted that my analysis of communicative affordances differs slightly from Hutchby's (2001) in a few respects. For one, I do not only deal with affordances of single tools, but also of complete environments including different tools, which in one sense is a return to Gibson's (1977, 1979) original use of the concept. Similarly, while I want to separate aspects of the material context from the dynamic context (cf. Gaver 1996), by including the dynamic context in my model, I invoke a broader ecological perspective than Hutchby. In addition, I propose to relate affordances to coherence and sense-making, not only to structural patterns.

Furthermore, in my view, a model of communicative affordances has very little in common with, for instance, social presence theory, as Nardi's (2005) analogy would suggest. A basic underlying assumption of the proposed model is that a focus on communicative affordances allows us to investigate specific environments in their own right. With regard to communication technologies, different media affect interactional patterns in different ways, yet, this is not to say that richer media are per default more efficient. Rich media have certain benefits and lean media have others.

In addition, I agree with Hutchby, in that an emphasis on communicative affordances can provide a valuable third stance in between radical forms of technological determinism and social constructivism, acknowledging the ways in which technological artefacts influence norms for interaction, while ascribing participants in interaction an important role in forming those norms.

The analytical model developed builds on previous conceptions of context. Jones (2004: 24) identifies three dimensions of context which are recurring in previous accounts on the topic, namely *setting*, *activity* and *participants*. These are represented in Figure 1.
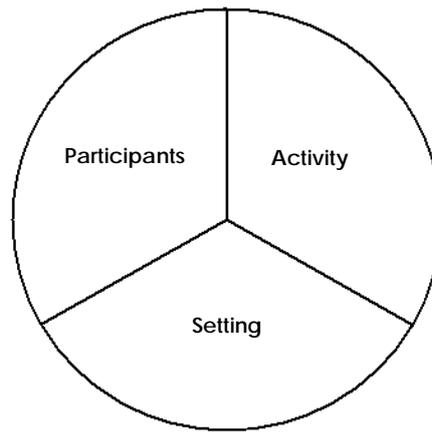
Figure 1: Contextual dimensions identified in Jones (2004)

Korolija (1998: 47) similarly summarizes previous research as proposing three kinds of major context types, but in identifying these contextual resources for coherence, she employs a slightly wider scope. The three dimensions she identifies are *cotext* (prior discourse), *situation* and *abstract background knowledge*, represented in Figure 2. Here it could be argued that the notion of situation might encompass all three dimensions included in Jones' (2004) summary.
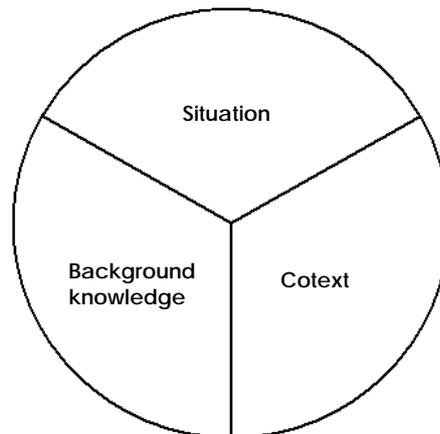
Figure 2: Contextual dimensions identified in Korolija (1998)

The models above illustrate dynamic contextual factors, where the environment (or the setting) is included alongside other contextual resources for sense-making, as is often the case in interactional approaches to context. As previously mentioned, I suggest that we should also include insights from ecological psychology and single out the material context, as it not only provides resources, but also sets constraints upon the available options for interaction. The model that I advocate is therefore a synthesis between interactional and ecological approaches.[20]

Figure 3 depicts the proposed analytical model. Here we can distinguish between figure and ground at two levels. First, the material, environmental context (physical or digital), which is stable, is seen as the ground, and the situationally invoked dynamic context is seen as the figure. Second, a further distinction can be made between the dynamic interaction itself and the complete context of interaction, where the former is the figure and the latter is the ground.[21]

---

[20] A similar synthesis has been suggested by Linderoth (2004), who combines concepts from Goffman (1974) with Gibson's (1977, 1979) notion of affordances. In short, he argues that affordances are the whole range of possible options for action available in an environment, whereas frames suggest which options are relevant in a particular situation.

[21] As previously mentioned, the figure – ground relationship has been criticized in previous research, as it can be seen to enforce dichotomies between text and context which some want to avoid (Goodwin 2000, Jones 2004). In the current study, however, such a distinction is seen as relevant in order to highlight the interrelationship between material context, dynamic context and strategies for coherence maintenance.
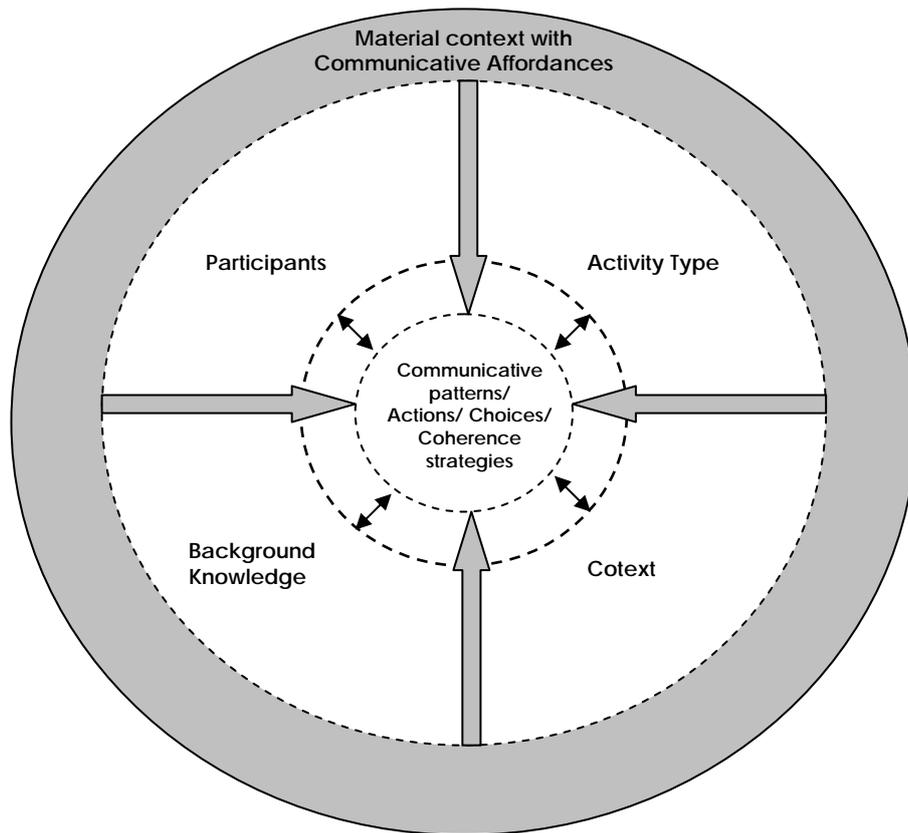
Figure 3: A model of coherence in context

The components of this model result from a combination of the models presented in Figures 1 and 2, complemented with elements from ecological psychology. Participants and activity type, which are included here, are both available in Figure 1, and are included in the notion of situation in Figure 2. The other two suggested dimensions of dynamic context, namely background knowledge and cotext, are taken from the model presented in Figure 2.

The current model illustrates that all interaction is part of a context with different *participants*, pursuing similar or diverse goals[22], and forming new social conventions in relation to other contextual dimensions, such as material context and activity type. By accommodating to the skills and goals of the

---

[22] This is in line with *activity theory* (see e.g. Nardi 1996) where the goal, or the object, is given central prominence.

other participants, and by interpreting contributions made by others in relation to their skills and goals, breakdown might be avoided (cf. Goodwin 1995). Participants have different *background knowledge* concerning the tool employed, the participants involved, the topic at hand and the activity type, and so when they approach a new situation, their cumulated previous experiences provide initial frames for action (cf. Goffman 1974). Furthermore, the activities in which the participants engage can be grouped into activity types (Levinson 1992), and consist of different tasks with different purposes (cf. Clark & Brennan 1991). By following the rules of the specific task, coherence can be created globally. On a more local level, interaction is also influenced by the multimodal cotext in which it is situated (cf. e.g. Korolija 1998).[23]

Acknowledging the view that each new action, for example, a turn at talk, is both *context-shaped* and *context-renewing* (Heritage 1984: 242), the model also shows that communicative actions both influence and are influenced by dynamic contextual dimensions. Also material context is dynamic in the sense that not all aspects of the material context are made relevant for interaction. This is in line with Goodwin's (2000) thoughts on contextual configurations. However, by separating material context from dynamic context, it is possible to distinguish how the environment influences the possibilities for action in relation to the participant.

This inclusive model can be applied both in relation to digitally-supported and face-to-face interaction, and can be complemented with an additional model portraying communicative affordances of the material context at different levels in technology-rich environments. Such a model is depicted in Figure 4.

---

[23] It should be noted that the description here is mainly concerned with situational context. Of course, all action is also part of larger institutional contexts which influence, for example, the goals of the participants and the norms associated with specific activity types (see e.g. Nardi 1996).

Modal constraints
(adapted from Clark and Brennan's
Constraints on Grounding)

| Co-presence | Audibility | Simultaneity | Reviewability |
| Visibility | Cotemporality | Sequentiality | Revisability |

Affordances of communicative tools
(Online environment)

Influences of the properties of the hardware and the software interface and modal constraints on:
- Perceptibility of own communicative options
- Perceptibility of others' communicative actions

Environmental affordances
(Physical + digital environment)

Tool affordances + affordances of the complete communicative ecology

Influences of the physical and digital surroundings, proximity, distance and spatial cues on:
- Perceptibility of own communicative options
- Perceptibility of others' communicative actions

Figure 4: Communicative affordances in different layers of the material context

The first layer concerns modal constraints, and here I use Clark & Brennan's (1991) categorization of constraints on grounding. They identify eight factors influencing grounding, summarized in Table 2.

45

Co-presence	A and B share the same physical environment.
Visibility	A and B are visible to each other.
Audibility	A and B communicate by speaking.
Cotemporality	B receives at roughly the same time as A produces.
Simultaneity	A and B can send and receive at once and simultaneously.
Sequentiality	A's and B's turns cannot get out of sequence.
Reviewability	B can review A's messages.
Revisability	A can revise messages for B.

Table 2: Modal constraints on grounding (adapted from Clark & Brennan 1991: 141)

Needless to say, designers of technological tools for interaction influence how their tools could be used by choosing which modes to include during the development process. Moreover, other features incorporated in systems and interfaces influence how we can express and perceive communicative action. Taken together, these are the affordances on the tool level. The combinations of different tools together with additional affordances in the physical environment are referred to as environmental affordances. The different layers of the model presented in Figure 4 will be further exemplified in relation to the case studies (Section 6).

5. Methodological considerations

The case studies which make up the basis for the current analysis deal with the following topics: IM interaction among a group of international students of interaction design, desktop video conferencing among a group of Swedish students of English at a distance, conversational multitasking from the perspective of one participant at her work computer and conversation initiation among a group of international students at a design school.[24] Below, I describe the methods employed for gathering and analyzing data for each of the four case studies, as well as discuss ethical concerns and different approaches to the role of context from a methodological perspective. For more detailed descriptions of the methods employed, see the corresponding papers.

5.1 Methods of gathering data

All of the case studies informing the current analysis were carried out in the ethnographic tradition, and a range of methods were employed. Whereas some define ethnography in relation to the methods used, such as observations, interviews and so on (Hammersley & Atkinson 1995), others are less specific concerning methodology:

> *Ethnography is not a particular method of data collection but a style of research that is distinguished by its objectives, which are to understand the social meanings and activities of people in a given 'field' or setting, and an approach, which involves close association with, and often participation in, this setting.*
>
> *Brewer (2000: 59)*

Apart from some of the log files in Study 1, which were given to me after the event, I was present as an observer in all the studies. In an effort to avoid affecting the interactional patterns studied, I took a passive role in all observations. However, as stated in the classical *observer's paradox* (Labov 1972), it is possible that my mere presence might have affected the patterns found. Nevertheless, whereas certain topics might have been avoided, I would argue that structural patterns would most likely not be influenced by my presence.

The methods employed to gather material varied both depending on the purpose of the specific study and on different ethical considerations that had to be met. An important ethical issue concerns informed consent. The main participants observed were informed about the purpose of the study in general terms, and were also told that they could withdraw from participating at any time if they would want to. All gave their approval to being

---

[24] It should be noted that Study 1 and Study 4 are based on materials from the same site of investigation, a design school, and that the participants as well as the log files analyzed are the same in both studies.

observed, whereas a few participants chose not to take an active part in my investigation.[25]

A more difficult question, however, concerns how to receive consent from remote participants with whom the main participants might interact. Ethical guidelines provided by the Association of Internet Researchers (Ess et al. 2002) state that when conducting ethnographic studies online, one should consider issues such as the sensitivity of the topic of investigation and the perceived privacy of the participants.[26] When observing a person at his or her computer, it is likely that private channels will be used, such as IM or email. Therefore, I sought legal advice concerning this issue, and was told that IM interaction could be compared with telephone conversations, which according to Swedish law can be recorded, but only by either party involved in the conversation. Details concerning how I came to terms with these issues will be provided below.[27]

Table 3 provides an overview of the contextual dimensions in focus in the different studies and also includes information about associated material.

---

[25] In the papers, I use fictional names and nicknames for all participants. Images shown (for instance the screen shot from FlashMeeting in Study 2) are arranged for inclusion in the papers and do not depict any of the original participants.

[26] Ågren (2000) provides another perspective, as he brings up copyright issues and discusses how instead of quoting anonymously, one might have the responsibility to give credits to the author in order not to break copyright laws.

[27] Concepts such as *virtual ethnography* (Hine 2000) have been used to refer to ethnography on the internet. This often implies making contact with and observing people with whom one has not had prior contact in other contexts. As this is not the case in my studies, further considerations pertaining in these situations will not be dealt with here.

	Material context	Dynamic context	Collected materials
Study 1	Communicative affordances of instant messaging (IM)	Cotext; material context	Log files
Study 2	Communicative affordances of FlashMeeting (FM)	Cotext; activity types; material context	Screen capture of online sessions
Study 3	Communicative affordances of IM and Internet Relay Chat (IRC) as well as the computer interface combining them	Cotext; activity types (Field); participants (Tenor); material context (Mode)	Log files; blurred screen capture
Study 4	Communicative affordances of physical and digital environments	Cotext; activity types; participants; material context	Log files; observational notes; diaries; interviews

Table 3: Overview of contextual foci and collected materials in the different studies[28]

As can be seen here, the data analyzed in the case studies include log files only (Study 1), video recordings of all modes of the online interaction (Study 2) and log files complemented with screen recordings and/or observations (Studies 3 and 4). As I suggest that the researcher should try to gain an emic perspective (that of the participants, cf. e.g. ten Have 2007), an important guiding principle in the interactional analyses included in my research is that gaining access to the same information that the other participants in interaction have access to is sufficient for analytical purposes. Whereas information about details concerning, for instance, edits made before sending a message off can be relevant when analyzing the linguistic production of one participant (see Study 3), it is not seen as necessary from an interactional viewpoint, as the other participant in the interaction does not have access to this information either. Nevertheless, if we want to learn more about how other simultaneous involvements influence interactional patterns, data concerning the individual context in front of the computer is needed.

As previously mentioned, in addition to interactional analyses, in some of the case studies I also make use of, for example, interview data. This is in line with the broader ethnographic scope advocated, where it is argued that

---

[28] Background knowledge, which is also included in my model, has not been the focus of any of the four studies. Nevertheless, it has been addressed, for instance, in relation to convention formation.

these data also help provide an emic perspective. Obviously, the views of the participants are in focus in the interviews, and this data is quite different from unedited moment-by-moment sense-making as available in recordings. Nevertheless, both types of data can be of aid as the researcher tries to gain insights into the activities of the participants.

The material for Study 1 was provided to me through log files from in-group interaction collected by a group of students at a design school.[29] A drawback with this material is that I do not have access to the information about when people start typing that is displayed to participants in interaction, nor do I have access to the exact timing of messages.

In Study 2, I participated in online meetings as a user called "Camera" to record interaction in the different modes available through a desktop video conferencing interface. I refrained from using a microphone, but could provide technical support in the text channel if I was asked to. In the sections included for analysis, however, I do not contribute to the conversation at all. Nevertheless, my use of the text for technology-related topics might have influenced the way in which the participants themselves used this mode. Furthermore, it should be noted that whereas the other studies took place in completely natural settings where the participants went about their everyday business, this case was slightly different. Here, I observed students of English at a distance, and for their first meeting they were neither acquainted with each other nor with the tool they were using. As I was able to follow them during their six sessions online (one test session and five discussion sessions), they had the time and opportunity to learn more about the medium and each other as time passed.

During Study 3, I was present as a passive observer for two hours which the participant spent working and interacting at the computer. For ethical reasons previously discussed, I refrained from capturing the detailed screen interaction on video, and instead taped it with a blurred focus. I asked the main participant to inform her interactional partners that she was being observed and to ask them if they would mind sharing the log files with me, and all agreed.

In Study 4, in-group interaction was investigated, thus making it possible to receive informed consent from all parties. My role varied from that of a participating observer during breaks, where I was involved in the conversations myself, to that of a shadowing observer as the students were located at their computers. Here, I followed three participants closely for two days each and tried to observe their complete communicative interaction. They were

---

[29] I observed some of these IM conversations as I was shadowing the participants, but as I do not have observational notes in relation to all log files, I have chosen not to include my observations in the current analysis. These are instead included in the analysis in Study 4.

encouraged to inform their communicative partners that they were being observed, and they were asked to share log files from in-group interaction with me. In this study, I never recorded the computer screens, and apart from some unsystematic video and audio recordings, the data I have access to are observational notes and log files from text-based interaction as well as diaries and interviews. A drawback with this material is that it is difficult to conduct detailed analyses of the moment-by-moment meaning-making in which the participants are involved.

5.2 Methods of analyzing data

Undoubtedly, analysis of ethnographic data begins already during the collection phase. This is especially noticeable if the data collection methods consist, in part, of observational notes, as is the case in Study 4, since important choices are made when deciding what to include in the field notes. Furthermore, if a video camera is used, the positioning of that camera will provide a specific perspective, influencing the view of reality to which the researcher gets access.

A second step of analysis has been to compose transcripts of the data. Even though much of my material consisted of typed log files, these files had to be merged. In Study 3, log files were matched with the time stamps from the blurred recording of the computer screen. Similarly, in Study 4, log files were matched with observational notes. In Study 2, I produced transcripts of the audio and video channel and combined this with transcripts of actions in the other modes.

As previously stated, I have strived to take an emic perspective in an attempt to have the same possibilities to make sense of the interaction as the participants themselves have. This is relevant when analyzing contexts, coherence and attention levels. However, it is important to remember that as a researcher approaching the material after the event, I have access to the complete transcribed sessions. This may affect my trying to investigate the moment-by-moment sense-making (cf. Korolija 1998, Linell 1998).

A related question concerns the notion of context itself, and more specifically whether the researcher should draw upon background knowledge in the analysis. There are different positions here, where strong proponents of CA claim that only the parts of the context that are being made relevant by the participants should be considered, and that no background information is needed. They refer to those taking in external contextual labels as using "bucket theory", where "[...] some preestablished social framework is viewed as 'containing' the participants' actions" (Drew & Heritage 1992: 19). Others claim that it does no harm to include information about, for instance, gender

and age in the analysis.[30] I have chosen to include general background information concerning the participants, such as age, gender and native language.

The methods employed for analyzing the empirical data in the current study are mainly qualitative, and draw upon insights from CA. For example, I have dealt with CA notions such as turn-taking and sequencing in most studies. However, since I do not have access to recordings of the unfolding interactions in all contexts, I have not been able to practice CA fully in all instances. I have also chosen to complement my qualitative CA inspired analyses with quantitative measurements on some occasions. For example, I have measured instances of intertwined threads and cohesive strategies (Study 1); participation rates based on number of messages, words included in typing and times holding the floor when speaking (Study 2); switching times, response rates and correlations between questions asked and switching times (Study 3); and topical distribution (Study 4).

An important assumption in CA and ethnomethodology is to avoid a priori conceptions about what will become relevant in the material. Whereas I have approached my material open-mindedly in all instances, I have also had theoretical considerations in mind. In some of the studies (1 & 2), I had quite clear questions in mind before approaching the data, and have looked mainly for certain types of phenomena. Of course, sometimes additional phenomena have emerged which I have also taken into consideration.

As will have become apparent here, I draw upon the reasoning of CA, but I do not follow the framework completely. Nevertheless, I would like to argue that it can be fruitful to refer to those aspects of this approach which are applicable in the current context, especially as they can shed light on the orderliness of seemingly incoherent interaction.

---

[30] See e.g. McHoul et al. (2008) for a discussion of these different views.

6. Summary of case studies

One basic assumption in CA is that when we study interaction, we should normally be able to identify "order at all points" (Sacks 1984: 22). This is to say that there is a logic to every detail in social interaction, both regarding minute actions, such as timing of gestures, and more official actions, such as turn-taking. The current study tests this claim by trying to identify structures at different levels of detail. The foci of the case studies were chosen to shed light on coherence in relation to communicative affordances on different levels. Figure 5 illustrates this (cf. also Figure 4, Section 4).
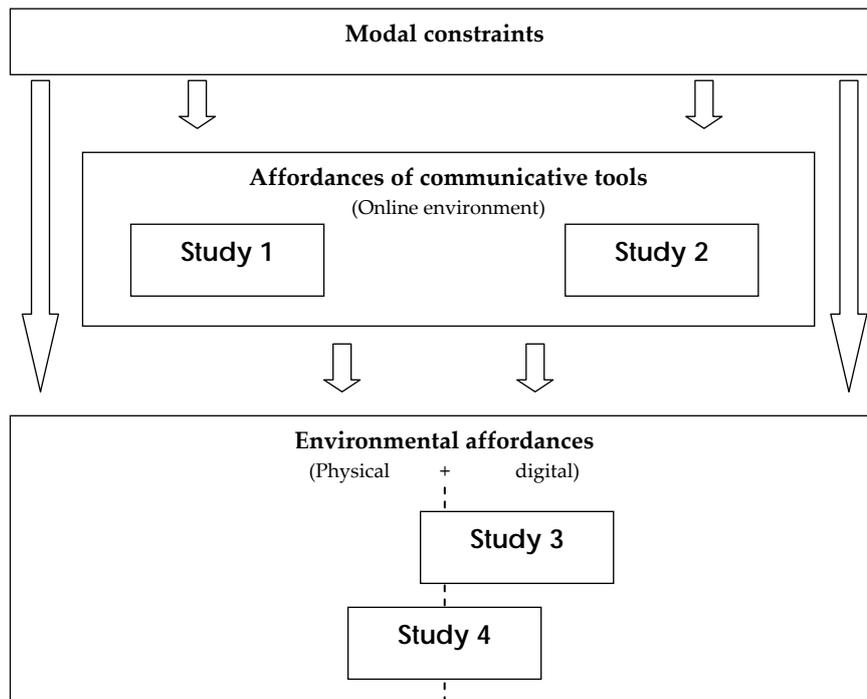
Figure 5: Mapping of the foci of the four case studies

Here, we see that Study 1 and 2 deal with affordances on the tool level, while Study 3 and 4 are mainly concerned with environmental affordances. Study 1 centres on text-based interaction only, whereas Study 2 has a somewhat broader focus and deals with interaction through a multimodal tool, supporting text, audio, video and graphical symbols. In Study 3, the scope is further broadened to include several simultaneous interactions through different

tools, and the emphasis here is mainly on digital options for interaction from the perspective of one individual. Study 4, which is based on group interaction data, combines the analysis of the employment of digital tools with an analysis of interaction in a shared physical space.

The model can be complemented with a summary of the practical problems which are investigated in the different studies as presented in Table 4.

How do participants establish and maintain coherence when…	
Study 1	… conversational threads are intertwined?
Study 2	… turn-taking is governed by the tool and when multiple modes are available?
Study 3	… multitasking leads to involvement in several conversations simultaneously?
Study 4	… interaction can take place both via digital tools and face-to-face?

Table 4: Initial practical problems investigated

In the following, the four case studies are described in more detail.

6.1 Case study 1: Disrupted turn adjacency and coherence maintenance in Instant Messaging conversations

This study deals with coherence maintenance in IM (instant messaging) interaction. The material consists of log files from IM interaction among a group of international students (25-30 years old) at a design school, who were all familiar both with the tools and with each other. In total, the log files comprise of 1689 messages, all produced during dyad interaction.

In order to investigate whether disrupted turn adjacency was an obstacle to coherence in the present data, instances where messages referred back to a message other than the one immediately previous were taken as the core of the analysis. The results show that in the material analyzed, disrupted turn adjacency did not cause coherence problems. Sometimes the participants made use of explicit cohesive devices to create clear links between utterances, but often they used general anaphoric reference despite the potential ambiguity of this strategy. This leads to the conclusion that disrupted turn adjacency was not seen as an obstacle to coherence creation by the participants themselves. Instead, most cases of potential ambiguity could be sorted out through the help of timing, replication of sequential structures and distinctions among different types of feedback. It is of interest to note that sequencing has been discarded as a sense-making strategy in text-based

CMC (Markman forthcoming). Contrarily, the findings presented here suggest that sequencing can reveal important cues, even though the structures are different from those in face-to-face interaction.

Additional signs of problematic coherence maintenance were also identified, and here, the focus was on explicit signs of miscommunication. The findings suggest that the few problems identified relate to difficulties in grounding on different levels. For example, excuses for what was perceived as belated replies often included descriptions of other simultaneously ongoing activities. Had it been possible to gain a better understanding regarding the availability of the other participant, this would not have been as problematic. Other explicit signs of miscommunication occurred when participants would refer to things mentioned in previous conversations or in interaction with other people, where common ground would mistakenly be taken for granted, and when misunderstandings due to for example typing errors could not be easily repaired in the textual medium.

Based on the results of this study, some suggestions for design are presented. One suggestion concerns a possibility to link more easily between messages in different conversations and different threads, for example by supporting simplified quoting. Another suggestion concerns possibilities to indicate availability for interaction clearly through automatic logging of current activities.

6.2 Case Study 2: Multimodal student interaction online: an ecological perspective

In this study, the focus is on interaction among a group of Swedish students of English at a distance as they take part in online discussions. There were between four and five participants in the analyzed sessions, and they were of different ages and had varying experiences with tools for online communication. The tool employed is called FlashMeeting[31], and is a multimodal desktop video conferencing environment, which at the time of investigation (2006) supported interaction via voice, video, text and graphical emoticons. A defining feature of this particular tool is the broadcasting system which ensures clear turn-taking as only one participant at a time can take the floor. If someone else is already broadcasting, you may line up and if more than one participant lines up there will be a queue where the first person in line will get the floor as soon as the current speaker steps down.

In total, five discussion sessions were recorded, and the first and the last session were chosen for further analysis. Two types of analyses were carried out. First, participation rates in the different modes were calculated and then the strategies employed to deliver conversational feedback were investigated.

---

[31] See http://www.flashmeeting.com for more information.

Results concerning participation rates show that it was the same people who were the most active in both the broadcasting and the text during both sessions, which may be seen as quite discouraging from a language learning point of view. The analysis of general participation patterns also reveals that both turns at talk and pauses between turns were unusually long here as compared to face-to-face interaction.

In the investigation of conversational feedback, examples of Clark & Brennan's (1991) three types of positive evidence in grounding were identified: *acknowledgements*, *relevant next turns* and *continued attention*. It was hypothesized that in this particular context, positive evidence was particularly relevant for three reasons:

> *1) the lacking possibilities for simultaneous audio feedback,*
>
> *2) the unusually long turns at talk, and*
>
> *3) the language learning setting.*
>
> <div align="right">Örnberg Berglund (2009: 198)</div>

However, the findings indicate that clear links between utterances were often missing. One explanation for this was sought in the design of the task, where students were to reply to a set of specified questions to receive their grades and they often started their contribution by returning to the main task, thus creating coherence on a global level. Another explanation could be found in the design of the tool, where the broadcasting device sometimes would make it difficult to create immediate links to the preceding contribution. As the students had to wait for their turn, they might have lined up in response to something mentioned early on or even in a previous turn, and if the current speaker chose to move on to another topic, it might have taken extra effort to create local coherence.

The resulting turn-taking patterns sometimes resembled disrupted turn adjacency in text-based CMC, where links are created through extended local coherence, or not at all. Here we also see similarities with patterns found in highly formalized seminars where there is a moderator present.

Furthermore, this study addresses how multimodal feedback and usage of the different modes could indicate display of attention. Suggestions relating to both task and tool design are presented in response to issues relating to coherence maintenance and to the fact that students had not learnt to master the communicative affordances of this multimodal tool fully.

6.3 Case Study 3: Multiplex conversations afforded by technology

The stated aim of this study does not explicitly deal with coherence maintenance, but it is possible to draw conclusions concerning coherence based on the results. Here, a 28-year-old participant with Swedish as her native lan-

guage was observed during two hours at her computer at her work in a computer-related field. During the observation, she was involved in conversational multitasking, participating in conversations over IM and Internet Relay Chat (IRC), as well as conversations in the physical space. At the most she was involved in five simultaneous conversations.

The theoretical basis for this analysis is Halliday & Hasan's (1985) model concerning influences on register. This model specifies three factors influencing register: *Field* (topic of interaction), *Tenor* (participants in interaction) and *Mode* (meditational means). The effects of these factors were investigated on two levels: register of the participant and switches between different modes and conversations.

The findings show that the main participant used different registers depending on conversational partner in the IM conversations, indicating that social conventions play an important role in coherence maintenance. The discussion concerning switching deals with sequencing, and specifically with so-called adjacency pairs (Schegloff 1968, see Section 3.1). One investigation concerned whether the participant was able to produce second pair parts in adjacency pairs where this was expected, despite the fact that she was involved in several simultaneous conversations. The results here show that she was always able to do so, even if sometimes there would be a lapse in time before her reply. The multitasking thus did not cause any conversational breakdown in that sense. Sequencing and adjacency pairs were also investigated in relation to switching patterns. Here the analysis explored whether the participant would switch faster to conversations where she had asked a question and thus was expecting a reply. The results show no such correlation, and instead the main reason for switching between conversations seems to have been the alerts that she received from the IM system when new messages had been sent. Based on the results here, a design suggestion for *unobtrusive conversational management* is introduced.

6.4 Case Study 4: Conversation initiation in multiplex communicative ecologies

This study deals with interaction among a group of international students at a design school. The aim of the analysis is to investigate strategies for coherent conversation initiation in a situation where several technologically-mediated tools were available alongside face-to-face interaction. The material was gathered during three weeks in the environment, and consisted of observational notes, diaries, log files and interviews. Three participants were shadowed during two days each, and also took part in interviews. The others were asked to keep conversational diaries, and all that agreed were asked to submit log files from their text-based in-group interaction. In the school, the students would sometimes work in the design studio, and sometimes in the

computer labs, resulting in a combination of interaction in the same room and in different rooms.

Research on pre-beginning sequences in face-to-face interaction (Mondada 2009) was taken as a point of departure, and coherent conversation initiation was defined as "[c]atching the attention of suitable people at a suitable time in a suitable manner" (Örnberg Berglund submitted: 3).

First, in order to investigate the available communicative options, the communicative ecologies of the three main participants were mapped, based on observations and interviews. These students used many of the same tools for interaction, such as IM, cell phones to send SMS and to make phone calls, email and social networking sites. It was then investigated which modes were employed for in-group interaction, and it was seen that with few exceptions, the students within the observed group only interacted with each other via IM and face-to-face. The remainder of the paper deals with how the students chose between these two modes when initiating conversations.

Building on results from interviews and observational data, it is argued that the choices made relate to three types of considerations which the students need to keep in mind, all drawing on different aspects of the notion of *attention*:

> - *Whose attention do you want and whose do you not want? What types of activities do you want to give public attention and which are private? (Privacy)*
>
> - *How soon do you need the others' attention and what kind of attention do you want? (Urgency)*
>
> - *Can you get the attention from the other through this particular mode, and is now a good time to do so? (Availability)*
>
> <div align="right">Örnberg Berglund (submitted: 3)</div>

The findings show that work-related discussions were more common in face-to-face interaction, whereas private plans were more often made in IM, indicating that private issues were often dealt with in IM. The results concerning urgency are not as straight-forward, as IM was used both for matters which did not need any immediate reply, and as a way of catching the attention of the others.

The set of considerations relating to availability shows that the presence indicators in IM were not reliable, as the participants did not provide detailed information concerning availability. Furthermore, it is shown that in same-room interaction, windows of opportunity for conversation initiation often related to adjoining activities. For instance, in one example, after a long period of silence, multiplex coordination took place in IM, whereas in another, after a short pause in face-to-face interaction, multiplex coordination continued in the face-to-face mode.

Despite the fact that participants in same-room interaction might be able to judge windows of opportunity more easily, one of the conclusions of this study is that same-room interaction does not necessarily result in coherent conversation initiation, as participants are involved in individual contexts in front of their computers. In response to the challenges identified, it is then discussed how coherence might be supported through a combination of *volume control* (Baron 2008) and *social translucence* (Erickson & Kellogg 2000). These concepts are further discussed and a suggestion for design is presented where these notions merge.

7. Discussion

The empirical studies included in the current work all address the interrelation between coherence, context and affordances, and in doing so they deal with different types of contexts and address different strategies for coherence establishment and maintenance. As the investigations are situation specific and concern relatively small materials, it is difficult to draw general conclusions. However, it is possible to identify some general tendencies across the studies, and these will be discussed in what follows.

A central theme which has emerged in the analysis is the contrasting of sense-making strategies identified in face-to-face interaction with those of digitally-supported interaction. It can be concluded that coherence builds on the same principles in both conditions. On a local level, a prerequisite for coherent interaction is that links between logically related utterances are available, and on a global level, interaction follows certain conventions relating to the tool, the activity type, the participants involved and their background knowledge. However, the strategies employed to achieve coherence are quite different. If we take IM as an example, we see that sequential structures often take other shapes here than in face-to-face interaction, due to disrupted turn adjacency, and that continued attention is not always a prerequisite. This means that participants need to find alternative ways of indicating and detecting links, whereby explicit and medium-specific cues become a valuable tool.

This discussion section is structured in correspondence with the questions explored in the current study. First, some of the characteristics of online and mixed-mode contexts are summarized and discussed. This is followed by a section on the strategies employed for coherence maintenance, and a related section on the relation between strategies and contextual factors. The discussion concludes with implications for design.

7.1 Characteristics of the analyzed communication situations

The results of the studies included here make it possible to identify two characteristics of technology-rich communication situations which are especially prominent: *polyfocality* (Scollon et al. 1999, Jones 2004) and *polycontextuality* (Engeström et al. 1995, Leander 2002, Saarenkunnas 2004). Polyfocality can be defined as the possibility to focus on and participate in different semi-simultaneous activities. Polycontextuality is closely related and concerns the fact that participants in technologically-mediated interaction are part of different contextual configurations simultaneously. It should be noted that these characteristics are not unique to situations involving digital tools for interaction. For instance, Scollon et al. (1999) when

coining the term *polyfocality* did not refer to technology-rich situations. Similarly, polycontextuality is not limited to digitally-supported interaction, since previous technological developments have also made it possible to bridge different contexts. As a matter of fact, Study 4 showed that poly-contextuality can also be a relevant concept in interaction in a shared physical space. However, digital tools for interaction add new affordances enhancing the possibilities for polyfocal and polycontextual interaction.

In the studies where participants were observed in their physical settings (Study 3 and 4), it became clear that polyfocality is a rule rather than an exception. Often, participants were involved in several different conversations semi-simultaneously and often in combination with trying to complete work-related tasks. For example, the observed participant in Study 3 activated each window on her computer screen for 26 seconds on average before switching to the next. These findings correlate with the observed popularity of IM in interesting ways. For instance, as reported in Study 4, it may well be that the participants chose to communicate via IM because they knew that the semi-synchronous interaction afforded by this tool would allow them to multitask. However, in Study 3 and 4, it was also seen that the IM alerts have great interruptive power, and it may also be that this particular affordance caused participants to multitask more than they would have if the alerts would not exist.

In the studies dealing with communicative affordances on the tool level, polyfocality was also apparent. In Study 1, it was seen as a potential explanation to the instances of belated reply, as oriented to by the participants, and in Study 2, the students divided their attention between the different modes available for interaction.

As was seen in Study 3 and 4, the fact that participants are part of different contexts simultaneously influences possibilities for polyfocal interaction. In previous research, a distinction has often been made between online and offline interaction, where online interaction has also sometimes been referred to as *virtual*, as opposed to the *real* interaction taking place face-to-face (see e.g. Bell 2000). Others have opposed this dichotomy (see e.g. Livingstone 2003, Jones 2004). Whereas I acknowledge that there might not be clear-cut boundaries, the results of the studies included here have shown that it can be useful to keep different types of contexts separate for analytical purposes. In addition to the proposed distinction between material and dynamic context available in the analytical model (Figure 3, Section 4), I believe that it is important to take both the context of the individual and the shared communicative context into account. However, in order to show that the distinction is not directly related to whether interaction is taking place online or offline and certainly not to whether it is virtual or real, the results of the studies have made me view polycontextuality in terms of *individual*

and *shared* contexts. The interrelation between different types of contexts in technology-rich environments is illustrated in Figure 6.
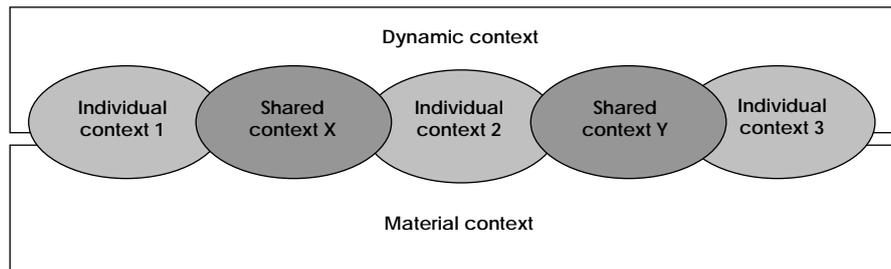

Figure 6: The polycontextuality of technology-rich communication situations

By individual context, I mean the complete context at a particular point in time in which the individual is situated, both physical and digital, which might include juggling between many different interactions. This means that even though people are physically located in the same room, they do not necessarily share the same individual context. The shared context can be either physical or digital or both, and it is the context which is made relevant in relation to the current interaction. Despite the linear depiction employed here, it is of course possible to maintain more than two semi-simultaneous shared contexts.

If we apply this model to the results of the studies, Study 1 shows that while it is possible to investigate coherence in light of the communicative affordances of the tool, individual contexts also have an influence on the patterns found. In Study 2, the participants were expected to take part in focused interaction and had scheduled a time to meet, so it can be assumed that the role of the individual context became less prominent. A concern here instead relates to ways of indicating which part of the shared multimodal contexts were in focus. Study 3 shows how the individual context of the observed participant had to be defined in relation to the several simultaneously active shared contexts in which she participated. Finally, Study 4 makes clear that awareness concerning individual contexts can be a valuable tool in coherent conversation initiation, and that shared and individual contexts are not only a question of online and offline, but that this distinction is relevant also when physically located in the same space.

7.2 Multimodal strategies for sense-making

A general conclusion which the results of the case studies confirm is that establishment and maintenance of mutual orientation is a multimodal endeavour. The main emphasis here has been on three types of strategies related to local coherence: lexical and grammatical cohesion, conversational structures and display of attention.

Creating coherence is about establishing connections between actions, and about establishing and maintaining mutual orientation. In face-to-face interaction, this can be accomplished in various ways. Links between utterances can be expressed through the use of verbal cohesive devices, such as reference and connectives, but they can also be manifested in extra-linguistic cues and in the structure of conversation (cf. Schegloff 1990). A key to coherent conversation in face-to-face interaction is shared attention. If we notice that the other person is not paying attention to what we are saying, this will be read as a sign that there might be little connection between actions and thus extra work will be needed to create coherence again. Ways of establishing shared attention include acknowledgements, relevant next turns and gaze (Clark & Brennan 1991, cf. also Goodwin 1981). Attention is highly important also during conversation initiation sequences. By determining what the other is paying attention to, we can draw conclusions concerning appropriate times to approach one another, and thus ensure coherent conversation initiation.

In studies of coherence, the main concern has typically been with lexical and grammatical cohesion in written texts and invented examples. The case studies illustrate how cohesive devices can play an important role also in technologically-supported authentic interaction. For example, Study 1 showed how links between utterances that were spatially or temporally separated became clear through extended local coherence. Here, important strategies included lexical repetition, lexical substitution and the use of conjunctions. Conversely, anaphoric reference and feedback seemed to be the greatest cause of potential confusion. However, as was noted, the participants seldom marked these instances as problematic. Similarly, in Study 2, we saw how links between utterances were created by using cohesive devices, acknowledgements and misplacement markers.

The remaining two strategy types identified in the previous literature on coherence in face-to-face interaction (conversational structures and display of attention), cannot be mapped upon the strategies found in the current studies in the same straightforward way, as both patterns and functions are dissimilar. Nevertheless, these areas are certainly of relevance in connection to how coherence is created here.

As was indicated in the discussion above, in technologically-mediated interaction, it is not always possible to establish mutual orientation by

assuming a logical connection between adjacent utterances. For example, Study 1 showed that logically adjacent turns were sometimes separated graphically, due to what has previously been referred to as disrupted turn adjacency (Herring 1999). However, the results of Study 1, Study 2 and Study 3 all pointed to the fact that sequencing is also relevant in digitally-mediated interaction, albeit in a different manner. Whereas adjacent contributions may not cohere, the basic mechanism of adjacency pairs is still in play: when a first pair part of an adjacency pair has been produced a second pair part is to be expected. This *conditional relevance* (Schegloff 1968) can help us create logical links between utterances separated in space and in time.

One example can be found in Study 3, where the observed participant always succeeded in replying when addressed, despite her involvement in conversational multitasking. Thus, by completing adjacency pairs, she was able to follow one of the most basic conversational patterns, and she was able to indicate that she was paying semi-synchronous attention. The results of Study 1 showed that sequencing can play an important role also during communication in intertwined threads, as here we saw how the sequencing structure of multiple first pair parts was replicated in the reply sequence.

Whereas IM interaction is not dependent on shared attention, in Study 2, which dealt with focused semi-synchronous interaction, ways of displaying attention were seen as crucial, and it was investigated how the different modes available were used towards this goal. We saw that feedback often did not occur in any of the available modes, possibly leaving participants in doubt as to where the others were directing their attention. The fact that some of the features relevant for communication were located on competing tabs in the tool interface might have lead to further complications.

The results from Study 4 concerned conversation initiation, and here we also saw how different considerations relating to attention are central. For instance, different ways of detecting windows of opportunity for coherent conversation initiation were identified, based on strategies for knowing where the others were paying attention and whether they were available for interaction both face-to-face and in IM.

These results indicate that sense-making in the situations analyzed depends on multimodal strategies relating to cohesive devices, conversational structures and display of attention, but that these strategies sometimes take different shapes in online and mixed-mode interaction than when interacting face-to-face. Some of the successful strategies identified in the case studies are summarized in Table 5.

Local coherence strategies	
Explicit cohesive devices - Repetition - Substitution - Conjunction	Implicit cohesive devices - Anaphoric reference - Elliptical feedback
Explicit change-of-state tokens	
Replication of specific sequential structures	
Explicit acknowledgements and media specific phatic strategies	
Replies when a reply is requested (adjacency pairs)	
Explicit signs to create modal density and indicate continued attention	
Explicit signs indicating availability	

Table 5: Successful local coherence strategies

Table 5 shows that with a few exceptions, the most successful strategies are explicit.[32] As far as cohesive devices are concerned, we saw that explicit strategies were less likely to cause ambiguities than implicit ones. Explicit *change-of-state tokens* (Heritage 1984), whereby participants indicate that a change in the current state of knowledge has taken place, were shown to provide additional cues. These tokens made it possible to differentiate between different types of feedback and to identify threads despite the use of elliptical feedback. Another way in which coherence was identified during instances of intertwined threads was through replication of specific sequential structures, where, for example, a sequence of two questions was replicated in the reply structure.

Explicit ways of indicating that one is paying attention were also seen as important in the sense-making process. This was especially clear in Study 2, where participants depended on shared attention, due to the ephemeral qualities of audio interaction. Successful strategies identified here included explicit acknowledgements and ways of expressing emotion (phatic strategies), employing the affordances of the tool and following the norms of the group. Other strategies to indicate continued attention included replying

---

[32] Deutschmann & Panichi (forthcoming) have also argued for the importance of explicit strategies in CMC. In their study of student and teacher interaction in Second Life, they conclude that participants need to learn new ways of expressing explicit supportive moves in line with the affordances of the tool employed.

when replies were requested, thus completing adjacency pair structures, and making use of the different modes available, creating modal density.

A final successful strategy identified concerns the use of explicit signs indicating availability to allow for participants to initiate conversation in a coherent manner. All of the above examples suggest that explicit cues are important in order to counter some of the practical problems that arise, as well as to ensure coherent interaction in online and mixed-mode contexts.

7.3 Context-dependent strategies for sense-making

The above discussion points to another general conclusion which the results of the case studies confirm, namely that the strategies employed are influenced by the context for their production. In previous research on the relation between context and coherence, a main distinction has been made between local and global sense-making strategies, depending on whether cotext or other contextual factors have been drawn upon as resources for sense-making (cf. Korolija 1998, Littlejohn 2002). In the current study, further dimensions have been added to the analytical model. For instance, it has been shown that it can be fruitful to distinguish between material and dynamic context, and between individual and shared context.

As far as conversational structures are concerned, we saw that communicative affordances of the material environment seem to have quite a strong influence. For example, disrupted turn adjacency, specifically addressed in Study 1 and Study 2, seems to correlate strongly with the communicative affordances of the tool employed. In Study 1, the fact that messages did not appear on screen until completed and sent off had an important impact, and in Study 2, it was clear that turn-taking in FlashMeeting was influenced by the broadcasting system, which allowed for only one speaker at a time.

However, despite the relatively clear evidence of the influence of communicative affordances in these cases, we have to consider how dynamic context and background knowledge may also play an important role. Let us first return to the results from Study 2. Despite the hypothesized need for indications of shared attention here, many instances were identified where speakers did not receive any acknowledgement whatsoever. Apart from the influence of communicative affordances, it could be argued that background knowledge and conventions might have had an influence. These participants were new to the tool (FlashMeeting) and to each other and might not yet have learnt to take advantage of all affordances provided, such as possibilities for multimodal feedback. Other parts of the dynamic context may also have played a role in coherence maintenance, not least the design of the task. Here, students were provided with questions ahead of time to which they should prepare and supply answers. In some of the cases where no local coherence was identified, we instead found examples of coherence created on this global level.

Similarly, the patterns found in relation to turn-taking in IM seem to be influenced by conventions. As previously mentioned, the main participant in Study 3 adapted her linguistic production depending on conversational partner. If we compare the patterns identified here with those of Study 1, we find even greater differences. The participants in Study 1 often engaged in intense and focused dyad IM conversations, typing several messages in quick succession, sometimes resulting in disrupted turn adjacency. Alerts that new messages had arrived were mainly relevant at the very beginning of new conversations. In Study 3, contrarily, the participants in the dyad IM conversations took clear turns and waited until they received a reply before responding. In one of the conversations, the participants even engaged in threaded conversations where each message consisted of replies of up to three different threads, separated into sequentially structured paragraphs. Here, the alerts became an important device throughout the conversation, providing indications that new contributions had been made. While this latter strategy may be beneficial in order to be able to multitask, it also resulted in continuous interruptions throughout the conversation.

Study 3 and Study 4, with their broad foci, showed the importance of also considering both the individual and the shared context. For instance, it was seen that the communicative affordances of the IM tool in combination with the polycontextuality of the situation supported multitasking. Furthermore, despite the fact that shared attention is not a prerequisite for IM interaction, the results of Study 4 concerning conversation initiation showed that upon receiving alerts, participants almost always immediately checked the message and replied. Here, we see an interesting dilemma in the use of IM. Whereas it provides control on behalf of the respondent as to how soon to reply, it also seems to demand immediate attention, and lead to constant interruptions.

In fact, if IM is used in a more unfocused way, as was the case in Study 3, it could be argued that each new message is a call for attention, and thus in some respects resembles an initiation of a new conversation (cf. Avrahami et al. 2008). This means that in some instances, knowledge about which part of the individual context the other is focusing on might be relevant not only in the pre-beginning sequence but also during subsequent interaction.

The above examples illustrate the importance of considering the influence of both the material context and of other contextual dimensions when identifying strategies for sense-making on a global level.

7.4 Supporting polyfocal and polycontextual interaction through design

Design is making sense of things.
Krippendorff (2006: xv)

An applied aim of this work is to relate my findings to interaction design. In workplace studies and CSCW, it is quite common that results from ethnographic or CA-inspired analyses are used to inform design of communication tools, and in that respect, the current work can be seen as a contribution to these fields of study.[33]

As was evident in the case studies, successful strategies for coherence establishment and maintenance were often explicit in character. I will now attempt to show how designing to help make explicit can increase conversational coherence on different levels. As will be seen, different components of the suggested design have been previously proposed and tested, but to the best of the author's knowledge, the resulting combination has not been previously suggested.

In Study 1, we saw the importance of making reference explicit. With this in mind, a refined system should support threading (cf. Herring 1999). By teaching the system basic cohesive strategies, threading structures could be automatically suggested. Each thread would be graphically separated from the others and, in ambiguous cases, the participants would be asked to choose a thread.

In Study 2, it became apparent that during focused interaction, it is important to explicitly indicate that one is paying attention. More specifically, the analysis concerned different types of conversational feedback. In order to convey emotions or other types of feedback during ongoing broadcasts, participants in the discussions analyzed needed to make a conscious effort and use pre-programmed emoticons or the text chat. In addition, the video representations of the other participants were not updated often enough to provide detailed feedback support. In order to support explicit feedback, a system used for focused interaction could make use of often available web cameras and microphones to capture emotion and convey this in the form of a continuously updated emoticon in connection with the text chat. This

---

[33] The relationship between ethnography and design has not been seen as completely unproblematic in the HCI community (see e.g. Dourish & Button 1998, Dourish 2006). Dourish (2006) distinguishes between the empirical and the analytical contribution of ethnography. In his view, the empirical contribution of ethnography is often given greater prominence in HCI than the analytical one. He proposes a shift in focus, but does not discard design implications completely. Rather, he suggests that a combination of empirical and analytical contributions might be useful.

would mean that the conscious effort on behalf of the users in making feedback explicit would be diminished.

Study 3 dealt with conversational multitasking, and here it was seen that a simple way of explicitly indicating that one is paying attention and that conversations are on track is to reply when a reply is expected. We also saw that the alerts received generally caused the main participant observed to interrupt her current activities and switch to the IM interface. As Hudson et al. (2002) remind us, interruptions can be a necessary component in finishing a task. Rather than eliminating them completely, they argue, we should design for more effective interruptions. By teaching the system to recognize first pair parts of adjacency pairs, it would be possible to make reply requests explicit, by assigning these messages more intrusive alert characteristics than the others.[34]

Study 4 illustrated the importance of making windows of opportunity for conversation initiation explicit. The participants often did not make the conscious effort to update their status messages to indicate current activities, and consequently it was difficult to find appropriate times to interrupt. One way of dealing with this would be to allow the computer to collect and display more information automatically about current activities. The IM systems employed in the studies already use some data to provide participants with insights concerning current activities, for example, whether the other has been idle at the computer for a while or whether the other is typing. My suggestion is to take this a step further and include also other types of data indicating availability for interaction. However, apart from ethical concerns relating to privacy (cf. e.g. Ackerman et al. 2001), this may also result in less control on behalf of the receiver, as it might be more difficult to avoid replying.

One solution could be to design tools that strive towards what I would like to call *balanced awareness*. These tools would be based on the idea of increasing awareness and translucence, "[...] making participants and their activities visible to one another" (Erickson & Kellogg 2000:1) while simultaneously ensuring control on behalf of the participants (cf. Baron 2008). This could be achieved through a combination of design for *interruptability* and *awareness* (cf. Danninger et al. 2008; see Section 3.2.2), incorporating both social cues which the computer automatically suggests based on current activities and cues which the participants themselves set.

In order to support polycontextual interaction, information concerning current activities could be automatically logged by the computer and trans-

---

[34] This is similar to a prototype implemented by Avrahami & Hudson (2004) where messages including questions or replies to previous questions are delivered immediately, while the delivery of other messages is postponed.

formed into a general status setting, indicating availability for interaction on a relatively detailed scale without giving away private information. Different types of information could be merged, such as physical location, calendar appointments, open windows and current conversations.[35] Previous research has identified opportune times for interruptions which could be used as a basis for suggested windows of opportunity (see e.g. Czerwinski et al. 2000, Hudson et al. 2002, Avrahami et al. 2008). However, the tool would also be user configurable so that the participant sharing information about current whereabouts could indicate under which conditions it would normally be acceptable to be interrupted. Furthermore, participants could choose what level of detail to convey to their different contacts concerning their current activities.

In order to facilitate polyfocal interaction despite increased awareness, it would be important not to give away information concerning whether a message has been read or not, as it could be hypothesized that the account-ability (Erickson & Kellogg 2000) resulting from such information makes it more difficult to ignore a message temporarily and focus on something else. This is in line with Nardi et al. (2000: 9) who suggest that "[p]aradoxically, an interface that provides *less* awareness information may be more success-ful because it addresses the problem of participant asymmetry" – not only the initiator, but also the person being contacted is given control over the interaction.[36]

In addition, it might also be useful to collect all incoming alerts and give the participants the possibility to choose level of intrusiveness depending on current activities and level of urgency of the message. Again, this intelligent device could give initial suggestions based on information about current activities, thus letting alerts through at convenient times while collecting them when interruptions are unwanted. Just as with the shared availability information, the participant could always configure the settings (cf. Bellotti & Edwards 2001). It would be possible to flag urgent messages, for instance reply requests, thus overriding some of the settings, but the receiver would always have the possibility to switch off all incoming alerts. This might lead to less interruption, as participants would know that all alerts would be available for them once they had the time to deal with them, while at the same time, distracting continuous alerts would be avoided. If there was a

---

[35] Similar solutions are introduced by Begole et al. (2004) in their system Lilsys and by ter Hofte et al. (2004) in their system Live Contacts. Live Contacts has the additional feature of suggesting suitable medium for interaction.

[36] In some cases, it could be valuable to receive notification once a message has been attended to by a conversational partner. For example, if a new message is sent while the other is typing, it might be difficult to know whether or not this message will be noticed. One solution would be to make such messages graphically distinct in the interface of the participant receiving them.

reflexive awareness concerning this feature, both parties could rest in the knowledge that replies would be received in due time.

In sum, by designing for explicitness and for balanced awareness, conversational coherence might be supported through a combination of increased awareness and individual control. In addition, it is hypothesized that balanced awareness may reduce information stress. Further steps are necessary, however, to develop a more complete design concept and to find the most accurate correlation between different types of ongoing activities and opportune times to interrupt. In addition, the hypotheses presented here would need to be tested further, something which falls outside the scope of the current study.

8. Concluding remarks

In this concluding section, I briefly point to some implications of the results from the current study in the areas of methodology and theory, as well as in some different applied contexts. I also point to areas of possible future research.

A methodological conclusion that can be drawn based on the current study concerns the importance of taking different contextual dimensions into account when describing technologically-mediated interaction. In two of the studies, the focus is on the shared online context, and whereas this is what participants in interaction themselves have access to when interacting in these situations, it becomes apparent that some aspects cannot be explained without also taking into consideration the individual context of the participants in front of their computers.

From a theoretical perspective, it can be concluded that by combining interactional and ecological approaches it is possible to explain the interrelation between material surroundings and interactional patterns. The analysis points to how technology has a very real influence on the interaction taking place. However, by focusing on affordances rather than material qualities per se, it is possible to avoid a model grounded in technological determinism. In addition, the notion of affordance emphasizes the active role of the participants in choosing what actions to take and appropriating the technology for their own needs. The different conventions identified in relation to IM turn-taking illustrate this.

One finding concerns the importance of explicit linguistic strategies for coherence maintenance. As is suggested, tools could be designed to help make explicit. However, as most tools employed today do not have this coherence support, participants need to be aware of the importance of explicit strategies. This could have bearing in many different situations where digital technologies for interaction are employed. In order to help facilitate efficient online interaction, participants could be educated and trained in successful strategies in relation to the affordances of the tool used. For example, in education it can be concluded that the instructor has a crucial role to play in stressing the importance of explicit cues and encouraging the development of medium-specific norms, as well as in designing interactive tasks which support an active exchange between students. In business contexts, similarly, it would be possible to support coherent interaction by raising awareness of the importance of certain strategies. For instance, if the tools employed do not support balanced awareness, it could be stressed that explicitly indicating availability might lead to fewer unwanted interruptions, which might make work more efficient.

Future research could apply and test the analytical model suggested here in relation to more recently introduced tools, in order to explore the communicative affordances that these tools add to our material surroundings. For example, much current digital interaction today takes place via mobile devices, adding dimensions to the interrelation between context and interactional patterns. It would also be possible to broaden the field of study to include other activity types than those under investigation here. From a methodological perspective, it might be useful to make use of video data to a larger extent than was the case in the current studies. This could add valuable details about the multimodal interaction taking place in the shared physical space.

In sum, the current research contributes mainly to two broad domains of study. First, to the field of interactional linguistics, as central concepts are discussed in relation to the affordances of technology-rich communication situations and a modified analytical model is presented. Second, this model also adds to previous research on computer-mediated communication by showing that log files are best understood in the "complete" context of interaction and in relation to both dynamic and material qualities and communicative affordances. Even though face-to-face interaction is taken as a point of reference, the findings indicate that the qualities of face-to-face interaction are not always optimal, but rather different contexts provide different types of affordances for interaction, influencing, for example, possibilities for polyfocal interaction. It is thus confirmed that situated analyses are invaluable when making sense of sense-making in technology-rich communication situations.

References

Ackerman, M. S., Darrell, T. & Weitzner, D. J. 2001. Privacy in Context. *Human-Computer Interaction* 16(2): 167-176.

Ågren, P-O. 2000. Forskningsetik i Cyberrymden. *HumanIT* 1: 139-156.

Arminen, I. & Weilenmann, A. 2009. Mobile Presence and Intimacy: Re-shaping Social Actions in Mobile Contextual Configuration. *Journal of Pragmatics* 41(10): 1905-1923.

Avrahami, D., Fussell, S. R. & Hudson, S. E. 2008. IM Waiting: Timing and Responsiveness in Semi-Synchronous Communication. *Proceedings of the 2008 ACM CSCW Conference*, San Diego, CA: 285-294.

Avrahami, D. & Hudson, S. E. 2004. QnA: Augmenting an Instant Messaging Client to Balance User Responsiveness and Performance. *Proceedings of the 2004 ACM CSCW Conference*, Chicago, IL: 515-518.

Baron, N. 2008. *Always On. Language in an Online and Mobile World.* New York: Oxford University Press.

Begole, J. B., Matsakis, N. E. & Tang, J. C. 2004. Lilsys: Sensing Unavailability. *Proceedings of the 2004 ACM CSCW Conference*, Chicago, IL: 511-514.

Bell, D. 2000. Introduction I: Cybercultures Reader: a User's Guide. In: Bell, D. & Kennedy, B. M. (eds.) *The Cybercultures Reader*. London: Routledge, 1-12.

Bellotti, V. & Edwards, K. 2001. Intelligibility and Accountability: Human Considerations in Context-Aware Systems. *Human-Computer Interaction* 16(2): 193-212.

Brewer, J. D. 2000. *Ethnography*. Milton Keynes: Open University Press.

Broth, M. 2009. Seeing through Screens, Hearing through Speakers: Managing Distant Studio Space in Television Control Room Interaction. *Journal of Pragmatics* 41(10): 1998-2016.

Button, G. (ed.) 1992. *Technology in Working Order. Studies of Work, Interaction and Technology*. London/New York: Routledge.

Clark, H. 1996. *Using Language*. Cambridge, UK: Cambridge University Press.

Clark H. H. & Brennan S. E. 1991. Grounding in Communication. In: Resnick, L. B., Levine J. M. & Teasley, S. D. (eds.) *Perspectives on Socially Shared Cognition*. Washington, DC: The American Psychological Association: 127-149.

Coates, J. 1995. The Negotiation of Coherence in Face-to-Face Interaction: Some Examples from the Extreme Bounds. In: Gernsbacher, M. A. & Givón, T. (eds.) *Coherence in Spontaneous Text*. Amsterdam: John Benjamins: 41-58.

Craig, R. T. & Tracy, K. (eds.) 1983. *Conversational Coherence: Form, Structure and Strategy*. Beverly Hills: SAGE.

Crystal, D. 2001. *Language and the Internet*. Cambridge, UK: Cambridge University Press.

Czerwinski, M., Cutrell, E. & Horvitz, E. 2000. Instant Messaging and Interruption: Influence of T*ask* Type on Performance. *Proceedings of the 2000 OZCHI Conference*, Sydney, Australia: 356-361.

Daft, R. L. & Lengel, R. H. 1986. Organizational Information Requirements, Media Richness and Structural Design. *Management Science* 32(5): 554-571.

Danninger, M., Robles, E., Sukumaran, A. & Nass, C. 2008. The Connector Service: Representing Availability for Mobile Communication. In: Waibel, A. & Stiefelhagen, R. (eds.) *Computers in the Human Interaction Loop*. London: Springer: 235-256.

Deutschmann, M. & Panichi, L. Forthcoming. Talking into Empty Space? Signalling Involvement in a Virtual Language Classroom in Second Life. *Language Awareness* 18(3-4): 311-329.

Dourish, P. 2004. What We Talk About When We Talk About Context. *Personal and Ubiquitous Computing* 8(1): 19-30.

Dourish, P. 2006. Implications for Design. *Proceedings of the 2006 ACM CHI Conference*, Montreal, Canada: 541-550.

Dourish, P. & Bly, S. 1992. Portholes: Supporting Awareness in a Distributed Work Group. *Proceedings of the 1992 ACM CHI Conference*, Monterey, CA: 541 - 547

Dourish, P. & Button, G. 1998. On "Technomethodology": Foundational Relationships between Ethnomethodology and System Design. *Human-Computer Interaction* 13(4): 395-432.

Drew, P. & Heritage, J. 1992. Analyzing Talk at Work: An Introduction. In: Drew, P. & Heritage, J. (eds.) *Talk at Work: Interaction in Institutional Settings.* Cambridge, UK: Cambridge University Press: 3-65.

Duranti, A. 1997. *Linguistic Anthropology.* Cambridge, UK: Cambridge University Press.

Engeström, Y., Engeström, R. & Kärkkäinen, M. 1995. Polycontextuality and Boundary Crossing in Expert Cognition: Learning and Problem Solving in Complex Work Activities. *Learning and Instruction* 5: 319–36.

Erickson, F. 2004. Origins: A Brief Intellectual and Technological History of the Emergence of Multimodal Discourse Analysis. In: Levine, P. & Scollon, R. (eds.) *Discourse & Technology. Multimodal Discourse Analysis.* Washington D.C.: Georgetown University Press: 196-207.

Erickson, T., Herring S. & Sack W. 2002. Discourse Architectures: Designing and Visualizing Computer Mediated Conversation. *Proceedings of the 2002 ACM CHI Conference*, Minneapolis, MN. Retrieved from http://hybrid.ucsc.edu/SocialComputingLab/Publications/erickson-herring-sack-chi01.pdf.

Erickson, T. & Kellogg, W. 2000. Social Translucence: an Approach to Designing Systems that Support Social Processes. *ACM Transactions on Computer-Human Interaction* 7(1): 59-83.

Ess, C. & the AoIR ethics working committee. 2002. *Ethical Decision-Making and Internet Research: Recommendations from the AoIR Ethics Working Committee.* Retrieved from http://aoir.org/reports/ethics.pdf.

Fillmore, C. 1976. Frame Semantics and the Nature of Language. *Annals of the New York Academy of Sciences: Conference on the Origin and Development of Language and Speech* 280: 20-32.

Garfinkel, H. 1967. *Studies in Ethnomethodology.* Malden, MA: Polity & Blackwell.

Garrett, R. K. & Danziger, J. N. 2007. IM=Interruption Management? Instant Messaging and Disruption in the Workplace. *Journal of Computer-Mediated Communication* 13(1): 23-42.

Gaver, W. W. 1991. Technology Affordances. *Proceedings of the 1991 ACM CHI Conference*, New Orleans, LA: 79-84.

Gaver, W. W. 1992. The Affordances of Media Spaces for Collaboration. *Proceedings of the 1992 ACM CSCW Conference*, Toronto, Canada: 17-24.

Gaver, W. W. 1996. Affordances for Interaction: the Social is Material for Design. *Ecological Psychology* 8(2): 111–129.

Gernsbacher, M. A. & Givón, T. (eds.) 1995. *Coherence in Spontaneous Text*. Amsterdam: John Benjamins.

Gibson, J. J. 1977. The Theory of Affordances. In: Shaw, R. E. & Bransford, J. (eds.) *Perceiving, Acting, and Knowing*. Hillsdale, NJ: Lawrence Erlbaum Associates: 67-82.

Gibson, J. J. 1979. *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin.

Goffman, E. 1961. *Encounters: Two Studies in the Sociology of Interaction*. Indianapolis: Bobbs-Merrill.

Goffman, E. 1963. *Behavior in Public Places*. New York: Free Press.

Goffman, E. 1974. *Frame Analysis: Essays on the Organization of Experience*. New York: Harper & Row.

Goffman, E. 1981. *Forms of Talk*. Philadelphia: University of Pennsylvania Press.

Goffman, E. 1983. The Interaction Order. *American Sociological Review* 48(1): 1-17.

Goodwin, C. 1981. *Conversational Organization: Interaction Between Speakers and Hearers*. New York: Academic Press.

Goodwin, C. 1995. The Negotiation of Coherence within Conversation. In: Gernsbacher, M. A. & Givón, T. (eds.) *Coherence in Spontaneous Text*. Amsterdam: John Benjamins: 117-137.

Goodwin, C. 2000. Action and Embodiment within Situated Human Interaction. *Journal of Pragmatics* 32(10): 1489-1522.

Goodwin, C. & Duranti, A. 1992. Rethinking Context: An Introduction. In: Duranti, A. & Goodwin, C. (eds.) *Rethinking Context: Language as an Interactive Phenomenon*. Cambridge, UK: Cambridge University Press: 1-42.

Goodwin, C. & Goodwin, M. H. 1996. Formulating Planes: Seeing as Situated Activity. In: Middleton, D. & Engeström, Y. (eds.) *Cognition and Communication at Work*. Cambridge, UK: Cambridge University Press: 61-95.

Grice, H. P. 1975. Logic and Conversation. In: Cole, P. & Morgan J. L. (eds.) *Syntax and Semantics, Volume 3: Speech Acts*. New York: Seminar Press: 225-242.

Grint, K. & Woolgar, S. 1997. *The Machine at Work*. Cambridge, MA: Polity.

Grudin, J. 1994. Computer-Supported Cooperative Work: History and Focus. *Computer* 27(5): 19–26.

Gumperz, J. J. 1982. *Discourse Strategies*. Cambridge, UK: Cambridge University Press.

Habermas, J. 1983. *Moral Consciousness and Communicative Action*. Cambridge, MA: MIT Press.

Halliday, M. A. K. & Hasan, R. 1976. *Cohesion in English*. London: Longman.

Halliday, M.A.K., & Hasan, R. 1985. *Language, Context, and Text: Aspects of Language in a Social-Semiotic Perspective*. Oxford: Oxford University Press.

Hammersley, M. 2003. Conversation Analysis and Discourse Analysis: Methods or Paradigms? *Discourse & Society* 14(6): 751-781

Hammersley, M. & Atkinson, P. 1995. *Ethnography: Principles in Practice*. London: Routledge.

Hård af Segerstad, Y. 2002. *Use and Adaptation of Written Language to the Conditions of Computer-Mediated Communication*. Doctoral thesis. Gothenburg University. Gothenburg monographs in linguistics, 22.

Heath, C. 1982. The Display of Recipiency: An Instance of Sequential Relationship Between Speech and Body Movement. *Semiotica* 42(2/4): 147–161.

Heritage, J. 1984. *Garfinkel and Ethnomethodology*. Cambridge, MA: Polity.

Herring S. 1999. Interactional Coherence in CMC. *Journal of Computer-Mediated Communication* 4(4). Retrieved from http://jcmc.indiana.edu/vol4/issue4/herring.html.

Hine, C. 2000. *Virtual Ethnography*. London: SAGE.

Hudson, J.M., Christensen, J., Kellogg, W.A. and Erikson, T. 2002. I'd Be Overwhelmed, But It's Just One More Thing to Do: Availability and Interruption in Research Management. *Proceedings of the 2002 ACM CHI Conference*, Minneapolis, MN: 97-104.

Hutchby, I. 2001. *Conversation and Technology. From the Telephone to the Internet*. Cambridge, MA: Polity.

Hutchby, I. & Wooffitt, R. 1998. *Conversation Analysis.* Cambridge, MA: Polity.

Hymes, D. 1974. *Foundations of Sociolinguistics: An Ethnographic Approach*. Philadelphia: University of Pennsylvania Press.

Jacobs, S. & Jackson, S. 1983. Speech Act Structure in Conversation: Rational Aspects of Pragmatic Coherence. In: Craig, R. T. & Tracy, K. (eds.) *Conversational Coherence: Form, Structure, and Strategy*. Beverly Hills: SAGE: 47-66.

Jones, R. H. 2004. The Problem of Context in Computer-Mediated Communication. In: Levine, P. & Scollon, R. (eds.) *Discourse & Technology. Multimodal Discourse Analysis*. Washington D.C.: Georgetown University Press: 20-33.

Kiesler, S., Siegel, J. & McGuire, T. W. 1984. Social Psychological Aspects of Computer-Mediated Communication. *American Psychologist* 39(10): 1123-1134.

Korolija, N. 1998. *Episodes in talk. Constructing Coherence in Multiparty Conversation*. Doctoral thesis. Linköping University. Linköping Studies in Arts and Science, 171.

Koschmann, T. & LeBaron, C. D. 2003. Reconsidering Common Ground: Examining Clark's Contribution Theory in the OR. *Proceedings of the Eighth European Conference on CSCW*, Helsinki, Finland: 81-98.

Krippendorff, K. 2006. *The Semantic Turn: A New Foundation for Design*. Boca Raton, FL: CRC Press.

Labov W. 1972. *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.

Lam, C. & Mackiewicz, J. 2007. A Case Study of Coherence in Workplace Instant Messaging. *Proceedings of the IEEE International Professional Communicators Conference*, Seattle, WA: 1-6. .

Lapadat, J.C. 2007. Discourse Devices used to Establish Community, Increase Coherence, and Negotiate Agreement in an Online University Course. *Journal of Distance Education* 21(3): 59-92.

Leander, K. 2002. Polycontextual Construction Zones: Mapping the Expansion of Schooled Space and Identity. *Mind, Culture and Activity* 9(3): 211–237.

Levine, P. & Scollon, R. (eds.) 2004. *Discourse & Technology. Multimodal Discourse Analysis*. Washington D.C.: Georgetown University Press.

Levinson, S. C. 1983. *Pragmatics*. Cambridge, UK: Cambridge University Press.

Levinson, S. C. 1992. Activity Types and Language. In: Drew, P. & Heritage, J. (eds.) *Talk at Work: Interaction in Institutional Settings*. Cambridge, UK: Cambridge University Press: 66-100.

Linderoth, J. 2004. *Datorspelandets Mening. Bortom Idén om den Interaktiva Illusionen.* Doctoral thesis. Gothenburg University. Göteborg Studies in Educational Sciences, 211.

Linell, P. 1978. *Människans Språk.* Malmö: Liber.

Linell, P. 1998. *Approaching Dialogue. Talk, Interaction and Contexts in Dialogical Perspectives.* Amsterdam: John Benjamins.

Linell, P. 2009. *Rethinking Language, Mind and World Dialogically. Interactional and Contextual Theories of Human Sense-Making.* Charlotte, NC: Information Age Publishing.

Littlejohn, S. W. 2002. *Theories of Human Communication.* 7th edition. Albuquerque, NM: Wadsworth.

Livingstone, S. M. 2003. Children's Use of the Internet: Reflections on the Emerging Research Agenda. *New Media and Society* 5(2): 147-166.

Luff, P. K. & Heath C. C. 1992. Towards the Social Organisation of Human Computer Interaction. In: Button, G. (ed.) *Technology in Working Order: Studies of Work, Interaction, and Technology.* New York: Routledge: 72-98.

Markee, N. 2000. *Conversation Analysis.* Mahwah, NJ: Lawrence Erlbaum Associates.

Markman, K. M. Forthcoming. Conversational Coherence in Computer-Mediated Team Meetings. Submitted to: Herring, S., Stein, D. & Virtanen, T. (eds.) *The Handbook of the Pragmatics of Computer-Mediated Communication.* Berlin: Mouton de Gruyter.

McGrenere J. & Ho W. 2000. Affordances: Clarifying and Evolving a Concept. *Proceedings of Graphics Interface*, Montreal, Canada: 179-186.

McHoul, A., Rapley, M. & Antaki, C. 2008. You Gotta Light? On the Luxury of Context for Understanding Talk in Interaction. *Journal of Pragmatics* 40(5): 827-839.

McIlvenny, P., Broth, M. & Haddington, P. 2009. Communicating Place, Space and Mobility. *Journal of Pragmatics* 41(10): 1879-1886.

Mondada, L. 2009. Emergent Focused Interactions in Public Places: A Systematic Analysis of the Multimodal Achievement of a Common Interactional Space. *Journal of Pragmatics* 41(10): 1977-1997.

Nardi, B. A. (ed.) 1996. *Context and Consciousness: Activity Theory and Human-Computer Interaction.* Cambridge, MA: MIT Press.

Nardi, B. A. 2005. Beyond Bandwidth: Dimensions of Connection in Interpersonal Communication. *Computer Supported Cooperative Work* 14(2): 91-130.

Nardi, B. A. & O'Day V.L. 1999. *Information Ecologies. Using Technology with Heart.* Cambridge, MA: MIT Press.

Nardi, B. A., Whittaker, S. & Bradner, E. 2000. Interaction and Outeraction: Instant Messaging in Action. *Proceedings of the 2000 ACM CSCW Conference*, Philadelphia, PA: 79-88.

Neubauer, F. (ed.) 1983. *Coherence in Natural-Language Texts.* Hamburg: Helmut Buske.

Norman, D. A. 1988. *The Psychology of Everyday Things.* New York: Basic Books.

Norman, D. A. 1999. Affordances, Conventions, and Design. *Interactions* 6(3): 38-41.

Norris S. 2004. *Analyzing Multimodal Interaction. A Methodological Framework.* New York/London: Routledge.

Oliver, M. 2005. The Problem with Affordance. *E-Learning* 2(4): 402-413.

Örnberg Berglund, T. 2009. Multimodal Student Interaction Online: An Ecological Perspective. *ReCALL* 21(2): 186-205.

Örnberg Berglund, T. Submitted. Conversation Initiation in Multiplex Communicative Ecologies.

Paltridge, B. 2006. *Discourse Analysis.* London: Continuum.

Rambusch, J & Susi, T. 2008. The Challenge of Managing Affordances in Computer Game Play. *Human IT* 9(3): 83-109.

Rappert, B. 2003. Technologies, Texts and Possibilities: A Reply to Hutchby. *Sociology* 37(3): 565-80.

Rawls, A. 2003. Harold Garfinkel. In: Ritzer, G. (ed.) *The Blackwell Companion to Major Contemporary Social Theorists.* Malden, MA: Blackwell: 545-576.

Rennecker, J. & Godwin, L. 2003 Theorizing the Unintended Consequences of Instant Messaging for Worker Productivity. *Sprouts* 3(3): 137-168.

Rutter, D. R. 1987. *Communicating by Telephone.* Oxford: Pergamon.

Rutter, D. R., Stephenson, G. M. & Dewey, M. E. 1987. Visual Communication and the Content and Style of Conversation. *British Journal of Social Psychology* 20: 41-52.

Saarenkunnas, M. 2004. *Multidimensional Participation in Polycontextual Computer-Supported Language Learning.* Doctoral thesis. University of Oulu, Faculty of Education,.

Sacks, H. 1984. Notes on Methodology. In: Atkinson, J. M & Heritage, J. (eds.) *Structures of Social Action: Studies in Conversation Analysis.* Cambridge, UK: Cambridge University Press: 21-27.

Sacks, H., Schegloff, E. A. & Jefferson, G. 1974. A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Language* 50(4): 696-735.

Sassenberg, K & Jonas, K. J. 2007. Attitude Change and Social Influence on the Net. In: Joinson, A., Postmes T. & McKenna, K. (eds.) *Oxford Handbook of Internet Psychology.* New York: Oxford University Press: 273-288.

Schegloff, E.A. 1968. Sequencing in Conversational Openings. *American Anthropologist* 70(6): 1075-1095.

Schegloff, E. A. 1979. Identification and Recognition in Telephone Openings. In: Psathas, G. (ed.) *Everyday Language.* New York: Erlbaum: 23-78.

Schegloff, E. A. 1987. Analyzing Single Episodes of Interaction: An Exercise in Conversation Analysis. *Social Psychology Quarterly* 50(2): 101-114.

Schegloff, E. A. 1990. On the Organization of Sequences as a Source of "Coherence" in Talk-in-interaction. In: Dorval, B. (ed.) *Conversational Organization and its Development.* Norwood, NJ: Ablex Publishing: 51-77.

Schegloff, E. A. 1991. Converation Analysis and Socially Shared Cognition. In: Resnick, L. B., Levine J. M. & Teasley, S. D. (eds.) *Perspectives on Socially Shared Cognition.* Washington, DC: The American Psychological Association: 150-171.

Scollon, R., Bhatia, V., Li, D. & Yung, V. 1999. Blurred Genres and Fuzzy Identities in Hong Kong Public Discourse: Foundational Ethnographic Issues in the Study of Reading. *Applied Linguistics* 20(1): 22-43.

Scollon, R. & Scollon, S. W. 2001. *Intercultural Communication: a Discourse Approach.* 2nd edition. Malden, MA/Oxford: Blackwell.

Scollon, R. & Scollon, S. W. 2003. *Discourses in Place. Language in the Material World.* London/New York: Routledge.

Short, J., Williams, E. & Christie, B. 1976. *The Social Psychology of Telecommunications.* New York: John Wiley & Sons.

Simpson, J. 2005. Meaning-Making Online: Discourse and CMC in a Language Learning Community. In: Méndez-Vilas, A., González-Pereira, B., Mesa González, J. & Mesa González J. A. (eds.) *Recent Research Developments in Learning Technologies*. Badajoz, Spain: FORMATEX: 175-179.

Spira, J. B. & Feintuch, J. B. 2005. The Cost of Not Paying Attention: How Interruptions Impact Knowledge Worker Productivity. Report from Basex.

Stahl, G. 2006. *Group Cognition: Computer Support for Building Collaborative Knowledge*. Cambridge, MA: MIT Press.

Stahl, G., Koschmann, T. & Suthers, D. 2006. *Computer-Supported Collaborative Learning: An Historical Perspective.* In: Sawyer, R. K. (ed.) *Cambridge Handbook of the Learning Sciences*. New York: Cambridge University Press: 409-426.

Stivers, T. & Sidnell, J. 2005. Introduction: Multi-Modal Interaction. *Semiotica* 156(1/4): 1-20.

Suchman, L. 1987. *Plans and Situated Actions: The Problem of Human-Machine Communication*. Cambridge, UK: Cambridge University Press.

Suchman, L. 1992. Technologies of Accountability. Of Lizards and Aeroplanes. In: Button, G. (ed.) *Technology in Working Order. Studies of Work, Interaction and Technology*. London/New York: Routledge: 113-126.

Tannen, D. 1993. *Framing in Discourse*. New York/Oxford: Oxford University Press.

Tanskanen, S. 2006. *Collaborating towards Coherence: Lexical Cohesion in English Discourse*. Amsterdam/Philadelphia: John Benjamins.

ten Have, P. 2007. *Doing Conversation Analysis*. 2nd edition. London: SAGE.

ter Hofte, G.H., Otte, R.A.A., Kruse, H.C.J. & Snijders, M. 2004. Context-Aware Communication with Live Contacts. *Conference Supplement from the 2004 ACM CSCW Conference*, Chicago, MA.

Voida, A., Newstetter, W. C. & Mynatt, E. D. 2002. When Conventions Collide: The Tensions of Instant Messaging Attributed. *Proceedings of the 2002 ACM CHI Conference*, Minneapolis, MN: 187-194.

Walther J. B. 1996. Computer-Mediated Communication: Impersonal, Interpersonal and Hyperpersonal interaction. *Communication Research* 23: 3-43.

Walther, J. B. & Parks, M. 2002. Cues Filtered Out, Cues Filtered In. In: Knapp, M. L. & Daly J. A. (eds.) *Handbook of Interpersonal Communication*. Thousand Oaks, CA: SAGE: 529-563.

Whittaker, S. 2003: Theories and Methods in Mediated Communication. In: Graesser, A., Gernsbacher, M. A. & Goldman S. R. (eds.) *The Handbook of Discourse Processes*. Hillsdale, NJ: Lawrence Erlbaum: 243-286.

Woerner, S. L., Yates, J. & Orlikowski, W. J. 2007. Conversational Coherence in Instant Messaging and Getting Work Done. *Proceedings of the 40th HICSS Conference*, Hawaii.