

Aufzeichnung technisch vermittelter Kommunikation - das Beispiel Videokonferenz

Olaf A. Schulte / Martin Friebe / Christian Klotzek

Abstract

Die Aufzeichnung technisch vermittelter Kommunikation stellt eine zweifache methodologische Herausforderung dar: Auf der einen Seite verlangt sie umfangreiche und detaillierte Kenntnisse des Equipments, auf der anderen Seite gilt es, eine den zentralen Fragestellungen angemessene Aufzeichnungstechnik zu entwickeln. Am Beispiel der Videokonferenz sollen die aus den Charakteristika dieser Kommunikationsform (getrennte Standorte, zeitliche Verzögerung) resultierenden Probleme erörtert und praktische Lösungsvorschläge vorgelegt werden.

Keywords: Aufnahmetechnik; Videokonferenz, Fernkommunikation, Synchronizität.

English Abstract

Recording computer-mediated communication poses two major methodological challenges: On the one hand, extensive and detailed knowledge of the equipment used is needed, on the other hand, the recording has to be adequate to the phenomena focused. Looking at videoconferencing, this article discusses problems arising from the main characteristics (separate locations, delay) and gives some practical advice to solve them.

Keywords: Digital Recording, Videoconferencing, Synchronisation.

1. Einleitung
2. Aufzeichnung als Instrument empirischer Forschung
3. Videokonferenz als Kommunikationsform
4. Aufzeichnungstechnik
5. Aufzeichnung der Wahrnehmungsbedingungen
 - 5.1. Aufzeichnungsgegenstand
 - 5.2. Kontextkamera vs. Monitoraufnahme
 - 5.3. Audio: Das externe Mikrophon
6. Verzögerung und Synchronizität
7. Fazit und Ausblick

"There is no such thing as a complete record of any event" (Grimshaw 1982:123).

1. Einleitung

Das gemeinsame, computerbasierte Arbeiten (CSCW) und die computergestützte zwischenmenschliche Kommunikation (CMC) haben ebenso wie der Umgang des Menschen mit dem Computer (CHI)¹ Eingang in den beruflichen und privaten Alltag gefunden. In einer kaum überschaubaren Vielfalt werden für die synchrone oder asynchrone Interaktion von Menschen mit Menschen technische Lösungen angeboten, die zudem in immer kürzeren Produktlebenszyklen mehr oder weniger grundlegende Modifikationen erfahren. Diese raschen Veränderungen sorgen nicht nur für eine Vielzahl verschiedenster Forschungsfoki und das zunehmende Problem der Vergleichbarkeit von Forschungsergebnissen über unterschiedliche Stufen der technischen Entwicklung. Auch in methodischer Hinsicht sehen sich die Forschenden mit zusätzlichen Problemen konfrontiert, wenn nämlich die Aufzeichnung als Instrument der empirischen Erfassung des Gegenstandes zur technischen und theoretischen Herausforderung wird. Auf der einen Seite bietet der technische Fortschritt zwar die Voraussetzungen einer einfacheren, flexibleren und detaillierteren Aufzeichnung: Immer kleinere Kameras mit immer besserer Bildqualität zu moderaten Preisen erleichtern den immer unauffälligeren Einsatz von Kontextkameras, Software- und Hardwarelösungen ermöglichen die Aufzeichnung der Inhalte des PC-Monitors auch ohne separate Kamera und parallele Logfileanalysen als detaillierte inhaltliche und zeitliche Protokolle der Arbeiten am PC erweitern und verfeinern das zur Verfügung stehende Instrumentarium (Holmer/Streitz 1999). Auf der anderen Seite wird diese Entwicklung konterkariert durch gestiegene Anforderungen an das praktische Technikwissen und die Notwendigkeit der theoretischen Ausdifferenzierung des eben auch immer komplexeren Aufzeichnungsinstrumentariums.

Ein ähnliches Bild bietet die Erforschung der vermeintlich synchronen audiovisuellen (Fern-) Kommunikation, die als Videokonferenz (*videoconferencing* bzw. *video-mediated communication*) nach Einschätzung ihrer Erfinder "the next best thing to being there" (Molnar 1969:135) werden sollte. Ihre Erforschung ist Gegenstand des DFG-Projektes "Audiovisuelle Fernkommunikation" an der Universität Essen,² das seit Februar 2000 unter der Leitung von H. Walter Schmitz und

¹ Computer Supported Cooperative Work, Computer-Mediated Communication bzw. Computer-Human-Interaction.

² <http://www.uni-essen.de/videokonferenz>.

Jens Loenhoff die Nutzung von Videokonferenzen aus kommunikationswissenschaftlicher Perspektive untersucht und dazu u. a. auch angemessene Methoden der Datenerhebung und –konstitution zu entwickeln hatte. Wenn nachfolgend die für diese Kommunikationsform vorhandenen Aufzeichnungsoptionen und die Frage nach deren Adäquanz dem eigentlichen kommunikativen Ereignis gegenüber diskutiert wird, dann zeigt sich darin exemplarisch die komplexe Wechselwirkung zwischen technischer Machbarkeit und (kommunikations-) theoretischen Anforderungen.

2. Aufzeichnung als Instrument empirischer Forschung

Die verschiedenen Möglichkeiten der Fixierung akustischer und/oder visueller Informationen als Ersatz oder Ergänzung der (teilnehmenden) Beobachtung haben die wissenschaftliche Perspektive insbesondere auf flüchtige Ereignisse verändert: "Sie gestatten es dem Sozialforscher, ein sich ereignendes soziales Geschehen in seinem realen zeitlichen Ablauf zu bewahren und gleichzeitig dessen Temporalstruktur in beliebiger und reversibler Weise zu manipulieren" (Bergmann 1985:304). Dabei hat die Forschung nicht nur von der grundsätzlichen Möglichkeit der Aufzeichnung profitiert; sie wurde nachgerade von der Entwicklung der Aufzeichnungstechnik beeinflusst, wenn nämlich gesprächsanalytische Feldforschung im Zuge der Entwicklung mobiler Audiorekorder und Studien zur Rolle nonverbaler Kommunikation oder des Körpers insgesamt im Gefolge der Entwicklung entsprechenden Videoequipments einen Aufschwung erfuhren (Heath 1997:183). Dies verwundert angesichts der im Vergleich oder in Ergänzung zur teilnehmenden Beobachtung größeren Möglichkeiten einer audiovisuellen Aufzeichnung kaum. In dem Maße, in dem sich die Forschung mit der Fülle kommunikativer Aktivitäten (Gestik, Mimik, Blick etc.) auseinandersetzt, wird die Fähigkeit des Forschenden eingeschränkt, diese in situ vollständig zu registrieren. Verschärft wird diese Problematik, wenn mehrere Personen am fokussierten sozialen Geschehen beteiligt sind. Die Aufzeichnung bietet hier mit zwei wesentlichen Merkmalen Abhilfe: Sie weist eine hohe *Informationsdichte* (density) auf und besitzt eine der Flüchtigkeit sozialen Geschehens entgegenwirkende *Dauerhaftigkeit* (permanence; Grimshaw 1982:122). So kann die Fülle der relevanten Informationen durch den Einsatz mehrerer Kameras mit dann unterschiedlichen Perspektiven (verschiedene Personen oder eine Totale in Ergänzung der Detailaufnahme) erfaßt werden. Forschende können dann in der wiederholten Betrachtung des Materials zuverlässigere Aussagen über die fokussierten Aktivitäten machen. Zudem können im Sinne einer intersubjektiven Überprüfung weitere Forschende das vorliegende Material sichten bzw. einer weiteren wissenschaftlichen Bearbeitung unterwerfen (z. B. Transkription) und die Zuverlässigkeit (Reliabilität) der Beobachtung, der zugrundeliegenden Kategorien und damit der gesamten Analyse verbessern. Hinzu kommt, daß - eine entsprechende Archivierung und Dokumentation vorausgesetzt - das Material auch Jahre später für alternative Forschungszwecke und Analysen zur Verfügung steht, falls es den dann veränderten Anforderungen an die Aufzeichnung genügt.

Gerade eine an (der Vielzahl der) Details der menschlichen Kommunikation interessierte Forschung wird zudem eine Aufzeichnungstechnik würdigen, die in der wiederholten Betrachtung des Materials technische Optionen für eine akribi-

sche Analyse bereitstellt: Wiederholung, Zeitlupe, Einzelbilddarstellung, Standbilder und Vergrößerungen bieten ebenso wie oszillographische Darstellungen der Audiospuren ein immer raffinierteres Repertoire für die Ex-post-Analyse. Mit seiner Hilfe können auch sublimale Aspekte des kommunikativen Geschehens wie die für dessen Organisation relevanten Mikroausdrücke (Augenbewegungen, Veränderungen der Lidspalte u. ä.) berücksichtigt werden: "Similarly, by viewing videotapes in slow motion or frame by frame, subtle changes in behavior may become apparent and, therefore, available for analysis" (Bottorff 1994:257). Ein entsprechendes Desiderat formulieren Storck und Sproull für die Erforschung der Videokonferenz: "Methods of measuring attention at a gross level – which monitor what speakers look at, what people remote from the speaker look at, where people local to the speaker look – need to be developed" (1995:214). *Last but not least* bietet die Aufzeichnung die Möglichkeit, Aussagen zur zeitlich-sequentiellen Organisation der Kommunikation zu machen: Über einen entsprechenden Timecode oder die Analyse der Einzelbilder können Aussagen etwa zum Rederechtswechsel mit der Genauigkeit einer 25stel Sekunde gemacht werden. Damit entspricht die Aufzeichnung insgesamt auch der Forderung Grimshaws (1982:129) nach "decomposability": Das Material sollte für jedes Analyseniveau aufzubereiten sein, sei es die framegenaue Darstellung einer Geste oder die vollständige Wiedergabe der Interaktion.

Obwohl die Aufzeichnungstechnik also Potentiale weit jenseits derjenigen eines menschlichen Beobachters bietet, bleibt die Aufzeichnung eben durch ihre technischen Charakteristika immer eine verlustbehaftete Abbildung des eigentlichen (kommunikativen) Ereignisses. Koerfer (1981:187) unterscheidet für das Verfahren der audiovisuellen Datendokumentation "generelle" und "spezielle" Verluste:³ Während erstere grundlegende Defizite benennen (Verlust der Dimensionalität sowie taktiler, olfaktorischer u. a. Verhaltensmodi), beziehen sich letztere auf die aus der Verwendung des Aufzeichnungsgerätes resultierenden Verluste. Auch für Grimshaw ist dies einer der Gründe für den fragmentarischen Charakter einer Aufzeichnung; er nennt als technisch-mechanische Defizite die Selektivität durch Kamerapositionierung und -zoom (Totale vs. Close-up und damit etwa Gestik vs. Mimik; vgl. auch Bottorff 1994:249) und verlustbehaftete Audioaufnahmen; hinzu kommt die Reduzierung um den subjektiven Gehalt und den historisch-ethnographischen Kontext: "Any notion of completeness is chimerical" (1982:143).

Diese Verluste sind im Einzelfall mit Blick auf das Forschungsinteresse gegenüber den genannten (Zu-) Gewinnen bei den Analysemöglichkeiten abzuwägen. Darüber hinaus ist zu berücksichtigen, daß die Aufzeichnung und eine nachfolgende Transkription ein ohnehin ergänzungsbedürftiges gesprächs- bzw. konversationsanalytisches Vorgehen darstellen: Mit Blick auf das für eine dem kommunikativen Ereignis adäquate Analyse notwendige Situations- und Sachwissen sind Aufzeichnung und Transkript um andere Formen der Datenerhebung bzw. -konstitution wie Beobachtung und Befragung zu ergänzen (Ingenhoff/Schmitz 2000a:149). Eine in diesem Sinne dem ethnographischen Holismus verpflichtete Forschung würde in erster Linie bedeuten, daß Forschende im Rahmen einer teilnehmenden Beobachtung zusätzliche Informationen über das beobachtete Ereignis

³ Hinzukommen die nicht speziell für die audiovisuelle Fixierung relevanten kulturellen Verluste.

erlangen, das dann in die spätere Datenkonstitution und -analyse einfließen sollte (Koerfer 1981:190). Auch die ergänzende Erhebung demo- und soziographischer Daten ist mit Blick auf das Hintergrundwissen zu erwägen (Bottorff 1994:247), ebenso jedoch die informelle Befragung. Darüber hinaus ist - im Sinne einer intersubjektiven Überprüfbarkeit - eine detaillierte und vor allem (selbst-) kritische Analyse (technisch bedingter) möglicher und tatsächlicher Defizite der Aufzeichnung vorzunehmen. Abgerundet wird die Aufzeichnung immer durch ein Aufzeichnungsprotokoll, das möglichst minutiös die jeweiligen Bedingungen der Aufzeichnung dokumentiert und die Grundvoraussetzung einer sinnvollen Arbeit am Transkript darstellt.

Insgesamt also sollte die Aufzeichnung das soziale Geschehen "as faithfully as is possible within the limits of technology and human error" (Grimshaw 1982:128) reproduzieren. Um aber überhaupt von einer dem ursprünglichen Geschehen gegenüber getreuen Aufzeichnung sprechen zu können, bedarf es zunächst einer Bestimmung des Gegenstandes, denn die Untersuchung "muß mit einer präzisen Bestimmung der zu untersuchenden Ereignisse und Prozesse angefangen werden, und der sind dann die Verfahren der Beobachtung und Beschreibung, der Aufzeichnung und der Transkription anzupassen, und nicht umgekehrt" (Ingenhoff/Schmitz 2000a:151).

3. Videokonferenz als Kommunikationsform

Videokonferenzen dienen der technisch vermittelten audiovisuellen (Fern-) Kommunikation, die über ISDN oder LAN⁴ erfolgt. Bei den Anlagen, die in der Grundausrüstung aus Kamera, Bildschirm, Mikrophon und Lautsprechern sowie einer der Verarbeitung der anfallenden Audio- und Videodaten dienenden Videokonferenzeinheit bestehen, werden PC-basierte Desktopsysteme, mobile Rollabouts und stationäre Großraumanlagen unterschieden. Für die Leistungsfähigkeit ist neben der Ausstattung der Anlagen (Bildschirmgröße, Kameraeigenschaften etc.) insbesondere die zur Verfügung stehende Übertragungsbandbreite relevant, die heute zwischen 128 kbit/s (für einen ISDN Basisanschluß) und 2 Mbit/s im LAN liegt und die Qualität der übertragenen Bilder, die meist deutlich unterhalb derjenigen des Fernsehens liegt, ebenso beeinflußt wie diejenige des Audiomaterials, das in den meisten Fällen mit der vom Telefon bekannten Sprachqualität zu vergleichen ist. Da diese qualitativen Merkmale aber von Anlage zu Anlage, in Abhängigkeit von der Raumgestaltung, den Aktivitäten der Teilnehmenden oder auch nur deren räumlicher Anordnung variieren, gehören vorübergehende oder dauerhafte Reduzierungen der akustischen und/oder visuellen Informationen zu den Merkmalen der Kommunikation per Videokonferenz.

Schon in den dreißiger Jahren des vergangenen Jahrhunderts prototypisch als Ableger der Fernsehtechnik vorgestellt, ist die Videokonferenz nach einer Renaissance in den siebziger Jahren bis heute ein "revolutionary concept on the brink of success" (Egido 1988:13). Einer der Gründe für den ausstehenden Durchbruch

⁴ Integrated Services Digital Network bzw. Local Area Network; Videokonferenzen in der ersten Variante werden oft nach dem zugrundeliegenden Protokoll auch als H.320-Videokonferenzen bezeichnet, während letztere dem Protokoll H.323 folgen und auch als „IP-basiert“ (Internet Protocol) firmieren.

dieser Technik mag sein, daß die idealtypische Beschreibung als "bidirektionaler vollduplexfähiger audiovisueller Echt-Zeit-Kommunikationsdienst zwischen Benutzergruppen in 2 oder mehr örtlich getrennten Standorten"⁵ wesentliche Charakteristika verdeckt, die die kommunikativen Aktivitäten maßgeblich beeinflussen. Zwei seien hier mit Blick auf die Aufzeichnungsproblematik genannt, die aus der räumlichen und (demzufolge) auch zeitlichen Trennung der beteiligten Personen resultieren: Zum einen handelt es sich bei der Videokonferenz eben nicht um eine Echt-Zeit-Kommunikation, wie sie die Vis-à-vis-Kommunikation darstellt. Vielmehr ergeben sich durch die Digitalisierung, Kodierung (Datenkompression auf Basis mathematischer Verfahren mit dem Ziel der Reduzierung des zu übertragenden Datenmaterials), Übertragung sowie Dekodierung und Analogisierung insbesondere des Videomaterials⁶ Verzögerungen zwischen den beteiligten Standorten zwischen 0,4 und 1 Sekunde, abhängig von der Leistung der für diese Aufgaben zuständigen Codecs⁷ der Videokonferenzanlage, der zur Verfügung stehenden Bandbreite und der Menge der jeweils zu übertragenden Daten. Letztere ergibt sich aus der Menge der Veränderungen zwischen den Einzelbildern⁸ und wird etwa durch Bewegungen der teilnehmenden Personen verursacht. Die durch die Verdopplung dieser Verzögerungen entstehenden Pausen können zu erheblichen Problemen in der Organisation der Kommunikation führen: "Transmission delay disrupts the pace of normal conversations, makes the appropriate timing of interruptions more difficult, and impedes the smooth resolution of simultaneous speech events" (Cohen 1984:292).

Zum anderen resultiert aus der Beteiligung *getrennter* Standorte eine Asymmetrie der Wahrnehmungsbedingungen: Nicht nur technisch unterschiedliche Anlagen liefern im Ergebnis unterschiedliche visuelle und akustische Informationen. Durch die Anordnung und Größe der verschiedenen Bildinformationen (Bild der Gegenstelle und Eigenbild), individuelle Einstellungen (Lautstärke, Bildqualität) und die jeweilige Konfiguration vor Ort führen auch identische Anlagen zu differierenden Wahrnehmungsbedingungen für die Beteiligten. Heath und Luff sprechen in diesem Zusammenhang von einer asymmetrischen Anordnung: "The technology [...] provides physically distributed individuals with incongruent environments for interaction" (1992:50).

Diese - für eine kommunikationswissenschaftliche Bestimmung des Gegenstandes zentralen - Aspekte standen bislang nicht im Blickfeld der Videokonferenzforschung. Über drei Jahrzehnte, verschiedenste Stufen des technischen Fortschritts (Videokonferenzstudios im Vermittelten Breitbandnetz der Deutschen Bundespost, ISDN-basierte mobile Videokonferenzen, IP-basierte Lösungen) und über unterschiedliche Forschungsansätze (technisch, psychologisch, betriebswirtschaftlich, arbeitssoziologisch) hinweg blieb die Frage nach solchen Unterschieden zur Vis-à-vis-Kommunikation und damit nach den spezifischen Charak-

⁵ Gerfen faßt so die Vorschriften der europäischen Regulierungsbehörde CEPT (Conférence Européenne des Administrations des Postes et des Télécommunications) zur Videokonferenz zusammen; Gerfen (1986:6).

⁶ Das Audiomaterial könnte fast verzögerungsfrei übertragen werden, wird aber mit dem Ziel der Synchronizität mit den Videodaten aufgehalten.

⁷ Codierung - Decodierung.

⁸ Die für die Videokonferenz unerlässliche Datenreduktion beruht in wesentlichen Teilen darauf, daß nicht das Gesamtbild, sondern die Veränderungen zum letzten Bild berechnet und übermittelt werden.

tika der Videokonferenzkommunikation relativ offen. Dabei hielt das Gros der Studien an einem (gerade angesichts des Gegenstandes der technisch vermittelten Kommunikation und den sich aufdrängenden Aufzeichnungsoptionen) recht konservativen Instrumentarium fest, bei dem Befragungen der Teilnehmenden und vergleichende Analysen von Arbeitsergebnissen über verschiedene Kommunikationsformen im Vordergrund standen (stellvertretend Schütze 2000:55f.). Erst Anfang der neunziger Jahre näherten sich einige Studien der Frage nach den Auswirkungen von Verzögerung und asymmetrischen Wahrnehmungsbedingungen. Diese wiesen - *cum grano salis* - zwei gemeinsame Merkmale auf: Zum einen handelt es sich um Studien, die auf methodischer Ebene mehr oder weniger explizite konversations- bzw. gesprächsanalytische Ansätze verfolgen und die zum anderen hierfür eine registrierende Aufzeichnung (und eine anschließende Transkription) als Instrumentarium wählen (O'Conaill et al. 1993; Sellen 1992; O'Malley et al. 1996; O'Conaill/Whittaker 1997).⁹ Oft beschränken die Studien sich dabei auf das akustische Material (Braun et al. 1999; Sellen 1992) - paradoxerweise aber ohne damit die Zielsetzung einer Erforschung der Charakteristika der *audio-visuellen* Kommunikation per Videokonferenz aufzugeben. Visuelle Informationen werden meist als Photos oder Screenshots berücksichtigt (Fussell/Benimoff 1995:233); Videoaufzeichnungen von Bild und Ton kommen erstmals bei Bergmann et al. (1999:8) zum Einsatz.

Eine Auseinandersetzung mit technischen und methodologischen Problemen der Aufzeichnung steckt demzufolge noch in den Kinderschuhen: Dort, wo Aussagen zur audiovisuellen Fixierung des empirischen Gegenstandes gemacht werden, sind diese entweder rein deskriptiv (Weinig 1996:116) oder vereinfachen aus ihrem Erkenntnisinteresse die zugrundeliegende Problematik, indem sie z. B. von einer "synchronen Kooperation" sprechen, in der die "Beteiligten an beiden Standorten mit den gleichen, technisch bedingten Handlungsproblemen konfrontiert sind und über die jeweils gleichen Ressourcen verfügen" (Meier 1998a:20 bzw. 28). Damit werden die für die Videokonferenz zentralen Probleme der zeitlichen Verzögerung und der differenzierten Wahrnehmungsbedingungen ausgeblendet.

4. Aufzeichnungstechnik

Für die Aufzeichnung von Videokonferenzen kommen in den meisten Fällen zwei Geräte in drei Aufnahmevarianten in Frage: Zum einen ein an die Videokonferenzanlage angeschlossener Videorekorder (VHS oder S-VHS), der gelegentlich vom Hersteller als Aufzeichnungsoption in die Videokonferenzsoftware integriert wird und die Video- und Audiodaten direkt von der Videokonferenzeinheit bezieht. Der Videorekorder bietet eine recht lange Aufnahmedauer (240 Minuten als Standard, in der Longplay-Funktion bis zu 610 Minuten), produziert allerdings nur analoges Material, das in der nachträglichen Bearbeitung nur sequentielle Zugriffe erlaubt; langwieriges Spulen ist oft die Folge. Zum anderen bietet sich ein (digitaler) Camcorder an; die beiden dominierenden Formate sind der Hi8-Nachfolger D8 sowie Mini-DV (Digitales Video), die sich in Technik und Handhabung mehr oder weniger entsprechen.

⁹ Als Vorreiter sollten zudem Périn (1983) und Cohen (1984) genannt werden.

Ein Camcorder kann zunächst als Kontextkamera eigene Audio- und Videosignale aufzeichnen, wofür neben dem eingebauten auch zusätzliche Mikrophone verwendet werden können. Alternativ kann er aber auch als (digitaler) Videorekorder fungieren, falls er über einen entsprechenden S-Video-Eingang verfügt: Dann nämlich kann der S-Video-Ausgang eines Monitors genutzt werden, wobei in der Regel der Einsatz eines zusätzlichen externen Mikrophons notwendig wird, da das camcorder-interne Mikrophon in der Betriebsart "Videorekorder" meist deaktiviert ist; dies erfordert u. U. zusätzlich - so zumindest die Erfahrung mit dem Sony DCR-TRV 20E - den Einsatz eines Verstärkers bzw. Mischpultes. Die letztgenannte Variante ermöglicht auch die Aufzeichnung PC-gestützter Videokonferenzen ("Desktop-Videokonferenz"), ebenso aber jeder anderen PC-Anwendung: Über den separaten Ausgang einer entsprechend konfigurierten Grafikkarte¹⁰ kann der Camcorder über diesen S-Video-Ausgang an der Karte oder ein VGA-S-Video-Kabel ("Kabelpeitsche") alle Vorgänge auf dem PC-Monitor ohne gravierende Qualitätseinbußen erfassen. Alternativen zu dieser Variante sind zum einen reine Softwarelösungen wie die für die Verwendung in PC-Schulungsvideos vorgesehene ScreenCam der Firma Lotus.¹¹ Da diese Programme aber für die Fixierung der Aktivitäten auf dem Monitor eine entsprechende Datei anlegen müssen und kaum Möglichkeiten der Kompression bestehen, müssen entsprechende Einschränkungen mit Blick auf die Auflösung und die Bildwiederholungsrate hingenommen werden; zudem ist die Kompatibilität gerade mit Videofenstern einer Videokonferenzanwendung unsicher. Zum anderen kommt der Einsatz eines Scanconverters in Frage, der die PC-spezifischen Videodaten in für konventionelle Aufzeichnungsgeräte lesbare Formate (S-Video, Composite) umwandelt.¹² Auch hier sind allerdings deutliche Defizite der Videoqualität hinzunehmen.

Im Einsatz dieser Geräte für die Aufzeichnung einer Videokonferenz kommt ein besonderer Vorteil der technisch vermittelten Interaktion hinzu: Im Vergleich zur Vis-à-vis-Kommunikation kann ihr Einsatz hier als minimalinvasiv bezeichnet werden. Zwar kann sich auch die Erforschung der Videokonferenz dem grundsätzlichen Gegensatz zwischen dem für die Beobachtung und Aufzeichnung sozialen Geschehens relevanten Beobachter-Paradoxon (Labov 1972:209) und der Einschätzung, daß zumindest unbewußte Aktivitäten der beobachteten Probanden wie Kopfnicken, Gestik, Blick und sogar Rededauer nicht auf die Aufzeichnungssituation reagieren (Wiemann 1981:309), nicht vollständig entziehen. Allerdings kann mit Blick auf die verwendete Technik die Registrierung des Geschehens zumindest in seiner Durchführung weitgehend unbemerkt vom Beobachtungsobjekt Mensch erfolgen.¹³ Mit Ausnahme einer Kontextkamera erfolgen alle für die Erfassung der Kommunikation per Videokonferenz notwendigen Aufzeichnungen versteckt: Der Videorekorder im Rollabout, der Camcorder hinter dem Monitor oder - im Falle der Desktop-Videokonferenz - dem PC: Alle registrierenden Abläufe erfolgen ohne direkte Störung der Beteiligten.¹⁴ Selbst die Kontextkamera

¹⁰ Z. B. http://www.ati.com/na/pages/products/pc/radeon_ve/index.html oder <http://www.matrox.com/mga/deutsch/dualhead/home.cfm>.

¹¹ <http://www.lotus.com/home.nsf/welcome/screencam>.

¹² <http://www.avermedia.de>.

¹³ Die wünschenswerte oder notwendige Teilnahme des Forschenden bleibt hier außen vor; für sie gelten die Probleme der Vis-à-vis-Kommunikation analog.

¹⁴ Ausnahmen sind das Konfigurieren und Starten der Aufzeichnungsgeräte sowie ein möglicherweise notwendiger Kassettenwechsel. Durch Starten der Aufnahme vor Beginn der Videokonferenz

und ein externes Mikrophon sind in den oft mit zusätzlichen (Decken-) Kameras, Mikrophonen und anderen technischen Geräten (Beamer, Dokumentenkamera) ausgestatteten Videokonferenzräumen weniger ein Fremdkörper als in der Vis-à-vis-Kommunikation. Hinzu kommt, daß gerade in der Videokonferenz die beteiligten Personen grundsätzlich um die Tatsache wissen, daß sie von einer Kamera beobachtet werden. Als mögliche Beeinflussungsfaktoren der teilnehmenden Personen bleiben dann lediglich die Laborumgebung (in experimentellen Settings) bzw. eine gewisse "Reaktivitätsproblematik" (Ellgring 2000:206), wenn die Teilnehmenden über die Aufzeichnung informiert bzw. hierzu um eine Genehmigung angegangen werden.

Mit Blick auf die Genauigkeit späterer Analysemöglichkeiten sollte auch die Qualität des Aufzeichnungsmaterials höchstmöglichen Anforderungen genügen: Das DV-Format ist sowohl in Qualität als auch in der Handhabung (Überspielen auf PC, digitale Schnittoptionen) jeder handelsüblichen analogen Technik überlegen, zumal es mit dem integrierten Timecode wichtige Informationen über die chronologisch-sequentielle Ordnung des Bildmaterials liefert. Die bislang vorgebrachten Gründe gegen diese Technik (Kosten, Aufnahmedauer; Meier 1998b:271) verlieren angesichts fallender Gerätepreise, der Möglichkeit digitaler Videoarchive und der verlustfreien Longplay-Funktion der Camcorder zunehmend an Bedeutung. Die Notwendigkeit einer optimalen Videoqualität wird vor allem dann größer, wenn die Aufnahmen zu Analyse Zwecken zusammengeführt werden sollen: Im Zusammenschnitt verschiedener Videoquellen sind deutliche Qualitätseinbußen zu beklagen. Meier (1998a:38) weist darauf hin, daß durch den für das Zusammenspiel verschiedener Kamerabilder (Gegenstelle, Eigenbild, Kontextkamera) in einem Bild die Darstellungsgenauigkeit leidet. Im geteilten Bildschirm (als "split screen" bzw. in der häufigsten Variante mit vier Bildquellen als "quarter split" bezeichnet) gehen durch die dann reduzierte Größe der Bilder und die bescheidene Qualität handelsüblicher Bildmischer (Multiplexer/Videosplitter) Detailinformationen verloren. Exemplarisch hier das Bild eines Quartersplit, der zwei Videokonferenzkameras (mit Eigenbild) und zwei Kontextkameras zusammenführt:



Abbildung 1: Quartersplit (mit drei Standorten)

ferenz und Verwendung der „Longplay“-Funktion für S-Video-Aufnahmen (mit dann bis zu 90 Minuten Aufnahmedauer) kann auch dieser Einfluß minimiert werden.

Für eine gesprächsanalytische Auswertung, die auch Wert auf Mimik, Blickkontakt oder sogar Mikroausdrücke legt, sollte möglicherweise auf das jeweilige Originalband zurückgegriffen werden, das dann nur für Demonstrationszwecke zusammengefügt wird. Allerdings wäre in diesem Zusammenhang auch die Frage zu stellen, ob die ex post diagnostizierten, oft subtilen Informationen in der für den Quartersplit typischen verminderten Darstellungsqualität¹⁵ durch die Beteiligten in situ überhaupt wahrgenommen werden (können).

5. Aufzeichnung der Wahrnehmungsbedingungen

Neben der zeitlichen Verzögerung spielt auch der Verlust des von den Beteiligten geteilten Raumes eine Rolle: Es handelt sich nicht mehr um die Interaktion mit dem physisch präsenten Gesprächspartner, sondern um diejenige mit seinem audiovisuellen Abbild. Damit stellt sich zunächst die Frage, was auf der jeweiligen Seite zum Gegenstand der Aufzeichnung werden sollte.

5.1. Aufzeichnungsgegenstand

Die Videokonferenz bietet den seltenen Vorteil, daß in der Eigenschaft ihrer technischen Vermitteltheit das Audio- und Videomaterial für eine Aufzeichnung des kommunikativen Geschehens schon bereitgestellt wird. Die Berücksichtigung der von den genutzten Kameras und Mikrofonen erfaßten Daten allein würde jedoch der Videokonferenz als kommunikativem Ereignis nicht gerecht werden. So ist das Bild der Person A, wie es von der Videokonferenzkamera am Standort A aufgenommen wird, nicht notwendigerweise die für Person B zugrundeliegende visuelle Information über Person A. Vielmehr muß hier immer die übertragene Variante des Bildes berücksichtigt werden, da diese durch die notwendige technische Transformation und mögliche Übertragungsstörungen qualitative Veränderungen erfährt. Zu diesen gehören Probleme in der Synchronisation von Audio und Video ("Lippensynchronizität"; Finger/Davis o. J.), der Verlust von Einzelbildern, der zu einem weniger flüssigen Bewegungsablauf oder sogar zu kurzzeitig "eingefrorenen" Bildern ("Stroboskopeffekt") führt sowie eine insgesamt verminderte Videoqualität (Johnson 1991:70). Das nachfolgende Bild zeigt die in Folge einer leichten Kopfbewegung typischerweise zu beobachtende "Kästchenbildung":

¹⁵ Durch die Zusammenführung mehrerer Standorte wird in der Multipointkonferenz, die meist mit einer Quartersplitdarstellung arbeitet, die Übertragungskapazität oft auf ein Minimum reduziert.



Abbildung 2: Qualitative Einbußen durch Übertragung (128 kbit/s)

Hinzu kommt, daß das Bild am Standort B möglicherweise manipuliert wird: Wird dort das in der Videokonferenz genutzte Eigenbild (das dann Person B zeigt) hinzugeschaltet, ist ein Teil der visuellen Informationen, die vom Standort A kommen, für Standort B nicht verfügbar, da das Eigenbild einen Teil des empfangenen Bildes abdeckt. Umfang und Position des Eigenbildes sind bei den meisten Videokonferenzsystemen individuell einstellbar; es kann bis zu 20% des Gesamtbildes einnehmen. Das nachfolgende Beispiel zeigt die durch das Eigenbild eingeschränkte Sichtbarkeit des Oberkörpers, der Arme und der Hände der Person rechts:



Abbildung 3: Bild der Gegenstelle und Eigenbild

Darüber hinaus können im Verlauf einer Videokonferenz möglicherweise alternative Bildinformationen zu sehen sein: Das kann der für beide Seiten vollzogene Verzicht auf Bildinformationen zugunsten einer Präsentation oder einer gemeinsam genutzten Anwendung ("Application Sharing") sein, aber auch das Umschalten auf eine parallel laufende Anwendung ("Multitasking") an nur einem Standort, wenn dort etwa kurzfristig in einer PC-Anwendung gearbeitet wird, die die Aufmerksamkeit des Gesprächsteilnehmers erfordert, z. B. durch den akustisch sig-

nalisierten Eingang einer E-Mail. Auch die Videokonferenzsoftware selbst verdeckt bzw. ersetzt oft das Monitorbild, so daß während der Nutzung derselben (um etwa die Lautstärke zu regeln) ebendiese auch aufzuzeichnen wäre. Im Ergebnis sollte die Aufzeichnung also jeweils den aktuellen Bildschirminhalt (und eben nicht nur die Videosignale der Videokonferenz) wiedergeben.

Ähnlich verhält es sich mit den Audiosignalen: Auch hier ist grundsätzlich der empfangene Ton zu berücksichtigen, weil dieser durch die technische Vermittlung (z. B. Einschränkung des Frequenzspektrums mit dem Ziel der Datenreduktion) und individuelle Einstellungen (Lautstärke, Klangbild) verändert wird. Allerdings gelten für die Aufzeichnung des Audiomaterials weitere Bedingungen (vgl. 5.3.).

5.2. Kontextkamera vs. Monitoraufnahme

Des weiteren stellt sich die Frage, ob die für die Videokonferenz genutzte Kamera alle für analytische Zwecke notwendigen oder gewünschten Daten zur Verfügung stellt. Ein erster Aspekt ist hier die Frage, was Person A tatsächlich sieht bzw. sehen könnte: Die Aufzeichnung nur der visuellen Informationen der Videokonferenz verstellt den Blick auf das potentielle Blickfeld der beteiligten Personen, das allein in der horizontalen Ebene etwa 170° ausmacht (Yamaashi et al. 1996:50) und in diesem Bereich Aktivitäten wahrnehmen kann, die dann möglicherweise mit einem fokussierten Blick gewürdigt werden. Der Blick der beteiligten Personen kann also auf das Bild ihres Gegenüber, auf ihr Eigenbild, auf die oberhalb des Monitors befindliche Kamera, aber eben auch auf jeden anderen Gegenstand (oder eine weitere Person) im Raum gerichtet sein. Eine entsprechende Aufnahme (oder eine detaillierte Skizze) würde die ex post Analyse beispielsweise der Blickrichtungen erleichtern.



Abbildung 4: Blickfeld vs. Monitorbild¹⁶

Außerdem ist die Verwendung zumindest einer Kontextkamera zu erwägen, die mit einem weiteren Blickfeld als die Videokonferenzkamera arbeitet. Denn obwohl die Videokonferenzkamera mit den genannten Einschränkungen und Anfor-

¹⁶ Auch dieses Blickfeld ist durch die Aufnahme mit der Kamera gegenüber dem menschlichen Blickfeld stark eingeschränkt.

derungen das für den Gesprächspartner relevante Material liefert (weil er letztlich mit diesem interagiert), können für analytische Zwecke auch Informationen relevant sein, die *nicht* übertragen werden: An erster Stelle sind hier Aktivitäten der an der Videokonferenz beteiligten Personen zu nennen, die zwar wahrnehmbar sind, aber durch die Eigenschaften von Kamera und Monitor und die unzulänglichen Möglichkeiten der Fokussierung des Blicks (Yamaashi et al. 1996:51) nicht en détail erkannt werden können, wie etwa die Ablenkung einer Person durch einen Gegenstand, den erst die Kontextkamera (oder die nachträglichen Erklärungen des teilnehmenden Beobachters) als mobiles Telefon ausmacht. Auch Informationen über Veränderungen im Raum (Eintreten einer Person o. ä.) und andere Aktivitäten jenseits des Blickfeldes der Videokonferenzkamera (möglicherweise sogar bewußt verheimlichte Handlungen) oder aber auch gescheiterte Versuche, kommunikative Signale zu übermitteln (etwa durch die Videokonferenzkamera nicht erfaßte oder fokussierte Embleme eines Teilnehmenden). Und schließlich kann die Kontextkamera der Kompensation einer speziell für Gruppenkonferenzen konfigurierten Kamerasteuerung dienen: Wenn die Kameras der Videokonferenzanlage manuell oder über akustische Steuerung die jeweils sprechende Person fokussieren, würde ohne Kontextkamera jegliche Information über die anderen, momentan nicht sprechenden Teilnehmenden verloren gehen. Die Kontextkamera liefert also - ähnlich der Funktion eines teilnehmenden Beobachters - weitergehende Informationen über das kommunikative Ereignis.

5.3. Audio: Das externe Mikrophon

Ob auch die akustischen Informationen in der gewünschten Form vorliegen, hängt entscheidend von der Technik der Videokonferenzanlage ab - sie sollte das gesamte an einem Standort anfallende Audiomaterial erfassen. Diese Forderung bedeutet für Standort A zunächst, daß sowohl verbale Äußerungen der Person A als auch die übertragenen Äußerungen von Person B zu erfassen wären. Darüber hinaus sollte dieses Material dann allerdings zumindest zwei Forderungen an Authentizität erfüllen: Zum einen sollte die Qualität derjenigen vor Ort entsprechen, d. h. daß Äußerungen von Person A am Standort A deutlich(er) zu hören sind - im Vergleich zu denen von Person B. Diese weisen durch die zum Zwecke der Datenreduktion durchgeführte Kompression (und in Abhängigkeit vom zugrundeliegenden Verfahren; Knoll/Stenger 1998:23), durch mögliche Fehler gerade in der Übertragung im LAN und möglicherweise auch durch Eigenschaften der Mikrophone und Lautsprecher der Videokonferenzanlage eine deutlich geringere Qualität auf (Finger/Davis o. J.). Im Ergebnis klingt die übertragene Stimme dumpfer, weist ein geringeres Frequenzspektrum auf oder wird durch einen Hall verändert. Die damit gerade in einer gesprächsanalytischen Bearbeitung des Materials wichtige Unterscheidung zwischen "hier" und "dort" wird dann erschwert, wenn die Aufnahmeoptionen der Videokonferenzanlage das Audiomaterial auch des eigenen Standortes in der für die Übertragung komprimierten Version aufnimmt und mit dem entsprechenden Material der Gegenstelle abmischt; im direkten Ver-

gleich zu einem mit externem Mikrophon arbeitenden Camcorder wird dies deutlich.¹⁷

Zum anderen sollten auch (Neben-) Geräusche erfaßt werden, wie sie typischerweise im Rahmen einer Besprechung per Videokonferenz anfallen. Auch hier zeigen die in der Videokonferenztechnik gängigen Mikrophone tendenzielle Schwächen: Sie besitzen in der Regel ein recht hohes Eingangsniveau, um nicht (durch Nebengeräusche) versehentlich aktiviert zu werden oder haben sogar eine entsprechende Option, die über die Videokonferenzsoftware konfiguriert werden kann. Dadurch können möglicherweise sogar (leise) Gespräche oder Bemerkungen der Aufzeichnung entgehen; die nachfolgenden Graphiken zeigen die Lautstärke der jeweiligen Aufnahmetechnik:

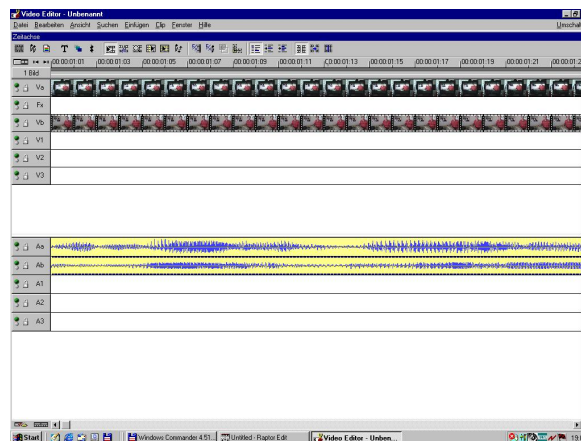


Abbildung 5: Audiospur Camcorder vs. Videorekorder (Lautstärke)

Die parallel zu den beiden Videoquellen (Camcorder und Videorekorder, Einzelbilddarstellung im oberen Drittel des Screenshots) verlaufenden Audiospuren im unteren Drittel des Bildes veranschaulichen die insgesamt geringere Lautstärke der Aufnahme über das Mikrophon der Videokonferenzanlage (untere Spur); außerdem wurde bei dieser experimentellen parallelen Aufzeichnung mit Videorekorder und Camcorder¹⁸ die während eines Turns der Gegenseite "[...] die Gesetzesänderung durchkommt, daß äh Eltern dafür strafbar gemacht werden können, wenn ihre Kinder [...]" gemachte Bemerkung "Warte mal - geh' mal n bißchen" nur durch das eingebaute Mikrophon des Camcorders erfaßt, während sie aufgrund der Charakteristika des Mikrophons bzw. der Abmischung der beiden Seiten durch die Videokonferenzeinheit der Aufzeichnung per Videorekorder entging - und damit einer weiteren Analyse entzogen würde.

Die Charakteristika des eingesetzten externen Mikrophons müssen ebenfalls zwei grundsätzliche Forderungen erfüllen: Zum einen müssen sie im Zusammen-

¹⁷ Videorekorder via Videokonferenzeinheit:
[http://www.uni-essen.de/kowi/schmitz/dfg-projekt/audiospurvcr\(hiervs.dort\).wav](http://www.uni-essen.de/kowi/schmitz/dfg-projekt/audiospurvcr(hiervs.dort).wav);
 Camcorder mit externem Mikrophon:

[http://www.uni-essen.de/kowi/schmitz/dfg-projekt/audiospurcc\(hiervs.dort\).wav](http://www.uni-essen.de/kowi/schmitz/dfg-projekt/audiospurcc(hiervs.dort).wav)

¹⁸ Videorekorder via Videokonferenzeinheit:
<http://www.uni-essen.de/kowi/schmitz/dfg-projekt/eingangsniveauvcr.wav>;
 Camcorder mit externem Mikrophon:
<http://www.uni-essen.de/kowi/schmitz/dfg-projekt/eingangsniveaucc.wav>.

spiel mit dem jeweiligen Camcorder bzw. Verstärker oder Mischpult funktionieren, was insbesondere eine Frage der entsprechenden Adapter ist. Zum anderen muß die Richtcharakteristik des Mikrophons den Anforderungen des Forschenden genügen: Die üblicherweise zur Ausblendung von Nebengeräuschen eingesetzten Mikrophone (z. B. "Niere") unterdrücken diese durch einen engen Aufnahmewinkel von etwa 80°. Sollen genau diese Nebengeräusche Teil der Analyse sein, muß dementsprechend ein Mikrophon mit Kugelcharakteristik (mit einem Aufnahmewinkel von bis zu 300°) eingesetzt werden.¹⁹ Die eingesetzten Sennheiser K6 konnten als "Kugelmikrophone" die gewünschte Wiedergabe der Raumakustik gewährleisten.

Die nachfolgende Skizze faßt die technischen Bedingungen für *eine* Seite (Person A in physischer Präsenz, Person B erscheint auf dem Monitor) zusammen. Die drei grau hinterlegten Felder kennzeichnen die Aufnahmeoptionen "VCR" (Videorekorder und Videokonferenzanlage), "KK" (Kontextkamera) und "Camcorder (VCR)" (S-Video-Signal des Monitors und externes Mikrophon).

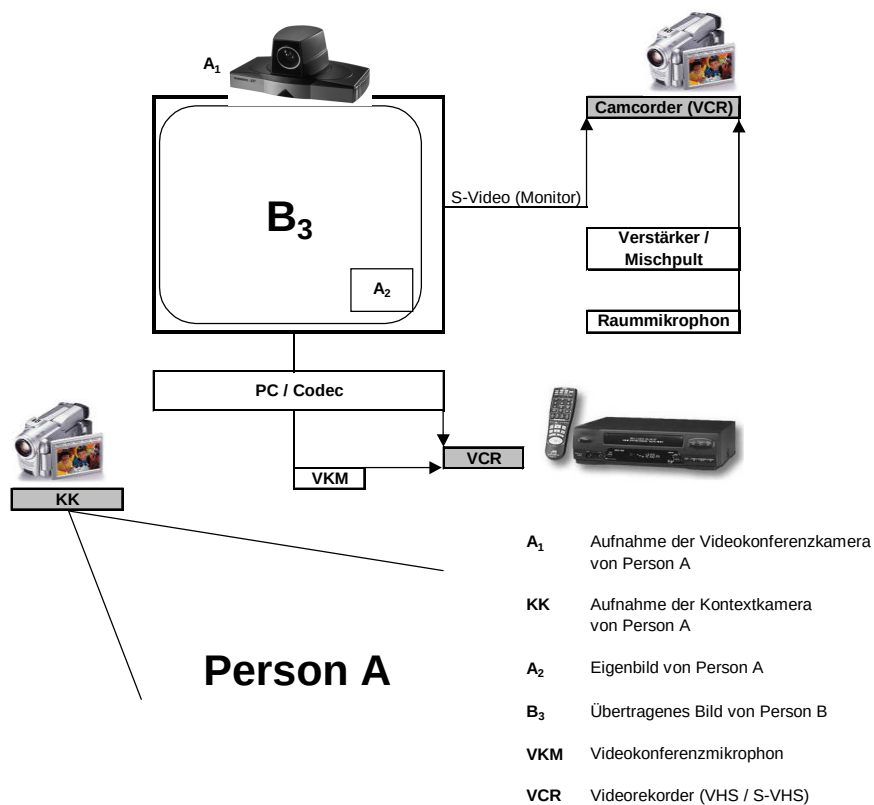


Abbildung 6: Aufzeichnungsoptionen

¹⁹ http://www.strauss-online.de/Video/Videotechnik/Videoton/body_videoton.html (Sep. 2001).

6. Verzögerung und Synchronizität

Eines der für die Kommunikation per Videokonferenz zentralen Probleme ist die geschilderte Verzögerung zwischen den beteiligten Standorten, die hier als zweites Merkmal der Videokonferenz beschrieben werden soll, das eine Aufzeichnung beider Seiten einer Videokonferenz verlangt. So würden etwa die monologischen Ausführungen von Person A, die Person B mit einer Reihe von verbalen Backchannel-Signalen ergänzt, in der Aufzeichnung an Standort B folgende Transkription ergeben:²⁰

```
A   if this is a register which has got too high a value [then=
B                                     [mm
A   =when I am testing this resistor [which should fail high
B                                     [mm
```

Am Standort A würden die Backchannel-Signale entsprechend verzögert auftreten:

```
A   if this is a register which has got too=
A   high a value then when [I am=
B                                     [mm
A   testing this resistor which should [fail high
B                                     [mm
```

Das Beispiel verdeutlicht, daß auch die zunächst minimal anmutende Verzögerung von weniger als einer Sekunde Relevanz für den Ablauf und die Organisation der Kommunikation per Videokonferenz besitzt. Die im vorliegenden Fall vom Sprechenden verzögert wahrgenommenen Backchannel-Signale ("mm") des Hörenden wirken durch die Bedingungen der technischen Vermittlung weniger affirmativ und Aufmerksamkeit signalisierend, sondern tragen eher zur Verunsicherung des Sprechenden bei; Studien (O'Conaill/Whittaker 1995) zu diesem Problem weisen auf eine resultierende Reduzierung der Backchannel-Signale hin. Aber auch andere, in ihrem Ablauf auf eine sehr genaue zeitliche Abfolge angewiesene kommunikative Ereignisse können durch die Verzögerung zwischen den Standorten gestört werden: Der Rederechtswechsel etwa leidet ebenfalls unter diesem spezifischen Merkmal der Videokonferenz, das in längeren Turns und Pausen zwischen ebendiesen resultiert (O'Malley et al. 1996; Braun et al. 1999). Und auch die Wahrnehmung kommunikativer Mikroereignisse, die als mimischer Ausdruck z. T. nur 200 Millisekunden dauern (Bruce 1996:171), wird von diesem Problem berührt.

Eine Analyse solcher Spezifika der Videokonferenz ist offensichtlich auf eine Aufzeichnung beider Standorte zwingend angewiesen. Zwar arbeitet eine ganze Reihe von Studien als Alternative auf der Basis nur einer Aufzeichnung mit einer entsprechenden konstanten Verzögerung, wie sie auch im vorliegenden Transkriptionsbeispiel angewandt wurde. Dies aber verlangt eine (in der Regel nicht vorgenommene oder zumindest nicht dokumentierte) Diagnose dieses Wertes und unterstellt zudem eine im Zeitablauf konstante Verzögerung. Um eine sol-

²⁰ Beispiel nach O'Conaill et al. (1993:13); Transkriptionssymbolik: Die rechteckige Klammer nach rechts (I) kennzeichnet den Beginn der Überlappung, das Gleichheitszeichen (=) am Ende bzw. Anfang der Zeile verweist auf eine fortlaufende Äußerung.

che aber handelt es sich bei der Videokonferenz nicht. Tatsächlich nämlich kann es insbesondere in IP-basierten, aber auch in ISDN-Videokonferenzen zu Schwankungen der Verzögerung kommen. Der technische Hintergrund ist dabei, daß die Schwankungen entweder durch externe Reduktion der zur Verfügung stehenden Bandbreite (insbesondere in IP-basierten Videokonferenzen, die keine entsprechende Bandbreitenreservierung aufweisen) oder durch eine plötzliche Zunahme des anfallenden Datenmaterials (z. B. durch größere Bewegungen der Teilnehmenden oder einen Kameraschwenk) verursacht werden können. Eigene Meßreihen, die aufgrund des experimentellen Designs schon relativ gute Bedingungen (geringe Videodatenmenge, kein Audiomaterial) für eine konstante Verzögerung besaßen, wiesen dennoch Schwankungen von bis zu einer Viertelsekunde auf; die Verzögerung zwischen den Standorten wich dadurch von dem als Mittelwert gemessenen Durchschnittswert um bis zu 30% ab.

Damit aber verlangt die Registrierung der Videokonferenz in letzter Konsequenz eine framegenaue Abstimmung der Aufzeichnungen bzw. eine entsprechende (nachträgliche) Synchronisation des in situ aufgezeichneten Materials. Eine solche, framegenaue Synchronisation verschiedener Aufzeichnungsgeräte wird in der professionellen Videobearbeitung mit Hilfe eines SMPTE- bzw. EBU-Timecodes²¹ bewerkstelligt: Auf den Aufzeichnungsgeräten wird von einem zentralen Timecodegenerator parallel zum Video- und Audiomaterial ein entsprechender Timecode geschrieben, der in der Bearbeitung dann ausgelesen werden kann. Als Alternative bietet sich im Labor an, über einen Videosplitter unmittelbar auf ein Band aufzunehmen. Dies verlangt allerdings entsprechende Räumlichkeiten und bringt die genannten Qualitätsverluste mit sich.

Da aber ohnehin nur in der Laborsituation von *einem* Standort ausgegangen werden kann, muß die Synchronisation in der Feldforschung mit dann zwei (oder mehr) getrennten Standorten auf Basis eines externen, für alle Seiten verfügbaren Signals erfolgen. Eine Lösung bieten hier Timecodegeneratoren, die ihrerseits die genaue Zeit über die atomuhrgenauen Funksignale der Physikalisch-Technischen Bundesanstalt oder über eine satellitengestützte Uhrzeitbestimmung per GPS²² beziehen und eine framegenaue Synchronisation ermöglichen.²³ Allerdings ist ein solches Vorgehen mit sehr hohem technischen Aufwand und hohen Kosten verbunden. Alternative Lösungen liefern der internen PC-Uhr über eine Schnittstelle oder eine einzubauende PC-Karte ein entsprechendes Signal, ohne jedoch die Genauigkeit der erstgenannten Lösungen zu erreichen; zudem verlangen sie einen Eingriff in die technische Anlage und sind daher in der Feldforschung nicht anwendbar. Eine (praktikable) Lösung jenseits des Labors bieten derzeit lediglich die akustischen Signale, die einige Radiosender (z. B. Deutschlandfunk) zur vollen und z. T. halben Stunde aussenden. Ihre Genauigkeit liegt im Bereich von 2-3 Frames (abhängig von der Entfernung des Standortes vom Sender und von der nachträglichen Bestimmung der relevanten Frames) und liegt damit noch deutlich oberhalb der handelsüblicher Funkuhren, die wesentlich seltener als die oben diskutierten Lösungen Signale erhalten und demzufolge eine geringere Präzision aufweisen.

²¹ Standardisierung der Society of Motion Picture and Television Engineers bzw. der European Broadcasting Union.

²² Global Positioning System.

²³ Z. B. <http://www.masterclock.com/masterclock/gps-200A.htm> (Juli 2001).

Ihren Abschluß findet diese Art der Synchronisation dann durch die Aufarbeitung am PC. Programme zur digitalen Bearbeitung des Videomaterials wie Adobe Premiere oder Ulead Media Studio Pro können mehrere Videoquellen in Einzelbilddarstellung vertikal anordnen,²⁴ so daß durch entsprechende Schnitte die jeweiligen Frames der unterschiedlichen Quellen mit dem entsprechenden akustischen Signal übereinander angeordnet werden können - die parallel verlaufenden oszillographischen Darstellungen der Audiospuren erleichtern diesen Vorgang bei der Verwendung akustischer Signale. Alternativ könnten die entsprechenden Videobänder mittels eines Videosplitters zu einem Band zusammengeführt werden, wobei die genannten qualitativen Defizite zu berücksichtigen wären; Ansätze einer Analyse dieses Materials finden sich bei Wulff (1991:281).

Mit Hilfe der skizzierten Lösungen wäre es möglich, einen gemeinsamen Startpunkt der getrennten Aufnahmen zu finden, so daß diese dann als zusammengeführte Dokumentation des kommunikativen Ereignisses "Videokonferenz" eine Grundlage für eine entsprechende Datenkonstitution bilden, die dann etwa im Rahmen einer Transkription erfolgt.

7. Fazit und Ausblick

Ob als Alternative oder Ergänzung der (teilnehmenden) Beobachtung: Die Aufzeichnungstechnik liefert gerade in der Erforschung computergestützter Kommunikationsformen zusätzliche Möglichkeiten, die ihre spezifischen Stärken in der Vielfalt der nachträglichen Analyseoptionen aufweisen. Allerdings werden diese Verbesserungen durch eine komplexere Aufzeichnungstechnik konterkariert. Die Verwendung entsprechender Geräte oder Softwarelösungen verlangt eine differenzierte Darstellung der zugrundeliegenden technischen Bedingungen, die zudem mit Blick auf ihre Angemessenheit dem eigentlichen Gegenstand der Forschung gegenüber kritisch zu hinterfragen sind.

Die weitere Entwicklung sowohl der Hardware als auch der für die Aufzeichnung besonders relevanten Codierungs- bzw. Kompressionsverfahren wird das Potential dieses Instrumentes weiter erhöhen. Zwar läßt dieameratechnik schon heute mit Blick auf Miniaturisierung und Qualität kaum noch (Forscher-) Wünsche offen, doch bietet gerade die Software für die Aufzeichnung computerbasierter Kommunikation noch deutliches Potential: Wo heute Programme zur Aufzeichnung der Bildschirminhalte in der Anzahl der Einzelbilder und der Bildauflösung deutliche Defizite aufweisen, wird zukünftig eine (fast) verlustfreie Fixierung "in Echtzeit" möglich sein. Aber auch dann muß zunächst die kritische Frage nach der Adäquanz des Aufzeichnungsverfahrens dem eigentlichen Beobachtungsobjekt gegenüber beantwortet und die genaue Bestimmung möglicher und tatsächlicher Defizite durchgeführt werden.

Eine dergestalt optimierte Aufzeichnungstechnik bietet das volle Repertoire der genannten Analyseoptionen allerdings nur dann, wenn dem Verfahren der Registrierung ein entsprechender Prozeß der Datenkonstitution folgt, der in Genauigkeit und Adäquanz der Kommunikation per Videokonferenz gegenüber der hier vorgestellten Aufzeichnungstechnik nicht nachsteht. In diesem Sinne ist eine weitere, nicht nur kommunikationswissenschaftliche Erforschung dieses Feldes auch auf

²⁴ Vgl. Abbildung 5.

die Entwicklung von Transkriptionsverfahren und –instrumenten angewiesen, die den Gegenstand der technisch vermittelten Kommunikation angemessen abbilden. Aus Sicht (nicht nur) der mit der Videokonferenz befaßten Forschung würde dies für ein Transkriptionssystem sprechen, das (u. a.) bei framegenauer Synchronisation mehrere Videoquellen zusammenspielen kann und eine hohe analytische Flexibilität (z. B. der Kategorien oder Symbole) aufweist; entsprechende Ansätze sind vorhanden (Ingenhoff/Schmitz 2000b im Druck).

Literatur

- Bergmann, Jörg R. (1985): Flüchtigkeit und methodische Fixierung sozialer Wirklichkeit. Aufzeichnungen als Daten der interpretativen Soziologie. In: Bonß, Wolfgang / Hartmann, Heinz (Hrsg.): Entzauberte Wissenschaft. Göttingen: Otto Schwartz & Co, 299-320.
- Bergmann, Jörg R. / Goll, Michaela / Meier, Christoph (1999): Abschlußbericht. Arbeitspapier Nr. 8 (Telekooperation. Strukturen, Dynamik und Konsequenzen elektronisch vermittelter kooperativer Arbeit in Organisationen). <http://www.uni-giessen.de/~g31047/bericht8.pdf> (Februar 2001).
- Bottorff, Joan L. (1994): Using Videotaped Recordings in Qualitative Research. In: Morse, Janice M. (eds.): Critical Issues in Qualitative Research Methods. Thousand Oaks, London, New Delhi: Sage Publications, 244-261.
- Braun, Sabine / Kohn, Kurt / Mikasa, Hans (1999): Kommunikation in der mehrsprachigen Videokonferenz: Implikation für das Dolmetschen. In: Gerzymisch-Arbogast, Heidrun / Gile, Daniel / House, Juliane / Rothkegel, Annely (Hrsg.): Wege der Übersetzungs- und Dolmetschforschung. Tübingen: Gunter Narr Verlag, 267-306.
- Bruce, Vicki (1996): The Role of the Face in Communication: Implications for Videophone Design. In: Interacting with Computers, 8. Jg., H. 2, 166-176.
- Cohen, Karen M. (1984): Speaker Interaction: Video Teleconferencing versus Face-to-face Meetings. In: Parker, Lorne, A. / Olgren, Christine, H. (Hrsg.): The Teleconferencing Resource Book. A Guide to Applications and Planning. Amsterdam, New York, Oxford: North Holland Elsevier Science Publishers B.V., 288-298.
- Egido, Carmen (1988): Video Conferencing as a Technology to Support Group Work: A Review of its Failures. In: ACM SIGCHI & SIGOIS (Hrsg.): Proceedings of the Conference on Computer-supported Cooperative Work: September 26 - 29, 1988 Portland, Oregon New York, N.Y.: ACM Press, 13-24.
- Ellgring, Heiner (1995): Audiovisuell unterstützte Beobachtung. In: Flick, Uwe (Hg.): Handbuch qualitative Sozialforschung. Weinheim: Beltz, 203-208.
- Finger, Roger / Davis, Andrew W. (o. J.): Measuring Video Quality in Videoconferencing Systems. www.wainhouse.com/article_j.html (Juli 2001).
- Fussell, Susan R., Benimoff, Nicholas I. (1995): Social and Cognitive Processes in Interpersonal Communication: Implications for Advanced Telecommunications Technologies. In: Human Factors, 37. Jg., H. 2, 228-250.
- Gerfen, Wilfried (1986): Videokonferenz. Alternative für weltweite geschäftliche Kommunikation - ein Leitfaden für Anwender. Heidelberg: R.v. Decker's Verlag).

- Grimshaw, Allen D. (1982): Sound-Image Data Records for Research on Social Interaction. Some Questions and Answers. In: *Sociological Methods & Research*, 11. Jg., H. 2, 121-144.
- Heath, Christian (1997): The Analysis of Activities in Face-to-face Interaction Using Video. In: Silverman, David (eds.): *Qualitative Research. Theory, Method and Practice*: London, Thousand Oaks, New Delhi: Sage Publications, 183-200.
- Heath, Christian / Luff, Paul (1993): Disembodied Conduct. International Asymmetries in Video-mediated Communication. In: Button, Graham (eds.): *Technology in Working Order. Studies of Work, Interaction, and Technology*. London, New York: Routledge, 35-54.
- Holmer, Torsten / Streitz, Norbert (1999): Neue Möglichkeiten der Analyse der Mensch-Computer-Interaktion zur Evaluation von computerunterstützten Gruppensitzungen. In: Arend, U. / Eberlerl, E. / Pitschke, K. (Hrsg.): *Software-Ergonomie '99. Design von Informationswelten*. Stuttgart: Teubner, 137-148.
- Ingenhoff, Diana / Schmitz, H. Walter (2000a): Über den Gegenstand gesprächsanalytischer Transkriptionen. In: Hess-Lüttich, Ernest W.B. / Schmitz, H. Walter (Hrsg.): *Botschaften verstehen: Kommunikationstheorie und Zeichenpraxis*. Festschrift für Helmut Richter. Frankfurt a.M., u.a.: Peter Lang, 143-154.
- Ingenhoff, Diana / Schmitz, H. Walter (2000b): ComTrans: A Multimedia Tool for Scientific Transcription and Analysis of Communication. In: Rector, Monica (ed.): *Gestures: Meaning and Use. Proceedings of the Conference at Fernando Pessoa University, Porto, 1. - 5. April 2000* (in press).
- Johnson, John Till (1991): Videoconferencing. Not just Talking Heads. In: *Data Communications*, November 1991, 66-68.
- Koerfer, Armin (1981): Probleme und Verfahren der Notation von Face-to-Face Interaktion. Anhand audio-visueller Aufzeichnungen von Seminardiskussionen. In: Lange- Seidl, Annemarie (Hrsg.): *Zeichenkonstitution. Akten des 2. semiotischen Kolloquiums Regensburg 1978*. Berlin, New York: Walter de Gruyter, Band 2, 187-197.
- Labov, William (1972): *Sociolinguistic Patterns*. Philadelphia: Univ. of Pennsylvania Pr.
- Meier, Christoph (1998a): Arbeit als Interaktion: Videodokumentationen als Voraussetzung für die Untersuchung von flüchtigen Telekooperationsprozessen. Arbeitspapier Nr. 3 (Telekooperation. Strukturen, Dynamik und Konsequenzen elektronisch vermittelter kooperativer Arbeit in Organisationen). <http://www.uni-giessen.de/~g31047/bericht2.pdf> (August 2001).
- Meier, Christoph (1998b): Zur Untersuchung von Arbeits- und Interaktionsprozessen anhand von Videoaufzeichnungen. In: *Arbeit*, 7. Jg., H. 2, 257-275.
- Molnar, Julius P. (1969): Picturephone Service - A New Way of Communicating. In: Bell Laboratories (Hrsg.): *Bell Laboratories Record*. New York: Bell Lab. Vol. 47, Nr. 5, May/June 1969, 134-135.
- O'Conaill, Brid / Whittaker, Steve (1995): An Analysis of the Spoken Aspects of Video Mediated Communication. Palo Alto, Calif.: Hewlett Packard Laboratories. (HP Laboratories Technical Report; 92-163, (R. 1)).
- O'Conaill, Brid / Whittaker, Steve (1997): Characterizing, Predicting, and Measuring Video-Mediated Communication: A Conversational Approach. In:

- Finn, Kathleen E. / Sellen, Abigail J. / Wilbur, Sylvia B. (Hrsg.): Video-Mediated Communication. Mahwah, New Jersey: Lawrence Erlbaum, 107-131.
- O`Conaill, Brid / Whittaker, Steve / Wilbur, Sylvia (1993): Conversations Over Video-Conferences: An Evaluation of Video-mediated Communication. Bristol: Hewlett Packard Laboratories. (HP Technical Report; 92-163)
- O`Malley, Claire / Langton, Steve / Anderson, Anne u.a. (1996): Comparison of Face-to-face and Video-mediated Interaction. In: Interacting with Computers, 8. Jg., H. 2, 177-192.
- Périn, Pascal (1983): Communication Interactive de Groupe et Médiatisation. In: Psychologie française, H. 28, 289-296.
- Schütze, Hans-Joachim (2000): Warum Videokonferenzen? Überlegungen und eine Experimentalstudie zur Medienvermittlung informeller Kommunikation. Frankfurt a. M., Berlin, Bern, u.a.: Peter Lang. (Europäische Hochschulschriften: Reihe VI; Psychologie).
- Sellen, Abigail J. (1992): Speech Patterns in Video-mediated Conversations. In: Bauersfeld, P. / Bennet, J. / Lynch, G. (eds.): CHI '92 Conference Proceedings ACM Conference on Human Factors in Computing Systems. New York: Association for Computing Machinery (ACM), 49-59.
- Weinig, Katja (1997): Wie Technik Kommunikation verändert. Das Beispiel Videokonferenz. Münster, Hamburg, London: LIT-Verlag. (Kommunikationsökologie, Bd. 2; Diss.)
- Wiemann, John M. (1981): Effects of Laboratory Videotaping Procedures on Selected Conversation Behaviors. In: Human Communication Research, 7. Jg., H. 4, 302-311.
- Wulff, Hans J. (1991): Split Screen: Erste Überlegungen zur semantischen Analyse des filmischen Mehrfachbildes. In: Kodikas / Code, 14. Jg., H. 3 - 4, 281-290.
- Yamaashi, Kimiya / Cooperstock, Jeremy R. / Narine, Tracy u.a. (1996): Beating the Limitations of Camera-Monitor Mediated Telepresence with Extra Eyes. In: Proceedings of CHI '96, ACM Conference on Human Factors in Computing Systems. New York: Association for Computing Machinery (ACM), 50-57.

Olaf A. Schulte, M.A.
Martin Friebe, M.A.
Christian Klotzek
DFG-Projekt "Audiovisuelle Fernkommunikation"
Universität Essen - R12 T03 E04
Universitätsstraße 12
45117 Essen
c/o Olaf.A.Schulte@uni-essen.de

Veröffentlicht am 20.12.2001

© Copyright by GESPRÄCHSFORSCHUNG. Alle Rechte vorbehalten.