Situated Chat Analysis as a Window to the User's Perspective: Aspects of Temporal and Sequential Organization

Michael Beißwenger
TU Dortmund University

urn:nbn:de:0009-7-15329

Abstract

This article addresses the temporal organization and the users' perception of conversational progression in ongoing chat conversations and presents a situated model of chat participation that focuses on how chat users individually manage their involvement in conversational interaction. It reports on the results of a case study on how chat users adapt their individual message production to the perceived momentary status of interaction. The study is based on multimodal data from chat user observations that capture both the users' onscreen activities (typing, editing, scrolling) and their gaze orientation toward the different visual targets that are relevant for producing and processing messages. The findings show that participating in and contributing to a chat is a highly individualized accomplishment. As a consequence, it is suggested, it is necessary to rethink the extent to which categories from the turn-taking paradigm can be adopted for the analysis of synchronous computer-mediated communication (CMC).

Introduction

This article aims to contribute to the topic of this special issue in two respects. First, it discusses the *adequacy of categories from the turn-taking paradigm* for the analysis of synchronous written CMC: Can concepts such as *turn, turn construction, control of the floor, turn negotiation, turn-holding*, and *turn-claiming* be used to describe the linear and sequential structure of exchanged text units and the progression of individual users' activities adequately? Or do these concepts actually obstruct the view of the crucial differences between oral discourse and synchronous CMC? Second, the article addresses the question of *what kind of data fit for what kind of research questions in computer-mediated discourse analysis (CMDA)*? In most empirical linguistic studies of synchronous CMC, logfiles, logfile excerpts, or logfile archives constitute the vital resource for empirical work. I will discuss if logfile data are sufficient for all research questions in CMDA or if there are research questions for which these data must be regarded as restricted. If the latter is the case, the question arises how the empirical basis for chat research may be expanded through the acquisition of other types of data.

I address these two questions in combination. First, I outline some features that are specific to chat, due to which the communicative conditions of chat must be regarded as considerably different from those of oral conversation. This concerns in particular the conditions of producing and "uttering" contributions and, relatedly, the lack of devices for real-time coordination between interlocutors. In order to examine empirically if the organization of chat should be regarded more as an individual accomplishment than a matter of interpersonal negotiation, more comprehensive data are needed than just logfiles. Logfile data do not provide information about the processual quality of message

production or of the individual users' communication-related activities—reading, typing, editing, waiting, or scrolling.

I offer an outline of how CMC research may benefit from integrating multimodal data from chat user observations, and I present an example to illustrate what these data can reveal about how users participate in chat. Finally, I present the results of a case study of a multimodal chat corpus in order to explore the phenomenon of deletions during the production of chat messages. These results lead back to the question of whether the turn-taking paradigm provides an appropriate framework for the analysis of interaction in synchronous CMC.

Coordination as an Individual Accomplishment

Because of the properties of the underlying communication technology, chat neither allows for processing of language production in real time, nor can the sender of a message be sure that the addressees perceive his/her message immediately after transmission. Specific to chat is a twofold detachment of the production and the mental processing of exchanged messages or—as Zitzen and Stein (2004, p. 989) describe it—a "spatiotemporal separation of *context of production* and *context of use,*" which in face-to-face conversations are concurrent. In addition to the lack of visual awareness, this causes a lack of simultaneity between generating behavioral output and taking in the data generated thus far. Specifically:

1. The *message exchange procedure specific to chat technology* is such that transmission in chat is not organized keystroke-by-keystroke. Instead, messages are transmitted en bloc, so that message *production* has to be entirely completed before message *transmission* or "uttering." Because of this characteristic of production, the point in time at which a chat user decides to become a message producer, as well as the whole process of message production, is detached from the point in time at which the complete message becomes visible on the addressees' computer screens.

2. Because of the *visual nature of writing*, perception and mental processing of messages become detached from the point in time at which the messages become visually available on the computer screen. In oral conversation it is almost impossible physically to ignore the occurrence of an utterance as an acoustic event. In contrast, messages in written discourse are not noticed until the reader directs his or her attention to a specific visual target on which textual information is displayed.

Owing to these two features, the users' individual mental models of the temporal and sequential structure of an ongoing conversation may differ quite significantly from one another. By *mental communication protocol* (or *mental protocol*, for short), I mean the individual mental representation of the ongoing conversation by a certain participant, which is constructed as the communication proceeds and which forms the mental resource on the basis of which he or she decides how and when to act and react in the ongoing conversation:

> Within the (mental) communication protocol, the speaker /communicant records the progression of a conversation up to that point. He updates this sequential information just like in a diary. (Herrmann & Grabowski, 1994, p. 33; Translation *MB*)[1]

For the analysis of chat, it is important to differentiate between the graphic protocol of exchanged text messages on the users' computer screens (henceforth: the *screen protocol*) and the individual *mental protocol*, the current state of which—due to the twofold detachment of the production and mental processing of exchanged messages—is not necessarily identical with the current state documented in the screen protocol. While the screen protocol is automatically updated whenever new messages are delivered to the users' computers by the chat server, the individual mental protocol is updated only when the user creates a new message (and, thus, intends to add a new move to an ongoing conversation) or when he/she directs his/her gaze and attention to the screen, detects new textual information (which has not yet been processed) delivered by the interlocutors, and, by processing this information, converts it into a new item of his/her mental protocol (Beißwenger, 2007, pp. 163-171).

To illustrate this process: While one user (*Pete* in the example given below) at a particular time is occupied with producing a message as an answer to a question posed by his interlocutor, the interlocutor himself (*Andy*), by interpreting Pete's previous utterances, may consider the requested answer as already given and therefore start initiating a new sequential pattern (for example, a new question–answer pattern). If Andy's question message is displayed before Pete finishes his answer, and if Pete does not notice this new message on his screen, the mere availability of the new message may have no effect on Pete's current activities. Instead, at the same point in time, Andy and Pete have different views of what is actually the "current state of sequential progression:" For Pete, the current state is "giving an answer to a question," whereas for Andy, it is "posing a new question after the last question was answered." In Herring (1999), phenomena of this kind are termed *interleaved exchange sequences*.

Example 1: Interleaved exchange sequences

> **Andy:** wie sieht denn die krise in der familie bei dir aus?
> 'tell me about the crisis in your family.'
>
> **Pete:** meine ma redet nicht mit meinemdad und andersrum
> 'my mom won't speak to my dad and vice versa'
>
> **Pete:** mein bruder redet nicht mit ihnen und andrsrum
> 'my brother won't speak to them and vice versa'
>
> **Pete:** ich rede wenig mit ihnen und sie gar nicht
> 'i hardly speak to them and they don't speak at all with me'
>
> **Pete:** meine eltern wollen sich scheiden
> 'my parents want to divorce'
>
> ⇨ **Andy:** darf ich fragen, wie alt du bist?
> 'may I ask how old you are?'

⇨   **Pete:**   meine mutter is hysterisch
                'my mother is going crazy'
       Pete: 12
                '12'

The example shows that chat users cooperate on the sequencing level; nonetheless, they have very little influence over the positioning of their messages in the linear order of the screen protocol. Real-time turn negotiation is not possible, because there is no mutual real-time perception between interlocutors. As mentioned before, it is characteristic of chat-based interaction that the users' mental representations of the current state of interactional progression do not correlate systematically with the current state of the screen protocol—instead, the mental communication protocols of two users at the same point in time can differ, depending on when each user last checked the screen protocol for new messages and consequently updated his/her mental protocol.[2]

Chatting as Self-Organization towards the Screen: A Situated Model of Chat Participation

Chat participation is based on self-orientation toward the screen. The communicative conditions of chat do not allow for real-time coordination of communication-related activity among the participants. Subsequently, in contrast to the co-construction of turns under face-to-face conditions, the construction of contributions to a chat dialogue in progress is not an *interpersonal*, but rather an *individual* task that each participant has to manage on his/her own. Since each participant can take up the production of a new message at any time without first having to negotiate for control of the floor, the textual units which a participant, invisible to his/her interlocutors, creates in his/her message entry box and the textual messages that he or she sends to the server do not have the status of turns, although they are textual representations of linguistic acts (or parts or sequences of linguistic acts). It is especially this technology-imposed *individuality* of *acting* and *reacting* that makes the organization of chat different from the organization of oral conversation. It follows that the challenges with which the individual has to cope while participating in a chat can best be described by adopting an individual-centered model of language production.

Figure 1 visualizes a *situated model of chat participation* that is based on the regulation model of speech production by Herrmann and Grabowski (1994). In this model, the individual chat participant is described as an information processing system that deals with its social (communicative) environment in order to make the latter comply with individually set standards. In order to fulfill its regulatory task (i.e., adapting to the specific requirements of the environment-as-is and producing output to transform the environment into a target state), the system is consecutively *evaluating input* about the as-is state of the environment (*processing mode*) and—if the as-is state is found to be suboptimal—*generating output* (*production mode*) to influence the environment-as-is.
The model is individual-centered, insofar as it does not describe *interaction* but rather the *activities of one individual* while being a participant in a communicative event; it is therefore a model for the description of chat *participation* (not of chat *communication*). The model is "situated" insofar as the communication-related activities of the individual

are described within their physical, temporal, medium-specific, and conversational contexts. The *physical context* is the real-life environment of chat participation: the computer workplace, a certain room with a certain interior, etc. The *medium-specific context* is the user interface and its functions, as well as the communication technology (chat) and its features. *Spatially*, the chat participant is situated individually, with no direct contact with his/her interlocutors; shared orientation therefore can only be established through textual information. *Temporally*, the chat participant is synchronously *co-situated* with his/her interlocutors, as they are occupied with exchanging messages at the same time.

Nevertheless, this co-situatedness is only *synchronous*; it lacks the possibility of processing utterances *simultaneous* with their production. The interlocutors are located in front of their computers during the same period of time (and, thus, interact *synchronously*), but at the same point in time they do not necessarily share the same knowledge about the current state of their conversation or the type of activity with which their interlocutor(s) actually is/are occupied (thus, they are mutually aware of each other's behavioral output *non-simultaneously*). For this reason, Garcia and Jacobs (1998) and Dürscheid (2005) characterize chat as a "quasi-synchronous" form. I retain the term "synchronous," but restrict the underlying concept of synchronicity to "doing something during the same period of time" and apply "simultaneity" as a separate concept and not as a part of the concept of synchronicity. This conceptual differentiation of "synchronicity" and "simultaneity" allows, in my view, for a more adequate distinction between chat and oral discourse than would be possible by means of a concept of synchronicity that includes simultaneity as a necessary feature (Beißwenger, 2007, pp. 35-37).

Owing to the visual nature of writing and the specific production/transmission procedure of chat (as described above), the processing and production tasks (typing in new messages and reading messages in the screen protocol) also cannot be fulfilled simultaneously. Instead, the chat user typically reads a message, updates his/her mental communication protocol, and evaluates whether the state of the conversation has changed in a way that makes it necessary to perform an action. If the parameter for the latter is set on "true," the action to be performed becomes defined, the language production apparatus becomes activated, and the transformation of the intended action into words begins. Because messages in chat can only be transmitted en bloc, the situation may arise that, after having started the production process and before completing a message, the user switches back into the processing mode in order to evaluate if the intended action is still up-to-date, i.e., if the current state of the screen protocol still complies with the last update of the users' mental communication protocol. If this is not the case, the verbalization of the intended action (i.e., the typing in of the message(s)) can be adapted to the current state (e.g., through text revision before transmission) or be given up and replaced by the verbalization of an alternative action that, according to the current state, seems to the user to be more efficient.
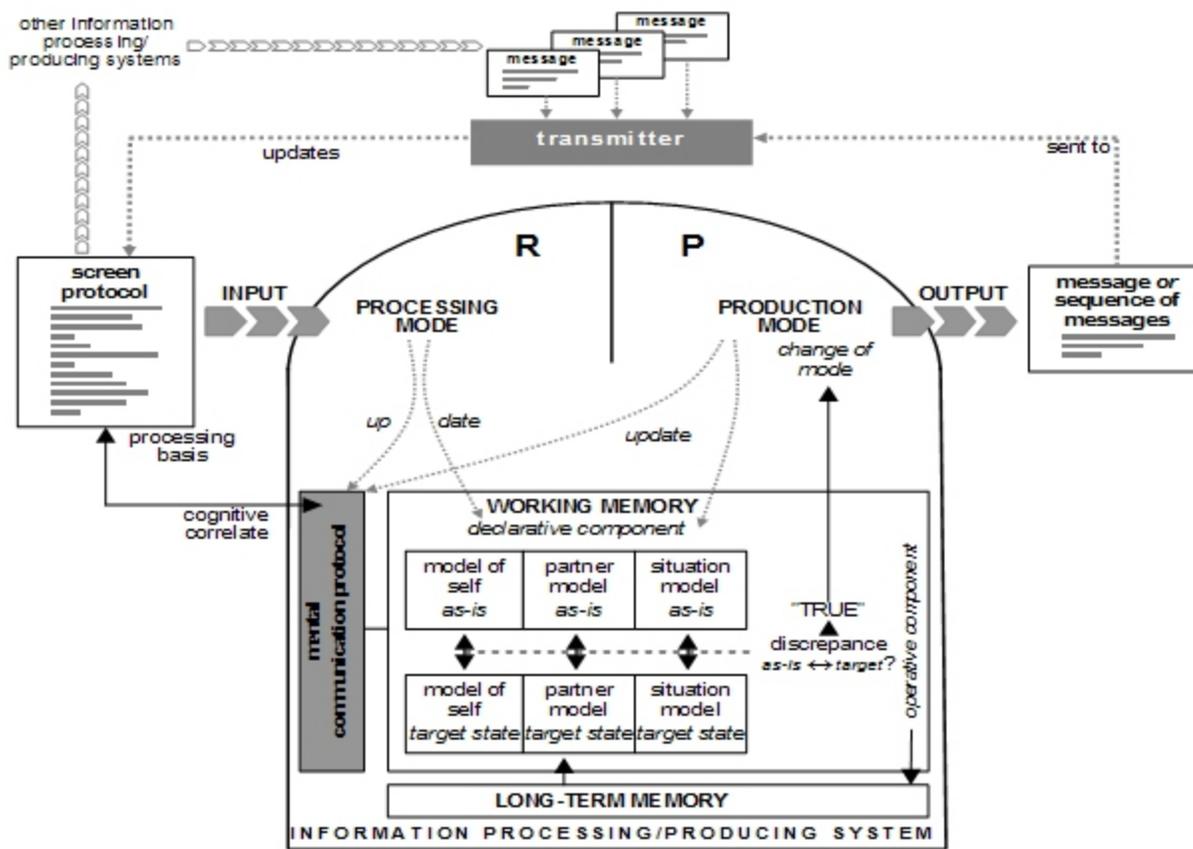
Figure 1. Situated model of chat participation

This model can help to describe the challenges of conversational participation that individual participants typically have to manage in order to coordinate efficiently their own message production with the perception and processing of their interlocutors' messages, which are displayed on their computer screen. While producing a message (entering and editing text in the message entry box) and therefore placing their main focus on executing and monitoring the text production process, participants run the risk that their mental communication protocol may lag behind the state of the screen protocol. Dedicating too much time and attention to monitoring the screen protocol runs the risk of prolonging the time needed for production; as a result, messages may be displayed in positions on the screen protocol that are at an unfavorable distance from the messages to which they directly refer. The case study section of this article presents a study on deletions of entered text, the findings of which suggest that many chatters try to adapt the intentions underlying their current message production fluidly to the current state of the screen protocol. Message production in chat is often not a continuous process in which visual attention is paid exclusively to the keyboard, the message entry box, and the text-produced-thus-far. Instead, interposed changes of gaze orientation toward the screen protocol in order to check the state of the conversational progression are not uncommon.

From Logfile Analysis towards a Situated Modeling of Chat Participation

To find out more about how chat users individually manage their participation, data are needed that go beyond logfiles and provide insight into what users do when chatting. As Marcoccia, Atifi, and Gauducheau (this issue) similarly point out, the situated dimension and the multimodal nature of the users' communication-related activities in CMC should be considered. Video data from user observations documenting individual behavior in a variety of modes—such as editing text on the screen, facial expression, gesturing, posture, and oral verbalization in front of the computer—may be a fruitful resource for this task.

"Situated Chat Analysis" on the Basis of Multimodal Chat Data

> Computer-Mediated Discourse Analysis (CMDA) ... applies methods from language-focused disciplines ... It may be supplemented by surveys, interviews, ethnographic observation, or other methods; it may involve qualitative and quantitative analysis; but what defines CMDA at its core is the analysis of logs of verbal interaction (characters, words, utterances, messages, exchanges, threads, archives etc.). (Herring, 2004, p. 339)

Until now, logfiles have constituted the main foundation for empirical studies of the distinctive linguistic and communicative features of chat-based communication (see also Androutsopoulos, this issue). Logfiles are the stored, static records of message sequences that have been put into their particular order by a server feature and that are displayed as a message protocol on the users' screens.

The variety of data within a logfile are constricted, since logfile data do not reveal

1. if, before or during the production of a message, a chat user has already acknowledged the previously displayed messages in the screen protocol;
2. how the addressees of a message immediately reacted—nonverbally or verbally—to a particular message read in the screen protocol, and to what extent their next message, which might be produced as a reaction to the received one, corresponds with this immediate reaction;
3. at which point in time a chat user decided to produce a message, how much time s/he needed for encoding, and whether or not s/he changed his/her original action plan during production (possibly more than once).

Currently, only a few studies systematically integrate into their analyses what happens on and in front of each chat user's screen and what is, accordingly, not documented in the message protocol (and thus also not in its record). An important milestone is the work of Garcia and Jacobs (1998, 1999), who were the first to employ screen-capturing methods for the analysis of chat discourse. *Screen capturing* means a videotaping of all activities that are visible on the individual user's computer screen.[3]

Screen movies are without a doubt an important resource for gaining insight into what users do on and in front of their screens. Another fruitful resource is the recording of the users' gaze orientation in the workspace, in order to collect information about what might be the earliest possible point in time at which a user could have recognized or processed new messages of his/her interlocutors. In Beißwenger (2007), I combined methods of both screen capturing and video observation of chat users in front of their computers.

In addition to screen capturing software, I used a video camera on a tripod and chose a chat system in which the message entry box was positioned above the display area for the message protocol (Figure 2). With this design, new entries in the protocol were displayed at the very bottom of the screen, while for monitoring one's own text input in the message entry box, the gaze focus had to be located at the top edge of the screen. The camera was set up at an angle that made it possible to differentiate the gaze orientation between these two visual targets (see the example freeze-frames in Figure 3).

In addition to gaze orientation, the video observation also allowed for the recording of data on facial expression, gesturing, posture, and verbal uttering in front of the computer.
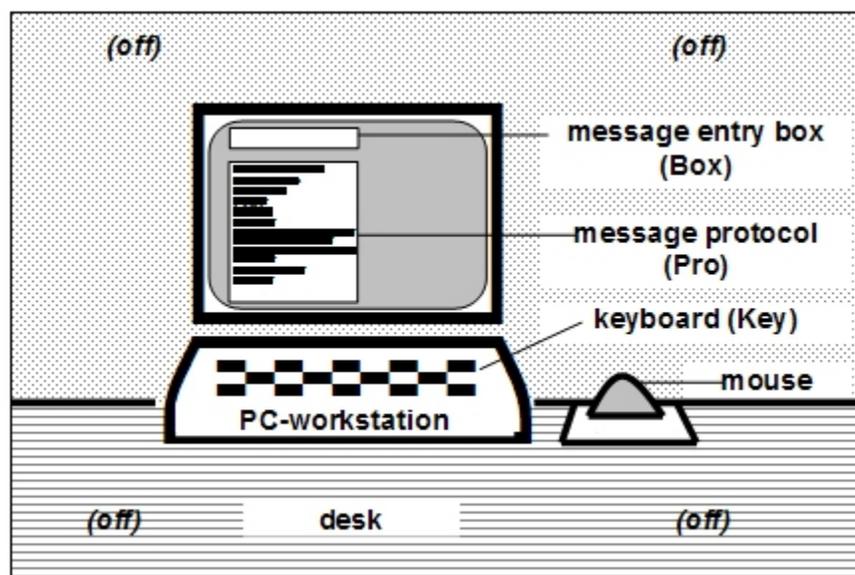

Figure 2. Visual targets relevant for chat participation

Figure 3. Examples of gaze orientation directed towards the visual targets *message entry box* (left), *message protocol* (center) and *keyboard* (right). The freeze-frames are taken from the head-on video recording of two test participants.

Since a "social chat" scenario seemed to be inadequate for observing chat under experimental conditions, I applied a scenario from the field of knowledge communication, namely a free online counseling session about "eBay and online auctions" with an expert on this topic. Overall, I conducted 18 chats of this type and recorded the screen and physical activities of a total of 32 test subjects.

The chats yielded 11 hours and 26 minutes of logfile records, 25 hours and 13 minutes of "screen movies," and 28 hours and 43 minutes of recordings with the video camera. Selected parts of these data were transcribed for the purpose of analysis. Since conversation-analytic systems for the transcription of speech data are not very appropriate for the transcription of data that convey writing processes and cases of subsequent revisions of entered text, I decided to create my own format for the transcription and visualization of the collected data. The format allows for a synoptic description of screen activities (especially activities related to message production), gaze orientation toward the screen, and the dynamic progression of the message protocol. In addition, it makes it possible to identify for each second of the respective chat episode:

- whether a new message appeared in the message protocol,
- whether, and if so, which production activities could be observed for the participant under observation, and
- which part of the workplace was the momentary target of the participant's gaze orientation.

A detailed description of the scenario and of the transcription format used is available in Beißwenger (2007, pp. 316-363; 2008a). An excerpt from one of the transcripts (with a short description of how it is constructed) is given below.

Case Study

Spoken language is permanent, whereas written is temporary. [...] Spoken language is permanent because once something is said, its impact cannot be erased, but something written can be crossed out, and it is as though it never was written. (Frank Smith, quoted in Tannen 1984, p. 29)

Text production in chat is a discontinuous process. Although aiming at creating textual units that will function as contributions within a form of exchange that develops dialogically and synchronously, the creation process contains revising and re-writing elements. Thus, it is not just a linear encoding activity that starts by entering the first character of the intended character string and ends by pressing the "send" button. As the screen capturing data described above show, deleting, substituting, and retrospectively adding textual items within the text produced up to that point (before being sent to the server) occur commonly. In my corpus, which consists of data documenting the chat participation of 17 test subjects, I identified 635 cases of revision in an overall total of 889 posted messages. Fifty-five percent of the revisions were carried out directly, i.e., without any delay after entering the revised expression/string, while 45% of the revisions were carried out retrospectively, i.e., after a certain delay and/or at a previous position of the text entered thus far (Beißwenger, 2008b).

Among the revisions carried out retrospectively, it is striking that *complete deletions* represent the most frequent type (73.2% of cases). In such cases, users do not just revise part of the text, rather they delete the complete text entered thus far, leaving the message entry box completely empty after the deletion process. As it increases the time and effort needed for message creation, complete deletion seems to be an uneconomic production strategy *prima facie*. Because of the synchronous setting and the "first come, first served" principle of the server-side sequencing of messages, quick production increases the chance that a message will still be relevant when posted. Therefore, the relatively high frequency of complete deletions requires explanation.

On the basis of transcripts from the data of my multimodal chat corpus, I conducted a case study in order to find out if complete deletions of entered text could be attributed to the fact that the producer, after having started editing text, took notice of a new message in the screen protocol that he/she had not acknowledged when initiating typing. My assumption was that if deletion occurs in such cases, it could be seen as a result of evaluating the action that is currently being verbalized against the recent update of the mental communication protocol. If such evaluation suggests that the intended action has become redundant or less relevant, an adaptation of the scheduled action according to the altered sequential conditions may follow. This may result in the deletion of the text already entered, and, subsequently, the initiation of the written realization of an alternative action plan.

I analyzed transcripts of 17 individual chat participation events with a total length of 11 hours and 24 minutes. The unit of analysis was the subjects' behavior within certain intervals of time, each starting and ending with the message entry box being completely

empty and in between being filled with text. I use the term *production sequences* for what was observable within these intervals.

I differentiate between two types of production sequences: The default type (type *A* in Figure 4) starts with the first character becoming visible in the message entry box, passes a state in which the entry box is maximally filled with text, and results in the user activating the "send" option of the user interface (by hitting the enter key or by clicking the "send" button with the mouse). The production process may contain pauses and revision. Cases in which the complete text entered into the message entry box is deleted are described as type *B*. In these cases, after the message entry box has been maximally filled with characters, instead of activating the "send" button, the user deletes all characters that have been entered so that the message entry box becomes completely cleared. Immediate new message entry may follow, but this is then analyzed as the beginning of a new production sequence.
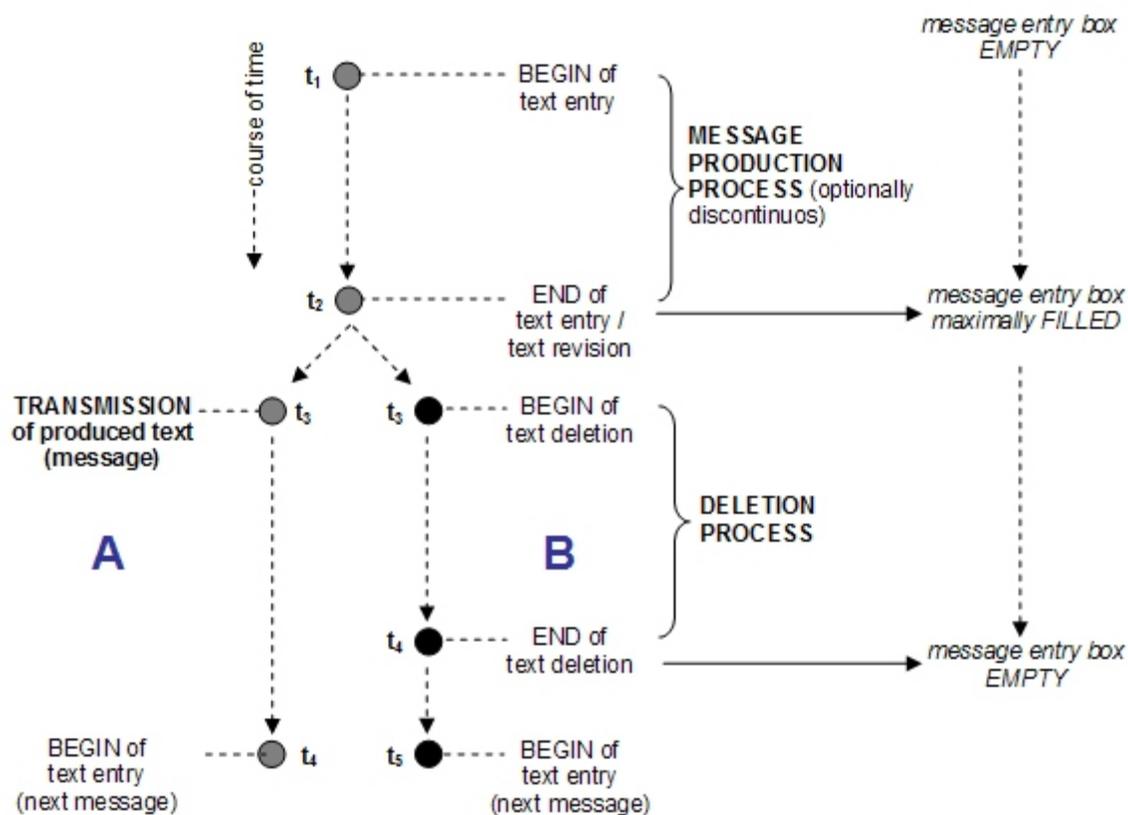

Figure 4. Types of production sequences

The study showed that 208 cases, or 18.96% of all production sequences (N=1,097), were finished with a deletion. To find out how often the decision to delete was made after noticing new messages in the screen protocol, I conducted a detailed analysis of a core corpus of four hours and 34 minutes of chat participation described in six transcripts. These transcripts included a complete transcription of the gaze behavior of the chat participants under observation. The ratio of production sequences finished with

deletion (N=80) to the total number of production sequences in the core corpus (N=427) was 18.74% and was thus similar to the overall corpus.

In 55% (N=44) of the cases of deletion (N=80), at least one new message had been displayed in the protocol between the beginning of text entry and the initiation of the deletion process, and the user under observation had his/her gaze focus for at least one second on the protocol area of the screen before switching from text editing to deleting. Seven cases could not be determined clearly on the basis of the video data alone.

Therefore, I performed a qualitative analysis of a subset of 38 cases of deletion in the core corpus to clarify the inconclusive cases. For the subset in question, the qualitative analysis determined that the percentage of complete deletions carried out as a result of noticing new messages is 71.05 %. This confirmed that in those production sequences in which deletion and perception of newly displayed messages co-occurred, there was a direct relationship between the decision to delete and an antecedent evaluation of the current action plan against the new state of conversational progression. An excerpt from one of the transcripts showing two cases of this kind is given below; it is designed as follows:

- The "TIME" column specifies points in time in an format. All events described in one of the other cells within the same row in the table are to be read as "happening at point in time X." All textual information given in one continuous cell, without a new row in between, represents either one punctual event which could be observed at the respective point in time or an event with a certain duration. For example: The cell in column 1 that is horizontally adjacent to the point in time 11:24:45 represents a punctual event (namely the appearance of a new message in the screen protocol of *jecom's* computer monitor), which could be observed at 11:24:45. The cell in column 3 that starts at 11:25:54 and finishes at 11:26:06 represents a continuous production activity by *jecom*, which could be observed on the screen between 11:25:54 and 11:26:06.

- Column 1 ("Screen protocol") has one or several entries for every point in time at which one or more new messages appeared in the screen protocol.

- Column 3 describes all message production that was observable in the screen movie. In case of deletions, the deleted text is repeated and crossed out. The expression "POST" indicates that the user performed a posting action, handing over the entered text to the server.

- Column 4 describes the main gaze focus of the user as observed in the video recording of the user's face. The abbreviation "box" stands for the message entry box, "pro" for the screen protocol, and "key" for the keyboard as visual target (see Figure 2). When the abbreviation is in standard size, it denotes the main target of gaze; when in small size, it describes short gaze deviations, after which the gaze orientation returns to the main target given within the same cell.

- Column 5 offers space for occasional comments on phenomena referring to other behavioral modes such as gesture or facial expression. Since these modes were

not systematically evaluated within my study, these descriptions were formulated intuitively and are in a non-standardized format.

Screen protocol	TIME	Message entry box: visible message production (*jecom*)	Gaze (*jecom*)	Other behavioral modes (*jecom*)
jecom: Genau das. Wie hätte ich mich denn am besten verhalten, wenn die Frau sich geweigert hätte zu bezahlen? 'That's it. What would have been the best way to react if the woman had refused to pay?'	11:24:45	[…]		
[…]				
bsommer: Normalerweise soll man immer versuchen 'You should normally always try'	11:25:39		Pro	
bsommer: mit dem Käufer zu reden 'to speak with the buyer'	11:25:43			
bsommer: um alles zu klären 'in order to clear up everything'	11:25:46			
	11:25:51			
	11:25:52		Key 1 Pro 2 Box	
	11:25:54	Aber auch wenn das Reden nichts gebracht 'But even if talking doesn't'		
bsommer: Sprich mehrere Mails schreiben und zur Bezahlung auffordern 'That means writing several e-mails and demanding that the person pay'	11:25:56			
bsommer: Falls er darauf jedoch nicht reagiert 'However, if he still does not react'	11:26:04			
	11:26:05			
	11:26:06		Pro 1 Box	
	11:26:09	~~Aber auch wenn das Reden nichts gebracht~~		
	11:26:10		Box	
	11:26:12	~~'But even if talking doesn't~~		
	11:26:13		Pro	*rests chin on one hand*
bsommer: können Sie bei eBay einen nicht bezahlten Artikel melden 'you can notify eBay of an unpaid article'	11:26:17			
bsommer: Der Käufer erhält daraufhin eine Email	11:26:24			

Screen protocol	TIME	Message entry box: visible message production (*jecom*)	Gaze (*jecom*)	Other behavioral modes (*jecom*)
'The buyer will then receive an e-mail'				
bsommer: in der er aufgefordert wird, die Zahlung vorzunehmen. 'in which he is told to pay the amount.'	11:26:32			
	11:26:36			
	11:26:37		Key	
	11:26:38	Soweit hatte ich das schon mal gemacht aber auch daraufhin 'I had already done everything up to that point but then'	1 Pro 1 Box 1 Pro	
bsommer: Falls er auch darauf nicht reagiert 'In case he still does not react'	11:26:43			
bsommer: können Sie vom Kaufvertrag zurücktreten 'you may rescind the contract'	11:26:52			
	11:26:53		Pro	
	11:26:54			
	11:26:55		Key	
	11:26:56			
	11:26:57		Box	
	11:26:59	~~Soweit hatte ich das schon mal gemacht aber auch daraufhin~~ '~~I had already done everything up to that point but then~~'	1 Pro	*rests chin on one hand*
bsommer: und die eBay-Gebühren zurückerstattet bekommen. 'and be reimbursed for the eBay fees.'	11:27:02			
	11:27:03		Pro	

Table 1. Transcript excerpt

At 11:26:09 and 11:26:59, the transcript shows two cases of complete deletion, each caused by the perception of new messages and an obvious misinterpretation of the interlocutor's (*bsommer's*) previous message sequence as a completed move. "Normalerweise soll man immer versuchen / mit dem Käufer zu reden / um alles zu klären" (messages 11:25:39, 11:25:43, 11:25:46) represents a series of utterance chunks which, taken together, form a syntactically complete unit and, semantically, a statement that follows coherently from the prior context. *Jecom* does not redirect her gaze focus from the screen protocol while these three messages appear on her screen, thus she must have taken notice of all three parts of this "splitting" sequence.[4] Since at 11:25:46 the sequence is obviously completed, she—according to the current state of her mental protocol—can start with a follow-up contribution without being un-cooperative or incoherent. Eight seconds after *bsommer's* last message was displayed,

she starts with the production of a message which encodes a response to the preceding sequence (11:25:54–11:26:06: "Aber auch wenn das Reden nichts gebracht").

At 11:26:06, *jecom* again turns her gaze focus to the screen protocol. In the meantime, two new messages from *bsommer* have been displayed which, due to the completeness of *bsommer's* prior contribution, could not have been anticipated before. These two messages seem to anticipate what *jecom*, with her own new message, intends to encode. The evaluation of the target state of conversational progression, which *jecom* intends to effect with her message under construction, against the as-is state given on the screen shows that in the as-is state, compared to the state at 11:25:51, the new topic "what to do when the buyer does not react?" has already been established, albeit by the interlocutor. Instead of finishing production and posting her intended contribution anyway, *jecom* deletes the text she has entered thus far and discards the respective contribution project from her mental agenda.

jecom's attempt to make a new contribution was not subject to interpersonal negotiation. On the one hand, due to the communicative conditions, *bsommer*, while producing, was not able to notice that *jecom* had started encoding a similar message; on the other hand, *jecom*, while producing her message, was not able to anticipate the same about *bsommer's* current activity. Therefore it would not be adequate to write off this type B production sequence as a case of failed turn-taking. The production and the deletion of *jecom's* text fragment between 11:25:54 and 11:26:13 is *not a joint activity* of both interlocutors. Instead, it is an individual attempt by jecom to make a contribution, which, at the time she starts producing and according to her individual assumption about the current state of conversational progression at this time, seems (to her) a good candidate for a coherent next move in conversation. When she notices *bsommer's* message displayed at 11:26:04, which encodes a similar move, her own plan becomes obsolete and therefore is given up.

From 11:26:17 until 11:26:36, *jecom* then tracks the further development of the screen protocol. From 11:26:38 on, she makes a next attempt to contribute and, again, begins producing a new message. Similar to the case before, *jecom* starts producing at a point at which *bsommer's* contribution—again distributed over a series of messages—can be interpreted as being complete. Again in this case, while *jecom* is encoding her plan, *bsommer* posts a follow-up message to her prior one that makes *jecom's* contribution obsolete. After taking notice of this (11:26:53), *jecom* again halts her encoding process and deletes the text entered thus far (11:26:59).[5]

Conclusion and Further Perspectives

The results of this study of deletions and the sample analysis of the transcript illustrate what multimodal data can reveal about how users individually organize their conversation-related activities when participating in a chat (processing textual input; producing textual output; evaluating the individual agenda against the assumed current state of conversational progression). They also demonstrate that how chat users plan

and realize their conversation-related activities differs considerably from oral conversation.

The study shows that entered text is not deleted due to interpersonal real-time negotiation, but rather as an effect of individual adjustment to relevance conditions, which—as is typical for *conversation*—change quickly, but whose progression—as is especially typical for *chat*—cannot be tracked continuously. Therefore, deleting is an indication of an individual restructuring of intended actions in cases in which a recent update of the individual mental communication protocol makes this seem essential to the participant.

The turn-taking paradigm, which was originally developed through the analysis of talk in interaction and which is described as the fundamental device for local coordination among interlocutors (Sacks, Schegloff, & Jefferson, 1974; Schegloff, 2007), is founded on communicative conditions that allow for mutual real-time perception. Since in written discourse the concurrence between the availability and the perception of utterances is broken up, turn organization should be regarded as a structuring device specific to *oral* conversation.

In contrast, *chat organization* is to a great extent based on the users' individual adaptation toward the screen. The perception of the temporal structure of the course of interaction differs among interlocutors: It depends on when users individually decide to gaze at the screen in order to acknowledge and process new messages and when they redirect their attention away from the screen in order to produce their own messages. When producing, as long as the users do not gaze at the screen again, the production process is based on a state of their mental communication protocol, which may lag behind the actual state of progression given on the screen.

With this in mind, I suggest reconsidering those categories used in CMDA which are custom-tailored for oral conversation but which conceal the distinct conditions that writing imposes on synchronous online conversation. Because of the lack of conditions that are fundamental for negotiating efficiently about when certain interlocutors finish producing their contribution to dialogue and which of their interlocutors subsequently becomes next in line to contribute, chat does not allow for the interactive construction of a linear sequence of the interlocutors' productive activities, each of them being planned and carried out in complete awareness of all prior contributions.
Instead, chat users can only gear their individual activities to their individual perception of what is happening on the screen and, thus, try to be coherent only to what *they* assume to be the current state of conversation (which might not correspond with what their interlocutors, at the same time, assume to be the current state). Thus, while the concept of the *turn* is not completely absent in chat, it is not interactional in nature. In Murray's (1989, p. 324)[6] words, it is a *psychological unit* that comprises what a certain user intends to contribute at a certain point in time and in view of a certain (individually-assumed) state of conversation. The interlocutors do not cooperate on a turn-taking level, since "any linear model for the organization of conversation [...] does not

adequately account for the structure of this mode of interaction" (Murray, 1989, p. 331). Instead, cooperation is organized exclusively on the action level:

> participants send utterances in the belief that they will contribute to the resolution of the task. To achieve this, they do not need to negotiate for turns but rather collaborate to get things done. (Murray, 1989, p. 319)

The individual-centered model presented in this article makes clear that when analyzing how users individually manage to be cooperative on the action level, researchers have their eyes not on the interactional dimension of chat conversation but on the individual user's perspective and, thus, on the question of how a given user manages the different subtasks he or she has to fulfill in order to participate actively in an ongoing chat conversation (i.e., constructing messages, waiting for messages, reading new messages; see Garcia & Jacobs, 1999). To investigate this individual dimension of chat conversation, capturing the users' onscreen activities (message encoding and editing, mouse and scrolling activity) and recording gaze behavior both provide valuable resources for analysis. Further studies and experiments using this "window to the user's perspective" may yield deeper insights into the fundamental differences between chat and oral conversation and thus provide designers of chat systems with a basis for the development of innovative features that support coordination among users.

Notes

1. German original: "Im (mentalen) Kommunikationsprotokoll hält der Sprecher/Kommunizierende den bisherigen Verlauf der sprachlichen Kommunikation fest. Er schreibt diese sequentiellen Informationen wie in einem Tagebuch fort" (Herrmann & Grabowski, 1994, p. 33).

2. Chat clients that provide audible signals—e.g., a simple beep—when new messages are displayed may mitigate this phenomenon to a certain degree. They may not compensate it completely, however, as perception and processing of new messages will still remain an activity which—due to the visual nature of the posted messages and to the message exchange procedure—is carried out retrospectively and not simultaneously with the production processes. Nevertheless, the divergence between the moment of visibility of a new message and the moment when an addressee starts processing may surely be decreased by automatically rendering a beep every time a new message is added to the screen protocol. Chat software that provides audible signals will not be addressed further in this article, although a comparative study of how users behave differently when using "standard" chat software and chat software of that kind could be enlightening for exploring further perspectives in the design of chat systems for e-learning and professional applications. As this is not the focus here, this article discusses only communication by means of "standard" chat systems.

3. For the use of screen capturing methods in chat analysis, see also Jones (2001) and Markman (2006). Furthermore, Vronay, Smith, and Drucker (1999) and Ogura and Nishimoto (2004) work with "typing histories" or with programs that automatically log any interface manipulation that is carried out by chat users through keyboard or mouse use.

4. "Splitting" sequences—which result from applying a production strategy according to which one dialogue contribution is not encoded in one message, but rather is distributed over a series of messages—can not be discussed here in detail; for details and example analyses,

see, e.g., Zitzen and Stein (2005, pp. 1004-1015) and Beißwenger (2007, pp. 245-253, 261-264).

5. The complete transcript is available at http://www.michael-beisswenger.de/sprachhandlungskoordination/. The study on deletions and its findings are described in detail in Beißwenger (2007, pp. 367-465).

6. By today's standards, Murray (1989) appears to have investigated a "primitive" messaging system that facilitated exchange between only two interlocutors and in which the exchanged messages were not archived in a screen protocol, but rather were "clicked away" after being read. However, Murray's position regarding turn-taking in computer conversation can still be applied to the communication in today's "standard" chat systems. The fundamental communicative conditions, as well as the lack of mutual real-time perception between interlocutors, are the same in the system investigated by Murray and the chat systems covered in this article.

References

Beißwenger, M. (2007). *Sprachhandlungskoordination in der Chat-Kommunikation*. Berlin, New York: de Gruyter (Linguistik—Impulse & Tendenzen 26).

Beißwenger, M. (2008a, in press). Multimodale Analyse von Chat-Kommunikation. In K. Birkner & A. Stukenbrock (Eds.), *Arbeit mit Transkripten in der Praxis: Forschung, Lehre und Fortbildung*. Verlag für Gesprächsforschung (to be published online at http://www.verlag-gespraechsforschung.de).

Beißwenger, M. (2008b, in press). Empirische Untersuchungen zur Produktion von Chat-Beiträgen. In T. Sutter & A. Mehler (Eds.), *Medienwandel als Wandel von Interaktionsformen—von frühen Medienkulturen zum Web 2.0*. Wiesbaden: VS Verlag für Sozialwissenschaften.

Dürscheid, C. (2005). Medien, Kommunikationsformen, kommunikative Gattungen. *Linguistik online*, *22*(1). Retrieved July 12, 2008 from http://www.linguistik-online.de/22_05/duerscheid.pdf

Garcia, A. C., & Jacobs, J. B. (1998). The interactional organization of computer mediated communication in the college classroom. *Qualitative Sociology*, *21*(3), 299-317.

Garcia, A. C., & Jacobs, J. B. (1999). The eyes of the beholder: Understanding the turn-taking system in quasi-synchronous computer-mediated communication. *Research on Language and Social Interaction*, *32*(4), 337-367.

Herring, S. C. (1999). Interactional coherence in CMC. *Journal of Computer-Mediated Communication*, *4*(4). Retrieved July 12, 2008 from http://jcmc.indiana.edu/vol4/issue4/herring.html

Herring, S. C. (2004). Computer-mediated discourse analysis: An approach to researching online communities. In S. A. Barab, R. Kling, & J. H. Gray (Eds.), *Designing for virtual communities in the service of learning* (pp. 338-376). Cambridge, New York: Cambridge University Press.

Herrmann, T., & Grabowski, J. (1994). *Sprechen. Psychologie der Sprachproduktion*. Heidelberg, Berlin, Oxford: Springer.

Jones, R. (2001, November-December). *Beyond the screen. A participatory study of computer mediated communication among Hong Kong youth*. Paper presented at the Annual Meeting of the American Anthropological Association, Washington D.C.

Retrieved   July   12,   2008   from
http://personal.cityu.edu.hk/~enrodney/Research/ICQPaper.doc

Markman, K. (2006). *Computer-mediated conversation: The organization of talk in chat-based virtual team meetings.* Ph.D. dissertation, University Texas at Austin.

Murray, D. E. (1989). When the medium determines turns: Turn-taking in computer conversation. In H. Coleman (Ed.), *Working with language. A multidisciplinary consideration of language use in work contexts* [Contributions to the Sociology of Languages 52] (pp. 319-337). Berlin, New York: Mouton de Gruyter .

Ogura, K., & Nishimoto, K. (2004, August). *Is a face-to-face conversation model applicable to chat conversations?* Paper presented at the Eighth Pacific Rim International Conference on Artificial Intelligence (PRICAI2004), Auckland University of   Technology.   Retrieved   July   12,   2008   from
http://ultimavi.arc.net.my/banana/Workshop/PRICAI2004/Final/ogura.pdf

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language, 50*(4), 696-735.

Schegloff, E. A. (2007). *Sequence organization in interaction. A primer in conversation analysis I.* Cambridge: Cambridge University Press.

Tannen, D. (1984). Spoken and written narratives in English and Greek. In D. Tannen (Ed.), *Coherence in spoken and written discourse* [Advances in Discourse Processes 12]] (pp. 12-41). Norwood: Ablex Publishing .

Vronay, D., Smith, M., & Drucker, S. (1999). Alternative interfaces for chat. *Proceedings of the 12th Annual ACM Symposium on User Interface Software and Technology* (CHI Letters 1,1), 19-26.

Zitzen, M., & Stein, D. (2004). Chat and conversation: A case of transmedial stability? *Linguistics*, *42*(5), 983-1021.

Biographical Note

Michael Beißwenger [michael.beisswenger@uni-dortmund.de] is a postdoctoral research fellow and lecturer in German linguistics in the Department of Cultural Sciences at TU Dortmund University. His research interests include computer-mediated communication, lexicology, orthography, and corpus linguistics. More information on his research interests and recent publications can be found at http://www.michael-beisswenger.de.