CHALLENGES WITH GOOGLE GLASS IN SOCIAL INTERACTION

BRIAN DUE
DEPARTMENT OF NORDIC
STUDIES AND LINGUISTICS.
UNIVERSITY OF COPENHAGEN.
BDUE@HUM.KU.DK

ABSTRACT

This paper discusses social implications of Google Glass use in multi-party and multi-activity participation frameworks in different Danish settings. As opposed to the often-employed HCI approach to computer and smart glass use, this paper addresses Glass through Conversation Analysis (CA). On the basis of quasi-experimental breaching experiments, the paper shows three interrelated social implications in interaction that may cause problems on the turn-taking level: 1) difficulties determining what the next relevant action is, 2) overlap, repair and the problem of intended, simultaneous talk and 3) participants' lack of understanding and the production of epistemic discrepancies. The paper ends by concluding that, if smart glasses continue to be controlled by verbal and gestural input in public, new communicational practices may need to be employed by participants. This has consequences for the future design.

INTRODUCTION

There are many different types of smart glasses on the market or in the pipeline today. To mention a few: Optinvent, Epson Moverio, Atheer One, Recon Jet, Vuzix, Meta spaceglasses, Weon Glasses, Chipsip, Laforge Optical, Epiphany Eyewear, Kopin, Sony smart glass, Glass Up, Lumus DK40, Lenovo smart glass, Xoeye Technologies, Innovega iOptik and 4iii SportIIII. Microsoft and Apple have also patented smart glass technology. However, Google Glass (Glass, in short) has taken the market lead and brought about the hype. The prototype version 1 is not on sale to consumers but used in enterprise. A new version of Glass will probably be on the market during 2015. Google Glass is a minicomputer attached to a frame, which is worn like ordinary spectacles. They project a screen in front of the user with information from Internet searches, calendar, time, weather, maps, news or whatever app is installed on the computer. In addition, Glass can take pictures and record video. In sum, it is like a smart-phone worn on the head: a wearable heads-up display computer. It is managed through talk and the touchpad on the side.

Figure 1. Google Glass, explorer edition.

This paper does not provide a test of the technology. Neither is it a reflection on whether the apps, camera or interface is working well or whether privacy issues are at stake, etc. Nor is it theoretically framed within Human Computer Interaction (HCI), User Experience or Interaction Design disciplines. Instead, Conversation Analysis (CA) is applied to explore the social organisation of talk and bodily conduct when using Glass.

Using wearables such as Glass when alone is different from using it in a social setting. Gesturing and talking to Glass when alone requires an intuitive and comprehensive interaction design (Rogers, Sharp & Preece 2011). Using Glass in a social setting requires - in addition to the useful UX - a coordination of relevant actions at relevant times that is meaningful to co-participants on a turn-by-turn level. This has been observed from almost the beginning of the development of wearables (Starner 1996; Starner et al. 1999). However, no systematic empirical studies of this phenomenon have been conducted.

Many of the *social* implications of Glass derive from its use of voice and gesture control and, hence, visible and audible actions in social interaction. This implies that there is no common or shared practice for how to behave in response to this kind of Glass use. This is not a surprise, since the technology is very new to people and behavioural systems and social scripts are often lacking behind the launch of new products (Schank & Abelson 1977; Norman 1983). This is, of course, also a question of the user's knowledge and practical experience with the technology as well as the co-participant's use of and practical experience with interaction with Glass-users. That is, the implications of using Glass in social interaction are relative to the

participant's epistemic stance with respect to the technology. At the moment, Glass is only shipped with English as the default language and with only some of its functionality working. Use in a Danish setting is, thus, restricted, and the users – at least, in my data – are novices with very limited experience with the technology.

Becoming socialised into the use of the technology and behaving accordingly takes time and demands, among other things, a broad distribution of the technology throughout society. Use among skilled experts and co-participants with symmetric epistemic stances will, naturally, be a different kind of situation than use among novices. However, at the moment, almost everyone using Glass and engaging in Glass-use interaction are novices.

METHOD AND THEORETICAL APPROACH

This paper takes as its departure a broad phenomenological and holistic sense of the human being-in-the-world (Merleau-Ponty 2002), which takes into consideration the uniqueness of each and every situation as a specific constitution of participants who use different modalities in semiotic ecologies. As has been stated within the Conversation Analytical (CA) and Ethnomethodological (EM) community, participants in talk-in-interaction are oriented towards previous turns and actions as relevant contexts for their contributions, when talking (or acting) participants project relevant next speakers and types of actions (Sacks, Schegloff & Jefferson 1974). For instance, when a speaker asks, "What time is it?", he projects the next relevant action to be an answer, such as: "It is 2 a.m.". But, in a Glass context, the big issue is: who or what is projected to do what in the next relevant position? This is particular an interesting question since Glass itself may also be defined as a non-human participant in the interaction – which will be shown in the analysis. Therefore, the use of smart glasses needs to be incorporated in one way or another into the "normal" organisation of turn-taking. This will be discussed throughout the paper.

So far, research on technology use has primarily been done from a HCI standpoint (Card et al., 1980). This research tradition has contributed many valuable insights. However, the approach naturally takes the computer as a starting point; and, when the human is incorporated, it is often done from a psychological or cognitive point of departure. HCI is, of course, a broad research tradition that encompasses different perspectives. It is often regarded as the intersection of computer science, behavioural sciences, and design. HCI involves the study, planning, and design of the interaction between people and technology (Jacko 2012). Attention to human-machine interaction is important because poorly designed human-machine interfaces can lead to many unexpected problems. But a more detailed method is also needed to understand how computers are used in social interaction (Hornecker & Buur 2006). Therefore, this paper will argue that a multimodal conversation analysis is needed to expand the range of HCI.

Multimodal CA is concerned with how people build actions by combining resources with diverse properties (e.g., modality of talking, gazing, touching, using gestures, and artefacts), which expands the repertoire of possible action available to participants (Streeck, Goodwin & LeBaron 2011). And CA-oriented work has provided a better understanding of the whole user situation in which the technology is embedded (e.g. Heath & Luff 1992; Heath, Luff & Sellen 1995; Hindmarsh et al. 1998; Fraser et al. 2000; Luff et al. 2003).

Every kind of multimodal action (e.g., talking, gesturing, moving around, gazing, and handling artefacts) is performed in time and space, thus occupying a slot in interaction at particular sequential moments. Thus, this paper builds on the sequential *turn-taking-system* (Sacks, Schegloff & Jefferson 1974) as a starting point. The sequential organization includes not only speech actions but also e.g. gesture-actions (Kendon 2005). And actions are embedded within a semiotic ecology of signs with relevant meanings relative to the amount of attention and orientation that participants give the action or artefact. Actions are also considered to be recipient designed. Hence, the meaning of an utterance is co-constructed by the response (Goodwin 1979). This is a fundamentally intersubjective architecture of interaction. Basically, everything about this turn-taking system and its description of how participants orient towards interaction as "normal" or "un-normal" is challenged when one is using Glass – at the moment. Thus, a deeper understanding of how Glass is actually used in conversation may help future product design.

DATA AND CONTEXT

It is difficult to collect naturally-occurring interactions with people wearing smart glasses, since very few people actually wear them at the moment. As an alternative, I have used quasi-experimental ethnomethodological methods with colleagues and students in order to construct situations that simulate naturally-occurring interactions and situations. They are quasi-experimental in the sense that they do not follow standard experimental design regulations with efforts to control causalities, possible sources of clinical bias, contamination, spurious inferences, etc., but simply invite participants to be at specific locations at specific times. And they are *breaching* (Garfinkel 1967) experiments in the sense that the very use of Glass during conversation seems to challenge the moral and social order, thus making the shared methods that participants use to construct the meaningful orderliness of social situations accountable and visible in interaction. I have collected data in three different settings in Denmark, and the participants are speaking Danish. However, I have translated the data transcriptions into English.

Case 1 is an experiment with students wearing Glass while interacting and talking about schoolwork and routine matters. This is a mundane, everyday type of interaction between four students who are specifically talking about an exhibition they went to see.

Case 2 is also with students. This was set up in an experimental lab with four cameras running from different angles. Again, the students were not told to do anything

specific but to talk about whatever they felt like - in this case, the completed study tour of one of the students. Use of Glass in these student-situations was embedded in the interaction as a resource for conducting searches on the Internet, note-taking, taking pictures, etc.

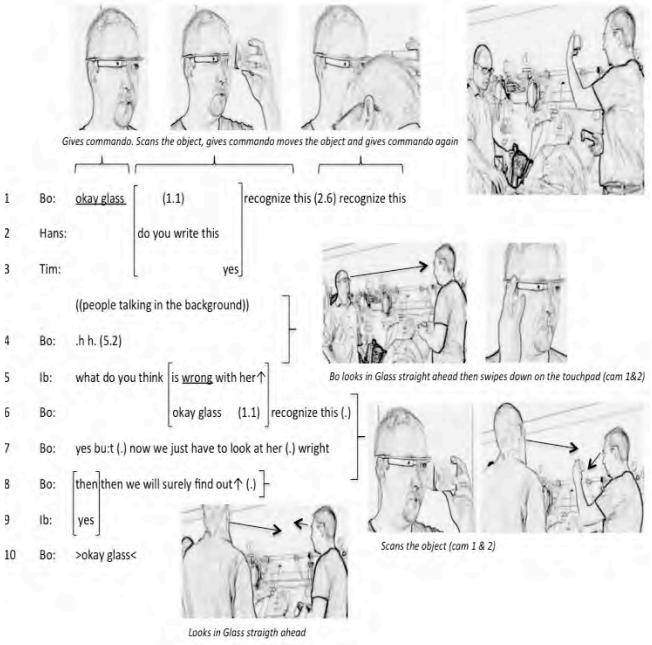
Case 3 is from an institutional setting in which participants (paramedics) use Glass during a simulated emergency at a hospital lab. This event was organised by a large Danish company in order to run a test case with possible implementation perspectives. In this situation, there are also four participants of whom two are paramedics and one of them is wearing Glass. Then, there is the patient, represented by a female doll, and another participant who acts as the patient's husband. The Glass-wearing paramedic uses a prototype medicine scan app - developed by two master's-level students (Kenny & Mathias) - as he enters the scene in order to determine what kind of medicine the patient has taken and what the implications might be. In this setting, there are several procedures and, thus, types of turn design that the paramedics are supposed to follow, e.g., examining and stabilising the patient (Nielsen et al. 2012). Thus, this kind of situation is somewhat more institutionalised with some kind of "turn-type preallocation" (Atkinson & Drew 1979) in which the activities of asking and answering (or responding to) questions are pre-allocated to the roles of being a paramedic (professional) versus being a patient (non-professional). In this situation, Glass is still a resource but in a different manner: the aim is to use a particular app for a particular purpose. However, the aim of this paper is not to point to differences between the data corpuses but to identify the overall social implications that cross the activity types.

RESULTS: THE SOCIAL IMPLICATIONS OF USING GOOGLE GLASS

Going through the video recordings over and over again, I have identified several different social implications of using Glass. Each type of social implication is comprehensive, and there are many examples in my data corpus of different situations of every type. However, the examples provided in this paper are based on single-case analyses (Schegloff 1988). This paper deals with *progressivity* (difficulties determining what the next relevant action is), *multiactivity* (overlap and repair in interaction and the problem of intended simultaneously talk), and *epistemics* (participants' lack of understanding and epistemic incongruities). I will go through the examples one by one, providing a descriptive analysis of the social interaction. In the discussion section, I shall discuss design implications.

PROGRESSIVITY: DIFFICULTIES DETERMINING WHAT THE NEXT RELEVANT ACTION IS

Why that now? Why do people do what they do at the exact time they do it? If participants depart from "normal actions" in any particular given situation, they are typically accounted for in various ways (Schegloff & Sacks 1973). The following example is from the institutional setting (case 3) in which the paramedic uses Glass to scan a medicine box. In this situation, the two paramedics are examining and stabilising the patient while, at the same time, reassuring the husband about the situation. There are, of course, many things going on in the situation, but we shall focus on lines 5 and 6.



Transcript (ex.1): Silence is treated as a transition relevant place. Overlap occurs. Transcript is from case 3

In lines 1-4, Bo gives a command to Glass and then waits silently while engaging in several embodied actions: gazing at the Glass optics/screen and swiping the touchpad. Thus, the silence in line 4 is not a pause in the interaction but filled with embodied and virtual actions. However, the silence is long and beyond the normal one-second standard maximum of silence before it is treated as an interactional problem – as registered in mundane interactions (Jefferson 1983). Hence, the co-participant (the husband (Ib)) treats the verbal silence as a transition and takes a turn in line 5. The husband does not treat the use of Glass as a legitimate turn in the sequential organisation of turn-takings but as a slot for something else to happen. Thus, he self-selects, which causes an overlap.

Normally, utterances create an interpretive environment that will be used by participants to analyse the interaction (Schegloff 1968). But this implies that participants know what is going on and show an understanding of the situation. When a user is wearing Glass (at the moment), co-participants do not seem to know how interaction with Glass fits into the turn-taking system, as shown in the example; and, consequently, they do not know what a relevant next action is on a simple turn-by-turn basis, e.g., how to treat silence and pauses. Thus, a lot of repair and overlap is also produced, which indicates interactional problems. This will be investigated in more detail below.

MULTIACTIVITY: OVERLAP AND REPAIR IN INTERACTION AND THE PROBLEM OF INTENDED SIMULTANEOUS TALK

The orderly distribution of opportunities to participate in social interaction is one of the most fundamental preconditions for viable social organisation. The most common default of speakership in interaction is the rule: one party at a time speaks (Sacks, Schegloff & Jefferson 1974). However, this is also one of the things that is challenged in conversation with a Glass-user. Simultaneous talk comes into being in different ways, e.g., by simultaneous beginnings or by treatment of silence as a transition relevant place. Participants are normally oriented towards an organised set of practices through which overlap is managed. But both the production and management of overlap are different when interacting with a Glass-user, because the interaction is not between a speaker and his co-speakers but among the speaker, the Glass, and the co-speakers.

Thus, *participation* is a key concept that needs to be taken into consideration by participants, because the Glass seems to function as a participant in interaction - not only through an orientation to the Glass as a thing or "dead" artefact but as a comprehensive participant with status and the possibility of affecting interactions. Thus, Glass can be seen as a non-human actant, meaning it can act in the world (Latour 1994). In this way, Glass is a kind of artificial intelligence, which can provide the user with feedback. Thus, Glass is a participant that can also produce actions that overlap with whatever is going on.

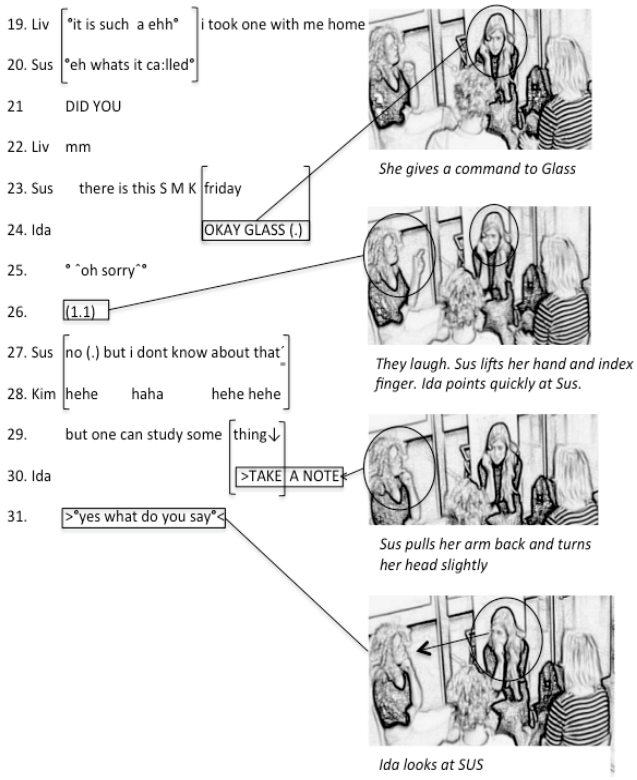
Now, take a look at transcript 1 again: Human overlap is produced in lines 5 and 6 in what seems to be a transition relevant place, but the speaker (Bo) does not treat it this way, and he continues his turn. Human overlap is, thus, a consequence of the co-participants' lack of orientation to the "correct" transition relevant place. Throughout the excerpt, Bo is performing a multiunit turn, which has no publicly displayed projections of next (human) speaker or possible completion points. The primary activity and projection of the next relevant thing are not towards the co-participants but towards the nonhuman participant (Glass). Thus, overlap is actually produced by Ib when he takes the turn in line 5, because he talks while the machine is giving feedback, thus stressing the Glass-user's cognitive system (Oulasvirta et al. 2005). This can consequently be framed as a *non-human overlap*. Non-human overlaps make the following question regarding participation status highly relevant: *who produces the overlap and consequently should withdraw and do repair?* I suggest that non-human overlaps can be produced in three ways:

1) A human co-participant acts in overlap with the (unrecognized) feedback from Glass to the user (ex. 1).

2) A human participant gives Glass a command in overlap with on-going talk as "non-human side sequences" (ex. 2).

3) A non-human participant provides information that affects social interaction because the Glass-user (suddenly) takes a turn with Glass-provided information (ex. 3).

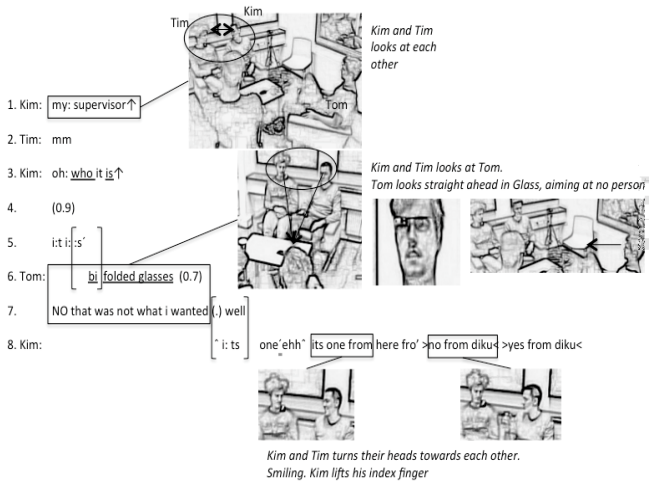
Having analysed ex. 1, let us now turn to ex. 2 and 3.



Transcript (ex. 2): A human participant gives Glass a command in overlap with ongoing talk. The transcript is from case 1.

In example 2, we see commands given to Glass as responses to ongoing "non-human side sequences" (the user's interaction with a running app) but in overlap with social interaction, e.g., line 24, which is remedied with a smiling repair (line 25). Sus, one of the students, projects herself as next speaker by lifting her index finger, and Ida (the Glass-user) responds by pointing quickly and indicating the same. Then, in line 30, another command to Glass is given as a response to a "non-human side sequence" - still in overlap with social interaction but quickly produced. Sus responds with embodied withdrawal, and Ida responds quickly with a projection of Sus as the relevant speaker. Thus, it seems that a human participant gives Glass a command in overlap with on-going talk.

In example 3, a non-human participant (Glass) provides information that affects social interaction because the Glass-user (suddenly) takes a turn with the Glass-provided information.



Transcript (ex. 3): A non-human participant (Glass) provides information that affects social interaction because the Glass-user (suddenly) takes a turn with the Glass-provided information. The transcript is from case 2.

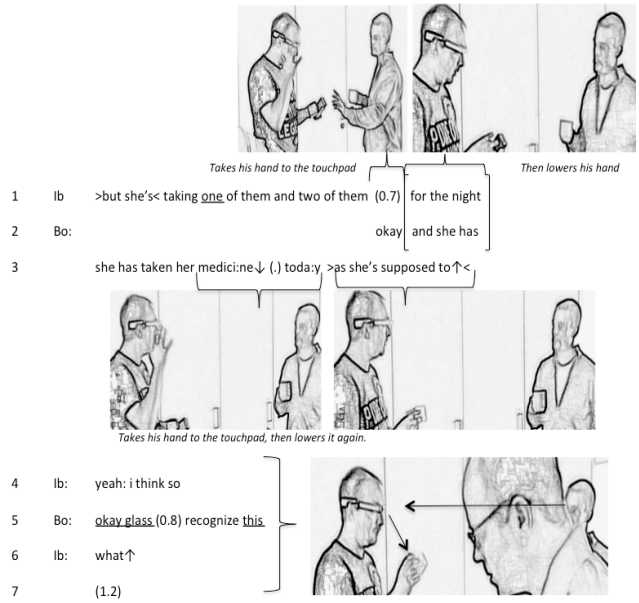
The students are talking about who one of them will have as supervisor. In line 4, there is a pause and the production of self-initiated repair indicators, as Kim (the student) seeks the supervisor's name (line 5). However, Tom (the Glass-user) takes a turn in overlap, thus treating the pause and hesitation or name search as a transition relevant place. In response, they gaze at Tom whose response in lines 6 and 7 is to private Glass information, as a "non-human side sequence". Before the closing "well", Kim starts producing the second pair part in line 8 (answer to Tim's initial question (not in transcript)), thereby treating Tom's actions as out of context and related to something unknowable. Thus, a non-human participant (Glass) provides information that affects social interaction because the Glass-user (suddenly) takes a turn with the Glass-provided information.

Glass interaction occupies a slot in conversation. This may be why overlap is produced in different ways and also why it is – at the moment – difficult for participants to interact pursuant to the "normal" rules of the turn-taking system. Non-human overlaps are difficult for (non-Glass-wearing) human participants to understand and project since they have no knowledge of the information provided by Glass.

Participants use several embodied markers to compensate for the lack of a verbally-produced context. The gaze of the Glass-user is awkward or, at least, unusual as he / she at times stares out into the open space in front of him / her. Co-participants may interpret these embodied semiotic systems (gazing and gesturing in specific ways) as actions oriented towards Glass as a non-human participant, but they are nevertheless – for the moment – not fully incorporated in the recipient design and production of relevant next actions. Hence, co-participants do not treat the silent Glass interaction as something that is sequentially relevant. This may produce confusion and a lack of understanding of the situation, which will be further discussed below.

EPISTEMICS: PARTICIPANT'S LACK OF UNDERSTANDING AND EPISTEMIC INCONGRUENCES

Constructing mutual understanding is often a complicated matter for participants in social interaction. The capacity for understanding the intentional, goal-directed behaviour of others is a fundamentally interactional process, one that cannot be extricated from the ongoing flow of social activity (Kidwell & Zimmerman 2007). Participants normally display their understanding of relevant objects, things, issues, etc., in the situation (Hindmarsh & Heath 2000). Research on objects in workplace settings has largely focused on the publicly available design of things such as computers, office items, screens, etc. (e.g., Goodwin & Goodwin 1996). However, Glass is a different kind of object - in part, because it can be seen, as stated before, as more than an object but also as a non-human participant and, in part, because the information, which is accessed during social interaction, is not socially displayed or publicly available (like, e.g., a common observable computer screen is) but a private experience for the Glass-user. Similarly, emotions or other kinds of experiences such as, for instance, tasting are private experiences that can be made public by displaying or producing accounts (Streeck 1996). The following example shows how the interaction between non-human and human participants is embedded in a social context of ambiguity. In this example from case 3, the paramedic (Bo) is trying to scan an object with Glass while interacting with the ill woman's husband (Ib).



Transcript 4. Difficulties in understanding the situation. The transcript is from case 3.

Bo has asked Ib which pills his wife has taken, and Ib answers this question in line 1. In line 3, Bo produces a question to which Ib responds in line 4 with a preferred second pair part. Both parties treat these adjacency pairs as common institutional interaction formats. However, Bo seems not only engaged in the conversation but also with actions oriented towards Glass. So far, no one has accounted for the fact that Bo is wearing Glass, although

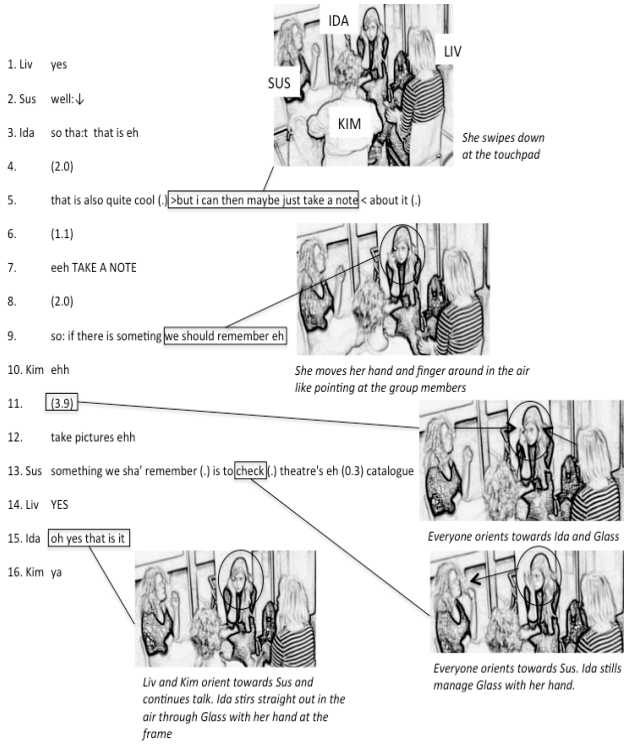
Bo' interactions with Glass have been visible and audible with his actions (touching the frame) and the gaze (looking with Glass at the box). After the response in line 4 in what has so far been a question-answer sequence, Bo quickly looks down at the medicine box and gives Glass a command: "okay, Glass (0.8), recognize this", which may be seen as preclosing the question-answer sequence. However, Ib responds in the next slot, which he treats as a transition relevant place with an open question marker and intonation going up: "what↑", which produces a gap of 1.2 seconds.

It is clear from this short example that, at first, the two human participants are engaged in a familiar question-answer sequence structure. But suddenly and without any context markers that otherwise might have projected a change of recipient design, Bo orients towards the non-human participant (Glass), thus producing interactional uncertainty. And the response from Ib is a repair initiator: "what↑". A regular sequence structure is turned into an ambiguous one.

Orientation is generally evident when speakers abort and restart units of talk in specific circumstances (cf. Maynard & Clayman 2003). Goodwin (1979), for instance, has shown that, when a speaker notices a recipient's gaze begin to wander, he will frequently make a cut off and restart the turn. Understanding or lack of understanding is, thus, displayed in interaction – just as Ib is displaying a lack of understanding in the situation by the "what↑" particle, due to 1) the sudden establishment of Glass as a relevant non-human participant and 2) the lack of any verbal context markers or the invocation of procedures (Nielsen et al. 2012) that might otherwise project Glass as a next relevant participant. The recipient could do a no-knowledge account towards the Glass interaction, but does not, which may be due to the general "norm": *display of no-knowledge* as a significant face-threatening social issue (Keevallik 2011). However, *epistemic incongruence* (Mondada 2011) is prevalent: When Ib responds with the "what↑", it functions as a *fishing device* (Pomerantz 1980), which is a resource for the pursuit of the interaction or it can even be interpreted as a practice for increasing participation. As a fishing device, the "what↑" does the job of responding to and accounting for the lack of understanding of the Glass-user's private experience and interaction with the non-human participant (Glass). Through interactively organised talk, gestures and postures, participants "normally" display relevant information about the sequential organisation of their joint participation. Participants orient towards a common ground and common understanding when trying to make sense of the situation. However, as shown in the analysis, the participant's sense-making and interpretation of the sequential environment may be troubled by unknown elements in the interaction.

One of the most obvious solutions with the present kind of smart glass technology is for participants to account for, explain, and define the situation and to use meta-language (Bateson 1972) to construct a common understanding and etiquette for how to behave "accordingly" with the non-human participant (Glass). From a classic HCI standpoint, Thad Starner (1999) has made an argument for the construction of a "shared mental model". However, as this cognitivist approach is inconsistent with CA, I propose a rather socio-practical solution, which is to make the private experience a public one through design of social interaction.

This task of making the private experience a public one did not succeed in the above example, producing the ambiguity in the situation. In the following example from case 1, participants are talking about an exhibition at the National Gallery of Denmark (SMK). In this situation, the Glass-user is accounting for her activity with the Glass, thus creating a basis for common understanding, which makes the situation a little more comprehensible.



Transcript 5. Constructing mutual understanding of the situation through accounts. The transcript is from case 1.

In line 7, Ida gives the command "eeh TAKE A NOTE", which is displayed by social markers: the code switching of language (from Danish to English) and high volume, thus making the action recognizable as directed towards the non-human participant (Glass). This is, of course, already interpretable by her co-participants as a relevant action projected by her in line 5, which may be the reason for the lack of uptake from the others as they may be interpreting her turn as a multi-unit turn, which has not yet reached a transition relevant place. So far, it has only been Ida who has taken a turn and, several times, produced a self-select at possible completion points (silence in lines 4, 6, 8), and there is no overlap or repair. This may be due to her explanation in line 5 about how she could take a note, which is treated tacitly as an understandable and relevant Glass action by the others. However, it is not until she includes them all in the Glass-directed action by using clearly marked *gestural drawings* (Due forth.) in the air that

she verbally invites them in line 9: "so: if there is something we should remember eh". When she does this verbal invitation, she also includes her co-participants by turning her index finger in the air, pointing at them all as next relevant speakers; and, in the pause in line 11 she is gazing at them. This clearly marked embodied invitation to participate in an interaction with her *and* the non-human participant (Glass) is a success, as Sus responds with a proposal in line 13: "something we sha' remember", to which the others align themselves afterwards.

This section has been about ambiguity and epistemics in social interaction with Google Glass. The analyses of the two examples have shown several important things: First, that social interaction with a Glass-user may produce ambiguity and difficulties, which are observable as embodied tacit orientation with shifts in gaze, repair initiations, lack of uptake, and the production of silence at possible completion points in the conversation; second, that this ambiguity may be mitigated by displaying and accounting for the Glass-directed interaction - most notably, by a clear invitation to participate in the Glass-directed interaction. Speaking to Glass as "nonhuman side sequences" may, at the moment, become an interactional problem, depending on the social involvement in the actions. Participants behave in accordance with the *Why that know*-question, and Glass interaction seems to be either relevant or irrelevant, depending on the meta-communication and social production of relevancies.

DISCUSSION

Today, the various functions of smart glasses are ambiguous to most participants, which is observable as, e.g., hesitation and repair initiations in social interaction. As Starner (1999) noticed, the affordances of portable devices in the past helped constrain their perceived use. As opposed, for instance, to the use of a smartphone, which is almost always observable in interaction, the use of Glass is a more uncertain activity. With Glass, there are no general observable markers for activity and even fewer markers for specific activities. The functionality of Glass is manifold: check emails, take photos, record videos, search the Internet, play games, or use other kinds of apps. However, none of these different activities are displayable in social context. When participants are using smartphones, their co-participants may have a clue about what is going on since they can see the display or know the context. This is not the case with Glass. Therefore, the experience and understanding of the situation are private. Co-participants may have no idea as to how the device is being used at the time and whether or not the user is interruptible or a relevant / non-relevant participant in interaction. This is, as has been shown, a complicated matter in social interaction. There seem to be two obvious solutions: the design of a new shared social praxis or etiquette and/or a new technological design.

It seems at the moment with the present technology that information cannot be provided fast enough to be presented at specific, relevant times in interaction, which inevitably makes almost every kind of Glass-provided information out of sequential context, thus producing the different overlaps as shown throughout the analysis. The proposed "rule" for *micro-interactions* with technology, which involves a maximum of two to four seconds of waiting time (Miller 1968; Oulasvirta et al. 2005), needs to be re-thought because it has only been investigated from an HCI perspective (Ashbrook 2010:8). As Ashbrook also concludes, micro-interactions are desirable because they may minimize interruption. However, as has been shown throughout this paper, Glass interactions in a social context often occupy slots in conversation. Two to four seconds may be a long time to wait for interaction with a computer, but it seems even longer time when displayed as pauses in social interaction. As shown, participants are, on one hand, uncertain about how to respond and when to respond in relation to Glass-based interactions and, on the other hand, they are not just waiting in silence for the Glass-user to come back into the conversation. Instead, co-participants move on in the interaction, making Glass interactions appear out of context and as inappropriate overlaps.

One of the most obvious methods for overcoming barriers to using Glass is not necessarily to develop the technology or user interface as it is to develop and design shared practices and understandings of what a non-human participant such as Glass is or should be in interaction. Some of the major questions that arise from this are how or in what ways people will accept or treat Glass or smart glasses in general as non-human participants in social interaction with rights and obligations like human participants.

Variations of "dual purpose speech" technology (Lyons 2005) in which Glass interprets ordinary talk and, on the basis of key words and phrases from the talk, provides relevant information (such as time, calendar, etc.) may help regarding input. One of the key product design challenges moving forward must, therefore, be to continue work on the speed of providing relevant information and the way commands (input) are given to Glass.

For the time being, Google Glass is not for sale on the consumer market, which may be because of the social complications. The future versions of Google Glass and other smart glasses may learn from micro-analytical studies to design high-speed, non-disruptive inputs and social markers.

ACKNOWLEDGMENTS

Thanks to Anne Lidell and Thomas Clausen from CA, E14, Thomas Prien-Larsen and his friends from Cognition & Communication, INSS, Mathias Andersen and Kenny from IT & Cognition, and KU for help with data collection. Also thanks for insightful comments from anonymous reviewers one and two and from Thad Starner on an earlier draft of this paper. Some of this research has been conducted by funding from the Synoptik Foundation.

REFERENCES

Ashbrook, Daniel L. 2010. *Enabling* M*obile* M*icrointeractions*. PhD from the College of Computing. Georgia Institute of Technology.

Atkinson, J. Maxwell & Paul Drew. 1979. *Order in Court: The Organisation of Verbal Interaction in Judicial Settings*. London: Macmillan.

Bateson, Gregory. 1972. *Steps to an Ecology of Mind: Collected Essays in Anthropology, Psychiatry, Evolution, and Epistemology.* San Francisco: Chandler Pub. Co.

Due, Brian L. forth. Co-construction of Imagination Spaces: A Multimodal Analysis of Idea Development. *Journal of Pragmatics*.

Fraser, Mike, Tony Glover, Ivan Vaghi, Steve Benford, Chris Greenhalgh, Jon Hindmarsh & Christian Heath. 2000. Revealing the Realities of Collaborative Virtual Reality. *Proceedings of the* T*hird* I*nternational* C*onference on Collaborative* V*irtual* E*nvironments*, 29–37. New York: ACM.

Garfinkel, Harold. 1967. *Studies in Ethnomethodology*. Englewood Cliffs, N. J.

Goodwin, Charles. 1979. The Interactive Construction of a Sentence in Natural Conversation. *G. Psathas (Ed.) Everyday Language: Studies in Ethnomethodology*, 97–121. New York, Irvington Publishers.

Goodwin, Charles & Marjorie H. Goodwin. 1996. Formulating Planes: Seeing as a Situated Activity. *Cognition and Communication at Work, (eds.) David Middleton and Yrjö Engestrom*, 61–95. Cambridge University Press.

Heath, Christian & Paul Luff. 1992. Media Space and Communicative Asymmetries: Preliminary Observations of Video-Mediated Interaction. *Human–Computer Interaction* 7(3). 315–346.

Heath, Christian, Poul Luff & A. Sellen. 1995. Reconsidering the Virtual Workplace: Flexible Support for Collaborative Activity. *Proceedings of the ECSCW 95 European Conference on Computer Supported Cooperative Work.* Dordrecht: Kluwer.

Hindmarsh, Jon, Mik Fraser, Christian Heath, Steve Benford & Chris Greenhalgh. 1998. Fragmented Interaction: Establishing Mutual Orientation in Virtual Environments. *Proceedings of the CSCW 98 Conference on Computer Supported Cooperative Work.* New York: ACM.

Hindmarsh, Jon & Christian Heath. 2000. Sharing the Tools of the Trade: The Interactional Constitution of Workplace Objects. *Journal of Contemporary Ethnography* 29(5). 523–562.

Hornecker, Eva & Jacob Buur. 2006. Getting a Grip on Tangible Interaction: A Framework on Physical Space and Social Interaction. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 437–446. (CHI '06). New York, NY, USA: ACM.

Jacko, Julie A. 2012. *Human Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications, Third Edition*. Boca Raton, FL: CRC Press.

Jefferson, Gail. 1983. Notes on a Possible Metric Which Provides for a "Standard Maximum" Silence of Approximately One Second in Conversation. *Tilburg Papers in Language and Literature*, vol. 42, 1–83. Tilburg.

Keevallik, Leelo. 2011. The Terms of Not Knowing. *The Morality of Knowledge in Conversation (eds.) Tanya Stivers, Lorenza Mondada, Jakob Steensig*, 184–206. Cambridge University Press.

Kendon, Adam. 2005. *Gesture: Visible Action as Utterance*. Cambridge University Press.

Kidwell, M. & D.H. Zimmerman. 2007. Joint Attention as Action. *Journal of Pragmatics* 39(3). 592–611.

Latour, Bruno. 1994. On Technical Mediation. *Common Knowledge Vol.3, n*o. *2*. 29–64.

Luff, Poul, Christian Heath, H. Kuzuoka, Jon Hindmarsh, K. Yamazaki & S. Oyama. 2003. Fractured Ecologies: Creating Environments for Collaboration. *Human-Computer Interaction* 18(1). 51.

Lyons, Kent. 2005. *Improving Support of Conversations by Enhancing Mobile Computer Input*. PhD. College of Computing Georgia Institute of Technology.

Maynard, D.W. & Steven Clayman. 2003. Ethnomethodology and Conversation Analysis. *Reynolds, L.T., Herman-*K*inney, N.J. (eds.) Handbook of* S*ymbolic* I*nteractionism*, 173–203. AltaMira Press.

Merleau-Ponty, Maurice. 2002. *Phenomenology of Perception*. London, New York: Routledge.

Miller, Robert B. 1968. Response Time in Man-computer Conversational Transactions. *Proceedings of the Fall Joint Computer Conference,* 267–277. New York, NY, USA: ACM.

Mondada, Lorenza. 2011. The Management of Knowledge Discrepancies and of Epistemic Changes in Institutional Interactions. *Stivers, T. Mondada, L. Steensig, J. (eds.) The Morality of Knowledge in Conversation*, 27–57. Cambridge University Press.

Nielsen, Mie Femø, Søren Beck Nielsen, Gitte Gravengaard & Brian L. Due. 2012. Interactional Functions of Invoking Procedure in Institutional Settings. *Journal of Pragmatics* 44(11). 1457–1473.

Norman, Donald A. 1983. Some Observations on Mental Models. *Mental Models*, *(eds.*) *D. Gentner and A.L.Stevens*. Psychology Press.

Oulasvirta, Antti, Sakari Tamminen, Virpi Roto & Jaana Kuorelahti. 2005. Interaction in 4-Second Bursts: The Fragmented Nature of Attentional Resources in Mobile HCI. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 919–928. (CHI '05). New York, NY, USA: ACM.

Pomerantz, Anita. 1980. Telling My Side: “Limited Access" as a ”Fishing" Device. *Sociological Inquiry* 50(3-4). 186–198.

Rogers, Yvonne, Helen Sharp & Jenny Preece. 2011. *Interaction Design: Beyond Human - Computer Interaction*. 3 edition. Chichester, West Sussex, U.K: Wiley.

Sacks, Harvey L., Emmanuel A. Schegloff & Gail Jefferson. 1974. A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Language* 50(4). 696–735.

Schank, Roger C. & Robert P. Abelson. 1977. *Scripts, Plans, Goals, and Understanding: An Inquiry Into Human Knowledge Structures*. 1 edition. Hillsdale, N.J., New York: Psychology Press.

Schegloff, Emmanuel A. 1968. Sequencing in Conversational Openings. *American Anthropologist* 70(6). (New Series). 1075–1095.

Schegloff, Emmanuel A. 1988. "On a Actual Virtual Servo-Mechanism for Guessign Bad News: A Single Case Conjecture. (Vol. 35, No. 4). (Social Problems).

Schegloff, Emmanuel A. & Harvey L. Sacks. 1973. *Opening Up Closings*. Semiotica 8.

Starner, Thad. 1996. Human-Powered Wearable Computing. *IBM Systems Journal* 35(3.4). 618–629.

Starner, Thad, Bradley Rhodes, Joshua Weaver & Alex Pentland. 1999. *Everyday-use Wearable Computers*.

Streeck, Jürgen. 1996. How to Do Things with Things. *Human Studies* 19(4). 365–384.

Streeck, Jürgen, Charles Goodwin & Curtis LeBaron. 2011. *Embodied Interaction: Language and Body in the Material World*. Cambridge University Press.