

An Ethnomethodologically-Informed Approach to Interface Design for Social Interactions around Video Online

Anna Zawilska

Department of Computer Science
University of Oxford
Parks Road OX1 3QD

Steven Albury

IT Services
University of Oxford
13 Banbury Road OX2 6NN

Abstract

With the emergence of video on community media sites such as YouTube or TED, there is a need to understand interactions by community participants around video in order to maximise the potential of these systems to support communities and create meaningful interactions online. In this paper, we illustrate how ethnomethodologically-informed studies of the social interactions among a small group of co-located participants around video may be used to inform the design of an online video interface which may support interactions by a large distributed group of participants. Our point of view is grounded in the idea that the interactional accomplishments of a co-located group remain relevant at a distributed level, thereby allowing an ethnomethodologically-informed approach to arrive at effective implications for design of online systems. A strength of this approach is the potential to create implications which are properly grounded in a set of observations of the precise ways in which people interact to accomplish social interaction around video. For the purpose of illustration, we perform an analysis of a fragment of video data collected from a quasi-naturalistic experiment of a co-located group collaboratively annotating a video, from which we propose a number of design implications for a video annotation interface.

Introduction

Many popular online community sites, such as YouTube and TED, primarily use video to present subject matter which becomes discussion points for online community members. In designing and developing these video-based sites many analytic lenses may be used. One category of analytic lens may be derived from sociological theory - such as the theory of social constructivism (Vygotsky 1978). However, these kinds of lenses may be critiqued for offering somewhat interactionally decontextualised analyses of social phenomena. Ethnomethodologically-informed studies represent a shift away from theory-based approaches. Rather than using social interaction as a *resource* for inquiry which leads to a rather structural view of interaction, ethnomethodology uses interactions as a *topic* of inquiry. For ethnomethodology, social life is produced 'from within' by members of the social setting, sociality is practically accomplished by the settings'

Copyright © 2015, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

members, and is of a locally situated character. With reference to interactions around video in community media sites, this means that ethnomethodology inquires how the natural 'facts' of sociality around video are produced in the first place rather than having them 'in place' and then theorising them. In this paper, we perform an ethnomethodologically-informed analysis of the interactions of a co-located group interacting around video: in particular, they are collaboratively annotating a video using a video annotation interface prototype. We use this to provide implications for the design of a video annotation interface online to support larger distributed groups by arguing that the interactional accomplishments of the distributed group will be the same as those of the co-located group even if the details of practice are different. A deep understanding of those accomplishments, as surfaced through the look at the details of the interactions of participants of the co-located group, will then provide implications for the improved design which offers functional accomplishment equivalence for the distributed group. Our aim in this paper is to provide a fragment of analysis which exhibits the possibility of usefully using this approach with respect to designing social interactions around video for online community sites.

Background

Ethnomethodology is grounded in Harold Garfinkel's dissatisfaction with conventional theory of social action which, according to Garfinkel, takes for granted the availability and intelligibility of social actions, social identities, and social settings (Garfinkel 1967). Where conventional sociological theory may impart on ordinary social life theoretical structures and factors seen to be played out, ethnomethodology draws attention to this taken-for-granted character of social life through which members of society produce their activities to be accountable for what they are. Thus, ethnomethodology draws the researcher's attention to the local, situated, real-time organisational specifics of social activity and insists on empirical studies grounded in data capturing real-world social phenomena. When used to provide implications for the design of computer systems, ethnomethodology cannot be used with traditional notions of abstraction and generalisation in design. In their seminal paper, Dourish and Button termed 'technomethodology' which considers how best to use an ethnomethodologically-informed ap-

proach to design (Dourish and Button 1998). The argument behind technomethodology is that the accomplishments of the members of a setting which have at their centre methodical and accountable practices of those members, remain relevant even when there is a change in form of setting, such as a change from the local to distributed. A deep understanding of the former allows for improved design of functional equivalence in the latter by giving attention to the outcomes of interaction. Studies of situated action have been valuable in other domains of computing, such as ubiquitous computing (Dourish 2004) and artificial intelligence (Suchman 1987), particularly in revealing new areas of exploration and breaking down preconceived notions of systems.

The idea of using ethnomethodologically-informed methods to examine social interaction among small groups as a way of informing the design of online systems to support larger distributed groups has been developed substantially in literature on groupware. Within this literature, most studies use as data the text produced over time among a small group of distributed users over text-based communication channels such as chat (e.g. (Stahl 2005; Arnseth et al. 2004)). The ways in which the text conversation unfolds over time is subject to ethnomethodologically-informed analysis and in particular Conversation Analysis (Sacks 1984). Stahl has suggested that, for the purpose of creating theory of social interactions among groups of distributed users, an effective approach may be to consider several layers of interaction: the individual human agent; the small group; and the community of practice, the linguistic community or the cultural community (Stahl 2009). He does this while admitting that these layers are not completely independent but may co-evolve. In this paper, we align somewhat with this prior work in groupware by considering how the social interactions within a group evolve over time and by applying an ethnomethodologically-informed approach to attempt to surface this. However, this paper remains more true to the traditions of ethnomethodology by not attempting to produce a theory as Stahl attempts to do. Furthermore, this paper aims to make a contribution by starting a discussion around how the social interactions seen among a co-located group (including the seen-but-unnoticed ways a group produces their social interaction) which are examined through an ethnomethodologically-informed lens, may be used to design systems to support distributed users. This paper makes a further contribution by applying this approach to the design of online video-based community sites more specifically as these studies are limited in existing literature, in order to see if this approach may be of value for analysis and design for these sites.

Sustained Participation of the Group During Video Annotation

Before presenting a fragment of data with analysis in this section, we provide some detail about the methodology used for eliciting and analysing the fragment data. The fragment is taken from a quasi-naturalistic study. Quasi-naturalistic study is where participants engage in an open-ended task or set of activities with a prototype technology (Monk 1985).

These experiments are frequently used to develop new technologies (eg. (Hindmarsh et al. 1998)).

In this study, three participants are asked to view a video on Big Data, to consider how the content may or may not be interesting to them, and to annotate any thoughts they may have using a prototype they use together. In this prototype, an iPad is mounted onto a whiteboard, and participants add annotations around the video using paper and magnetic strips. The participants are three Computer Science students. The experiment was allowed to run for as long as the participants wished, and the participants ended up spending forty-five minutes interacting with the video. The participants characterised each other as work colleagues (and so they knew each other as members of the Department of Computer Science) but were not familiar with each other's current work or prior academic, professional, or personal experiences. The participants were told they would watch a video on Big Data as part of the experiment, and all of them responded that they were somewhat interested in this topic. However, the participants were not given any more information about the video content and had not seen the video previously.

Figure 1 provides a transcript of the fragment, with notation taken from Heritage (Heritage 1984). Figure 2 provides screenshots of camera footage of the group captured at an angle in front of the group, with the whiteboard not visible to the right of the footage. Just before the fragment, three learners (V, K, and J) have paused the video they are watching, and are discussing the value of different kinds of data. At the beginning of the fragment, the group form one conversational circle; their bodies and gazes are orientated towards each other. In line 3, J shows his support for the point that has been made by K ('that's a good point'), and turns away from the group, reaching for paper and pencil to add an annotation to the video. As J does this and moves to add the annotation, V continues to talk in line 4. K turns her gaze and body orientation back towards V. In lines 4 to 20, V and K have their bodies and gazes oriented towards each other, closing the conversational circle around themselves, no longer including J with visual cues such as eye contact or body orientation. During this time, J has his back turned to this new conversational circle, and is occupied writing up the annotation alone and in silence. In line 21, J stops writing the annotation and re-orientates his body so that he begins to turn back toward V and K. J's talk in line 21 ('that's a good point because') leads from and builds on the talk which occurred between V and K while J was writing. K and V's gazes and body orientations shift away from each other exclusively, to include J back into the conversational circle. K also acknowledges J's re-entry as a ratified member of the conversational circle by verbally agreeing with J, in line 24 ('Ya'). In line 29, at ('to (0.5) senso:r'), J orientates his body and gaze away from the conversational circle again, turns his back on V and K, and adds text to the annotation he created previously. V and K, rather than closing off the conversational circle around themselves and continuing their talk, shift their gazes toward the annotation to view what J is adding.

The fragment illustrates how J tacitly shifts in and out of

1 K: Based on patterns that they see they (0.2) could be
 2 really wrong (0.5) sometimes
 3 J: =Actually that's a good point
 4 V: Surely a GP a lot of it is (0.1) I mean the reason we
 5 don't have just a robo:t (0.2) deciding
 6 K: [Ya true
 7 V: what you've got or why you don't just look on Google
 8 K: [Ya [Ya
 9 V: (0.5) symptoms or whatever
 10 K: [Ya
 11 =Exactly
 12 V: =Is because there's also a judgement that the doctor
 13 makes that's not necessarily from
 14 K: [Not necessarily ya
 15 V: From the data in a way that is measurable↑ (0.3) in
 16 that way↑ I mean it's more like
 17 K: [From the data ya
 18 V: Just a value judgement (0.5) You look at someone and
 19 you (0.5) you recognise something that's intangible
 20 K: =Ya:: (0.5) that's true
 21 J: (2.0) Ya (2.0) that's a good point because ofte::n (0.1)
 22 when you go to a doctor it starts with how you feel it's
 23 like very (1) ya
 24 K: =Ya
 25 J: (1.0) Ah:: (1.5) not measurable in a sense (0.5) and also
 26 yeah especially (0.2) it's very subjective how you fee::l
 27 and the pain sca::le and all these things (0.5) so I think
 28 (0.5) there are ya (0.2) certain limitations (0.1) to what
 29 we can measure (2.0) to (0.5) senso::r or (0.3) or input
 30 in general

Figure 1: Fragment

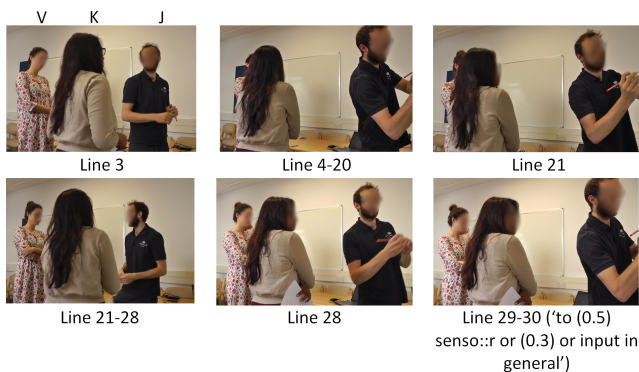


Figure 2: J moving in and out of the conversational circle maintained by V and K, illustrating continued involvement of the annotation-creator in the group during video annotation

the conversational circle including V and K. Even while J is outside of the conversational circle, annotation-writing is not his only involvement, rather he also takes on the role of the 'overhearer' of the talk between V and K. This fragment

shows that annotation creation is not a neat and bounded event driven by the annotation creator during which the annotation creator allocates their full attention to creating the annotation thereby temporarily isolating themselves from any interactions with other learners. Rather than only being an outcome of interaction, the annotation is part of a contingent, unfolding, and situated process that is the continued negotiation of the group about the content of the annotation. It is the inquiry into the group's interaction which have revealed annotation creation to be a rich process.

Design Implications

Although it is not possible to immediately functionally represent all of the rich interactional cues seen in the fragment in one interface tool for distributed learners online, we suggest this fragment does provide some substantive implications for the design of the video annotation interface for video-based community sites. The implications have been derived by formulating some of the interactional accomplishments seen in the fragment, and making them visible in the system design (Button and Dourish 1996). With regard to video annotation, this fragment illustrates:

1. the possible need for mutual availability of the work of annotation to all parties involved.
2. the concurrent need for the availability of ongoing interaction outside of the annotation task.
3. the need to support the ongoing revision of annotation in the context of interaction.

These interactional accomplishments would have been difficult to surface using other methods as the practices of the participants were not pre-planned or explicitly agreed upon by the group in advance, but rather emerged during and through their engagement with the prototype. These accomplishments emerge from the seen-but-unnoticed practices at play. Furthermore, documenting the situated practice and allowing the interactional accomplishments to emerge from an analysis of this practice, allows us to not only know create a list of interactional accomplishments, but to also understand in detail the centrality and purpose that these accomplishments play in collaborative video annotation. Even if these interactional accomplishment could somehow be deduced using an alternative methodology, such as through interview, this important and rich understanding of the centrality of accomplishment in the collaboration would be lost. These broad elements of the group's interactions may be translated into the following implications for the video annotation interface:

1. the interface for creating annotations should not block the annotator's view of other learner interactions while creating the interface, so as to continue to allow the annotator to directly participate in, or witness ('overhear'), other learner interactions.
2. annotations should be editable, even after they are first submitted by the creator, so as to support the kind of iterative building up of the information state of the annotation seen in the fragment.

3. it may be valuable to allow learners to link annotations to other learner discussions or annotations, so that the interactions on which the information state of the annotation draws may be seen.

Our future work will include a study of use of a video annotation prototype, informed by these requirements, by a group of distributed learners. In this study, we will have three participants sitting simultaneously in different rooms using the same system and we will study any social interaction that occurs between the participants through the system. Additionally, we will ask these participants to ‘think aloud’ (Van den Haak, de Jong, and Schellens 2004) while using the system so we get some understanding of how they make sense of the prototype and orient towards it. There is a possibility that social activity may be transformed to such a dramatic extent that the interactional accomplishments of the co-located group of participants in the quasi-naturalistic study may not apply to a distributed group of users. If the interactional accomplishments seen for the co-located group are not applicable at the distributed level, then the researcher would see many indicators of ‘breakdown’ (Douglas 1995) for the distributed group which occurs when users can’t find expected objects or they repeatedly select the wrong objects in the interface. In this case, an ethnomethodologically-informed approach as suggested here would have at least provided traction in the investigation into the design of the interface of community media sites, even if it is to draw contrast between online and offline interactional accomplishments.

Conclusion

In this short paper, we have shown how an ethnomethodologically-informed approach may be used to inform the design of online video interfaces for community media sites where content is delivered using video. In particular, an ethnomethodologically-informed analysis of the interactions of a small co-located group around video may be used to inform the design of an online interface supporting larger distributed groups by considering the same interactional accomplishments of the co-located group to be relevant to larger distributed groups. An understanding of these accomplishments is rooted in the detailed analysis of the local, situated, and emergent interactions of the co-located group which ethnomethodology provides. This is in contrast to traditional social theory-centric approaches to analysing interactions of online community media site participants which may provide a largely structural, interactionally de-contextualised view of interactions. This paper uses an example fragment of a small co-located group collaboratively annotating a video to provide implications for the design of an online video annotation interface.

References

Arnseth, H.; Ludvigsen, S.; Morch, A.; and Wassen, B. 2004. Managing Intersubjectivity in Distributed Collaboration. *PsychNology* 2(2):189–204.

Button, G., and Dourish, P. 1996. Technomethodology:

paradoxes and possibilities. In *Proc. SIGCHI conference on Human factors in computing systems*, 19–26. ACM.

Douglas, S. 1995. Conversation analysis and human-computer interaction design. In *The Social and Interactional Dimensions of Human-Computer Interfaces*. Cambridge, UK: Cornell University Press. 184–203.

Dourish, P., and Button, G. 1998. On ‘‘technomethodology’’: Foundational relationships between ethnomethodology and system design. *Human-computer interaction* 13(4):395–432.

Dourish, P. 2004. What we talk about when we talk about context. *Personal and Ubiquitous Computing* 8(1):19–30.

Garfinkel, H. 1967. *Studies in ethnomethodology*. NJ: Englewood Cliffs.

Heritage, J. 1984. *Garfinkel and ethnomethodology*. Cambridge: Polity Press.

Hindmarsh, J.; Fraser, M.; Heath, C.; Benford, R.; and Greenhalgh, C. 1998. Fragmented interaction: establishing mutual orientation in virtual environments. In *Proceedings of the 1998 ACM conference on Computer supported cooperative work*, 217–226. ACM.

Monk, A. 1985. *Fundamental of Human-Computer Interaction*. London: Academic Press.

Sacks, H. 1984. Notes on methodology. In Maxwell, J., and J., H., eds., *Structures of Social Action: Studies in Conversation Analysis*. 21–27.

Stahl, G. 2005. Groups, group cognition and groupware. In *Groupware: Design, Implementation, and Use*. Berlin Heidelberg: Springer. 1–16.

Stahl, G. 2009. Collaborative information practices in virtual math teams. In *Proc. Conference for Groupware*.

Suchman, L. 1987. *Plans and Situated Action: The problem of human-machine communication*. Cambridge, UK: Cambridge University Press.

Van den Haak, M.; de Jong, M.; and Schellens, P. 2004. Employing think-aloud protocols and constructive interaction to test the usability of online library catalogues: a methodological comparison. *Interacting with Computers* 16(6):1153–1170.

Vygotsky, L. S. 1978. *Mind in Society: the development of higher psychological processes*. Harvard University Press.