CHRISTOPH MEIER

Neue Medien – neue Kommunikationsformen?

Strukturmerkmale von Videokonferenzen¹

Abstract

Obwohl kombinierte Ton-Bild-Verbindungen seit etwa 60 Jahren technisch möglich sind, ist noch längst nicht geklärt, inwiefern Videokonferenzen sich von herkömmlichen Besprechungen unterscheiden. In diesem Kapitel werden besondere Merkmale der Interaktionsorganisation und -dynamik von Videokonferenzen beschrieben, indem vier kurze Passagen vorgestellt und analysiert werden. Dabei werden Ausschnitte der untersuchten Materialien gegenstandsnah in Form von kurzen Videoclips präsentiert und auf diese Weise die Analyse nachvollziehbar und überprüfbar gemacht. Zu den besonderen Merkmalen von Videokonferenzen gehören unter anderen die veränderte Ökologie des Besprechungsraums, die veränderte Funktionalität von Blickzuwendung und Lautstärke des Sprechens für das Etablieren von Beteiligungskonstellationen sowie die eingeschränkte Transparenz von Handlungen „hier“ für die Beteiligten „dort“. Dazu gehören aber auch die zeitweise holprige Interaktionsqualität aufgrund der technisch bedingten Zeitverzögerung, die Schwierigkeiten, sich von „hier“ aus in eine kontroverse Diskussion „dort“ einzuschalten, sowie schließlich Probleme mit Aktivitäten wie etwa Lachen und Sich-Amüsieren, die eine hochgradige Koordination der Beteiligten erfordern.

Vorbemerkung

Dieser Aufsatz beschäftigt sich nicht nur inhaltlich mit den sogenannten Neuen Medien, er ist auch selbst ein Versuch, neue Medien für wissenschaftliches Arbeiten zu nutzen. Das heißt in diesem Fall: für die Präsentation von Materialien, auf denen die vorgestellten Analysen beruhen. Ein

¹ Die Fallstudie, über deren Ergebnisse hier berichtet wird, wurde im Rahmen eines größeren Forschungsprojekts durchgeführt, das nach den spezifischen Qualitäten von kooperativen, über digitale, interaktive Medien realisierten Arbeitsprozessen fragte. Eine Darstellung des Projekts sowie Arbeitsberichte zu den beiden im Rahmen des Projekts durchgeführten Fallstudien finden sich unter <http://www.uni-giessen.de/~g31047>. Für Anregungen, Kommentare und Verbesserungsvorschläge bedanke ich mich bei Jörg Bergmann und Michaela Goll. Ich danke auch den Mitgliedern des *Customer Services Council* der Firma *Technics* für ihr Vertrauen und ihre Unterstützung.

Problem im Zusammenhang mit der Veröffentlichung von Untersuchungen zum Sprechen-beim-Interagieren bestand bislang darin, dass die untersuchten Phänomene in eine Schriftform überführt werden mussten, um sie in nachvollziehbarer Weise einem Publikum präsentieren zu können. Das World Wide Web und verschiedenste Software-Anwendungen machen mittlerweile neue Formen der Explikation von nicht-schriftbasierten Forschungsgegenständen möglich. So können zum einen audiovisuelle Primärmaterialien (zumindest in Auszügen) für KollegInnen verfügbar gemacht werden. Zum anderen kann auch die Präsentation von Ergebnissen eine andere Form annehmen. Ich nutze diese Möglichkeiten für meinen Beitrag, um anonymisierte Ausschnitte aus den untersuchten Materialien in der Form von Videoclips on-line verfügbar zu machen. Diese Clips stehen unter der Adresse „<http://www.ids-mannheim.de/pub/jahrbuch/jb99m.html>“ zum Herunterladen bereit. Um sie anschauen zu können, benötigt man den Real Player G2, dessen Basisversion kostenlos über die Web-Seite „<http://www.real.com>“ verfügbar ist. Darüber hinaus empfiehlt sich ein Rechner mit einem schnellen Prozessor (z. B. Intel Pentium II), da ansonsten die hochkomprimierten Videoclips nicht flüssig abgespielt werden.

1. Einleitung

Die technische Realisierung einer Bild- und Sprechverbindung zu einem entfernten Gegenüber gelang schon vor mehr als sechzig Jahren. Die Deutsche Reichspost betrieb zwischen 1936 und 1940 einen öffentlichen Bildfernspredienst zunächst zwischen Berlin und Leipzig, später auch nach München (Reuter 1990, S. 210 f.). 1964 wurde dann auf der Weltausstellung in New York ein Videokonferenzsystem von AT&T vorgestellt („Picturephone“), das in der Folge vermarktet wurde – zunächst allerdings ohne großen Erfolg (Hart et al. 1995, S. 396). Während mit jeder neuen Generation von Videokonferenz-Systemen der endgültige Durchbruch dieser Technologie prophezeit wurde, ist die tatsächliche Nutzung bislang stets hinter den optimistischen Voraussagen zurückgeblieben (z. B. Stoetzer 1994, S. 16). Dies hat wohl nicht nur mit technischen, finanziellen und organisatorischen Aspekten der Einführung solcher Systeme zu tun oder mit dem Erreichen von kritischen Schwellenwerten bei Nutzung und Verbreitung. Es liegt vermutlich auch daran, dass Videokonferenzen eine neue Kommunikationsform darstellen, eigene Merkmale und eine besondere Dynamik aufweisen und damit nicht für alle Besprechungsaktivitäten gleichermaßen geeignet sind. Worin diese Merkmale bestehen und wie die Interaktionsdynamik in Videokonferenzen zu charakterisieren ist, das ist Gegenstand dieses Beitrags.

2. Untersuchungen zu video-vermittelter Kommunikation und Interaktion

Die vorliegende Literatur zu video-vermittelter Kommunikation und Interaktion (vvK) entstammt verschiedensten disziplinären Zusammenhängen und ist entsprechend heterogen in bezug auf methodische Ansätze und Fragestellungen. So fragen etwa Untersuchungen aus dem Forschungsbereich Computer-Supported Cooperative Work (CSCW) danach, wie gemeinsames Arbeiten technisch unterstützt werden kann und wie wichtig dafür jeweils der Audio- und der Videokanal sind (etwa Olson et al. 1995). Demgegenüber interessieren sich Organisationsforscher mehr für die Folgen von Videokonferenzen auf organisationsinterne Prozesse der Informationsverarbeitung und Entscheidungsfindung (zum Beispiel Bronner 1996). Sozialpsychologen stellen die Auswirkungen von technisch vermittelten Kommunikationsprozessen auf die Bewältigung von kooperativen oder konfliktbehafteten Aufgaben, auf die interpersonelle Wahrnehmung und die sich ergebende Gruppendynamik in den Mittelpunkt (klassisch hierzu Short et al. 1976). Kommunikationswissenschaftler schließlich fragen unter anderem danach, worin die Besonderheit video-vermittelter Kommunikationsprozesse zu suchen ist (etwa Weinig 1997). Im folgenden werde ich einige Ergebnisse dieser Untersuchungen anführen, da sich darüber die Relevanz des hier gewählten Ansatzpunkts deutlicher herausstellen lässt.

In bezug auf den Interaktionsverlauf wird berichtet, dass in Videokonferenzen – verglichen mit face-to-face (f-t-f) Besprechungen – Phasen der Gesprächsorganisation und des Gesprächsabschlusses länger ausgedehnt sind, während Diskussionsphasen eher kürzer sind (z. B. Barefoot/Strickland 1982, S. 563; Olson et al. 1995, S. 366; Weinig 1997, S. 159). Dazu passt auch die Beobachtung, dass es in video-vermittelter Interaktion schwieriger ist, das Geschehen zu steuern (Sellen 1992, S. 57). Ergebnisse verschiedener Untersuchungen deuten darüber hinaus darauf hin, dass in Videokonferenzen eine Polarisierung und Blockbildung zwischen „uns hier“ und „denen dort“ zu beobachten ist (z. B. Short et al. 1976, S. 137; Weinig 1997, S. 140 und 166).

Während die eben genannten Ergebnisse weitgehend im Einklang miteinander stehen, ist dies in bezug auf andere Aspekte video-vermittelter Kommunikation und Interaktion nicht der Fall. So wird etwa in bezug auf Beteiligungschancen angeführt, dass Rollendifferenzierungen (und damit auch Führungsrollen) in Videokonferenzen weniger ausgeprägt sind und dass sich eine etwas ausgeglichene Beteiligung als bei face-to-face Besprechungen findet (Williams 1977, S. 970; Weinig 1997, S. 140 und 167). Demgegenüber konnten Barefoot/Strickland (1982, S. 563) in bezug auf die Diskussionsanteile von Vorgesetzten keine signifikanten Unterschiede ausmachen. Und hinsichtlich der Qualität der Ergebnisse gemeinsamen

Arbeitens zeigen sich sogar krasse Widersprüche in den Aussagen verschiedener Untersuchungen:

Problemlösende Gruppen Gruppen erzielen in f-t-f Interaktion bessere Ergebnisse als in vvK.	In bezug auf die Qualität der Arbeitsergebnisse zeigen sich keine signifikanten unterschiede zwischen f-t-f Interaktion und vvK.	Problemlösende Gruppen er- zielen in vvK bessere Ergebnisse als in f-t-f Interaktion.
(Cohen 1984, S. 297; Bronner 1996, S. 111)	(Olson et al. 1995, S. 365)	(Rosetti/Surynt 1985, S. 28 f.; Weinig 1997, S. 159)

Tabelle 1: Aussagen zu Arbeitsergebnissen in Videokonferenzen

Aber nicht nur bei solchen höher aggregierten Merkmalen video-vermittelter Interaktion wie der Ausgeglichenheit der Beteiligung oder dem Arbeitsergebnis finden sich inkompatible Ergebnisse. Ähnliches gilt auch in bezug auf die Details körperlichen und sprachlichen Handelns. Zwar konvergieren die Ergebnisse in bezug auf nicht-sprachliche Aspekte (etwa Gesten, Haltungswechsel, Kopfbewegungen oder mimische Ausdrücke) dahin, dass diese in Videokonferenzen einerseits an Häufigkeit zunehmen (Weinig 1997, S. 158), gleichzeitig aber viel von ihrer handlungskoordnierenden Kraft einbüßen (Heath/Luff 1993, S. 45 f.). Uneinheitlich bis widersprüchlich sind dagegen die Ergebnisse in bezug auf Zahl, Länge und Verteilung von Äußerungen:

In f-t-f Interaktionen fin- den sich fast zwei Mal so viele Sprecherwechsel wie in vvK.	Es findet sich kein Unter- schied zwischen f-t-f Inter- aktionen und vvK in bezug auf die Zahl und Länge von Redezügen.	In vvK zeigt sich eine größere Zahl von Äußerungen, die zu- gleich kürzer sind als dies bei einer f-t-f Interaktion der Fall ist.
(Cohen 1984, S. 296)	(Sellen 1992, S. 56)	(Weinig 1997, S. 158)

Tabelle 2: Aussagen zum sprachlichen Austausch in Videokonferenzen

Auch in bezug auf die Frage, wie zufrieden die Nutzer von Videokonferenzsystemen mit dem Arbeits- und Interaktionsprozess sind, sind die Ergebnisse problematisch. Einerseits zeigen sich die Nutzer zufriedener, wenn sie sich zusätzlich zu einer Tonverbindung (Audiokonferenz) auch noch sehen können (Videokonferenz) und sie bewerten die Zusammenarbeit im gleichen Raum (d. h. face-to-face) am zufriedenstellendsten. Andererseits finden sich aber keine signifikanten Unterschiede in bezug auf den Arbeits- und Interaktionsprozess zwischen diesen drei verschiedenen Bedingungen (Olson et al. 1995, S. 366; ähnlich auch Isaacs/Tang, 1997, S. 195).

Ich sehe vor allem zwei Ursachen für diese widersprüchlichen Ergebnisse. Zum einen besteht ein Problem der Forschung zu video-vermittelter Interaktion darin, dass sich die untersuchten Gruppen (ad-hoc gebildete Laborgruppen oder reale Arbeitsgruppen) und Aufgabenstellungen (meist Rollenspiele) ebenso wie die untersuchten Videokonferenzsysteme (dedizierte Studios, Rollabout- oder Desktop-Systeme mit unterschiedlichen technischen Spezifikationen) unterscheiden, diese Unterschiede aber nicht systematisch im Blick gehalten werden (vgl. Olson/Olson 1997). Zum anderen scheinen diese Untersuchungen entweder auf einem zu hohem Niveau der Aggregation (Arbeitsergebnisse) oder auf zu niedrigem Niveau (Verteilung sprachlicher Äußerungen) anzusetzen. Die Besonderheit von Videokonferenzen liegt aber vielleicht eher auf einem mittleren Niveau und betrifft Beteiligungskonstellationen und Beteiligungschancen einerseits sowie die Realisierung bestimmter Aktivitäten andererseits.²

3. Die Fallstudie „Videokonferenzen bei Technics“

Das Customer Services Council (CSC) ist eine Besprechungsrunde, deren Mitglieder den Bereich Dienstleistungen innerhalb von Technics-Deutschland³ repräsentieren. Die Runde, der rund ein Dutzend Personen angehören, tagt regelmäßig und soll der Koordinierung der beteiligten Manager dienen, von denen einige für bestimmte Produktlinien zuständig sind, andere für interne Dienstleistungen wie Marketing, Controlling oder Personal. Bis Januar 1996 trafen sich die Mitglieder alle 14 Tage entweder in Karlsruhe oder in Wesseling. Schließlich wurde beschlossen, die Besprechungen als Videokonferenzen durchzuführen. Diese Konferenzen beginnen meist um 8:30 Uhr und dauern – inklusive Pausen – vier bis fünf Stunden. Sie beginnen in der Regel mit dem Durchsprechen des Protokolls der letzten Sitzung. Ein weiterer fester Tagesordnungspunkt ist die Diskussion der Auftrags- und Ertragslage des Unternehmensbereichs. Darüber hinaus berichten Mitglieder des CSC oder Gäste zu verschiedenen Punkten und schließlich wird über eine ganze Reihe von Fragen, die von Arbeitszeitmodellen bis hin zu Werbeaufträgen reichen, diskutiert und entschieden.

Bei Beginn der Fallstudie war die Videokonferenzanlage⁴ bereits seit ei-

² Möglicherweise schlagen sich diese Besonderheiten nicht deutlich wahrnehmbar in den jeweiligen Arbeitsergebnissen einer Sitzung nieder. Allerdings ist anzunehmen, dass sie längerfristige Auswirkungen auf Arbeitsbeziehungen, wechselseitiges Vertrauen und Kollegialität haben und damit indirekt auch auf die Effizienz kollaborativen Arbeitens innerhalb von Teams und Organisationen.

³ Alle in diesem Text verwendeten Namen sind Pseudonyme.

⁴ Eine PictureTEL Concorde 4510 mit elektronischem Tonmodul (Voll-Duplex, Echo-Cancelling, Rauschunterdrückung). Die Verbindung zwischen den beiden Gegenstellen besteht aus zwei ISDN-B-Kanälen mit einer gebündelten Band-

nem halben Jahr im Einsatz. Nach einer dreimonatigen Probephase, in deren Verlauf mehrere Videokonferenzen teilnehmend beobachtet wurden, stimmten die Beteiligten auch einer audiovisuellen Dokumentation zu. Im Verlauf der nächsten fünf Monate wurden dann vier Videokonferenzen und eine face-to-face Besprechung aufgezeichnet und einige informelle Gespräche mit verschiedenen Beteiligten geführt.⁵

Abbildung 1: Der Videokonferenzraum in Wesseling

Eine audiovisuelle Dokumentation muss – sofern sie eine tragfähige Grundlage für eine detaillierte Analyse des interaktiven Geschehens sein soll – möglichst alle die Aufmerksamkeitsfoki erfassen, die für das interessierende Geschehen und die daran beteiligten Personen relevant sind. Im Fall von Videokonferenzen müssen also nicht nur alle beteiligten Personen im Bild erfasst werden, sondern beispielsweise auch Präsentationsfolien. Gleichzeitig ist es wichtig, im Feld minimalinvasiv vorzugehen, um eine Transformation der Situation zu vermeiden. Videokonferenzen sind in dieser Hinsicht ein besonders glücklicher Untersuchungsgegenstand, da hierbei – unabhängig von einer wissenschaftlichen Untersuchung – eine audiovisuelle Aufzeichnung des Handelns der Beteiligten über Kamera und Mikrophon vorgenommen wird. Damit bietet sich auch die Möglichkeit, die in der Situation selbst erzeugten Prozessdaten für eine wissenschaftliche Dokumentation anzuzapfen, etwa indem Verbindungskabel an

breite von 128kb/Sekunde. Die Signalverzögerung zur Gegenstelle beträgt ungefähr 0,25 Sekunden in jede Richtung.

⁵ Eine Diskussion der mit diesem Vorgehen verbundenen methodischen und technischen Fragen findet sich in Meier (1998). Die Besprechungen wurden – abgesehen vom Tagesordnungspunkt „Auftrags-/Ertragslage“ – vollständig aufgezeichnet.

Abbildung 2: Die Struktur der audiovisuellen Dokumentation

die Fernsehgeräte angeschlossen werden, die Bestandteil des Videokonferenzsystems sind (vgl. Abbildung 2, unten und Meier 1998, S. 270 ff.). Die im Untersuchungsfeld selbst generierten Prozessdaten wurden über einen Bildteiler an einen Videorekorder geleitet und dort in den beiden Quadranten (1) und (2) aufgezeichnet (vgl. Abbildung 2). Da es für die Rekonstruktion des interaktiven Geschehens an den beiden Standorten und seines Zusammenhangs notwendig ist, beide Teilgruppen durchgängig im Blick zu halten, wurde oberhalb der Fernsehgeräte in Wesseling eine weitere Kamera positioniert. Das Bildsignal dieser zusätzlichen Kamera wurde ebenfalls durch den Bildteiler geschickt und im Quadranten unten links (3) aufgezeichnet.

Wenn ich in den folgenden Abschnitten einzelne Ausschnitte aus den audiovisuellen Dokumentationen vorstelle, dann sehen diese Ausschnitte

anders aus, als aufgrund von Abbildung 2 zu erwarten wäre. Der Grund dafür liegt darin, dass die Materialien für Präsentationszwecke digitalisiert und beschnitten wurden. Sofern im Quadranten 2 keine Folie eingeblendet ist, werden die relevanten Bereiche der Quadranten (1) und (2) gezeigt, anderenfalls die relevanten Bereiche der Quadranten (1) und (3) (vgl. Abbildung 3).

Abbildung 3: Die für die digitalen Videoclips verwendeten Ausschnitte

4. Beobachtungen zur Interaktion in Videokonferenzen

4.1 Vorbemerkungen zur Analysemethodik

In den folgenden Abschnitten werden kurze Szenen aus dem vorliegenden Material vorgestellt und Beobachtungen daran ausführt. In den meisten dieser Szenen steht die Organisation der Besprechung im Mittelpunkt, etwa wenn darüber gesprochen wird, wer Protokoll führt, ob ein Bericht auf eine spätere Sitzung verschoben werden soll oder ob jetzt eine Pause gemacht wird. Es gibt gute Gründe, zunächst solche Episoden zu analysieren, in denen nicht inhaltlich über ein auf der Tagesordnung stehendes Thema gesprochen wird. So kann die inhaltliche Diskussion eines Tagesordnungspunkts eine besondere Beteiligungskonstellationen implizieren, die in den organisatorischen Arrangements, nicht aber in der Situationslogik des Er-

eignisses Videokonferenz begründet ist. Dies ist etwa dann der Fall, wenn das Thema den Verantwortungsbereich eines bestimmten Teilnehmers tangiert, nicht aber den einer anderen Person. Natürlich ist die Beobachtung, dass bei einer Äußerung zum Thema X zwar Person A, nicht aber Person B zu einer Entgegnung ansetzt, unabhängig von ethnographischem Kontextwissen möglich. Aber die möglichen Implikationen für die Beteiligten selbst („B kennt sich nicht aus“ oder „Warum sagt B denn nichts?“) sind dem externen Beobachter nicht zugänglich. Dies kann insofern zu einem analytischen Problem führen, als diese Implikationen in nächsten Handlungen der Beteiligten bearbeitet werden können („Ich gebe B eine Vorlage, damit auch er was dazu sagt“). Es besteht also die Gefahr, dass der Analytiker die (Nicht-)Beteiligung von bestimmten Personen auf Merkmale der Videokonferenzsituation zurückführt (etwa die räumliche Konfiguration von „hier“ und „dort“), wohingegen den Beteiligten selbst eine solche (Nicht-)Beteiligung über bekannte Zuständigkeiten und Verantwortlichkeiten verständlich und plausibel wird. Demgegenüber ist umfangreiches ethnographisches Wissen bei der Analyse von Episoden, in denen es um die Organisation des Ereignisses selbst und dessen weiteren Verlauf geht, weniger wichtig. Zwar gibt es auch in bezug auf die Organisation der Besprechung selbst unterschiedliche Grade an Zuständigkeit unter den Beteiligten. So ist offensichtlich, dass Johannes, der Leiter des CSC, zugleich als Moderator in diesen Besprechungen auftritt. Aber Entscheidungen über den weiteren Verlauf der Besprechung gehen tendenziell alle Teilnehmer (qua ihres Teilnehmerstatus) etwas an. Daher kann in diesen Episoden die ausgeglichene Beteiligung erwartet werden und sie eignen sich dementsprechend in besonderer Weise für die Analyse der Beteiligungsstrukturen und -dynamiken, die für Videokonferenzen spezifisch sind. Wenn es überhaupt Momente gibt, in denen das Interaktionssystem als solches sichtbar wird, dann in diesen gesprächsorganisatorischen Phasen.⁶

Im Verlauf der Arbeiten an den Materialien erwies sich ein Vorgehen entsprechend dem klassischen konversationsanalytischen Zugriff mit einer Fokussierung auf die Systematik der Organisation der Redezugverteilung (vgl. Sacks et al. 1974; Schegloff 1987) als wenig ertragreich. Fruchtbarer war dagegen die Beschäftigung mit den von Goffman (1981) in den Blick genommenen und als „participation frameworks“ eingeführten Beteiligungsphänomenen. Darüber hinaus wurde deutlich, dass sich bei der Arbeit mit komplexen audiovisuellen Materialien die Rolle von Transkriptionen als einer grundlegenden Technologie für die Analyse von Sprechbeim-Interagieren (z. B. Edwards/Lampert 1993) wandelt. Dies nicht nur, weil diese Transkriptionen sehr aufwendig und unübersichtlich werden,

⁶ Dies gilt übrigens auch für Arbeitsbesprechungen, die face-to-face stattfinden.

sondern auch deshalb, weil sie eine wechselseitige Verfügbarkeit von (Äußerungs-)Handlungen unterstellen, die in Videokonferenzen nicht unbedingt gegeben ist. Im Verlauf der analytischen Arbeit wurde es nämlich notwendig, von einer naiven Unterstellung unmittelbarer und wechselseitiger Verfügbarkeit des Handelns an den beiden Standorten Abstand zu nehmen. Wieviel und welche Aspekte des lokalen Handelns für die Beteiligten an der Gegenstelle verfügbar ist, bleibt nicht nur für den Analytiker, sondern häufig auch für die beteiligten Personen selbst unklar.

4.2 Episode 1: Wer macht das Protokoll?

Die folgende Szene (vgl. dazu den Videoclip 1: „sc4_too_late“) ereignet sich zu Beginn einer der Videokonferenzen. Johannes hat kurz nach seinem Eintreffen im Konferenzraum in Karlsruhe die Tonverbindung zur Gegenstelle nach Wesseling (wo bereits alle Teilnehmer versammelt waren) aktiviert und schließlich dort auch Aufmerksamkeit für seine Eröffnung und Begrüßung erlangt. Er kündigt an, dass er die Tagesordnung gerne umstellen möchte, damit ein Gast (Otmar), der einen weiteren Termin hat, seinen Bericht gleich zu Beginn liefern kann. Otmar ist gerade dabei, seinen Bericht einzuleiten, als Jürgen fragt, wer denn an diesem Tag mit Protokollführen dran ist und an dieser Stelle setzt auch der Clip ein (Abbildung 4 zeigt die Beteiligten zu Beginn der Szene).

Abbildung 4

Johannes wendet sich zum Videokonferenzsystem (und mithin zu den Kollegen in Wesseling) und fragt: „wer macht protokoll heute“. Aus Wesseling erfolgt allerdings keine deutliche Stellungnahme. Michael R. äußert sich ganz leise (und vermutlich für die Kollegen in Karlsruhe nicht hörbar) dahingehend, dass er das letzte Protokoll gemacht habe. Die anderen Teilnehmer äußern sich gar nicht und packen weiterhin ihre Unterlagen aus (Michael D. und Holger), schauen in die vor ihnen ausgebreiteten Unterlagen bzw. Notebooks (Michael R.), schauen unbeteiligt vor sich hin (Frank-Ernst) oder schauen zu den Fernsehgeräten (Ernst, Holger). Demgegenüber stellt Jürgen in Karlsruhe die Frage, wer das letzte Protokoll geschrieben hat und beteiligt sich damit aktiv an der Lösung des von ihm

ursprünglich aufgeworfenen Problems. Während Johannes in seinen Unterlagen blättert, schaut sich Hartmut in Wesseling um, macht sich plötzlich klein, lehnt sich etwas zurück, beugt sich zu seinem Nachbarn Holger und sagt leise „glaub ich hab noch NIE protokoll geschrieben“. Mindestens zwei weitere Kollegen in Wesseling zeigen, dass sie dies registrieren, indem sie sich zu Hartmut und Holger umwenden (Frank-

Abbildung 5

Ernst) und leise Kichern (Michael R. – vgl. Abbildung 5). Drei weitere Teilnehmer in Wesseling (Michael D., Ernst und der Beobachter) zeigen dagegen keine Anteilnahme an diesem Austausch.

Johannes äußert dann, dass Michael R. das letzte Protokoll gemacht habe und daraufhin sagt Michael D. „R S“, womit er seinen Sitznachbarn Hartmut als potentiellen Kandidaten ins Spiel bringt, da dieser mit Nachnamen Streck heißt. Hartmut wendet sich zu Michael D. und sagt vernehmlich „shit“, woraufhin zunächst Michael D. und Hartmut laut zu lachen beginnen und nach und nach alle anderen Teilnehmer in Wesseling einstimmen. Die KollegInnen in Karlsruhe bekommen diesen Lachausbruch offensichtlich mit, zeigen sich zugleich aber durch ihre langen Blicke zu den Fernsehgeräten in einer Weise desorientiert und verwundert, die Hartmut zum Anlass für eine Erläuterung nimmt (vgl. Abbildung 6).

Abbildung 6

Während Hartmut eine Erläuterung für die KollegInnen in Karlsruhe liefert, quittieren dies einige der Kollegen in Wesseling mit weiterem Lachen. Dabei zeigt sich auch, dass weder das Lachen noch die Erläuterung von den Beteiligten in Karlsruhe erkennbar oder hörbar aufgegriffen werden.

4.2.1 Raumökologie und Positionierung

Eine erste Beobachtung, die sich anhand dieser Episode anstellen lässt, betrifft die Positionierung der Personen im Raum. Es fällt auf, dass sich der Besprechungsleiter Johannes nicht am symbolisch aufgeladenen und in herkömmlichen Besprechungen interaktionslogisch vorteilhaften (dazu Meier 1997, S. 157) Platz am Kopfende der Runde in Karlsruhe positioniert (d.h. dort, wo in Abbildung 5 die Kollegin Carla Platz nimmt). Vielmehr sucht er sich seinen Platz an der Seite, ganz in der Nähe des Videokonferenzsystems. Dies ist auch in anderen Besprechungen zu beobachten (vgl. die Abbildungen 7 und 8).

Abbildung 7

Abbildung 8

Mars (1994, S. 102 f.) berichtet in seiner Feldstudie zu Videokonferenzen ebenfalls über eine solche Positionierung der „wichtigen“ Teilnehmer. Damit stellt sich die Frage, ob diese Praxis als Lösung für ein Strukturproblem der Interaktionsorganisation in Videokonferenzen verstanden werden kann.

Indem Johannes sich an die Seite und damit in die Nähe des Videokonferenzsystems setzt, positioniert er sich an einem Ort, an dem er für die Teilnehmer an der Gegenstelle groß im Bild und damit vergleichsweise gut wahrnehmbar ist – im Unterschied etwa zu den KollegInnen, die am Kopfende plazierte sind.⁷ Gleichzeitig positioniert er sich damit an zentra-

⁷ Mantei et al. (1991, S. 205) weisen darauf hin, dass die Größe der Darstellung die Präsenz und den Einfluss einer Person in video-vermittelter Interaktion beeinflusst.

ler Stelle. Wer immer zu den KollegInnen am anderen Standort schaut, hat zugleich auch den Besprechungsleiter mehr oder weniger deutlich im Blickfeld. Und schließlich, drittens, werden durch diese Positionierung die Orientierungswechsel von Johannes besonders auffällig, weil er sich fast um 180° drehen muss wenn er sich zwischen den Beteiligten in Wesseling (d. h. den Fernsehgeräten) und den KollegInnen in Karlsruhe hin- und herwendet. Diese Positionierung kann damit als eine systematische Lösung für das in der Literatur immer wieder angeführte Problem gelten, dass es in einer Videokonferenz schwieriger ist als in herkömmlichen Besprechungen, das Geschehen zu steuern.

4.2.2 Beteiligungskonstellationen

Weitere Beobachtungen betreffen die Beteiligungskonstellationen (Goffman), die während des Geschehens realisiert werden. Im Verlauf einer Interaktion sind ja nicht immer alle beteiligten Personen gleichermaßen eingebunden und gefordert. Insbesondere über eine Blickzuwendung oder eine vorangestellte Adressierungsform („Peter, was hältst Du von den Zahlen?“) ist es den Beteiligten möglich, ganz gezielt bestimmte Personen anzusprechen. Auf diese Weise können sie deutlich machen, dass eine Handlung (beispielsweise eine Stellungnahme zu einer Frage oder zu einem Vorschlag) von dieser und nicht von einer anderen Person erwartet wird (Goffman, 1981, S. 133).

In der betrachteten Szene finden sich zum einen Stellen, in denen Sitznachbarn für kurze Zeit in einen dyadischen Austausch miteinander eintreten. So etwa, als sich Hartmut zu Holger beugt. Dann finden sich Momente, in denen einige der Teilnehmer an einem Standort einen gemeinsamen Fokus etablieren, andere dagegen nicht daran beteiligt sind. Dies war zu beobachten, als sich Frank-Ernst dem Austausch zwischen Hartmut und Holger zuwendete und Michael R. diesen leise kichernd kommentierte. Es finden sich Momente, in denen alle Teilnehmer auf einer Seite einen gemeinsamen Fokus aufrechterhalten oder in eine gemeinsame Aktivität eingebunden sind. Etwa nach Hartmuts drastischem Kommentar, als die Beteiligten in Wesseling gemeinsam Lachen. Schließlich gibt es Momente, in denen alle Beteiligten an beiden Standorten ihre Aufmerksamkeit gemeinsam auf einen einzigen Fokus ausrichten. In der gezeigten Szene war das beispielsweise zu beobachten, als Hartmut seine Erläuterung für die KollegInnen in Karlsruhe lieferte.

Für das Etablieren und Aufrechterhalten dieser Beteiligungskonstellationen spielt die Lautstärke, mit der gesprochen wird, eine wichtige Rolle. Unabhängig vom Inhalt einer Äußerung können die Beteiligten über die Lautstärke ihrer Äußerungen andere KollegInnen ein- oder ausschließen, und zwar sowohl KollegInnen am eigenen Standort, als auch KollegInnen am entfernten Standort (vgl. Abbildung 9).

Abbildung 9: Ebenen der Inklusion und Exklusion aufgrund der Lautstärke einer Äußerung

Wenn wir uns mit anderen Personen im gleichen Raum befinden, haben wir meist ein gutes Gefühl dafür, wie laut wir sprechen müssen, damit uns andere (nicht) verstehen können. In einer Videokonferenzsituation kommt für die Beteiligten erschwerend hinzu, dass der aktuelle Regelungszustand des elektronisch gesteuerten Mikrophons (d.h. welche der sich überlagernden und in verschiedene Richtungen weisenden Keulen des Mikrophons gerade aktiviert ist und Tonsignale verstärkt) nicht zu erkennen ist. So kommt es wiederholt zu Situationen, in denen die Beteiligten davon ausgehen, dass sie auf der anderen Seite (nicht) verstanden werden, sich dann jedoch das Gegenteil herausstellt.

4.2.3 Das Geschehen am anderen Standort ist nur eingeschränkt transparent

Anhand von Clip 1 lassen sich aber noch weitere Beobachtungen anstellen. Ich hatte bereits darauf hingewiesen, dass das Geschehen in Wesseling und die Genese des lauten Lachens für die Beteiligten in Karlsruhe anscheinend nicht nachzuvollziehen sind. Darüber hinaus ist zu sehen, dass die Kollegin Carla, die zuvor in Karlsruhe Platz genommen hat, kurz nach der Feststellung von Johannes („letzte protokoll hat der michael rimbach gemacht“) zu sprechen beginnt. Es ist nicht zu verstehen, was sie sagt, aber sie spricht auch in der Folge, während die Beteiligten in Wesseling laut lachen, erkennbar zu Johannes. Hier wird ganz deutlich, dass es sich zwar um ein Zwei-Wege Tonsystem handelt, gleichzeitig aber leisere Äußerungen von der Gegenseite entweder gar nicht oder nur sehr unvollständig zu hören sind. Damit wird aber ein Teil des Geschehens am anderen Standort intransparent.

4.2.4 Gemeinsame Aktivität

Schließlich zeigt sich in dieser Episode, dass das Sich-Amüsieren als gemeinsame Aktivität nicht gelingt. Während die Beteiligten in Wesseling laut lachen, zeigen sich die KollegInnen in Karlsruhe zunächst desorientiert. In der Folge, während Hartmut seine Erläuterung liefert und einzelne Teilnehmer in Wesseling diese mit weiterem Kichern quittieren, ist zwar zu sehen, dass Johannes grinst. Aber darüber hinaus findet sich unter den Beteiligten in Karlsruhe kein hörbares Aufgreifen des Amüsements. Lachen ist aber eine Aktivität, die typischerweise gemeinsam vollzogen wird und damit eine hochgradig koordinierte Beteiligung mehrerer Personen erfordert (vgl. Jefferson et al. 1987). Dies scheint in einer Gruppen-Videokonferenz nur schwer möglich zu sein und in Clip 1 wie auch in anderen Episoden (vgl. Meier im Erscheinen a) ist zu sehen, dass das Sich-Amüsieren eher unter den Beteiligten an einem Ort stattfindet als gemeinsam zwischen allen Beteiligten an beiden Standorten.

4.3 Episode 2: Einen Bericht verschieben?

Die nächste Episode (Clip 2: „csc3_gibts_termine_kurz“) stammt aus einer anderen Besprechung. Kurz zuvor ist eine Präsentation von Vertretern einer Werbeagentur zu Ende gegangen. Den Beteiligten stellt sich jetzt die Frage, wie sie angesichts der fortgeschrittenen Zeit mit dem als nächsten Tagesordnungspunkt vorgesehenen Bericht von Holger umgehen wollen (Abbildung 10 zeigt die Beteiligten zu Beginn dieser Szene).

Abbildung 10

Johannes fragt, indem er sich zu den Kollegen nach Wesseling orientiert, ob man den Bericht vertagen oder ob man noch eine halbe Stunde dafür spendieren solle. Michael D. wendet daraufhin ein, dass der Bericht schon einmal vertagt wurde. Holger zeigt sich leicht verschnupft, indem er zunächst nach der Zielgruppe für diesen Bericht fragt und dann anmerkt, dass er den Bericht ja dem abwesenden Kollegen Michael R. auch allein vorstellen könne. Johannes weist diese Möglichkeit zurück und macht deutlich, dass alle Mitglieder des CSC mitbekommen sollen, wie der Stand der Dinge bei diesem Projekt ist. In der Folge entwickelt sich eine kontroverse Diskussion in Wesseling. Frank-Ernst spricht sich dafür aus, das Thema noch durchzuziehen. Hartmut ist dagegen und plädiert angesichts der Abwesenheit zweier Kollegen dafür, den Bericht zu verschieben. Daraufhin entgegnet Holger, ob es denn überhaupt Termine gebe, wo man vollständig sei und Michael D. weist nochmals darauf hin, dass man den Bericht schon einmal verschoben habe. Während die Beiträge von Frank-Ernst und zunächst auch Hartmut mit einer deutlichen Orientierung zu den KollegInnen in Karlsruhe gesprochen werden, orientieren sich die Teilnehmer in Wesseling im weiteren Verlauf dieser Diskussion deutlicher zueinander (vgl. Abbildung 11).

Abbildung 11

Abbildung 12

Hartmut wechselt in der Folge zwischen einer Orientierung zu den Kollegen in Wesseling und einer Orientierung zu den Fernsehgeräten und macht deutlich, dass es ihm egal sei, ob er jetzt oder später davon erfahre, dass er aber gern über das Thema diskutieren wolle. Fast gleichzeitig mit dem

Beginn von Hartmuts Äußerung macht Johannes einen ersten Anlauf, in die Diskussion einzugreifen (vgl. seine große Geste in Abbildung 12, unten). Dann bricht er allerdings ab und unternimmt kurz darauf einen zweiten Anlauf.

Mit seiner Intervention bemüht sich Johannes um eine Refokussierung der Diskussion und reaktualisiert zugleich seine Frage. Wenn Holger bereit ist, dem abwesenden Michael R. noch einmal separat zu berichten, dann könne man den Bericht in dieser Runde jetzt trotz dessen Abwesenheit hören. Im Verlauf der Intervention von Johannes macht Hartmut deutlich, dass sich für ihn damit die Voraussetzungen seiner Stellungnahme geändert haben („ach NOCHma“). Er produziert darüber hinaus ein „okeh“ und eine Geste die so aussieht, als würde er hier etwas zurücknehmen. In der Folge von Johannes' deutlich akzentuierter Klausel („WENN wer noch die halbe stunde zeit habn“) findet sich allerdings nur ein leise geäußertes Zustimmungszeichen von Michael D. und eine weitere Äußerung von Hartmut, mit der dieser herausstellt, dass dies jetzt andere Voraussetzungen sind („(das is) was anderes“). Im Anschluss daran entsteht ein Moment der Desorientierung. Johannes ist nach wie vor zu den Kollegen in Wesseling orientiert und zeigt damit, dass er auf eine Stellungnahme von dort wartet. Demgegenüber verhalten sich die Kollegen in Wesseling nicht so, also ob weitere Stellungnahmen von ihnen erwartet werden könnten. Schließlich wendet sich Johannes von den Fernsehgeräten ab und den KollegInnen in Karlsruhe zu, wobei er – während Hartmut dann doch noch eine Äußerung beginnt – seine Frage ein weiteres Mal stellt („mach wer die [d. h. die halbe Stunde, C. M.] jetzt noch“).

Unmittelbar nachdem sich in Karlsruhe Zustimmung dafür abzuzeichnen beginnt, den Bericht jetzt zu hören, äußert sich Holger dahingehend, dass er seinen Bericht auch in einer Viertelstunde schaffen könne. Er greift damit möglicherweise die kurz zuvor von Richard in Karlsruhe angesprochene Verschiebung weiterer Termine („wenn wer statt um eins um viertel zwei anfangen“) auf und präsentiert sich damit zugleich als kooperativ und kollegial. Über dieses Angebot gerät nun offensichtlich Johannes in Karlsruhe aus dem Takt. Er beginnt eine Äußerung, mit der er die Formulierung von Holger aufzugreifen scheint („(du kriegst“) und in deren Verlauf er seine Hände in eine neue Haltung bringt. Dann bricht er seine Äußerung ab und friert die Geste ein. Nach einer Schweigephase produziert er ein knappes „ne“, macht mit seinen Händen eine Bewegung, die wie ein Neu-Ansetzen oder Anlauf-Nehmen wirkt und äußert schließlich die Meinung, dass der Bericht jetzt noch durchgezogen werden sollte. Holger seinerseits hatte mit Beginn seines Angebots die Hände über dem Kopf verschränkt. In der Folge von Johannes' „vielen dank“, mit dem dieser die Diskussion abschließt, löst er diese Haltung wieder auf, wobei die Art und Weise wie er dies tut, Enttäuschung oder Unzufriedenheit vermittelt.

Ich werde jetzt wiederum einige Beobachtungen, die sich anhand dieser Episode machen lassen, ausführen. Dabei werde ich zunächst auf die mit digitalen Videokonferenzsystemen verbundene Zeitverzögerung und die daraus resultierenden Probleme der Handlungskoordination zu sprechen kommen, dann auf die Trägheit des einmal debalanzierenden Interaktionssystems und schließlich auf die Frage, ob die sequenzielle Organisation des Handelns in Videokonferenzen gelockert ist.

4.3.1 Zeitverzögerung und beeinträchtigte Handlungskoordination

Den ersten Teil der Episode will ich anhand von zwei Transkriptausschnitten etwas genauer betrachten. Der Transkriptausschnitt <1> zeigt eine Passage aus der kontroversen Diskussion unter den Teilnehmern in Wesseling. Die Schnelligkeit und Kompaktheit der Interaktion in dieser Phase wird anhand der wiederholten Überlappungen beim Übergang von einem Sprecher zum nächsten deutlich (vgl. Zeilen 104–105, 117–118 und 125–126).

<1> „gibts termine“, Ausschnitt, vereinfacht⁸

	103	Frank-Ernst	hh ich müßt mich dann (.) spätestens
->	104		in ner halben stun[de]ausble[ndn;]
->	105	Hartmut	[<<p>ne'>] [aber]
	106	Hartmut	(.) aber von dem (.) vo(r) dem hintergrund
			.
			.
	115	Hartmut	weil heute sind der Michael und der Arnulf
	116		nich dabei; = un ich mein wenn DIE des nich
->	117		(- -) [MIT]kriegn =
->	118	Holger	[gi'] = gibts termine wo wir
			.
			.
	122	Michael D.	wir[hams schonMAI verscho]bn. ne, =
	123	Hartmut	[<<p>(das schon eher)>]
	124	Holger	= <<p> hm.;>
->	125	Michael D.	AUCH aus dem [(grund)()]
->	126	Hartmut	[ja (-) ich sach das] auch

⁸ Die verwendete Transkriptionsnotation orientiert sich an Selting et al. (1998). Bei durchgängiger Kleinschreibung repräsentieren Großbuchstaben Akzentuierungen. Überlappungen sind durch eckige Klammern markiert, Interpunktionszeichen repräsentieren Intonationsverläufe (z.B. „“ stark fallender Intonationsverlauf; „“ fallend-steigender Intonationsverlauf). Kommentierungen wie „<<all>>“ oder „<<p>>“ stehen für schnelleres bzw. leiseres Sprechen und markieren zugleich die Reichweite dieser Veränderungen. Die Rahmenmarkierungen am linken Rand zeigen an, welche Zeilen aufgrund von Überlappungen oder schnellen Anschlüssen zusammengelesen werden müssen. Äußerungen aus Karlsruhe sind schattiert dargestellt (vgl. Ausschnitt 2).

Diese schnellen Anschlüsse tragen dazu bei, dass es für Teilnehmer „dort“ schwierig wird, sich in die Diskussion „hier“ einzubringen. Verschärft wird diese Schwierigkeit noch dadurch, dass bei ISDN-basierten Videokonferenzsystemen aufgrund der digital-analog Wandlungen und der Signallaufzeit zwischen den Standorten eine Verzögerung von etwa 0,25 Sekunden je Richtung auftritt. Das heißt, ein Geschehen „hier“ wird erst mit einer Verzögerung von 0,25 Sekunden „dort“ sicht- und hörbar. Eine Intervention, die dann einsetzt, wenn „dort“ zu hören ist, dass „hier“ jemand mit seinem Beitrag geendet hat, läuft daher systematisch Gefahr, mit einer „hier“ inzwischen begonnenen Fortsetzung zu kollidieren. Dies ist auch anhand der Intervention von Johannes zu beobachten, die ich jetzt anhand von zwei Versionen des Geschehens genauer betrachten werde.

Aus Wesselingers Sicht (und das entspricht auch der audiovisuellen Dokumentation, die ja in Wesseling durchgeführt wurde (stellt sich das Geschehen wie folgt dar (vgl. Ausschnitt 2): Während Hartmut seinen Beitrag liefert, beginnt Johannes eine eigene Äußerung, und zwar an einer Stelle, die weit von einer übergangsrelevanten Stelle (Sacks et al. 1974, S. 703) entfernt ist (Zeile 135). Seine Intervention wirkt dadurch unterbrechend.

<2> „gibts termine“, Ausschnitt, vereinfacht, „Wesseling-Perspektive“

	132	Hartmut	ob ich das in sechs WOchen erfahre
	133		oder HEUte; das is mir
->	134		ega:l:. <<all>aber ich würd ge[rm noch drüber>]
->	135	Johannes	[() hartmut]
	136	Hartmut	[<<all+p> diskutie() >]
	137	Johannes	[die frage war j]a wenn der holger
	138		bereit is es dem michael nochmal sepaRAT zu

Stellt man nun die erwähnte Signalverzögerung in Rechnung, dann lässt sich rekonstruieren, dass sich das Geschehen für die Beteiligten in Karlsruhe anders dargestellt hat:

<3> „gibts termine“, Ausschnitt, vereinfacht, „Karlsruher-Perspektive“

	132	Hartmut	ob ich das in sechs WOchen erfahre
	133		oder HEUte; das is mir
->	134		ega:l[. <<all>aber ich würd gern noch drüber>]
->	135	Johannes	[hartmut die frage war]
	136	Hartmut	[<<all+p> diskutie() >]
	137	Johannes	[ja wenn der] holger bereit is es
	138		dem michael nochmal sepaRAT zu

In dieser Rekonstruktion ist zu sehen, dass Johannes an einer Stelle zu sprechen beginnt, an der – aus seiner Perspektive – die laufende Äußerung von Hartmut (aufgrund des Intonationsverlaufs, der syntaktischen Kon-

struktion und der Tatsache, dass eine erkennbar vollständige Handlung vorliegt) ein potentielles Ende erreicht („das ist mir egal“, Zeile 134). Allerdings setzt Hartmut seine Äußerung weiter fort und so kommt es zu einem parallelen Sprechen der beiden Teilnehmer (Zeilen 134–137).

Die Signalverzögerung gefährdet also die Handlungskoordination insbesondere bei Sprecherwechseln zwischen den beiden Standorten. Eine Äußerung, die aus der Perspektive der Teilnehmer „dort“ ohne Überlappung oder Pause an eine Äußerung von „hier“ anschließt, erscheint aus der Perspektive der Teilnehmer „dort“ entweder leicht verzögert oder aber sie unterbricht eine „dort“ inzwischen begonnene Fortsetzung.⁹

4.3.2 Trägheit des einmal debalanzierten Interaktionssystems

Zu Beginn der Episode war zu sehen, dass sich die Teilnehmer in Wesseling nicht aktiv an der Lösung des Problems („wer macht das Protokoll?“) beteiligen, sondern sich vielmehr abwartend verhalten. Im weiteren Verlauf der Episode verlagert sich der Schwerpunkt der Diskussion auf den Standort Wesseling und dann sind es die Beteiligten in Karlsruhe, die sich nicht einbringen, sondern vielmehr zuschauen und zuhören. Darüber hinaus ist zu sehen, dass in dem Moment, in dem die Beiträge der Diskutanten in Wesseling deutlicher nur noch zueinander adressiert sind, einige der KollegInnen in Karlsruhe in ihrer Aufmerksamkeit für das Geschehen in Wesseling nachlassen (vgl. Abbildung 11 und Clip 2 im Bereich der Zeiteinblendung 12:16:02). Hier, wie auch in anderen Szenen zeigt sich, dass es eine Tendenz dazu gibt, dass sich ein Schwerpunkt der Aktivitäten an einem der beiden Standorte ausbildet.

Ist erst einmal – beispielsweise im Verlauf einer kontroversen Diskussion an einem der beiden Standorte – ein solcher debalanzierter Zustand eingetreten, dann wird es für die Beteiligten am anderen Standort offenbar schwierig, sich einzubringen. Das hat zum einen mit der schon angesprochenen Schnelligkeit der Interaktion und der Zeitverzögerung zu tun. Dazu kommen aber noch weitere Faktoren wie die lokale Fokussierung der Aufmerksamkeit. Wenn es kontrovers wird, sind die Diskutanten in der Regel zueinander orientiert und haben das Geschehen auf dem Fernsehschirm nicht mehr im Blick (vgl. Abbildung 11). Obwohl Johannes eine große Geste macht und anscheinend auch zu sprechen beginnt (zumindest hat er den Mund geöffnet – vgl. Abbildung 12), gelingt es ihm an dieser Stelle nicht, den Rederaum für sich zu erlangen. Schließlich spielt hier

⁹ Eine wichtige Frage für weitere Forschungen ist die, ob diese Verzögerung von den Beteiligten selbst in Rechnung gestellt wird, ob sie auf die Interpretation von einzelnen Handlungen (etwa als „ungeduldig“ oder aber als „skeptisch“) und auch auf die Organisation etwa von Entscheidungsprozessen durchschlägt.

noch die Tatsache hinein, dass sichtbare Aktionen wie etwa große Hal- tungswechsel oder Bewegungen auf eine zweidimensionale, flache und ei- genartig „entkörperte“ (Heath/Luff) Darstellung auf den Fernsehschirmen reduziert sind. Heath/Luff haben darauf hingewiesen, dass die Beteiligten in video-vermittelter Interaktion einerseits die Wirksamkeit ihres körper- lichen Handelns unterstellen, diese Wirksamkeit aber durch die Art der Repräsentation auf dem Bildschirm deutlich verringert ist und Gesten und Blickzu- oder Abwendungen viel von ihrer handlungskordinierenden Kraft einbüßen (1993, S. 46 ff.).

Die beiden genannten Beobachtungen – die wenig aktive Beteiligung von den Beteiligten an der Gegenstelle sowie das erschwerte Sich-Ein- bringen von der Gegenstelle bei kontroversen Diskussionen (lassen sich mit Beobachtungen an anderen Passagen zusammenführen und ergeben für die Videokonferenzen bei Technics das Bild einer vergleichsweise trä- gen Aufmerksamkeits- und Interaktionsorganisation (zumindest dann, wenn sich ein Aktivitätsschwerpunkt an einem der beiden Standorte aus- gebildet hat. So findet die Schwierigkeit, sich von „hier“ aus einzubrin- gen, wenn „dort“ kontrovers diskutiert wird, ihr Gegenstück darin, dass Initiativen von „dort“ „hier“ leicht versanden können. Besonders deutlich wird dies in der nächsten Szene (vgl. Clip 3: „csc2_wann_solln_wer_da_ reviewn“ und Abbildung 13).

Abbildung 13

In der Folge einer Meinungsverschiedenheit zwischen Johann und Mi- chael R. fragt Johannes, wann denn ein Thema im CSC besprochen wer- den soll. Michael R. entgegnet, dass er zunächst gern einen Vorschlag von einem anderen Kollegen zu diesem Thema hätte. Aber Johannes insistiert und stellt seine Frage erneut. In Wesseling äußern daraufhin Michael D. und Michael R. unterschiedliche Vorschläge dazu. Sie beginnen zueinan- der zu sprechen, liefern aber keine abschließende Stellungnahme für Jo- hannes, der in Karlsruhe erkennbar gerade darauf wartet. Schließlich wen- det sich Johannes von den Fernsehgeräten ab und stellt seine Frage noch einmal an die Kollegen in Karlsruhe.

4.3.3 Gelockerte sequenzielle Organisation?

Sowohl im Verlauf der Episode über eine mögliche Verschiebung des Berichts von Holger als auch in der Szene zu der Frage „wann solln wer das reviewn“, ist zu beobachten, dass Handlungen, die von „dort“ eingefordert oder relevant gemacht werden, „hier“ nicht geliefert werden. In Episode zu Holgers Bericht findet sich in der Folge von Johannes' Vorschlag (den Bericht heute hören) und seiner nachgeschobenen Klausel („WENN wer noch die halbe stunde zeit haben“) zunächst keine Stellungnahme aus Wesseling. Darüber hinaus wird das Angebot von Holger, seinen Bericht kürzer zu halten, von Johannes nicht aufgegriffen und gewürdigt. Beide Male entsteht der Eindruck einer momentanen Desorientierung der Beteiligten in bezug darauf, wo es jetzt weitergeht. Und in der Episode zur Frage „wann solln wer das reviewn“ ist schließlich zu sehen, dass der von Johannes eingeforderte Terminvorschlag von den Kollegen in Wesseling nicht geliefert wird.

Aufgrund dieser Beobachtungen kann man die Frage stellen, ob die sequenzielle Organisation des Handelns in Videokonferenzen grundsätzlich in Mitleidenschaft gezogen oder gelockert ist. Ich gehe allerdings nicht davon aus, dass dies der Fall ist. Vielmehr sind diese Momente der Desorganisation meiner Ansicht nach darauf zurückzuführen, dass sprachliche Handlungen nicht allein über eine Blickzuwendung zum Videokonferenzsystem eindeutig an einzelne Personen an der Gegenstelle adressiert werden können. Die hier angesprochenen Äußerungen sowohl von Johannes als von Holger enthalten keine sprachlich realisierte, eindeutige Adressierung an bestimmte Personen. Die Beteiligten am jeweils anderen Standort verhalten sich in allen drei Fällen als „nicht-persönlich-angesprochen“ und als „nicht-persönlich-mit-einer-Stellungnahme-gefordert“. Es scheint in Videokonferenzen mithin möglich zu sein, von der Gegenseite etablierte aber nicht punktgenau adressierte Handlungsverpflichtungen und -erwartungen auszusitzen.

4.4 Episode 3: Eine Pause machen?

Die Episode, die ich abschließend vorstellen will (vgl. Clip 4: „csc4_bilaterales_thema“), folgt auf einen Bericht von Carla. Sie hatte für die Dauer ihres Berichts und der Diskussion darüber die Moderation der Besprechung übernommen und fragt jetzt, ob man fünf Minuten Pause machen solle (Abbildung 15 zeigt die Beteiligten zu Beginn dieser Szene).

In Wesseling spricht sich Michael D. mit einem knappen, kräftigen „ja“ für eine Pause aus. Von den anderen Teilnehmern, sowohl in Wesseling als auch in Karlsruhe, ist dagegen keine deutliche Stellungnahme zu hören. Kurz darauf bemüht sich Johannes mit einer Fokussierungsaufforderung („äh Hartmut“, vgl. Levinson 1983, S. 345 ff.) um die Aufmerksamkeit seines Kollegen in Wesseling. Nachdem dieser mit einem „ja“ seine Auf-

Abbildung 15

merksamkeit signalisiert hat, beginnt Johannes eine als Bitte deklarierte Äußerung. Noch bevor er allerdings deutlich gemacht hat, worum es ihm geht, unterbricht er seine Äußerung und sagt, dass dies jetzt kein „generelles“, sondern nur ein „bilaterales“ Thema sei. Während Johannes seine zuvor begonnene Äußerung fortsetzt, stehen in Karlsruhe die Kollegen Burkhard und Jürgen auf (kurz darauf auch Carla) und schicken sich an, den Raum zu verlassen. In Wesseling dagegen bleiben alle Beteiligten zunächst sitzen. Sie schauen entweder in ihre Unterlagen (Holger, Michael R.), auf den Gang hinaus (Frank-Ernst) oder aber zu den Fernsehgeräten (Michael D.). Erst nach einer ganzen Weile beginnt Michael D. aufzustehen. Holger folgt ihm unmittelbar nach und über seine nur bruchstückhaft verstehbare Äußerung „(s is) (glaub ich die fünf) (3w) pause“ liefert er sowohl eine Interpretation des Vorgehens in Karlsruhe, als auch eine Begründung für das Aufstehen und Sich-Entfernen in Wesseling.

Anhand dieser kurzen Szene ist zum einen zu sehen, wie eine bestimmte Person am anderen Standort in einen Austausch eingebunden wird. Indem Johannes seiner Äußerung eine Fokussierungsaufforderung mit Adressierungsform („äh hartmut“) voranstellt, sichert er sich dessen Aufmerksamkeit in einem Moment, in dem Hartmut nicht zu den Fernsehgeräten orientiert ist. Prä-Sequenzen aus Fokussierungsaufforderungen mit nachfolgender Fokussierungsbestätigung sind in meinem Material aller-

dings eher selten. Viel häufiger finden sich dagegen Adressierungsformen, die im Verlauf einer Äußerung eingestreut werden (vgl. Ausschnitt 4).

<4> csc3, „wenn das käme“, Ausschnitt

1		((Richard befindet sich in Karlsruhe, Frank-Ernst in Wesseling))
2	Richard	was war noch mal das volu:men das: (.)
-> 3		frank-ernst das hinter dieser dresdner

Gleichzeitig stellt sich bei der Durchsicht des Materials der Eindruck ein, dass solche Adressierungsformen vor allem dann Verwendung finden, wenn Äußerungen an die Gegenstelle gerichtet werden. Werden Äußerungen an einen anderen Teilnehmer am gleichen Ort gerichtet, sind Adressierungsformen vergleichsweise seltener. Dies spricht dafür, dass im zuletzt genannten Fall andere Ressourcen (etwa Blickzuwendungen) erfolgreich eingesetzt werden (vgl. Weinig 1997, S. 163).

Eine weitere Beobachtung, die anhand des zuletzt gezeigten Clips möglich ist, betrifft die Bemerkung von Johannes, dass jetzt kein generelles, sondern ein bilaterales Thema komme. Die Teilnehmer in Karlsruhe scheinen diese Bemerkung schnell und unproblematisch als entweder ein Hinauskomplimentieren oder aber als eine implizite Ratifikation der Frage von Carla nach einer Pause zu verstehen. Sie stehen sofort auf und machen Anstalten, hinauszugehen. Demgegenüber bleiben ihre Kollegen in Wesseling zunächst sitzen und zudem rekonstruiert Holger später beim Aufstehen, dass dies wohl die zuvor angesprochene Pause sei. Der implizite Anteil der Äußerung von Johannes („dieses Thema betrifft nicht alle, ihr könnt/solltet in die Pause gehen“) wird also auf der Gegenseite erst deutlich verspätet aufgegriffen. Darüber hinaus wird dieser implizite Anteil auch zum Gegenstand einer expliziten Rekonstruktion und damit für eine Korrektur durch andere Teilnehmer verfügbar gemacht. Auch wenn eine solche Korrektur in diesem Fall ausbleibt, so deutet sich hier doch ein Bedarf an Absicherung bezüglich der Interpretation des Geschehens „dort“ an.

5. Eine neue Kommunikationsform?

In diesem abschließenden Abschnitt möchte ich nun noch einmal die zuvor ausgeführten Beobachtungen zusammenführen und dann auf die mit dem Titel dieses Beitrags gestellte Frage zu sprechen kommen. Zunächst eine zusammenfassende Übersicht der Strukturhypothesen zur Interaktion in Videokonferenzen, die sich anhand des hier vorgestellten Materials formulieren lassen:

1) *Mit dem Durchführen einer Videokonferenz verändert sich die Ökologie des Besprechungsraums.*

Ist bei einer herkömmlichen Besprechung die Position am Kopfende eines Tisches hervorgehoben und interaktionslogisch vorteilhaft für dieje-

nigen, die steuernd in das Geschehen eingreifen wollen, so sind es in einer Videokonferenz die Plätze in der Nähe des Videokonferenzsystems.

- 2) *Im Verlauf von Videokonferenzen finden sich Besonderheiten in bezug auf das Etablieren und Aufrechterhalten verschiedener Beteiligungskonstellationen.*

Zum einen ist die Lautstärke, mit der gesprochen wird, ein wichtiger Hinweis darauf, ob sich eine Äußerung eher an die Teilnehmer „hier“ oder an die Teilnehmer „dort“ richtet. Zum anderen ist über eine Blickzuwendung zum Konferenzsystem kein punktgenaues Adressieren bestimmter Personen an der Gegenstelle möglich. Hierzu müssen sprachliche Adressierungsformen eingesetzt werden. Dadurch, dass die Beteiligten einerseits die Wirksamkeit ihrer eigenen Blickzuwendungen zu unterstellen scheinen, diese jedoch andererseits in ihrer Funktionalität entkräftet wirken, ergeben sich Störungen im Interaktionsverlauf und Phasen der Desorientierung in bezug darauf, wer jetzt eigentlich „dran“ ist. Schließlich sind die etablierten Beteiligungskonstellationen aufgrund der eingesetzten Tontechnik störungsanfällig, da für die Beteiligten der aktuelle Regelungszustand des elektronischen Tonsystems nicht durchschaubar ist.

- 3) *Das Interaktionssystem Videokonferenz zeichnet sich, sobald sich ein Schwerpunkt der Aktivität an einem Standort ergeben hat und es damit debalanciert ist, durch eine gewisse Trägheit der Aufmerksamkeits- und Beteiligungsorganisation aus.*

Wenn sich im Verlauf einer Konferenz ein Schwerpunkt der Diskussion „dort“ herausbildet, dann wird es für die Beteiligten „hier“ schwer, diese Verlagerung aufzubrechen und sich einzubringen. Gleichzeitig versanden „hier“ gelegentlich Impulse, die von „dort“ aus eingebracht wurden. Darüber hinaus zeigt sich, dass bestimmte Aktivitäten nicht gemeinsam, sondern eher „hier“ oder „dort“ vollzogen werden.

- 4) *Das Geschehen „hier“ ist für die Beteiligten „dort“ nur eingeschränkt transparent.*

Es finden sich wiederholt Situationen, in denen zu sehen und zu hören ist, dass „dort“ etwas geäußert wird, ohne dass die Inhalte dieser Äußerungen „hier“ nachvollziehbar werden. Gelegentlich (aber insgesamt nicht sehr häufig) nehmen die Beteiligten dies zum Anlass, um eine lautere Wiederholung solcher Äußerungen einzufordern. Darüber hinaus ist zu beobachten, dass implizite Aspekte von Äußerungen an der Gegenstelle verspätet oder gar nicht aufgegriffen werden.

Da ich auf einige der Konsequenzen, die sich aus diesen Besonderheiten der Interaktion in Videokonferenzen für die Beteiligten wie für Teams und Organisationen ergeben können, an anderer Stelle hingewiesen habe (Meier, im Erscheinen, b), möchte ich abschließend nur noch kurz auf die

Frage eingehen, ob wir es bei Videokonferenzen mit einer neuen Kommunikationsform zu tun haben.

Neue kommunikative Formen entstehen nicht über Nacht, sondern sind das Ergebnis von Gewöhnung, Standardisierung und Institutionalisierung. Von einer sozialen Form kann bei Videokonferenzen sinnvoll erst dann gesprochen werden, wenn sich Konventionen deutlich herausgebildet haben. Vermutlich wird aber weniger eine einzige soziale Form „Videokonferenz“ auskristallisieren als sich vielmehr (wie bei herkömmlichen Besprechungen auch – eine Ansammlung von Formen entwickeln, die Familienähnlichkeiten beispielsweise in bezug auf die räumliche Ökologie, die Verfahren der Eröffnung und Beendigung, der Aufmerksamkeitsfokussierung oder des Managements von Beteiligung aufweisen. Wann sich solche Konventionen herausgebildet haben werden, ist nicht abzusehen. Aber der Nachweis, dass dies (schon oder noch nicht (der Fall ist, kann nur über vergleichende interaktionsanalytische Untersuchungen in der hier vorgestellten Weise erfolgen.

Literatur

- Barefoot, J. C. & L. H. Strickland (1982). Conflict and dominance in television-mediated interactions. *Human Relations*, 35, S. 559–566.
- Bergmann, J. R. & C. Meier (im Erscheinen). Elektronische Prozeßdaten. In: *Qualitative Forschung: Ein Handbuch*, hrsg. von U. Flick/E. v. Kardorff/I. Steinke. Hamburg: Rowohlt.
- Bronner, R. (1996). *Entscheidungs-Prozesse in Video-Konferenzen. Eine empirische Untersuchung der Leistungsfähigkeit moderner Kommunikationstechnik zur Bewältigung komplexer Management-Aufgaben*. Frankfurt a. M.: Peter Lang.
- Cohen, K. M. (1984). Speaker interaction: video teleconferences versus face-to-face meetings. In: L. A. Parker/C. H. Olgren (eds.), *The teleconferencing resource book: A guide to applications and planning* (S. 288–298). Amsterdam: North-Holland.
- Edwards, Jane A. & Martin D. Lampert (1993). *Talking data: Transcription and coding in discourse research*. Hillsdale, N. J.: Lawrence Earlbaum.
- Goffman, E. (1981). *Forms of talk*. Philadelphia: University of Pennsylvania Press.
- Hart, P., L. Svenning & J. Ruchinskas (1995). From face to face meeting to videoconferencing: potential shifts in the meeting genre. *Management Communications Quarterly* 8 (4), S. 395–423.
- Heath, C. & P. Luff (1993). Disembodied conduct. Interactional asymmetries in video-mediated communication. In: G. Button (ed.), *Technology in working order: Studies of work, interaction and technology* (S. 35–54). London/New York: Routledge.
- Isaacs, E. A. & J. C. Tang (1997). Studying video-based collaboration in context: from small workgroups to large organizations. In: K. A. Finn, A. J. Sellen, S. B. Wilbur (eds.), *Video-mediated communication* (S. 173–197). Manwah, NJ: Lawrence Earlbaum.
- Jefferson, G., H. Sacks & E. A. Schegloff (1987). Notes on laughter in the pursuit of intimacy. In: G. Button/J. R. E. Lee (eds.), *Talk and social organization* (S. 152–205). Clevendon: Multilingual Matters.

- Levinson, S. C. (1983). *Pragmatics*. Cambridge: Cambridge University Press.
- Mantei, M. M., R. M. Baecker, A. J. Sellen, W. S. Buxton, T. Milligan & B. Wellman (1991). Experiences in the use of a media space. In: *Proceedings of the CHI '91 Conference* (S. 203–208). New York: ACM.
- Mars, F. O. (1994). *Business communication via videoconferencing. A case study of media effects*. Unpublished M. A. thesis, Graduate School of Arts and Sciences, George Washington University.
- Meier, C. (1997). *Arbeitsbesprechungen. Interaktionsstruktur, Interaktionsdynamik und Konsequenzen einer sozialen Form*. Opladen: Westdeutscher Verlag.
- Meier, C. (1998). Zur Untersuchung von Arbeits- und Interaktionsprozessen anhand von Videoaufzeichnungen. *Arbeit. Zeitschrift für Arbeitsforschung, Arbeitsgestaltung und Arbeitspolitik*, 7, S. 257–275.
- Meier, C. (im Erscheinen, a). Doing groupness in a spatially distributed work group: the case of videoconferences at Technics. In: L. R. Frey (ed.), *Group communication in context: studies of bona fide groups*. Manwah, NJ: Lawrence Earlbaum Associates.
- Meier, C. (im Erscheinen, b). Videokonferenzen (Beobachtungen zu Struktur, Dynamik und Konsequenzen einer neuen Kommunikationsform. In: M. Boos/K. Jonas & K. Sassenberg (Hrsg.), *Virtuelle Kommunikation in Organisationen. Theorie (psychologische Technologie) (Praxiserfahrungen)*. Göttingen: Hogrefe.
- Olson, J. S., G. M. Olson & D. K. Meader (1995). What mix of video and audio is useful for small groups doing remote real-time design work? In: *Proceedings of the CHI '95 Conference* (S. 362–368). New York: ACM Press.
- Olson, G. M. & J. S. Olson (1997). Making sense of the findings: common vocabulary leads to the synthesis necessary for theory building. In: K. A. Finn, A. J. Sellen, S. B. Wilbur (eds.), *Video-mediated communication* (S. 75–91). Manwah, NJ: Lawrence Earlbaum.
- Reuter, M. (1990). Telekommunikation: Aus der Geschichte in die Zukunft. Heidelberg: R. v. Decker's/Schenk.
- Rosetti, D. K. & T. J. Surynt (1985). Video teleconferencing and performance. *Journal of Business Communication*, 22, S. 25–31.
- Sacks, H., E. A. Schegloff & G. Jefferson (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, 50, S. 696–735.
- Schegloff, E. A. (1987). Between micro and macro: contexts and other connections. In: *The micro-macro link*, ed. by J. C. Alexander/B. Giesen/R. Muench/N. J. Smelser. Berkeley: University of California Press, S. 207–234.
- Sellen, A. (1992). Speech patterns in video-mediated conversations. In: P. Bauersfeld, J. Bennet, G. Lynch (eds.), *Proceedings of the CHI '92 Conference* (S. 49–59). New York: ACM.
- Selting, M., P. Auer, B. Barden, J. Bergmann, E. Couper-Kuhlen, S. Günthner, C. Meier, U. Quasthoff, P. Schlobinski & S. Uhmman (1998). Gesprächsanalytisches Transkriptionssystem (GAT). *Linguistische Berichte*, 173, S. 91–122.
- Short, J., E. Williams & B. Christie (1976). *The social psychology of telecommunications*. London: Wiley.
- Stoetzer, M. (1994). Neue Telekommunikationsdienste: Stand und Perspektiven ihres Einsatzes in der deutschen Wirtschaft. *ifo Schnelldienst*, 47 (7), S. 8–19.
- Weinig, K. (1997). *Wie Technik Kommunikation verändert. Das Beispiel Videokonferenz*. Münster/Hamburg/London: LIT-Verlag. A review. *Psychological Bulletin*, 84 (5), S. 963–976.