Observing Mobile Multimedia[1]

June 16, 2013

Ilpo Koskinen
Aalto University, Helsinki

Abstract

This paper describes how mobile multimedia was observed in two early studies in Helsinki, *Mobile Image* and *Radiolinja*. It first details theoretical reasons for methodic choices. It then details research designs, how the Internet was used as a tool for observation in these studies, discusses some limits of the method, and describes some of the ethical problems in collecting actual multimedia messages.

1. Introduction

How do people use mobile multimedia? This was a seemingly simple question we posed in Helsinki early 1999 after getting funding for a research project for the future of digital imaging. Even though multimedia messaging was not yet on the market, it had figured prominently in industry concepts for years, and we knew it was going to be reality in near future. The question for us was, however, how best to study it. This question led us to a series of questions about research methods and methodology, but more importantly, it also led us to several decisions on theory. The result of these investigations were pioneering studies on mobile media using a methodology that is still relevant for the research community.

This paper describes two studies we did in Helsinki between 1999 and 2002. The first of these, *Mobile Image*, saw daylight in March 1999 and went on for about slightly over a year. The second study, called *Radiolinja* after the mobile carrier that funded the project, was done in (northern) Summer 2002. Four researchers and numerous research assistants participated in these projects. The main researchers in *Mobile Image* were Ilpo Koskinen and Turo-Kimmo Lehtonen, two sociologists and Esko Kurvinen, industrial designer; in *Radiolinja*, Koskinen and Kurvinen teamed up with Katja Battarbee, who is industrial designer. These studies led to several follow-up projects. By 2007-8, these studies had led to six books (including four PhD theses) and several articles. In this paper, we will not deal with these follow-up studies, though.[2]

Although we have been talking about design research, the studies were done in a design university, and our main inspiration for these studies was the idea of ready-mades, this paper shows that there was more than art in these studies.

This more can be found in their theoretical roots. In terms of methods, these studies built on two sources. First, ethnomethodology came into picture when we saw the first messages and realized that they are organized locally in a sequential context, much like conversation analysis claims.[3] Messages 1-2 show how one sequence works.

Message 1. Greetings

July 2 2002 19:16 Tom to Ann Marie
Greetings.

Message 2. Return Greeting

July 2 2002 19:33 Ann Marie to Tom
Greetings back to you! Also from mom!

We also soon realized that looking at sequences of messages gave us a robust method for explication: we saw the same sequences again and again. There were greetings, questions and answers, teases and replies, riddles, and so on. In the language of ethnomethodology, multimedia messages were, if not quite like conversation, designed with a very robust set of ethnomethods that had a moral basis as well. A greeting with no reply prompted demands for reply, or accounts for not replying very much like in conversation. Second, symbolic interactionism came into picture when we wanted to understand how experiences are co-construed. For instance, a sign like “Paris” functions in messaging not only locally, as

ethnomethodology taught us, but also as a symbol used to define qualities not evident in messaging.[4]

2. Studying Messages as they Happen

Both frameworks posed similar demands for research methods. Namely, they told us to study messages as they happen. This is obvious for any ethnomethodologist; when we followed Blumer we wanted to focus not so much on how people construct experiences together, but how they modify and twist these experiences.[5]

This background thinking meant that we had to find a way to observe multimedia very closely. Things crucial to ethnomethodology required that we needed to get individual messages. Also, we had to see how recipients treat these messages. Furthermore, we needed to see how the senders of the original messages treat these responses – and so on. Things crucial to interactionism set similar requirements. To see how people create shared definitions of situation, we needed to see how they observe things, agree on these observations, and sink these agreements into an unquestioned background resource of action. Also, we had to see how these definitions are maintained and how people work their way out of them.

We needed to find ways to capture messages in sequence, which was our essential resource for interpretation.

For example, a picture of a crying baby can be designed in many ways. It can become an endearment in the family, but it can also be a tease sent to a whining teenager or little brother. Whether it becomes an endearment or a tease, on the other hand, depends on how the recipient takes it. A response like “he is soooo CUTE when he is crying!!!” establishes a very different definition of the situation than “I’m NOT

a baby." Without seeing the actual messages we quite simply do not know what happens in messaging, our background theories told us.

In terms of methods, the key implication of our theoretical choices was that it rules out things like interviews and observations. People are notoriously bad at observing their own action as it happens, and most folk theories of interaction are barely more than bad psychology. Observing people sending messages was out of the question for even better reasons. By observing – designers usually talk about "shadowing – people, we might have been able to see what they do when they capture something and how they design a message. What would have been impossible, however, was seeing what the recipient does. By relying on ethnographic observations, we would have lost a resource that was essential for interpretation, messages as they happen in their naturally occurring sequences.

Our theories also gave us analytic protocols. Both traditions rely more on less on "analytic induction," as the sociologists Thomas and Znaniecki once called a method that builds up from observations into more abstract categories and finally to conceptual systems.[6] This method starts from individual instances of data, sets them in sequence, describes these sequences, and then poses the question of what people achieve with these sequences. For example, our crying baby picture sent to a teenager may become an efficient tease if the teenager is insulted, leading to an exchange that actually strengthens the emotional bond between the sender and the receiver. It may also become entertainment to others who are CC'ed into the message sequence. After describing these sequences and their functions, the analysis proceeds into classifying them. Thus, we described sequences like greetings, good mornings, question-answer pairs, teases, riddles and fights.

As this brief introduction shows, our theoretical choices led us away from the then-dominant methods of studying phones. Most early work on mobile phones was based on interviews or observations.[7] One particularly important precedent had, however, collected actual messages. In a study of text messages Elisa Kasesniemi and her colleagues had collected actual messages and analyzed how they were performed, but their analysis built on ethnology rather than ethnomethodology, and was interested in studying the culture of text messaging rather then their delivery.[8] For them, the sequence of messages was not important; for us, it was the key to understanding messages.

Our focus on messages in their naturally occurring sequential context was, as far as we know, the first of its kind in mobile media studies. By now, it is my no means unique in literature. Other researchers who deserve mention are Ditte Laursen, whose PhD thesis looked at text messages; Ilkka Arminen and Sanna Raudaskoski, who have focused on Wireless Application Protocol, an early version of Mobile Internet (); Christian Licoppe, who has focused on video in interaction with Marc Relieu, who has studies several experimental technologies like mobile drawing platforms (); and Nancy van House and Marc Davis, whose study of how a community of students used its camera phones was partly inspired the Helsinki studies. To our knowledge, in addition to us, only Licoppe, Relieu and van House's group have had access to actual messages by observing them through the Internet.[9]

3. Research Design

In *Mobile Image*, we gave Nokia Communicator (9110) and a Casio digital camera for four groups for roughly two months each to be used freely. Radiolinja (today

Elisa), a Helsinki-based mobile operator, provided telephone service. In 1999, the Communicator was the only phone on the market capable of storing, sending, and receiving photographs. With the Casio camera, people could take pictures, beam them into the phone with infrared, and send them as attachments in e-mail messages. Recipients could view these messages either in e-mail, or download them to their phone to be viewed on the screen, which used grey scale colors at that time. The whole process took place over a wireless connection (GSM).

We recruited three groups of friends and acquaintances, each group having five members. Table 1 describes the groups and gives details of their composition and the study period. Before the study, we organized a 2-hour training session to teach people how to use their communicators and the camera, and how to troubleshoot issues like IP addresses. We also interviewed them. After the study, we did a follow-up interview. During the study the male and the female groups sent a total of 371 e-mail messages (258+113). A single message contained 1-16 photographs.

Table1. Groups in *Mobile Image*

1. Pilot group: 30.6.1999-31.12.2000
Researchers of the University of Art and Design Helsinki, four men and one woman (three industrial designers, two sociologists) (*1961-1973)

2. Male group: 2.3.-8.5.2000
A group of friends, university-level business and engineering students/graduates already in jobs (*1973-1975)

3. Female group: 17.5-3.7.2000
A group of friends, students of social sciences at the University of Helsinki (*1973-1976)

4. Control group: 14.12.2000-15.3.2001
Friends and work acquaintances, designers working in a U.S. owned new media company (training in design and technology) (*1977-1979)

In *Radiolinja*, we followed three groups in the Finnish mobile phone operator Radiolinja's (now Elisa) technology and service pilot, which took place in July 11-20,

2002, and lasted about 5 weeks. Three mixed-gender groups with 7, 11, and 7 members were studied. There were 15 men and 10 women in these groups. The average age was 28 (Md=30). Each user was given a multimedia phone. Seventeen (17) participants had Nokia 7650 with an integrated camera, and eight (8) SonyEricsson T68i with a plug-in camera. Out of the Radiolinja pilot, we selected groups to take into account gender difference, terminal types, and the city-countryside axis. Most participants lived in the Helsinki metropolitan area.

Exact numbers are confidential, but the following figures point the scale of messaging in the pilot. In all, users sent over 4000 messages during the pilot. Over 2000 were unique (the rest being duplicates in group messages, or recycled messages). These data were produced through the Radiolinja system automatically. The service was free of charge.[10]

4. Using the Internet for Observation

In both studies, we observed messages through the Internet. In *Mobile Image*, the method was very simple. Esko Kurvinen's e-mail address was added to the default list of recipients. We also requested that when the participants sent photographs to people outside the research group, they would simultaneously send the photograph to his e-mail. The process was not automated, though, as this gave people an opportunity to keep some messages out of our sight. In Kurvinen's mail, messages were saved as Microsoft Outlook files, and could be browsed with any Web browser that could import messages from Outlook.

In *Radiolinja*, observations were done by tapping into Radiolinja's pilot study. Radiolinja was getting prepared for a new service early 2002. As multimedia requires

far more bandwidth than phone calls and text messages, Radiolinja had to build extra capacity for new devices. The company, however, did not know how much was needed, and to learn this, is did a pilot study. It recruited seven groups of friends and acquaintances from its customer base for the study, gave them phones, and opened its network for messaging in summer. Though most reasons for the pilot were technological, this research design was partly inspired by *Mobile Image*. It had been reviewed on the economy pages of the largest newspaper of the country, and this review got into the attention of a product manager at Radiolinja. He read the book and contacted us, asking whether we would like to participate in the pilot.

Our observations were done using a variation of the method developed in Mobile Image. When one of the participants in the study captured an image, designed a message, and sent it, it went through Radiolinja's network that transmittted it to recipients. We built a Web link to this data stream, and through this link, we captured messages while they were being sent. From messages, we captured text, sound and images (but not video). We hired a research assistant who went to the Web site once a day to record the day's messages. She did screen captures of messages and saved them as Adobe Photoshop files. Sound snippets were saved in .amr format, which could be opened with most multimedia and sound programs. (Figure 1).

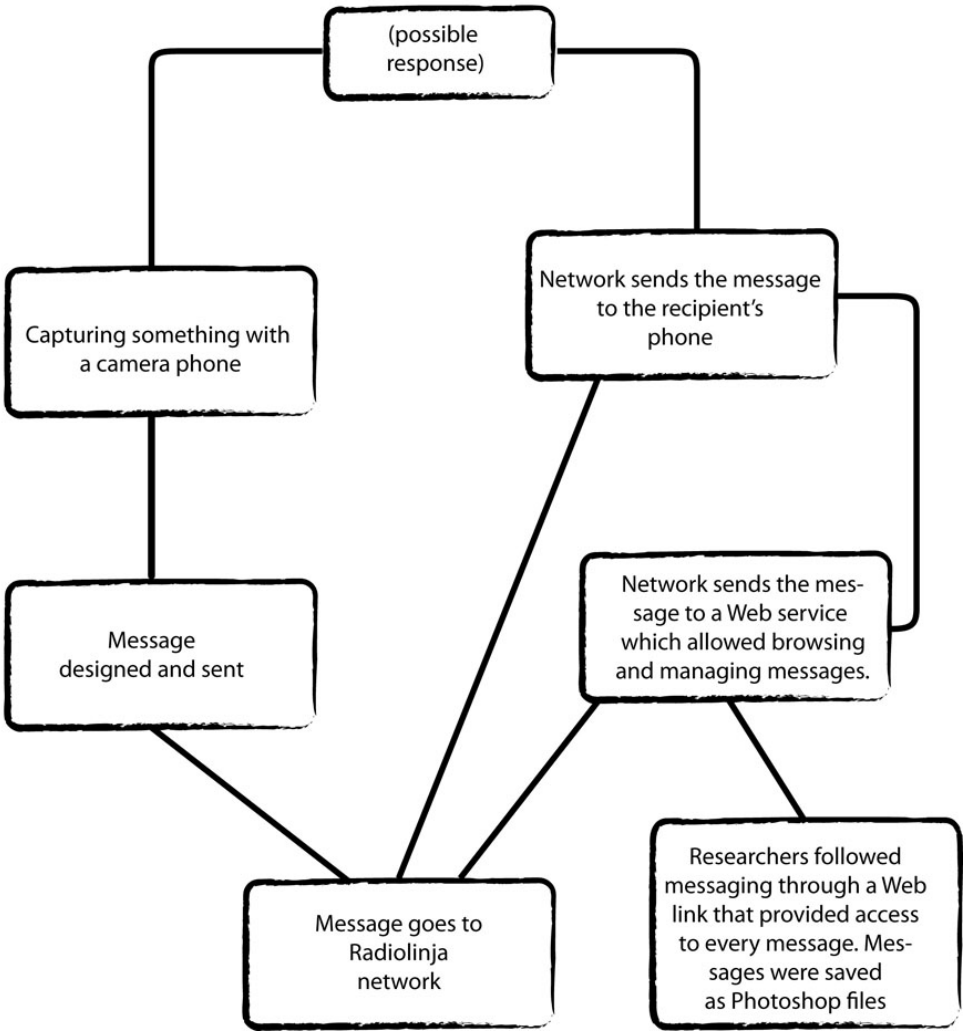

Figure 1: Observing multimedia messaging through the Web in *Radiolinja*

It goes without saying that having a corpus of thousands of messages made many kinds of other analyses possible. Most interestingly perhaps, we did many types of time series and breakdowns to compare differences between various groups and things like weekends. Statistical studies, however, were not in our interest, as it was unclear what the population for sequences like riddles would have been.

5. On the Limits of Observation

Like any method, ours had some obvious limitations. The most common criticism we have heard is that our method does not give us access to the actual moment in which some message is designed. Who took a photo? Why was it taken – what was the specific reason for taking it? Why was text added to the message, and how was it compiled? What was not said? Why?

Getting at this aspect of messaging, however, was not out aim. We were interested in social action in the virtual domain. When someone gets a multimedia message, he has to make sense of it and think about what it means and what kind of response it requires. When he responds, he shows a part of this analysis to the sender of the first message. When that person, in turn, gets to see the response, he learns how the recipient understood the first message. He can also analyze what kind of action the recipient did. He does not have access to what went on in when the recipient got the message. This work can sometimes be inferred from the message, but in the main, the only thing the sender can build on is the response. Everything else is guesswork that may build on the message, but may take a more imaginative direction just as well.

As far as we know, a few studies have focused on the in-situ work of compiling messages. For example, as early as 2002, Marc Relieu integrated video camera and eyeglasses to capture how people observed their whereabouts, captured parts of it, and designed their messages. The “glass cam” gave him a glimpse of what went on around the sender when he was working on a message. Relieu was able to show how messages resulted from joint work: how people discovered photo-worthy things around them together, interpreted them, and how they designed messages together. In contrast, we were able to read bits of this work only indirectly by looking closely at messages. Thus, a picture from a man basking in the sun in the Italian Alps is obviously taken by someone, and the relaxed pose was rehearsed, but we did not

get access to this in-situ work. For us, this was a separate question and posed no challenge to our method: recipients had access to this work only through the same information as we did.

6. Research Ethics

A method like ours raises a few ethical questions. Obviously, every person we studied knew they were being studied. In *Mobile Image*, we made it clear that they knew their messages were being recorded and that we would be using their images and texts in our studies. We also told that whenever we publish our work, we will mask the identities of the participants the best we can, and will not show their faces without permission. The baseline of our ethics was informed consent in its standard form. This was the baseline in *Radiolinja* as well, but with a strict promise that whenever we publish messages, we specifically ask permission from people in images, as well as people involved in taking them. One major difference between the studies was the quality of the camera, though. In the first study, pictures were taken with a mid-market camera with good controls and resolution, which made many details of the participants' lives open. The latter study was done with the first generation of camera phones, and many images were so smudgy that details did not matter that much, unless text explained what was going on in them. This made the latter study a bit easier from our standpoint.

The problem with informed consent was that in observational studies, people tend to forget they are being studied surprisingly quickly. Both ethnographers and conversation analysts have repeatedly noted that, people soon forget that they are being studied.[11] Quite simply, when people are busy doing something, they focus on

their line of action, not on staging a performance for researchers. At the workplace, people have work to do. In informal settings, people commit themselves to socializing. Indeed, this was the case in both studies. We witnessed family fights, drinking sprees, and even home-made porn. We have written about these messages, but have never published them in an attempt to be tactful. In both studies, we also saw messages in which senders asked others to delete some of the silliest messages they had received, and not to circulate them. If we noticed such messages, we kept these messages out of our analysis.

Part of the problem was that they didn't believe we would actually use the data, maybe because we were based in a design school. Thus, when we published the Finnish version of *Mobile Image* in 2001, we gave a copy of the book to every participant. When Kurvinen met some of the women studied for the book, they were surprised, telling that our intention to write about the study had not been real.

Our ethics, however, went beyond informed consent. In *Mobile Image*, the most important ethical procedure was the method we used. All messaging went through the university's server. What was more important, though, was that people could choose whom they wanted to add to the list of messages. This gave them an option not to share their messages with researchers. Indeed, this happened: we saw messages in which there were references to messages we had not seen. This, we thought, was a fair price we had to pay for ethics.

Radiolinja, on the other hand, introduced two types of complications to *Mobile Image*. First, people could not hide their messages from our research assistant. As soon as a message was sent, it was in the network and available to her. Second, we were on a legally complicated terrain of telecoms law. The right to privacy is one of the legal cornerstones of communications industry, and there was no way around this

fact. However, in legal terms *Radiolinja* was a technology pilot organized by a mobile operator, which funded the pilot, recruited participants, gave people phones, and paid their telecom bills during the pilot. In exchange, participants agreed to waive some of their rights to privacy. They knew that they were being studied beyond any reasonable doubt. Research contract rather than telecom law defined their right to privacy, which meant that informed consent, a procedure that guaranteed that participants could block us from publishing images and ordinary tactfulness were enough to secure our rights to data.

7. Observing Mobile Multimedia through the Internet

As this paper has shown, the seemingly simple question we posed in 1999 was not that simple. It led us to a journey to theory, methodology, technology, and research ethics, and produced a significant document of how ordinary people used the first generation of camera phones.

We believe that our analyses are still valid, with one qualification. A good deal of multimedia content captured today with camera phones is downloaded to social media rather than sent directly to another phone. We explored this option in both studies as well, making people aware of services line Zing and Kodak, but at the turn of the century, it took some technical mastery to use these services fluently. It took even more technical mastery to browse content on these services. In one sense, however, we see many things that form a continuous line between what we saw then and what we see now. First, when people capture and send multimedia content to another phone, they still use methods we described. Second, most interaction models in social media are not at all different from what we saw. People share content,

comment it, ask questions, tease and poke others, and design riddles – just like in our studies. The distribution mechanism differs rather than the interaction order. A good deal of the interaction order, also is commercialized and gamified rather than organized on a person-to-person basis.

In methodic terms, our studies are relevant even today. When looking at current research on mobile media in this book, our studies of course look clumsy. They were, however, pioneering studies, and our methods gave us data most researchers working today would still be envious of.

Still, if this paper has one key message, it is probably a classic old university adage. Theory and method must go hand in hand. For us, our theoretical choices told us those kinds of data we needed, and which methods to use. After seeing the first messages and realizing that they are sequentially organized, we built our analysis on ethnomethodology and conversation analysis (and in the case of Katja Battarbee, on symbolic interactionism).[12] This choice gave us marvelous tools to identify, describe and classify methods people use – ethnomethods, in brief – in working with multimedia phones.[13] As soon as we decided to focus on describing ethnomethods, we had to get naturally occurring data with sequential context intact. It is good to keep in mind, though that our work was never pure theoretically; one of our researchers, the sociologist Turo-Kimmo Lehtonen, construed his interpretations around Georg Simmel's theories of interaction. However, he mostly worked with interviews and used actual messages only as a way to illustrate his arguments, while in Kurvinen's, Koskinen's and Battarbee's work, actual messages were the starting point of explication, and also provided the test for whether their explications made sense.

When looking back at research on mobile telephony around 2000, our studies constructed their object in a very different way from the then-popular cultural studies

perspectives. Even Timo Kopomaa, a sociologist interested in interaction, relied on interviews and casual observations rather than tried to get at the actual content of calls.[14] In contrast to Kopomaa's, Elisa Kasesniemi's and Rich Ling studied the uses and meanings of phones and multimedia through statistics and interviews, our approach explicated them by looking at the actual activity of capturing, sending and responding messages.[15] Slightly later, we got company from Marc Relieu and Christian Licoppe, Ilkka Arminen, as well as Rich Ling and Tom Julsrud in tapping into the content of messages.[16] What we won was detail; what we lost was an ability to make sweeping generalizations to society and culture.

The interests of designers and social scientists tend to be years apart in that the former work usually with products of imagination and are interested in things that may become reality one day, while the latter prefer to study things that are already taking place in society. In our study, the line was barely noticeable. Our data, methods and analytic protocols heavily borrowed from sociology, but the implications of our studies were written for engineers and designers. We never faced any demands from the design side to change our tools either. In hindsight, the reason was right in front of our eyes. Our tools were familiar for designers from art history. In design school terms, both studies probed possible futures through ready-mades rather than more typical design alternatives like systems or experience prototypes. We never found it difficult to explain our studies to designers; they were able to see what made our studies useful just as easily as sociologists and communication scholars.

7. References

[1] We would like to thank Finland's Ministry of Trade and Transportation, Radiolinja and the industrial design program of the former University of Art and Design for support for projects described in this paper in 1999-2002. In Mobile Image, the tip for using Nokia Communicator with Casio camera came from Dr. Hannu Nieminen, then at Nokia, Tampere. Special thanks to Radiolinja's Seppo Väkevä.

Early versions of this paper have been presented in "Fieldwork Untethered,"a workshop organized by Mirjana Spasojevic, Mizuko Ito, Nancy Van House, Ilpo Koskinen, Fumitoshi Kato, Daisuke Okabe in November 2005 at Keio University in Tokyo; and in a seminar organized by Larissa Hjorth at the Royal Melbourne Institute of Technology in December 2010.

[2] Ilpo Koskinen, Esko Kurvinen and Turo-Kimmo Lehtonen, *Mobile Image.* (Helsinki: IT Press, 2002). (In Finnish as *Mobiili kuva* by IT Press, 2001); Esko Kurvinen, *Prototyping Social Action* (Helsinki: University of Art and Design, 2007); Ilpo Koskinen, *Mobile Multimedia in Action.* (New Brunswick, NJ: Transaction Publishers, 2007).

[3] For a statement of classic ethnomethodology, see Harold Garfinkel, *Studies in Ethnomethodology* (Englewood Cliffs, NJ: Prentice-Hall, 1967). This paper is not a place to go into details about ethnomethodology or conversation analysis. Suffice it to say that it is debatable how compatible they are; ethnomethodology tends to push analysis to unique details of how social organization is achieved, while conversation analysis has universalistic tendencies. For discussion of conversation analysis and ethnomethodology, see Michael Lynch, *Scientific Practice and Ordinary Action. Ethnomethodology and Social Studies of Science* (Cambridge: Cambridge University Press, 1993).

[4] In this analysis, we followed Herbert Blumer, *Symbolic Interactionism: Perspective and Method* (Berkeley: University of California Press, 1969). See Katja Battarbee and Ilpo Koskinen, "Co-Experience: User Experience as Interaction," *Co-design* 1(1) 2004: 5-18.

[5] Esp. Erving Goffman, *Behavior in Public Places. Notes on the Social Organization of Gatherings (*Glencoe, Ill.: The Free Press, 1963).

[6] William I. Thomas and Florian Znaniecki, *The Polish Peasant in Europe and America, 1-2* (New York: Octagon Books, 1974).

[7] For example, Timo Kopomaa, *The City in Your Pocket. Birth of the Mobile Information Society* (Helsinki: Gaudeamus, 2000); and Daisuke Okabe and Mizuko Ito, Everyday Contexts of Camera Phone Use: Steps Toward Technosocial Ethnographic Frameworks. In Joachim Höflich and Maren Hartmann (eds.) *Mobile Communication in Everyday Life: An Ethnographic View* (Berlin: Frank & Timme, 2006).

[8] Eija-Liisa Kasesniemi, *Mobile Messages. Young People and New Communication Culture.* (Vammala: University of Tampere Press, 2003).

[9] Ditte Laursen, Please Reply! The Replying Norm in Adolescent SMS Communication. In Richard Harper, Leysia Palen and Alex Taylor (eds.) *The Inside Text. Social, Cultural and Design Perspectives on SMS* (Dordrecht: Springer, 2005); Sanna Raudaskoski, *Tools and Machine. The Affordances of the Mobile Phone*. (Tampere: Tampere University Press, 2009); Ilpo Koskinen and Petteri Repo, "WAP and Accountability: Shortcomings of the Mobile Internet as an Interactional Problem," *Journal of Usability Studies* 2(1), 2006: 22-38; Christian Licoppe and Marc Relieu (eds.) *Mobiles.* (Paris: *Hermes Science revue Reseaux* N. 112-113, 2002); Nancy van House and Marc Davis, "The Social Life of Cameraphone Images," presented at the Workshop Pervasive Image Capture and Sharing: New Social Practices and Implications for Technology at *Ubicomp'05*, Shinagawa, Tokyo, Sept. 11-14. Retrieved Sept. 20, 2005, from http://ubicomp.org/ubicomp2005.

[10] More details of both studies are in Koskinen, op. cit.

[11] Howard S. Becker, "Field Work Evidence," in Howard S. Becker, *Sociological Work. Method and Substance*. (New Brunswick, NJ: Transaction Books, 1970), 42-44; Douglas W. Maynard, *Inside Plea Bargaining: The Language of Negotiation*. (New York: Plenum, 1984), 21-22.

[12] Katja Battarbee, *Co-Experience. User Experience as Interaction* (Helsinki: UIAH, 2004).

[13] Kurvinen, op. cit.

[14] Kopomaa, op. cit.

[15] Kopomaa, op. cit.; Kasesniemi, op. cit.; Rich Ling, *The Mobile Connection.* (San Francisco: Morgan Kaufmann, 2004).

[16] Licoppe and Relieu, op. cit.; Marc Relieu, "The "Glasscam" as an observational tool for studying screen-based mobile phone uses and management of parallel activities," presented at *The International Conference on Conversation Analysis ICCA-2002*, Copenhagen, Denmark, June, 2002; Ilkka Arminen, *Institutional Interaction. Studies of Talk at Work*, (Aldershot: Ashgate, 2005); Rich Ling and Tom Julsrud, "The Development of Grounded Genres in Multimedia Messaging Systems (MMS) among Mobile Professionals," in Kristof Nyíri (ed.) *A Sense of Place.* (Vienna: Passagen Verlag, 2005).