

# Towards Fine-Grained Usability Testing: New Methodological Directions with Conversation Analysis

Marko Nieminen<sup>1</sup>, Sari Karjalainen<sup>2</sup>, Sirpa Riihiahho<sup>1</sup>, and Petri Mannonen<sup>1</sup>

<sup>1</sup> Helsinki University of Technology, Department of Computer Science and Engineering,  
Espoo, Finland

<sup>2</sup> University of Helsinki, Department of Speech Sciences, Helsinki, Finland  
{Marko.Nieminen, Sirpa.Riihiahho, Petri.Mannonen}@tkk.fi,  
Sari.Karjalainen@helsinki.fi

**Abstract.** We examine the possibilities of conversation analysis (CA) in usability testing. The goal is to examine how usability test setups serve as source for the CA analysis. We used video data from two earlier usability tests. Our results indicate that traditional test setup does not serve as a sufficient source for CA. The actions in the user interface were unclear, the user's facial reactions were not visible, and the user is occasionally having more conversation with the moderator than the system. CA approach can be taken towards two separate directions regarding usability tests: Analysis can be focused to the dialogue between the moderator and the user or on the user-system interaction. There is a need to fine-tune data gathering with detailed level recording of keypresses and system outputs. However, CA-enhanced usability testing allows in-depth analysis of usability problems as well as analysis of holistic interaction between user and system.

**Keywords:** Usability test, usability evaluation, conversation analysis, situated action, dialogue structures.

## 1 Introduction

Usability testing is perhaps the most widely applied practical usability evaluation method [1] in the development of interactive applications. Nowadays, the practical utility of the method is evident. It has long been considered as the cornerstone method for all companies that hesitate whether to apply user-centred approach in their development work. In his concept of "Discount Usability Engineering" (DUE), Nielsen [2] already promoted the method to be a good starting point: According to the idea of DUE, it is pretty easy to employ some users to test software and see how they perform with it. In our own research, we have applied usability testing both in research [3] and teaching [4] in the usability laboratory of Helsinki University of Technology. The core method has been studied and developed further in variations called Visual Walkthrough [5], Informal Walkthrough [6] and Contextual Walkthrough [3].

The main focus in traditional usability testing has been in the actions on the computer screen and the recorded speech. A key element in most usability tests is the

thinking aloud method. The core of the method has remained rather fixed with tiny variations applied to different study setup types (e.g. Apple Guidelines [7]) During the last decade, several development and consulting companies have adopted and applied the method successfully. Through the method, the researchers and practitioners have been able to discover usability problems and suggest improvements to the interface design.

The thinking aloud method reveals the user's cognitive reasoning and motivational basis behind his actions with the interface. Reflecting user's behaviour indirectly, the thinking aloud method supplements other usability evaluation methods, such as cognitive walkthrough and heuristic analysis. So far, user's other behaviour than explicitly expressed conscious thinking has not been addressed explicitly with supporting conceptual framework in the analysis of the results of the usability test. In order to find out deeper insights into user's understanding and actions, new approaches are required.

There is a need to move from straightforward summarizing type of analysis towards more interpretive and hermeneutic analysis. In searching for the new analytical insights, we are interested in dealing with the whole interaction between the user and the interface, and the variety of interactional resources used in interaction. We expect that Conversation Analysis (CA) will provide new means for analyzing the human-computer interaction in more detail and for answering practical questions, such as: Why (and how) do users end up with problems in the interaction? Why (and how) are some features fluent to use?

### 1.1 Conversation Analysis Method

Ethnomethodological conversation analysis (CA) is a qualitative and data driven analysis approach. Usually CA is done by analysing video data that is gathered in natural situations. During the CA, the selected data relevant for the research objectives are transcribed and analysed. The analysis proceeds without theoretical speculations of the nature of the interaction. The CA reveals how interaction is locally organized and sequentially structured of both verbal and nonverbal utterances. In spite of its name, CA is concerned with the understanding of multimodal interaction including speech and body language [8, 9].

In practice, the details of the internal structure of individual sequences are examined, for example by following Have's [10] guidelines. First, successive utterances are constructed as turns, and then the turns are organized as sequences through the recurrent structures. Characterizing the actions in the sequence include for example the participant's selection of utterance, timing and taking of turns, and how these support participant's understanding and meaning-making of the actions performed. Thus, the organized interpretations that participant himself employ, are described.

CA provides qualitative findings including detailed descriptions of the systematic structural characteristics of particular interactional phenomena. The different phenomena, then, are grouped in collections of different sequences in shorter or longer sequences of interaction. The findings concern the concepts and structures such as turn taking, repair, topic, opening and closing.

*Turn-taking* is a unit of conversational exchange including an initiation by A as an item which begins anew conversation and sets up an expectation of a response

followed by a response by B. There may also be optional elements of exchange structure in the follow-up turns.

*Repair* is related to some trouble in interaction during a conversation. It includes the sequence of actions and procedures relating and following the error or misunderstanding. The repair is essential in interaction since problems are likely to arise and must be corrected if the interaction is to be successful.

*Topic* means the matter dealt with in the interaction. Especially the ways each participant introduces (new) topical material and provides opportunities for the other participants to introduce items of interest. The topics flow from one to another usually through stepwise progression.

*Opening* and *closing* are turns where flows of interaction are initiated or brought to an end. In human-to-human conversations participants usually introduce themselves during openings and the closing statement calls for goodbye from the participant who is about to leave.

Generally the findings concern the ways in which the actions are recognized and performed by the participants in the selected sequences. Including the detailed analysis of context, CA findings offer deep micro-analytical insight of the actual phenomena of situated human action (for example [11]).

Recently, CA has been successful in exploring early human interaction with a variety of multimodal resources [12]. The analytical challenges to be dealt with reflect those arising in the HCI-data whereby the computer is examined as another participant in interaction and treated as a 'social agent'. Practically, the aim is to explore how the user makes the action through his embodied actions and how the application displays its orientation to the user's actions through certain (iconic) means that are depending on the nature of the application and its resources.

The CA approach has been contributing to both practical and theoretical issues from different aspects on the interactional settings of human-machine interface to emerging re-conceptualizations of social/material relations within HCI field (see [13]). CA has been used to study and design new technology in a variety of fields (see [14], [15], [16], [17]).

## 1.2 Aim of the Study

In this paper, we examine the possibilities of CA in usability testing. The main goal here is to examine, how well the traditional test setups may serve as a source of information for the CA analysis. We also try to assess the extra value that CA gives compared to traditional evaluation methods, such as heuristic evaluation, cognitive walkthrough and usability tests with thinking aloud protocol.

The video recordings of usability tests were selected as analysis material for CA since utilizing existing data would be an easy and inexpensive way to introduce CA to product development projects. The goal is not to compare the individual results, i.e. found usability problems, of different methods but to analyze whether CA could provide a useful and important additional viewpoint to usability evaluation methods in product development projects.

## 2 Our Experiment: Conversation Analysis with Usability Test Material

We examined the possibilities of CA by analyzing videotapes recorded in our earlier usability studies. The evaluated systems were a training application that helps in planning and monitoring goal-directed training, and a new electric payment system in a gaming slot machine. In addition to the usability tests, we had conducted a cognitive walkthrough and a heuristic evaluation to the systems.

### 2.1 Video Data

Since CA usually focuses on certain interaction sequences and not to whole episode, we needed to select the most prominent parts of usability test recordings for our analysis. The identified usability problems were selected for the analysis for two reasons: first, the goal of usability evaluation is to reveal usability problems and to discover the reasons, why it is a problem, and, second, the user-system interaction is “most natural” in the problem situations, since the user often stops thinking aloud. Selecting only those usability problem situations that had been also covered in the cognitive walkthroughs and heuristic evaluations further narrowed the focus of the CA.

Two examples of selected video sequences are:

- Training software usability evaluation: 3.20 min sequence of ‘Analyzing the exercises’

In the selected sequence, the user is trying to compare two training sessions. The results are illustrated graphically in a frame. The user has difficulties in finding the command for comparing. Moreover, she has difficulties in identifying which session is which in the common frame.

- Electronic payment system usability evaluation: 2.20 min sequence of ‘Completing the payment’

In the selected sequence, two test users push a button on the slot machine to finish an operation in the paying procedure. The action is quite correct, but they are not sure, where the won money went, since it was electronic money and not coins, as usual.

### 2.2 Observations

The original usability testing video data were not gathered Conversation Analysis in mind. Due to this, some questions concerning mainly the actions of the interface remain without answer. As one of the aims of this study is to reveal new types of findings in a usability test, the role of CA is of informative nature. Thus, instead of the detailed conversation analysis procedure, we will present here some observations regarding the ways in which such an analysis may be accomplished in terms of CA and ways in which it may contribute to an understanding of the wider issues of the CA approach.

It is observable from the video data that the users activate the test moderator to participate in the on-going sequence and conversation. In problem situations, the users

often pose an explicit question expressing their thoughts and confusions. The user in the payment system poses a question that is accompanied with a gaze shift towards the moderator. The confirmative question (originally in Finnish: 'ja ne kaikki meni tilille vai?'; translated into English: 'And they all went to the account, right?') displays uncertainty about the destination of the money in the paying procedure she just accomplished.

The user's question in the training software ('mistä mä tiedän kumpi on kumpi'; 'how can I know which is which') reveals the difficulty in unclear visualization of the objects the user is interested in. The question is not accompanied with a gaze shift. The question is not necessarily directed to the moderator, but may be rather motivated by the thinking aloud procedure.

The more detailed analysis of the user's situated action in relation to the interface would give more insight into those observations. Moreover, bearing on the CA framework, the data from both sources evoke some questions that have a bit different focus than the traditional usability testing.

In the training software case, the user assumes that the two sessions are displayed within the same frame. The 'traditional' analysis of usability tests is interested in the fact that the user cannot identify which data is which in the visualization. The CA approach is also interested in how the user makes the assumption, i.e. what kind of preceding activity has resulted the user to make the assumption. Naturally the CA is also concerned with how the user finally ends up with the right operation, i.e. what kind of dialogue the user has with the system in order to accomplish the task.

In the electronic payment system case, CA opens up questions, such as: How does the application express what the next relevant action in relation to the preceding action is? Does the interface include simultaneous or competing elements in directing the user, such as 'Cash Out' and 'Bet' buttons flashing simultaneously at the current moment? Does the application provide any help or extra information to support solving potential problems in use?

Having provided some observations and simple examples to demonstrate the ways in which CA-approach may be accomplished, we now turn to discuss both our findings and more general issues concerning CA-enhanced analysis in usability testing.

### **3 Discussion: Towards CA-Enhanced Analysis in Usability Testing**

We discuss here our findings regarding the ways in which CA may be accomplished and ways in which it may contribute to an understanding of the wider issues of the approach.

Our small experiment showed that the traditional test set up does not serve as a sufficient source for conversation analysis for two reasons. First, there is the technical problem that either the actions in the user interface were unclear or the user's facial reactions were not visible, as the user was not facing the camera. Secondly, the setting in the test is problematic, because the user is occasionally having more conversation with the moderator than the system. Therefore, we recommend some changes to the usability test settings concerning the moderator's role and data collecting.

The moderator's role in a test is to be a gracious host, be in control of how the session goes and to be a neutral observer [18]. It is important to establish and maintain a rapport with the test participants. Therefore, in our tests, the moderator practically participates in completing the tasks – in spirit, but not in practice. The moderator asks questions as a follow-up to user's expressions of problems, and gives prompts or assistance when needed. Within conversation analytic framework, the active role of the moderator contributes to more natural 'thinking aloud' by the user, being a natural conversation rather than talking aloud by oneself, which is far from being natural.

In general, conversation analysis approach can be taken towards two separate directions regarding the usability tests. Either the CA analysis can be focused to the dialogue between the moderator and the user, or the focus of the CA can be the user-system interaction.

In the first case, it would be valuable to assess the possibilities of developing the conversational interchange between the moderator and the user towards even more structured interaction. This could be realised by questions, that were prepared in advance to process some critical and potentially problematic points (interviewing-like conversation) or posing spontaneous but user-oriented clarification requests as follow-up to the user's own topical offers that are arising from his own actions with the interface.

Regarding the user-system interaction, the potential of CA is especially in situations where thinking aloud method cannot be used. For example contextual walk-through [3] is developed for these kinds of situations and CA should be tested with it. Other way to utilize CA could be to setup laboratory tests where a user could be let to test and experiment the product without a moderator or other external persons. This would ensure that the interaction happens solely between the user and the product and make the analysis easier.

In addition, the CA could fit to earlier phases of user-centred design. CA could be used in user research to analyze the users' interaction with current devices and systems, or in design and prototyping to help designers to take into account the conversational aspects of the interaction.

In order to make use of CA as a complementary method in usability tests, there is a need to fine-tune the data gathering. The normal usability test leans heavily to users capability of explaining his or her actions by thinking aloud method. CA requires that all actions are visible to the analyser. Thus, all users key-presses and mouse gestures need to be visible in the recording. In addition CA could benefit from gaze tracking as it could bring more insight into users' activities and intentions. The same goes also for the tested system. All outputs of the system, i.e. beeps, animations and blinking of the cursor etc., need to be visible for the analyzer. All in all the CA requires a more rigorous recording of the test sessions than the normal usability test.

CA-enhanced analysis also allows for making specific collections of the video data demonstrating certain types of problems encountered in the usability testing which is often important for communication of the results.

Specifically, CA can contribute to elaborating usability test findings more fine-grained, for example as to the notion of lacking feedback that was one topic in the usability testing material we used. CA can provide us with the notion that the interaction includes also many conversational actions that are not directly related to completing the task at hand. For example repair, i.e. participants ways to correct misunderstandings and error that have happened before, is not usually used as a concept

in human-computer interaction. The user interfaces include elements and functions such as undo, cancel and exit but they effectively reset the situation instead of starting a corrective dialogue with the user. CA also offers the view of the ways in which repairs are constructed (see e.g. [19]). Equally the concepts of turn-taking, topic, opening and closing allow to analyse the relationship between the system and the user, the complexity of the interaction regarding what kinds of and how many topics can be covered and how turns and roles are allocated, and the ways the interaction is initialized and ended.

CA would also contribute to supplementing the results from other methods, such as cognitive walkthrough and or heuristic analysis. The sequence examined in detail through CA-approach would also allow for more exact comparison of the empirical findings after synchronizing the current sequence with the findings from cognitive walkthrough and heuristic evaluation.

Generally, usability testing enhanced with CA would allow for in-depth analysis of found usability problems as well as the holistic interaction between the user and the system. However the benefits are not free of charge as the CA also requires stricter recording of the test sessions and takes more time than normal usability test.

## 4 Conclusions

In this paper, we examined the possibilities of CA in usability testing applied in the traditional usability test type of settings. CA offers an appealing and interesting framework for deeper understanding of the interaction between the user and the application.

The major question was how well the traditional test setup may serve as a source of information for the new analysis in order to find and figure out the fundamental interactional phenomena concerning the user's actions with the interface.

The current original testing data does not allow accomplishing a detailed CA procedure, but it allows for examining the possibilities and requirements for using CA and suggesting some changes to usability testing. Implementation of CA-approach into usability testing would result in different solutions in constructing usability testing set up. Most importantly, CA-approach can contribute to the nature of video data and data collecting techniques.

CA can contribute to elaborating usability test findings more fine-grained, as a single method or supplementing the results from other methods. Analysis is time consuming but it can be harnessed as a method for very narrow focusing on certain sequences extracted from the usability testing data. In the future, it should be examined how to develop CA for meeting the specific needs in the usability testing set up. CA offers also one framework for unifying concepts used in human-computer interaction.

More generally, there is no one way of 'doing' CA. The analyst may follow one's own particular interests in choosing the analytical focus. Traditionally, the usability testing has been interested in problematic phenomena. However, to contribute wider to the current needs and trends in HCI field, it may be valuable to explore also the phenomena that reveal the unproblematic interaction between the user and the application.

Promising uses for CA are situations in which thinking aloud is not possible. These test setups (Contextual walkthrough [3]) are often arranged so that the review of the

walkthrough is done afterwards when the actual test event has passed. Such situations are e.g. tests that are arranged in real work settings without the possibility to interact in the situation. The organisers of the test do not interfere the progress of the workflow but analyze it later with or without users. The joint interpretation of the events with users provides an interesting way of ascertaining the interpretation made by the researcher or analyst improving the accuracy of the results.

Altogether, using CA provides interesting extensions to usability tests but not directly without modifications to the test arrangements. The recording of the test needs to be adjusted to fit the needs of both analyses. This can be achieved with reasonable amount of additional test instrumentation and work, though.

Additionally, the uses of CA are not restricted to testing and evaluation only. We expect CA to provide means for analysing and improving the existing interaction and dialogue structures towards repairing ones for improved flexibility and adaptability in less constrained interaction situations. This provides interesting directions for more fundamental consideration of human-computer dialogue.

It is our intent to further research on the inclusion of CA in our forthcoming research projects both for the improvement of usability testing and for the basic research on interaction structures.

**Acknowledgments.** We are grateful to Lasse Lumiaho, Ville Toivonen and Mari Tyllinen for their input to re-evaluating and analysis of the original usability data.

## References

1. Bias, R., Mayhew, D.: Cost-justifying usability. Academic Press, Boston (1994)
2. Nielsen, J.: Usability Engineering. Academic Press, London (1993)
3. Riihihaio, S.: Experiences with usability evaluation methods. Licentiate's thesis. Helsinki University of Technology, Department of computer science and engineering (2000)
4. Koivunen, M., Nieminen, M., Riihihaio, S.: Launching the Usability Approach: Experience at Helsinki University of Technology. SIGCHI Bulletin 27(2), 54–60 (1995)
5. Nieminen, M., Koivunen, M.R.: Visual Walkthrough. In: Allen, G., Wilkinson, J., Wright, P. (eds.) HCI 1995, People and Computers, Adjunct Proceedings. The School of Computing & Mathematics, The University of Huddersfield, pp. 86–89 (1995)
6. Nieminen, M.P.: Designing user interface concepts for multimedia services. Master's thesis, Helsinki University of Technology (1996)
7. Gomoll, K., Nicol, A.: Discussion of guidelines for user observation. User Observation: Guidelines for Apple Developers (1990)
8. Goodwin, C., Heritage, J.: Conversation Analysis. Annual Review of Anthropology 19, 283–307 (1990)
9. Goodwin, C.: Action and embodiment within situated human interaction. Journal of Pragmatics 32(10), 1489–1522 (2000)
10. Have, P.T.: Doing conversation analysis: a practical guide. Sage, London (1999)
11. Atkinson, M., Heritage, J.: Structures of social action: Studies in conversation analysis. Cambridge University Press, Cambridge (1984)
12. Karjalainen, S.: Multimodal Resources in Co - Constructing Topical Flow: Case of "Father's Foot". In: Esposito, A., Bratanic, M., Keller, E., Marinaro, M. (eds.) Fundamentals of verbal and nonverbal communication and the biometrical issue. IOS press, The Netherlands (2007)



13. Suchman, L.: *Human-Machine Reconfigurations: Plans and Situated Actions*, 2nd edn. Cambridge University press, Cambridge (2007)
14. Luff, P., Gilbert, G.N., Frohlich, D.M. (eds.): *Computers and conversation*. Academic Press, London (1990)
15. Button, G., Sharrock, W.: On simulacrums of conversation: toward a clarification of the relevance of conversation analysis for human-computer interaction. In: Thomas, P. (ed.) *The Social and Interactional Dimensions of HumanComputer Interfaces*, pp. 107–125. Cambridge University Press, Cambridge (1995)
16. Luff, P., Hindmarsh, J., Heath, C.C. (eds.): *Workplace Studies: Recovering work practice and informing system design*. Cambridge University Press, Cambridge (2000)
17. Woodruff, A., Aoki, P.M.: Conversation analysis and user experience. *Digital Creativity* 15(4), 232–238 (2004)
18. Dumas, J., Loring, B.: *Moderating usability tests: Principles and practices for interacting*. Morgan Kaufmann/ Elsevier (2008)
19. Raudaskoski, P.: Repair work in human computer interaction. In: Luff, P., Gilbert, G.N., Frohlich, D.M. (eds.) *Computers and conversation*, pp. 151–172. Academic Press, London (1990)