

Erschienen in: Mondada, Lorenza/Schmitt, Reinhold (Hrsg.):
Situationseröffnungen. Zur multimodalen Herstellung fokussierter Interaktion.
Tübingen: Narr, 2010, S. 7-52.

Lorenza Mondada/Reinhold Schmitt

Zur Multimodalität von Situationseröffnungen

1. Einleitung

Der vorliegende Band ist das Ergebnis einer seit Oktober 2006 bestehenden offiziellen Kooperation zwischen Mitarbeiter(inne)n des Laboratoire ICAR (Centre National de Recherche Scientifique und Université de Lyon) sowie Mitarbeiter(inne)n der Abteilung Pragmatik des Instituts für Deutsche Sprache.¹ Der Zusammenarbeit liegt ein gemeinsames Forschungsprojekt mit dem Titel „Vergleichende Analysen von Situationseröffnungen“/„Analyses comparées d'ouvertures“ zugrunde. Das Projekt, das seinen Fokus im Vergleich von französisch- und deutschsprachigen Situationseröffnungen in unterschiedlichen sozialen Kontexten hat, setzt die langjährigen Arbeitskontakte vor allem von Lorenza Mondada und Reinhold Schmitt als zentrale Initiatoren und Koordinatoren fort, die seit Ende der 1990er Jahre bestehen. Ersten organisatorischen Ausdruck fand diese deutsch-französische Kooperation in den seit 2003 regelmäßig am IDS stattfindenden halbjährlichen „Arbeitstreffen zu Fragen multimodaler Kommunikation“ (Schmitt 2004, 2009).

Die gemeinsamen Analysen und Diskussionen im Rahmen dieser Arbeitstreffen zielen auf die Entwicklung eines Ansatzes, der auf der empirischen Grundlage audiovisueller Daten Interaktion in ihrer komplexen multimodalen Qualität methodisch kontrolliert untersuchbar macht. Getragen werden diese Diskussionen von der Einsicht, dass die faktische Komplexität von Interaktion nur durch eine methodische und theoretische Fokussierung auf den Gesamtzusammenhang von Sprache, Intonation, Mimik, Blick, Gestikulation, Körperpositur, Präsenzmodi (Sitzen, Stehen, Gehen), der Manipulation von Gegenständen, der Konstellation der Beteiligten zueinander, der Positionierung im Raum sowie den praxeologischen Gegebenheiten (Interaktion als Bestandteil anderer Aktivitätszusammenhänge) erfasst werden kann.

Die Arbeitstreffen führten zu einem zweitägigen Kolloquium, das im Oktober 2005 am IDS abgehalten wurde und der gemeinsamen Diskussion von Koordination als einem neuen Untersuchungsgegenstand der videogestützten Interaktionsanalyse einen Rahmen zur Verfügung stellte. Die Ergebnisse dieser Diskussionen wurden 2007 in dem Band „Koordination. Analysen zur multimodalen Interaktion“ publiziert (Schmitt (Hg.) 2007). Nach dem Kolloquium

¹ Siehe Mondada/Schmitt (2007) für eine Darstellung dieses Projekts.

fiel die Entscheidung, die deutsch-französische Komponente als von Anfang an zentraler Bestandteil dieser multimodalen Initiativen zu einem eigenständigen Analysezusammenhang mit Projektstatus zu entwickeln. So entstand das Projekt „Vergleichende Analysen von Situationseröffnungen“/„Analyses comparées d'ouvertures“, das in methodisch-methodologischer Hinsicht eine unmittelbare Weiterführung des multimodalen Diskussionszusammenhangs mit einer neuen inhaltlich-thematischen Schwerpunktsetzung darstellt.

Die Wahl genau dieses Aspekts reflektiert dabei zentrale Einsichten der gemeinsamen Diskussionen, wie beispielsweise die folgende: Es ist notwendig, die im konversationsanalytischen Forschungszusammenhang entwickelten – primär verbal definierten – Konzepte im Lichte der neuen, audiovisuellen Qualität der Videodokumente zu überprüfen. Die zentrale Frage lautet dabei: Wie wirkt die Tatsache, dass bei Videodooaufzeichnungen auch die visuellen Anteile im interaktiven Verhalten der Beteiligten der Analyse zur Verfügung und im Fokus stehen, auf die etablierten konversationsanalytischen Konzepte zurück?

Die Wahl, die „multimodale Herstellung fokussierter Interaktion“ zum Gegenstand systematischer Analysen zu machen, reagiert auf diese durch die neue materiale Grundlage notwendig gewordene Konzeptreflexion. Da die Analyse von Gesprächseröffnungen für die Konversationsanalyse eine ganz zentrale Rolle gespielt hat – und auch heute noch spielt –, war es naheliegend, diesen Untersuchungsgegenstand aus einer multimodalen Perspektive aufzugreifen. Dass dies nicht zur „Analyse der multimodalen Herstellung von Gesprächseröffnungen“ führte, sondern „die multimodale Herstellung fokussierter Interaktion“ als Fokus ausweist, macht Folgendes deutlich: Es handelt sich tatsächlich um die Konstitution eines neuen Gegenstandes. Die Beiträge verstehen sich folglich nicht in erster Linie als eine nur um Visualität erweiterte Konversationsanalyse, sondern als Analysen in einem eigenständigen interaktionstheoretischen Erkenntniszusammenhang, der durch den Begriff „Multimodalität“ definiert ist.

Gleichwohl stellt die Konversationsanalyse für alle Beiträge einen zentralen Bezugspunkt in methodologischer Hinsicht dar, und professionsbiografisch sind alle Beiträger(innen) mehr oder weniger deutlich konversationsanalytisch „sozialisiert“. Daher skizzieren wir – bevor wir die einzelnen Analysen vorstellen – die konversationsanalytischen Wurzeln, die in der konstitutionsanalytischen Perspektive in allen Beiträgen deutlich werden (Kap. 2). Wie die Studien zeigen, beziehen sich die einzelnen Autoren jedoch in unterschiedli-

cher Deutlichkeit und Ausprägung auf diese methodologischen Grundlagen und damit auch thematisch auf die konversationsanalytischen Arbeiten zu Gesprächseröffnungen.

Dies spiegelt zum einen die unterschiedliche Nähe der Autor(inn)en zur konversationsanalytischen Methodologie wider. Zum anderen ist dies jedoch auch der Spezifik der analysierten Daten geschuldet. So legt die Analyse technisch vermittelter Interaktion, wie sie von Lorenza Mondada am Beispiel von Videokonferenzen durchgeführt wird, aufgrund materialer Beschaffenheit und der dabei konstitutiven Rolle der Verbalität einen stärkeren Bezug zu den konversationsanalytischen Ergebnissen und Konzepten zu Telefonkommunikation nahe. Bei der Analyse der gestreckten Eröffnungsphase eines Gottesdienstes mit Glockenläuten, Orgelspiel und der weitgehend stillen Kopräsenz der einzelnen Gottesdienstbesucher, wie sie von Heiko Hausendorf und Reinhold Schmitt vorgestellt wird, ist dies quasi „naturwüchsig“ etwas anders. Hier spielt vor allem die Sozio-Semiotik des speziell für die Realisierung des religiösen Rituals hergerichteten und mit unterschiedlichen Klängen erfüllten Raums eine größere Rolle als verbale Aktivitäten.

Im Anschluss an die Darstellung unserer konversationsanalytischen Wurzeln skizzieren wir unter dem Begriff „Multimodalität“ den für die Studien gemeinsamen interaktionstheoretischen Rahmen. Dabei erheben wir nicht den Anspruch, die sich aktuell entwickelnde interaktionstheoretische Konzeption, auf die Multimodalität verweist, vollständig darzulegen. Vielmehr fokussieren wir diese Darstellung auf unser spezifisches Erkenntnisinteresse an der Herstellung fokussierter Interaktion (Kap. 3). Diese Fokussierung führt dann im nächsten Schritt dazu, dass wir Konsequenzen und Implikationen präsentieren, die sich aus den vorherigen Ausführungen für die Analyse multimodaler Situationseröffnungen ergeben (Kap. 4). Die Einleitung endet schließlich mit einer Kurzbeschreibung der Beiträge (Kap. 5).

2. Konversationsanalytische Wurzeln

2.1 Die Analyse von Eröffnungssequenzen

Eröffnungssequenzen (*openings*) gehören zu den klassischen Themen der Konversationsanalyse. Sie sind Gegenstand der Dissertation von Schegloff (1967) und sind seitdem in den unterschiedlichsten Kontexten immer wieder analysiert worden. Die Untersuchung von Eröffnungssequenzen folgt dabei der allgemeinen Entwicklung der Konversationsanalyse: Als erstes dienten

aufgenommene Telefongespräche als Analysebasis (Kap. 2.2), dann folgte die Analyse von Gesprächen im institutionellen Rahmen, die ebenfalls auf der Grundlage von Audioaufnahmen durchgeführt wurden (Kap. 2.3). Erst seit geraumer Zeit wurden auch Videoaufnahmen herangezogen, um die multimodalen Aspekte des interaktiven Handelns in Phasen der Herstellung und der Eröffnung von Interaktion genauer untersuchen zu können (Kap. 2.4). Im Folgenden skizzieren wir die geschichtliche Entwicklung der Erforschung von Eröffnungssequenzen und schenken dabei der Erweiterung der technischen Möglichkeiten besonderes Augenmerk. Es wird deutlich werden, dass die Analyse von Eröffnungssequenzen stetig revidiert wurde. Diese permanente Revision ist zum einen in der technologischen Veränderung im Bereich der Aufnahmetechniken begründet, ist jedoch auch durch die generelle konzeptionelle Entwicklung der Konversationsanalyse motiviert.

2.2 Die Eröffnung von Telefongesprächen

Eröffnungssequenzen von Telefongesprächen spielen in der Geschichte der Konversationsanalyse eine wichtige Rolle. Sie sind nicht nur ein ausgezeichnetes Beispiel dafür, wie sich die 'Maschinerie' der sequenziellen Gesprächsentwicklung in Gang setzt; ihre Analyse diente vielmehr ganz grundlegend dazu, aufzuzeigen, dass Paarsequenzen, bestehend aus initiativen und reaktiven Sequenzteilen, als Aufforderung-Antwort-Sequenz (*summons-answer-sequence*) konzeptualisiert werden können und als prototypisches Sequenzpaar einen grundlegenden Teil der gesprächskonstituierenden Maschine darstellen.

Im Falle eines Telefongesprächs bildet das Klingeln des Telefons die Aufforderung und damit den ersten Teil der Paarsequenz. Schegloff (1968, S. 1080) hat von Anfang an darauf hingewiesen, dass in unterschiedlichen Kontexten den Teilnehmer(inne)n unterschiedliche Verfahren zur Verfügung stehen, um die Aufmerksamkeit anderer zu gewinnen. Er erwähnt beispielsweise verbale Verfahren wie die Anstandsformel „Entschuldigen Sie bitte...“ aber auch körperliche Verfahren wie „jemandem auf die Schulter klopfen“ und gestikulatische Verfahren wie „mit der Hand winken“ oder „das Aufhalten der Hand eines Publikumsteilnehmers“ etc. Das Sequenzpaar im Kontext der Telefoneröffnung war von Anfang so konzeptualisiert, dass es einerseits den technologischen Besonderheiten eines Telefongesprächs gerecht wird, andererseits jedoch generell für die unterschiedlichsten interaktiven Bedingungen Gültigkeit besitzt.

Die Paarsequenz ‘Aufforderung-Antwort’ ist fundamental, weil sie den Teilnehmer(inne)n erlaubt, auf koordinierte Art und Weise in die Organisation der Interaktion einzusteigen. Die Antwort, die auf die Aufforderung folgt, symbolisiert ganz grundsätzlich die Verfügbarkeit des Angerufenen. Sie erlaubt es auch, die Unterscheidung zwischen Ko-Präsenz und gegenseitiger Verfügbarkeit zu verstehen. Die Paarsequenz ‘Aufforderung-Antwort’ bildet somit eine Vor-Sequenz für die folgenden Sequenzen. Als solche enthält sie die Eigenschaft der ‘*non-terminality*’: Nach Beendigung des Antwortteils durch den Angerufenen hat der Anrufende die Verpflichtung, wieder zu sprechen und somit das Gespräch weiterzuführen (Schegloff 1968, S. 1081). Gleichzeitig steht der Angerufene in der Pflicht, zuzuhören (ebd., S. 1083).

Die mit der *non-terminality* zusammenhängenden gesprächsorganisatorischen Implikationen, die sich primär als jeweils spezifische Partizipationsverpflichtungen der beiden Beteiligten zeigen, verweisen ihrerseits auf das allgemeine interaktionskonstitutive Prinzip der konditionellen Relevanz (*conditional relevance*): Die Realisierung eines ersten Teils einer Paarsequenz etabliert die normative Erwartung auf einen spezifischen, d.h. korrespondierenden zweiten Teil. Im Falle der Nichterfüllung dieser Erwartung ist die Abwesenheit des zweiten Teiles als relevantes Phänomen sichtbar (Schegloff/Sacks 1973) und in solchen Fällen wird die Aufforderung wiederholt.

Sobald der koordinierte Einstieg der Teilnehmer(innen) in das Telefongespräch organisiert ist, erfolgt die Eröffnung des Gesprächs, in der in der Regel die wechselseitige Identifikation der Gesprächspartner erfolgt (Schegloff 1979). Die Gesprächseröffnung findet normalerweise in folgender Reihenfolge statt: die Begrüßung, der Austausch „wie geht es dir?“ sowie die Erörterung der Gründe für den Anruf (Schegloff 1986).

Schegloffs Analysen der Eröffnungssequenzen haben eine Vielzahl von Studien zu Telefongesprächen inspiriert. Die Dominanz gerade dieses Gesprächstyps für konversationsanalytische Untersuchungen hat seine Gründe primär in der technologischen Entwicklung der frühen Entwicklungsphase der Konversationsanalyse, die Tonbandaufzeichnungen zum Standard der empirischen Grundlage machte.

Bezogen auf die Reduktion der interaktiven Ursprungssituation und ihren Tondokumenten stellt die Telefonkommunikation einen Sonderfall dar (Mondada 2008c). Hier gibt es eine weit(er)gehende Übereinstimmung hinsichtlich der Komplexität der Ursprungssituation und des Dokuments als Analysegrund-

lage: Interaktionsbeteiligte und Analysierende haben in etwa den „gleichen“ Zugang zu den Interaktionsbeiträgen des Gesprächspartners: Beide hören, was gesprochen wird. Diese weitgehende Übereinstimmung von Ursprungsdaten und Interaktionsdokumenten ist – im Sinne der Lösung des grundsätzlichen methodologischen Problems, das dem Verhältnis dieser beiden Datentypen inhärent ist – ein Grund, mit dem Schegloff das analytische Interesse an Telefonkommunikation motiviert hat.

... in those days when there were substantial limitations of technology and cost of gathering the visually accessible data of interaction – limitations barely imaginable in today's world – these telephone materials appealed on other grounds as well. For studying copresent interaction with sound recording alone risked missing embodied resources for interaction (gesture, posture, facial expression, physically implemented ongoing activities, and the like), which we knew the interactants wove into both the production and the interpretation of conduct, but which we as analysts would have no access to. With the telephone data, the participants did not have access to one another's bodies either, and this disparity was no longer an issue. (Schegloff 2002, S. 288)

Interessanterweise finden sich in Schegloffs ersten Artikeln zur Sequenzkonstitution in Telefongesprächen zahlreiche Referenzen zu interaktiven Face-to-Face-Situationen. Diese Referenzen bleiben jedoch auf der Ebene des Anekdotischen und basieren ausschließlich auf ethnographischen Beobachtungen. Spätere Studien haben sich dann auch Eröffnungssequenzen gewidmet, die im Rahmen faktischer Wahrnehmungswahrnehmung (Hausendorf 2001) und somit mit direktem Blickkontakt zwischen den Beteiligten stattfinden. Interessanterweise wurden auch diese Untersuchungen zunächst auf der Grundlage von Audioaufnahmen von Interaktionen im institutionellen Rahmen durchgeführt. Erst zu einem späteren Zeitpunkt, als sich Video als dominantes Dokumentationsmedium bereits in der gesellschaftlichen Öffentlichkeit etabliert hatte, fand auch im Kontext der Konversationsanalyse eine Erweiterung der Gegenstandsdefinition hin zur Analyse der multimodalen Ressourcen statt. Dies führte dazu, dass das Modell der Telefongesprächseröffnung noch lange, nachdem die Videotechnologie bereits einsatzbereit war, als Standardwerk galt und dass nur wenig Arbeiten die faktische multimodale Komplexität von Interaktionen untersuchten.

Schegloffs Pionierarbeit zu Eröffnungen in Telefongesprächen hat zu zahlreichen Arbeiten über Telefongespräche in anderen Sprachen, Kulturen und Kontexten geführt, was zur Erweiterung der Diskussion über die universelle Natur ihrer Organisation beigetragen hat. Gesprächseröffnungen am Telefon wurden beispielsweise

- hinsichtlich ihrer kulturellen Variation untersucht (siehe beispielsweise Hopper/Koleilat-Doany 1989, Hopper et al. 1990/1991, Pallotti/Vacasia o.J., Luke/Pavlidou (Hg.) 2002),
- im Rahmen des konversationsanalytischen Ansatzes und der interkulturellen Soziolinguistik analysiert (zu Telefonanrufen in Frankreich siehe Godard 1977; in Holland, Houtkoop-Steenstra 1991; in Schweden, Lindström 1996; in Taiwan, Hopper/Chen 1996; in Samoa, Liddicoat 2000; in Hong Kong, Luke 2002; in Korea, Lee 2006; in Japan, Park 2002; in Italien, Vacasia 2003 und Bowles/Pallotti 2004),
- aber auch unter der Perspektive des Spracherwerbs fokussiert (zu Telefongesprächen von *non-natives* siehe Taleghani-Nikazam 2002, Bowles 2006).

Eine zweite Gruppe von Arbeiten hat dazu beigetragen, die Analyse von Telefonanrufen im Rahmen mediatisierter Situation weiterzuentwickeln. Hierzu gehören beispielsweise die Untersuchungen von

- Gesprächen mit Mobiltelefonen (Relieu 2002; Arminen 2002, 2005; Hutchby/Barnett 2005),
- Gesprächen mit Bildtelefonen (Fornel 1994),
- Radiosendungen (Hutchby 1999),
- Videokonferenzen (Meier 1998, Licoppe/Dumoulin 2007, Mondada 2007d und
- Gesprächen mit *virtual walls* (Bonu 2007, Relieu 2007); hierbei handelt es sich um eine Technik, bei der man die Beteiligten an anderen Orten nicht auf einem Bildschirm sieht, sondern auf einer virtuellen Wand, die in den Raum integriert ist.

Generell haben diese Arbeiten die visuelle Dimension der Interaktion in ihre Analysen integriert, indem sie sich auf Kontexte konzentriert haben, die den Vergleich mit Telefoninteraktion notwendig machten. Der Vergleich hat dementsprechend dazu geführt, auf medien- und situationsbedingte Variationen des Ursprungsmodells hinzuweisen und Abweichungen zu formulieren.

Ein dritter Typus von Arbeiten hat sich auf Telefongespräche im institutionellen Kontext konzentriert. Obwohl auch Schegloffs Aufnahmen selbst einem institutionellen Kontext entstammen, galt seine analytische Aufmerksamkeit nicht in erster Linie deren institutionellem Charakter. Er war vielmehr daran interessiert, fundamentale und generelle Mechanismen der Sequenziellen Interaktionskonstitution herauszuarbeiten. Den Arbeiten der 1980er und

1990er Jahre war hingegen daran gelegen, aufzuzeigen, wie der institutionelle Kontext durch spezifische Formate von Telefongesprächen widerspiegelt und konfiguriert wird.

2.3 Die Eröffnungen institutioneller Interaktion

In den 1980er Jahren führte die Untersuchung von „Kommunikation in Institutionen“ dazu, institutionelle Interaktion als einen spezifischen Gesprächstyp zu konzeptualisieren, der im Vergleich zu gewöhnlichen Gesprächen durch einen begrenzten Gebrauch von Formaten gekennzeichnet ist. Der Nachweis dieser Unterschiede erbrachte gleichzeitig den Beweis, dass beiden Gesprächsarten dieselbe Organisationsmatrix zugrunde liegt (siehe beispielsweise die Einführung von Drew/Heritage (Hg.) 1992).

Die Analyse der Telefoneröffnungen in institutionellen Kontexten zeigt, dass im Rahmen von Telefonanrufen an die Notfallnummer 911 der Polizeidienste in Amerika eine Spezifizierung und Reduktion der zentralen Eröffnungssequenz („specification and reduction of the core opening sequence“, Zimmerman 1992a, S. 435; 1992b) zu beobachten ist. Tatsächlich zeichnen sich diese Eröffnungen durch ein kompaktes Format aus: Indem die Begrüßung und der Austausch von „Wie geht es dir“ ausgelassen wird, wird die Antwort auf die Aufforderung, die Identifizierung und den Grund für den Anruf in einem einzigen Redebeitrag des Anrufs vereint (Whalen/Zimmerman 1987, S. 175ff.). Die möglichst rasche Einführung des Grundes für den Anruf ist charakteristisch für die Eröffnung von Telefongesprächen in einem institutionellen Kontext. Zudem zeigt die Analyse der institutionellen Eröffnungen, dass die Interagierenden in den ersten Redebeiträgen ihre Identitäten so aufeinander ausrichten, dass eine Fortsetzung des Gesprächs gewährleistet ist. So gesehen spiegelt das spezifische Format der Eröffnungen von Telefongesprächen in Institutionen nicht einfach den institutionellen Kontext. Das spezifische Format stellt vielmehr den Ort der reflexiven Realisierung dieses institutionellen Kontextes dar: Durch die wechselseitige Ausrichtung der Interaktionsbeteiligten und durch die Sichtbarmachung der Produktion und des Erkennens dieses Kontexts, (re-)produzieren sie gemeinsam die spezifische Form von Institutionalität (Heritage/Greatbatch 1991, Zimmerman 1992b). Seit Zimmermans grundlegenden Analysen sind zahlreiche Untersuchungen zu Gesprächseröffnungen in verschiedenen sozialen, medizinischen und polizeilichen Institutionen durchgeführt worden (Baker/Emmison/Firth 2001, Bergmann 1995, Danby/Baker/Emmison 2005, De Gaulmyn 1994 und Emmison/Danby 2007).

Ein anderes Phänomen der Eröffnungssequenzen von Telefongesprächen, das bei Analysen von Interaktionen im institutionellen Rahmen wiederholt untersucht wurde, sind *pre-beginnings*. Damit bezeichnet Zimmerman eine vor der Aufforderung als erstem Sequenzteil platzierte Phase (Zimmermann 1992a, S. 432; 1992b, S. 38ff.), auf die schon Schegloff aufmerksam gemacht hat (Schegloff 1979, S. 27). Diese Phase umfasst die Aktivität des Auswählens und Eintippens der Telefonnummer und setzt für den Anrufenden selbst eine Reihe von Erwartungen in Gang und etabliert eine spezifische Form der Vororientierung. Diese bezieht sich beispielsweise auf die Person, die den Anruf entgegennehmen wird und auf mögliche erwartbare Gründe für den Anruf. Die Mitarbeiter des angerufenen Dienstes haben hingegen Erwartungen dahingehend, dass der Anrufer spezifische Gründe hat; sei es, dass er Unterstützung braucht oder sich aktuell in einer Notfallsituation befindet. Der Vor-Beginn projiziert dementsprechend eine Ausrichtung, die während der Identifizierungssequenz der Eröffnung ratifiziert wird. Diese Phase ist jedoch auf den Audioaufzeichnungen, die die Basis dieser Analysen bilden, nicht dokumentiert (siehe jedoch Mondada (2008c) für eine auf Videoaufnahmen basierende Analyse dieser Phase). Gleichwohl beschäftigt sich Zimmerman (1992a, 1992b) analytisch mit solchen *pre-beginnings* auf der Basis von abweichenden bzw. auffälligen Beispielen. Zum einen handelt es sich um Fälle, in denen der Anrufer die Nummer versehentlich gewählt hat und schweigt, zum anderen um Fälle, bei denen es sich um Scherzanrufe handelt. Bei der Analyse dieser Anrufe zeigt Zimmerman, dass sich der Angerufene an Elementen der *pre-beginnings* orientiert, um mögliche Gründe für den Anruf und dessen Verstehbarkeit zu rekonstruieren.

Analysen von institutionellen Interaktionen weisen regelmäßig auf die Phase des Vor-Beginns hin. Dies trifft beispielsweise auf die Analysen von medizinischen Interaktionen zu, in denen Eröffnungssequenzen untersucht wurden, die einen wichtigen institutionellen Kontext darstellen (Bergmann 1979, Heath 1981, ten Have 1988 und Ruusuvoori 2000). Heritage/Robinson (2006, S. 57f.) weisen darauf hin, dass die Verabredung eines Termins beim Arzt und der Weg zur Praxis einen Vor-Beginn darstellen, durch den der Patient sich auf den Arztbesuch als legitimes und aus medizinischer Hinsicht gerechtfertigtes Unterfangen ausrichtet, und so sein „Patient-Sein“ aufbaut.

Die Eröffnung von Arztkonsultationen gehört zu den Bereichen, in denen sehr früh Videodokumente ausgewertet wurden. Dies gilt vor allem für die Arbeiten von Heath (1981, 1986). Seine Analysen haben gezeigt, dass der Aufbau von Wechselseitigkeit und Verfügbarkeit für einen gemeinsamen Austausch

schrittweise erfolgt. Heath berücksichtigt das Eintreten des Patienten in den Konsultationsraum, die Art und Weise wie der Patient gebeten wird, sich hinzusetzen und wie der Arzt den Übergang von der Einsichtnahme der Patientenakte hin zum Anfang der eigentlichen Konsultation organisiert. Dabei unterstreicht Heath die Wichtigkeit der gegenseitigen Blickkontakte und die Blickorganisation des Patienten, um den Arzt in seiner Eröffnung zu unterstützen. Die gegenseitige Bereitschaft zu beginnen („mutual readiness to start“) wird somit durch das Zusammenspiel beider Parteien erreicht, wobei sie visuelle, vokale und verbale Ressourcen einsetzen. Obwohl schließlich der Arzt die Konsultation mit einem „Was kann ich für Sie tun?“ initiiert, ist diese Initiation das Ergebnis einer gemeinsamen Herstellung. Hierzu trägt der Patient mit seinen Blicken und seiner Haltung bei und lockt den ersten Themenbeitrag des Arztes gewissermaßen aus diesem heraus (Heath 1986, S. 24ff.).

Die Rolle der Blicke und der Körperhaltung in der Übergangsphase der Eröffnung hin zur Initiierung der eigentlichen Konsultation wurde auch von Robinson (1998) und Modaff (2003) erwähnt. Andere Arbeiten zur Eröffnung von Arztkonsultationen rücken die Präsentation des Problems und die Beschreibung des Grundes für den Arztbesuch („accountable reason for the visit“) ins Zentrum ihres Erkenntnisinteresses. Diese Aspekte sind für den Aufbau der Glaubwürdigkeit und Legitimität des Patientenproblems und des Arztbesuches als ein ‘durch den Arzt behandelbares’ Problem zentral (Heritage/Robinson 2006). Der Begriff „doctorable“ ist durch Mehans Arbeiten zu Anrufen an die Polizei inspiriert. Mehan (1989) charakterisiert die für den Anruf angegebenen Gründe als „policeable“. Die Analyse des Grundes für einen Anruf und die Einführung des ersten Themas (Schegloff 1986), das die Phase der Eröffnung beendet, erlauben zweierlei: Zum einen kann die Ausrichtung beider Parteien beschrieben werden, zum anderen kann auch die Konvergenz beziehungsweise mögliche Divergenz zwischen der Präsentation des zu behandelnden Problems aus Patientensicht und der Eröffnungsfrage des Arztes rekonstruiert werden (Robinson 2006, Heritage/Robinson 2006).

Neben den ärztlichen Konsultationen haben sich konversationsanalytische Untersuchungen auch im Kontext von Arbeitssitzungen für die Struktur der Eröffnungssequenz inklusive des Vor-Beginns in systematischer Weise interessiert. In einem bahnbrechenden Artikel weist Turner (1972) darauf hin, dass die Art und Weise, in der die Beteiligten am Sitzungsort eintreffen, sich setzen und zusammenkommen, bereits ihr Wissen um die Eigenheiten der kommenden Sitzung und ihre Ausrichtung auf dieselbe deutlich macht. Nach Turner ist bereits auch die kategoriale Zugehörigkeit der Beteiligten erkennbar, in der sie

an der Sitzung teilnehmen werden. Der Vor-Beginn bildet in diesem Sinne eine Phase, während der sich die Beteiligten auf die kommende Aktivität und gleichzeitig auf spezifische Aktivitäten der Phase, die der Eröffnung vorangehen wird, ausrichten. Der Beginn („beginning“) der Sitzung selbst kann aufgrund eines vordefinierten Zeitplans geschehen oder das Resultat einer Ko-Präsenz von gewissen Beteiligten sein (Turner 1972, S. 371). In diesem zweiten Fall richtet sich die Eröffnung der Sitzung auf die für die kommende Aktivität relevanten Kategorien. Turners Analyse zeigt, dass die Herstellung der Eröffnung einer Therapiesitzung einen Wechsel in der Art und Weise mit sich bringt, wie die Beteiligten verbale Beiträge behandeln. Von dem Moment an, in dem die Sitzung beginnt, bilden Gesprächsbeiträge der Patienten keine Konversationsbeiträge mehr, sondern gelten als Gegenstände („data“), die vom Therapeuten analysiert werden. Die Eröffnungen von Sitzungen sind von Boden (1994, S. 92ff.) als „premeeting talk“ analysiert worden: Es sind Aktivitäten, die vor dem eigentlichen Beginn der Sitzung stattfinden, aber auf diesen ausgerichtet sind. Mirivel/Tracy (2005) unterscheiden im Verlauf des *premeeting talk* folgende Phasen: *small talk*, *work talk*, *preparatory work* und *shop talk*. Oft ist der Übergang zum Beginn der Sitzung durch Markierungen (z.B. ‘so’, ‘gut’, ‘uh’) und der Initiierung des ersten Themas, das als Grund der Sitzung präsentiert wird, erkennbar (Boden 1994, S. 97f.).

Auch für die Mehrheit dieser Arbeiten gilt, dass sie auf der Basis von Audiodokumenten durchgeführt wurden. Die Überlegungen zum Vor-Beginn (*pre-beginning*) sind um einiges älter als die allgemeine Verfügbarkeit von Video als Dokumentationsmedium. Interessanterweise wurden sie aber nicht mehr weiter entwickelt, obwohl in der Zwischenzeit die Dokumentation dieser interaktiven Momente in Form von Videoaufnahmen grundsätzlich möglich ist.

Für die Analyse und Konzeptualisierung von Interaktion als multimodalem Gesamtzusammenhang müssen also bestimmte historische Bedingungen gegeben sein. Hierzu zählt ganz wesentlich – das ist bereits hervorgehoben worden – die technologische Entwicklung im Bereich der Videodokumentation und der Videoanalyse. Doch es bedarf weitaus mehr als nur dieser technischen Voraussetzungen, um wissenschaftlich eine neue Sicht auf Interaktion zu entwickeln. Es müssen vielmehr auch die wissenschaftsinternen Bedingungen und die wissenschaftliche Relevanz gegeben sein, diese vorhandene Technologie unter der Perspektive systematischer Erkenntnisgenerierung einzusetzen. In dieser Hinsicht sind Verzögerungsphasen in der wissenschaftlichen Adaption technischer Entwicklungen interessant, die sich – so hat man fast den Eindruck – mit einer gewissen Systematik wiederholen:

Sowohl die „flächendeckende“ Adaption der technischen Möglichkeiten der Audiodokumentation als auch die der audiovisuellen Dokumentation hinken dem technisch Machbaren sowohl bei der Entwicklung der Konversationsanalyse als auch der „Multimodalität“ erkennbar hinterher (Mondada 2006, 2008c und 2009).

2.4 Die Eröffnungen in Face-to-Face-Situationen

Konversationsanalytische Untersuchungen, die sich in Form einer expliziten Gegenstandskonstitution mit der faktischen multimodalen Komplexität von Situationseröffnungen beschäftigen, sind auch nach wie vor eher eine Seltenheit.

Etwas anders stellt sich diese Situation hinsichtlich des Aspektes „Begrüßung“ im Bereich der Soziolinguistik, der Ethnographie der Kommunikation und der anthropologischen Linguistik dar. Hier gibt es durchaus Arbeiten, die bei der Analyse von Begrüßungen die Wichtigkeit des räumlichen Kontextes, der körperlichen Nähe, der Gestikulation und der Blickorganisation als integrale Bestandteile der Eröffnungsorganisation zumindest theoretisch hervorheben und den Aspekt der Multimodalität diesbezüglich als Grundlage formulieren. Aber auch hier werden die multimodalen Ressourcen als für die Interaktionskonstitution zentrale Grundlagen analysepraktisch kaum ernsthaft berücksichtigt (beispielsweise Ag-Youssouf/Grimshaw/Bird 1976, Collett 1983, Duranti 1992, Irvine 1974, Salmond 1974, Caton 1986 und Goody 1972 zu außereuropäischen Kulturen). Andererseits haben Wissenschaftler, die sich für Gestik und andere modale Verhaltensweisen interessieren, eine generelle Beschreibung der Begrüßungssequenz geliefert, die typisierenden Charakter besitzt und konventionelle Haltungen mit Sprache und Kultur verbindet. Beispielsweise beobachten Greenbaum/Rosenfeld (1980) Begrüßungen in Flughäfen, die sie geschlechtsspezifisch typisieren. Argyle (1975, S. 78) bemerkt wiederum, dass die körperlichen Verhaltensweisen der Begrüßung zahlreiche Variationen aufweisen. Aufgrund dieser Beobachtung schlägt er eine Typologie der Begrüßungs-Haltungen und Begrüßungs-Gestiken vor, die quer durch die unterschiedlichen Kulturen am häufigsten vorkommen und erstellt eine Liste der in den verschiedenen Kulturen am häufigsten auftretenden Aspekte. Hierzu zählen beispielsweise:

- der *eyebrow flash* (auf dessen Einsatz für Begrüßungszwecke Eibl-Eibesfeld (1972) zum ersten Mal hinweist),
- Lächeln,

- wechselseitige Blicke,
- räumliche Nähe,
- körperlicher Kontakt (der für Kulturen, die den Körperkontakt ansonsten nicht privilegieren, in diesem Zusammenhang charakteristisch ist),
- die Präsentation der Handinnenfläche (*presenting palm*) oder
- der *head toss*.

Aufgrund der existierenden Literatur identifiziert Duranti (1997) sechs universelle Eigenschaften der Begrüßungsorganisation, zu denen auch nicht-verbale gehören. Obwohl der Autor diese im Verhältnis zur verbalen Dimension ansonsten vernachlässigt (ebd., S. 68), insistiert er jedoch beispielsweise darauf, dass die Begrüßung dazu beiträgt, einen gemeinsamen Orientierungsrahmen zu schaffen, den die Beteiligten teilen (ebd.). Das Anzeigen des Erkennens/der Identifikation kann durch visuelle oder verbale Ressourcen realisiert werden. Wichtig ist dabei, dass die Form und die Modalität der Begrüßungen bereits Ausdruck der wechselseitigen Identifikation und der sozialen Kategorisierung der Beteiligten sind (ebd., S. 71). So ist es durchaus wichtig zu wissen, ob die Person gegenüber ein „Feind“ ist oder aber ein Hausangestellter oder ein Kind. Die adäquate Einschätzung ist dann die Voraussetzung dafür, das eigene Verhalten entsprechend anzupassen: beispielsweise dem Feind ausweichen oder sich angemessen gegenüber dem Kind verhalten.

In Samoa kontextualisiert die Form der Begrüßung und deren Position im Verhältnis zum Aktionsablauf den Kontext und die Aktivität. So zeigen Durantis Analysen zu den Begrüßungsaktivitäten (1992, 1997), dass gewisse Begrüßungen innerhalb der Begegnung relativ spät stattfinden. Die Begegnung beginnt mit einem Austausch von Witzen, Fragen, Kommentaren und sieht vor dem Beginn der offiziellen Zeremonie keinen Austausch von Begrüßungen vor. Sherzer (1983) macht bei den Kunas ebenfalls die Beobachtung einer verspäteten Begrüßung. Generell lassen diese ethnographischen Beobachtungen (Durantis Arbeit basiert auf Videoaufnahmen) auf eine Art von Vor-Beginn-, Vor-Sitzungs-Gespräch schließen, das den formellen, zeremoniellen Aktivitäten eigen ist.

Wie die Literatur zeigt, lassen sich auch im ethnographischen Forschungskontext nur wenig detaillierte Analysen finden, die auf der Grundlage von Videoaufzeichnungen durchgeführt worden sind. Eine Ausnahme bildet in diesem Zusammenhang jedoch die wegweisende Arbeit von Kendon/Ferber (1973), die von Goffmans Untersuchung (1963, S. 91f.) zum Blickaustausch als *clear-*

ance signal innerhalb des Übergangs von der *unfocused* zur *focused* Interaktion inspiriert ist. Indem Kendon und Ferber auf der Grundlage von Filmmaterial und multimodalen Transkriptionen arbeiten, ist es ihnen möglich, die Details von Begrüßungen in Kontexten, in denen die gegenseitige visuelle Wahrnehmung der Beteiligten aufgrund der räumlichen Distanz der Annäherung vorangeht, als gestalthaften interaktiven Moment („complete interactional event“, Kendon 1990b, S. 20) zu erfassen.

Die Begrüßungen kann durch wechselseitige Wahrnehmung (*sighting*) oder durch die explizite Präsenzankündigung eines Beteiligten initiiert werden (beispielsweise durch Husten, durch Klopfen an der Tür oder indem ein Beteiligter den anderen beim Namen ruft; Kendon 1990d, S. 163). Hierbei handelt es sich um Verfahren, die denen in der Konversationsanalyse beschriebenen Anforderungen gleichkommen. Sobald sich die Beteiligten visuell identifizieren (*clearance*), initiieren sie ihre Annäherung, die in ihrer Realisierung auf die räumlichen Gegebenheiten des Ortes reagiert (die analysierten Geburtstagsgäste überschritten beispielsweise den zentralen Raum der Party nur, wenn sie vorher vom Gastgeber begrüßt wurden – dies traf vor allem auf Gäste zu, die mit dem Gastgeber nur weitläufig bekannt waren). Während der Annäherung kann durch ein Kopfnicken oder eine Handbewegung eine Begrüßung auf Distanz ausgetauscht werden. Charakteristisch für die Annäherungsphase bis zum Moment der bevorstehenden Begrüßung ist das Vermeiden von Blickkontakt und Auto-Kontakten (*grooming*). Von dem Moment an, an dem die Vereinigung der Beteiligten stattgefunden hat, ist die Begrüßung durch einen intensiven Blickaustausch, durch Lächeln, Handpräsentation (*palm presentation*) als Vorbereitung für den Handschlag gekennzeichnet. Zudem kann die Begrüßungsphase auch Körperkontakt einschließen.

Die Studie von Kendon/Ferber (1973) wird ein Jahr vor Sacks/Schegloff/Jefferson (1974) zum *turn taking* publiziert, bleibt aber im konversationsanalytischen Kontext weitgehend unbeachtet. Kürzlich wurden jedoch gewisse Elemente dieser Perspektive vertieft. So analysiert Pillet-Shore (2008) auf der Basis von Videoaufzeichnungen die multimodalen Praktiken des Zusammenkommens (*coming together*) in der Eröffnungsphase von Face-to-Face-Interaktionen. Sie zeigt, dass sich die Beteiligten bei der Organisation der Eröffnung an sichtbaren und symbolhaltigen Grenzen (z.B. Türen) orientieren: Sie erfragen den Zutritt (*admission gained from pre-present parties*) oder bewilligen ihn sich selber (*self-admission*). Ist der Zutritt gewährt, orientieren sie sich auf die Struktur des Raumes und erst, wenn diese erkannt ist, richten sich die Neankömmlinge wieder auf den Eingang aus und realisieren ‘darf-ich-

eintreten'-Aktivität („‘may-I-enter’ action“, Pillet-Shore 2008, S. 20). Eine solche Aktivität kann beispielsweise durch eine zögerliche Körperhaltung erfolgen, die dem Warten auf Einlass Ausdruck gibt. Die folgenden Phasen umfassen

- die Begrüßungen in ihrer multimodalen Realisierung,
- die Vorstellungen,
- die Verfahren, mit denen sich die Beteiligten setzen und im Raum einrichten (hierzu zählen die Übergabe der Mitbringsel, der Austausch von Geschenken, das Wiederherrichten der Kleider und die Positionierung im Raum) und anhand derer sie ihre Präsenz markieren (*doing 'how I'm coming here'*),
- der Austausch von „Wie geht es dir?“,
- der Austausch von Scherzen und Formulierungen, welche die Verbindung mit früheren Begegnungen wieder herstellen und explizieren.

Die genaue Beschreibung der verschiedenen Phasen einer Annäherung oder der Ankunft der geladenen Gäste zeigt das Ineinandergreifen der verschiedenen multimodalen Verhaltensweisen. Es wird deutlich, dass diese Verhaltensweisen sequenziell organisiert sind und sich auf die Eigenheiten des Raumes, der persönlichen Beziehungen und der Begegnung ausrichten und sich ihnen anpassen. In diesem Sinne unterscheidet sich der Besuch von Freunden zu Hause (Traverso 1998, Pillet-Shore 2008) erkennbar von einer öffentlichen Veranstaltung mit Unbekannten (Mondada i.Dr.) oder einer Situation, die der von Kendon/Ferber (1973) analysierten Geburtstagsparty gleicht. Hier erfolgt die Annäherung jeweils schrittweise nach und nach aus der Bewegung heraus durch einen ersten Blickkontakt, gefolgt von einem ersten Redebeitrag, der die Positionierung und die Koordination der Körper im Raum und den Einstieg in die Unterhaltung begleitet.

Die Eröffnungen von Face-to-Face-Interaktionen werden somit einerseits mit Hilfe ähnlicher sequenziell organisierter Aktivitäten vollzogen wie die eines Telefongesprächs. Andererseits kommen aber bevor das erste Wort gesprochen wird auch systematisch körperliche, nur visuell wahrnehmbare Ressourcen zum Einsatz. In den nachfolgend präsentierten Studien sind es gerade diese visuellen Ressourcen, die zur Anbahnung und Herstellung einer Interaktion eingesetzt werden – wie Kopfbewegungen, Richtung des Blickes, wechselseitige Wahrnehmung, Orientierung des Körpers und die Koordination der wechselseitigen Annäherung – die neben und vor allem vor dem Einsatz von Verbalität bei den Analysen systematisch berücksichtigt werden.

3. Der interaktionstheoretische Rahmen: Multimodalität

Im Folgenden wollen wir – zumindest ansatzweise – die vielfältigen methodisch-methodologischen und theoretischen Implikationen skizzieren, die mit einer Konzeption von Interaktion als multimodalem Ereignis verbunden sind. Wir konzentrieren uns dabei vor allem auf diejenigen Implikationen, die sich insbesondere im Hinblick auf unseren zentralen Gegenstand „Eröffnung von fokussierter Interaktion“ ergeben.

3.1 Dokumentierte Situation und Interaktionsdokument

Geht man davon aus, dass die technisch vermittelten Interaktionsdokumente, die wir als empirische Basis unseren Analysen von Interaktion zu Grunde legen, der dokumentierten Ursprungssituation möglichst nahe kommen sollten, dann wird Folgendes deutlich: Wenn man Interaktion untersucht – sei es als komplexe soziale Aktivität oder im engeren Sinne als Sprachgebrauch – hat man es zwangsläufig mit Multimodalität als nicht hintergehbare Qualität der dokumentierten Ursprungssituation zu tun. Interaktion als multimodal konstituiertes soziales Ereignis ist sozusagen der empirische „Normalfall“ und die Bezeichnung ‘multimodale Interaktion’ in diesem Sinne ein „weißer Schimmel“.

Die multimodale Konzeption von Interaktion ist kein neuer Ansatz, der aus dem Nichts auftaucht. Sie hat vielmehr tiefe Wurzeln in der qualitativen Erforschung von Interaktion und der soziologisch begründeten Interaktionstheorie. Eine zentrale Rolle für die empirische Untersuchung und Konzeptentwicklung des Zusammenhangs unterschiedlicher Modalitätsebenen spielt der Forschungsansatz der *context analysis* (Birdwhistell 1970, Schefflen 1972 und Kendon 1990b, c). Der Ansatz spielte zwar für die konversationsanalytische Entwicklung kaum eine Rolle, wird jedoch im Zuge der multimodalen Perspektive auf Interaktion wieder entdeckt (Müller/Bohle 2007).

Multimodalität ist also keine neue Analysemethode oder eine neue Methodologie, sondern bezeichnet eine konstitutive Qualität von Interaktion. Wissenschaftsgeschichtlich treten zu dem bislang dokumentierten Hörbaren der Tonbandaufzeichnungen nun auch die sichtbaren Anteile des interaktiven Verhaltens der Beteiligten hinzu und werden somit zum Untersuchungsgegenstand, teilweise als eigenständige Untersuchungsaspekte (beispielsweise Koordination als zentrale Interaktionsvoraussetzung und kontinuierliche Leistung der Beteiligten; vgl. Schmitt (Hg.) 2007), teilweise in systematischem Bezug

zu bereits etablierten Untersuchungsaspekten der Konversationsanalyse (beispielsweise die Rolle visueller Verhaltensweisen für die Turn-Taking-Organisation; Mondada 2007b, Schmitt 2005).² Im engeren Sinne kann man also auch nicht von einer multimodalen Analyse sprechen. Multimodalität ist vielmehr ein theoretischer Bezugsrahmen, der – in unserem Verständnis – auf eine Form der Interaktionsanalyse verweist, die auf der Basis von Videodokumenten authentischer Interaktionssituationen auf der Grundlage der konstitutionsanalytischen Methodologie der Konversationsanalyse durchgeführt wird.³

In dieser Hinsicht unterscheidet sich unsere auf der konversationsanalytischen Methodologie aufsetzende interaktionistische Konzeption von Multimodalität von der im englischsprachigen Raum entstandenen *multimodal discourse analysis*.⁴ Die *multimodal discourse analysis* beschäftigt sich mit der Materialität unterschiedlicher Interaktionskanäle, sei es in Texten (als Ineinandergreifen von Text und Bild), sei es in der Interaktion (als Alignment von Gesten, Gegenständen und Sprache), sei es in der technisch vermittelten Interaktion (Organisation der Chat-Bildschirme und Webseiten etc.). Diese Fokussierung auf den jeweiligen Kanal behandelt die konkrete Materialität als eine notwendige Voraussetzung und als einen semiotischen Rahmen, der mit spezifischen Eigenschaften ausgestattet ist, nicht jedoch wie eine Ressource aus Plastik, deren Benutzung jeweils lokal durch die Beteiligten/Benutzer definiert wird. Im Verständnis dieses Ansatzes bezieht sich ‘Multimodalität’ als Oberbegriff auf die unterschiedlichsten Veränderungen im Bereich der Kommunikation, die mit dem technologischen Wandel und den durch ihn hervorgebrachten neuen Kommunikationsmedien zusammenhängen.

Die für die Konversationsanalyse lange Zeit vorherrschende analytische Konzentration auf das Verbale stellt demgegenüber eine weitgehende Reduktion der Ausgangsdaten dar. Von der multimodalen Komplexität der dokumentierten Situation bleibt im Interaktionsdokument bei Audioaufzeichnungen nämlich nur das Verbale oder besser gesagt: das Tonale, das Hörbare erhalten. Verlässt man die nicht nur technisch, sondern auch methodologisch

² Das Konzept „multimodale Kommunikation“ ist bereits in Schefflen (1972, S. 230) terminologisch vorgeprägt, der von Körperpositur als einer „modality of communication“ spricht.

³ Eine vergleichbare interaktionistische Ausrichtung zeichnet beispielsweise den Sammelband von Sidnell/Stivers (Hg.) (2005) aus.

⁴ Siehe hierzu beispielsweise den von LeVine/Scollon (Hg.) (2004) herausgegebenen Sammelband „Discourse Analysis and Technology“ und O'Halloran (Hg.) (2004); weiterhin Taylor (1989, 2000), Kress/Van Leeuwen (2001), Jewitt/Kress (2003) sowie Norris (2004).

motivierter Selbstbeschränkung auf Telefonkommunikation und betritt den Bereich der Face-to-Face-Situation, in der die Interaktionsbeteiligten unter Bedingungen wechselseitiger Wahrnehmungswahrnehmung interagieren, wird die strukturelle Reduktion des technisch vermittelten Transfers situierter Ursprungsdaten in Interaktionsdokumente zum Gegenstand methodologischer Aufmerksamkeit.

Die Implikationen der strukturellen Reduktion spielten jedoch bei der Erweiterung der empirischen Datenbasis für konversationsanalytische Untersuchungen, die ihre anfängliche Fokussierung auf Telefongespräche aufgab und sich verstärkt der faktischen multimodalen Komplexität von Face-to-Face-Interaktion zuwendete, keine nennenswerte Rolle. Die strukturelle Reduktion der Ursprungssituation wurde eher beiläufig erwähnt, selten jedoch hinsichtlich der damit verbundenen Implikationen reflektiert. Verbalität wurde auch dann noch als zentrale, eigenständige und de facto unabhängige Ausdrucksmodalität behandelt, als sie in der multimodalen Komplexität der Ursprungssituation nur noch ein Teil des interaktiven Austauschs war.

Man braucht sich nur einmal einen beliebigen Videoausschnitt anzusehen, um einen Einblick in die Vielfalt der unterschiedlichen, aufeinander bezogenen und miteinander koordinierten Aktivitäten auf verschiedenen Modalitätsebenen zu bekommen. Man begreift schlagartig die Gleichzeitigkeit und Komplexität der an der Interaktionskonstitution insgesamt beteiligten Ausdrucksmöglichkeiten der Beteiligten. Wenn wir in Face-to-Face-Situationen interagieren, soziale Bedeutung konstituieren, individuelle Ziele verfolgen und uns selbst und andere dabei in sozial relevanter Weise positionieren und als mit bestimmten sozialen Eigenschaften ausgestattete Personen präsentieren, tun wir das immer auf der Grundlage multimodaler Ausdrucksressourcen.

3.2 Modalitätsebenen

Zu den unterschiedlichen Dimensionen oder Modalitätsebenen, die bei der Interaktionskonstitution eine Rolle spielen und die in der grundsätzlichen Körperlichkeit und Raumgebundenheit der „Inter-Aktion“ basieren, gehören beispielsweise:

- Verbalität (Prosodie inbegriffen) und ihre Körperlichkeit
- Vokalität
- Blick

- Kopfbewegung
- Mimik
- Gestikulation
- Körperpositur
- Bewegungsmodus
- Präsenzform
- Proxematik: Verteilung im Raum: Positionierung der Beteiligten (*face to face, side by side, back to face ...*), Nähe/Distanz
- Räumlichkeit/Materialität: Manipulation von Objekten (Tools, Instrumente etc.)

Aus einer multimodalen Perspektive auf Interaktion werden die unterschiedlichen Dimensionen oder Modalitätsebenen theoretisch als grundsätzlich gleichwertig konzipiert. Keine der Dimensionen ist demnach für die Analyse der Interaktionskonstitution zentraler oder wichtiger als andere: Blickkontakt oder Körperpositur kommen als Modalitäten theoretisch, methodisch und konzeptuell der gleiche Status zu wie Verbalität. Konkret tritt eine Modalität nur dann in den Vordergrund, wenn sie von den Beteiligten selbst zur situationsensitiven Realisierung eines bestimmten Handlungszusammenhangs erkennbar relevant gesetzt wird.

Dies ist eine Sicht, die mit zentralen Annahmen der *context analysis*, wie sie von Kendon vertreten wird, übereinstimmt. Kendon (1990b, S. 16) betont „the importance of an integrated approach to the study of interaction“ und „refuses to assume that any particular modality of communication is more salient than another“.

Das Primat der Verbalität bei der Untersuchung von Kommunikation zu Gunsten eines Ansatzes aufzugeben, der alle Modalitätsebenen bei der Analyse und theoretischen Konzeption von Kommunikation berücksichtigt, bedeutet nun jedoch nicht, die empirische Faktizität außer Acht zu lassen, dass Interaktionsbeteiligte zur Bearbeitung bestimmter interaktiver Anforderungen spezialisierte Modalitäten einsetzen. Diese spezialisierten Modalitäten tragen jedoch nie die Gesamtlast der Interaktionskonstitution, sondern tun dies immer im koordinierten Zusammenspiel mit anderen Modalitäten. Sie müssen folglich in genau diesem koordinierten Zusammenhang rekonstruiert werden.

3.3 Konsequenzen der materialen Spezifik

Aus dem grundlegenden materialen Unterschied zwischen Audioaufnahmen und den audiovisuellen Videoaufzeichnungen ergibt sich eine zentrale und weitreichende Konsequenz für die Interaktionsanalyse, die sich zunächst als methodische Anforderung zeigt. So lange man sich ausschließlich in der Modalität verbaler Aktivitäten bewegt, ist es durchaus ausreichend, den Prozess der Interaktionskonstitution als sequenzielle Abfolge wechselnder Sprecher(innen)beiträge zu konzipieren. Mit Videoaufzeichnungen als empirische Basis tritt nun jedoch neben der Sequenzialität als primäres Rekonstruktionsprinzip die Anforderung, der Gleichzeitigkeit sehr unterschiedlicher modalitätsspezifischer Aktivitäten analytisch gerecht zu werden. Aus einer multimodalen Analyseperspektive sind beide Konstitutionsprinzipien konstitutiv für die Herstellung interaktiver Ordnung: Sequenzialität *und* Simultaneität.

Die strukturell nicht aufhebbare Gleichzeitigkeit sehr unterschiedlicher modaler Aktivitäten und die theoretische Gleichwertigkeit aller Modalitätsebenen sind der Ausgangspunkt für eine Vielzahl von Konsequenzen, die sich im Vergleich zur verbalen Analyse von Interaktion für Analysen im theoretischen Rahmen einer Konzeption von Multimodalität ergeben.

Wir werden im Folgenden – ohne Anspruch auf Vollständigkeit – exemplarisch Implikationen darstellen, die sich zwangsläufig aus einer multimodalen Konzeption von Interaktion ergeben. Diese Implikationen beziehen sich auf die Konstitution neuer Untersuchungsgegenstände (Kap. 4.1), die Entwicklung gegenstandsadäquater Analysemethoden (Kap. 4.2) und die Reflexion zentraler methodologischer Aspekte (Kap. 4.3).⁵

3.3.1 Die Konstitution neuer Untersuchungsgegenstände

3.3.1.1 Raum

Dass Interaktion immer auch ein raumbezogenes und durch Raum definiertes Unternehmen ist, wurde bereits früh erkannt und führte beispielsweise bei Kendon (1990c) zur Formulierung so genannter *F-formations*. Raum stellt jedoch nicht nur eine wesentliche Konstituente für Interaktion dar. Raum wird auch als Ressource bei der Lösung spezifischer interaktiver Probleme eingesetzt (LeBaron/Streeck 1997). Dies ist beispielsweise der Fall, wenn Interaktionsbeteiligte in das Blickfeld anderer treten, um diese zu bestimmten Handlungen zu veranlassen (Deppermann/Mondada/Schmitt i.Dr.). Und Raum

⁵ Siehe auch Schmitt (2007a, b).

wird von den Interaktionsbeteiligten nicht als etwas Gegebenes betrachtet. Raum als für die Interaktion wichtiger Relevanzrahmen muss vielmehr für die jeweiligen Zwecke von Interaktion in Form von „Interaktionsräumen“ situativ erst hergestellt werden (Schmitt/Deppermann 2007, Mondada 2007c und Müller/Bohle 2007).

Interaktionsräume werden konstituiert durch das Zusammenspiel von physikalischen Gegebenheiten, die auf Grund ihrer Beschaffenheit bestimmte Implikationen für die Strukturierung von Interaktion haben, und interaktiven Herstellungsleistungen, bei denen Beteiligte diese Gegebenheiten für ihre thematisch-pragmatische Praxis als Ressource nutzen. Interaktionsräume sind mit bestimmten Relevanzstrukturen verbunden, die beispielsweise in der Symbolisierung von Inklusivität und Exklusivität zum Ausdruck kommen. Das Konzept 'Interaktionsraum' beschreibt dynamische, sich stetig verändernde Konstellationen, die teilweise klare räumliche Konturen aufweisen. Es verweist nicht auf statische, gegenständlich-territoriale Gebilde.

3.3.1.2 Koordination

Neben der von der Konversationsanalyse bislang primär fokussierten verbalen Handlungsebene laufen noch eine Vielzahl zumeist gleichzeitiger Aktivitäten ab. Sie sind zwar auf die Handlungsebene bezogen und für diese funktional, können jedoch in ihrer interaktiven Bedeutung weder automatisch noch systematisch mit einer Handlungsanalyse erfasst werden. Solche Aktivitäten fassen wir mit dem Begriff 'Koordination'. Koordination stellt eine permanente Anforderung an alle Interaktionsbeteiligte dar und ist in diesem Sinne interaktionskonstitutiv. Aus multimodaler Perspektive ist Koordination ein Untersuchungsgegenstand *sui generis* (Deppermann/Schmitt 2007, Schmitt (Hg.) 2007).

Koordinative Aktivitäten sind solche Verhaltensweisen, die im Zusammenhang und zeitgleich mit verbalen Kooperationsbeiträgen und als deren Voraussetzung in den unterschiedlichen Modi körperlichen Ausdrucks realisiert werden. Sie selbst stellen jedoch keine zielorientierten und handlungsschematisch bezogenen Beiträge dar. Im Gegensatz zu Kooperation zielen koordinative Aktivitäten von Beteiligten nicht auf die Herstellung eines gemeinsamen Produkts und haben in diesem Sinne auch keinen produktspezifischen Handlungscharakter. Es sind vielmehr Anforderungen, die bei der Analyse als Voraussetzung inhaltlicher Kooperationsbeiträge in den Blick kommen.

Koordination als interaktionskonstitutive Anforderung wirkt sich teilweise auch unmittelbar auf die verbalen Aktivitäten aus. Dies ist beispielsweise der Fall, wenn eine Sprecherin ihre begonnene Äußerung solange unterbricht, bis ein anderer Beteiligter seine Unterlagen so weit geordnet hat, dass sie sich sowohl gestisch als auch verbal darauf beziehen kann (siehe Mondada 2004). Solche Verzögerungen sind in der Konversationsanalyse formal-strukturell beschrieben und zumeist mit der interaktiven Behandlung „schwieriger“ Sachverhalte in Zusammenhang gebracht worden.

3.3.1.3 Materialität

Materialität spielt für die Strukturierung von Interaktion und für Koordinationsleistungen in vielfältiger Weise eine Rolle. Das wird beispielsweise bei der Manipulation von Gegenständen deutlich (vgl. etwa Kallmeyer/Streeck 2001; Streeck 1996; Schmitt 2001; Mondada 2007d, 2008d). Das interaktionsstrukturierende Potenzial von Gegenständen manifestiert sich besonders dann, wenn es sich um signifikante Objekte handelt. Signifikante Objekte sind Gegenstände (wie beispielsweise ein Videomonitor und eine Filmkamera), um die herum und unter Bezug auf diese durch Interaktion zwischen verschiedenen Beteiligten temporär ein Interaktionsraum etabliert wird (Schmitt/Depermann 2007).

Signifikante Objekte sind gegenständliche Mitkonstituenten von Interaktionsräumen, deren Zentrum und koordinativen Bezugspunkt sie temporär bilden. Sie besitzen auf Grund ihrer eigenen Spezifik koordinative Implikationen (ein Videomonitor impliziert auf Grund seiner semiotischen Eigenschaften eine gewisse Orientierung derjenigen, die das laufende Video betrachten wollen, und eine gewisse Positionierung zu dem Objekt). Signifikante Objekte erhalten ihre Relevanz jedoch erst auf Grund der Bedeutung, die sie für Beteiligte zu einem gegebenen Zeitpunkt für die Realisierung bestimmter Aktivitäten besitzen.

3.3.1.4 Wahrnehmung

Ein weiterer Aspekt, der bislang in seiner zentralen interaktionskonstitutiven Bedeutung ebenfalls nicht gesehen werden konnte, ist die Rolle der Wahrnehmung und Wahrnehmungswahrnehmung (Hausendorf 2001) der Interaktionsbeteiligten. Videoaufzeichnungen machen deutlich, dass Interaktionsbeteiligte auf sehr unterschiedliche Weisen der Interaktionsentwicklung folgen, dass sie permanent – wenn auch manchmal nur aus den Augenwinkeln heraus – verfolgen, was gerade passiert.

Videoaufzeichnungen ermöglichen es, unterschiedliche Formen von Wahrnehmung zu untersuchen und deren Rolle für die lokale Interaktionskonstitution zu reflektieren. Deppermann/Mondada/Schmitt (i. Vorb.) untersuchen, wie sich Monitoring-Aktivitäten der Moderatorin eines Meetings in Reaktion auf die Bemühungen einer Mitarbeiterin ändern, die versucht, von der Pause wieder zur inhaltlichen Arbeit zurückzukehren. Der Aspekt der Wahrnehmung lenkt notwendigerweise den Fokus auf die Blickorganisation als eine Form der empirischen Manifestation von Wahrnehmung. Blickorganisation als ein Teil grundsätzlicher koordinativer Anforderungen und als für die Konstitution von Interaktion wichtige Anforderung in ihrer empirischen Vielfalt und lokal-spezifischen Funktionalität systematisch zu fokussieren, stellt für die multimodale Analyse eine wichtige zukünftige Aufgabe dar.

Wir wissen im Moment noch relativ wenig darüber, wie in Mehrpersonen-Konstellationen verbal nicht aktive Beteiligte der Interaktion folgen (Heidtmann/Föh 2007). Videoanalysen der unterschiedlichen Weisen, in denen die Beteiligten dem Interaktionsgeschehen folgen, ermöglichen es, die interaktionskonstitutive Bedeutung von Wahrnehmung systematisch zu untersuchen. Dabei gelangen körperliche Verhaltensweisen erstmalig als zeitlich gestreckter empirischer Ausdruck von Interpretationen der Beteiligten in den Blick, die bislang nur lokal und auf die Sprecher-Hörer-Dyade bezogen als *display* beschrieben worden sind (Goodwin 1981, Heath 1982).

3.3.2 Methodische Implikationen

Auch in methodischer Hinsicht wirkt sich die Veränderung der empirischen Basis für Interaktionsanalysen unmittelbar aus. Wir werden hier kurz exemplarisch auf zwei Aspekte eingehen: die Mehrebenenanalyse und die visuelle Erstanalyse.

3.3.2.1 Mehrebenen-Analyse

Die an der Strenge der konversationsanalytischen Methodologie orientierte konstitutionsanalytische Rekonstruktion multimodaler Interaktionsstrukturen verlangt die Entwicklung und Methodisierung spezifischer, auf die Eigenschaften des Untersuchungsgegenstandes bezogener Analyseverfahren.

Die Analyse von Videodokumenten macht aufgrund des komplexen Zusammenspiels unterschiedlicher Modalitätsebenen und der theoretischen Egalität der verschiedenen Modalitätsebenen Analyseverfahren erforderlich, die neben

dem Prinzip der Sequenzialität (d.h. der Vor- und Nachzeitigkeit) auch dem Prinzip der Gleichzeitigkeit in adäquater Weise Rechnung tragen. Da man bei der konkreten Analyse von Videoausschnitten jedoch nicht alle Aktivitäten auf den unterschiedlichen Modalitätsebenen gleichzeitig erfassen kann, ist man gezwungen, sich bei wiederholten Durchgängen durch den zu analysierenden Ausschnitt auf jeweils einzelne Ausdrucksebenen zu konzentrieren.

Konkret bedeutet dies, dass man entscheiden muss, mit welchem Modalitätsaspekt man die Analyse beginnt und mit welchem Aspekt man sie auf der Grundlage des dabei erarbeiteten Wissens fortführen will. Es ist ebenso naheliegend wie problematisch, grundsätzlich mit der Analyse des verbalen Geschehens zu beginnen und die Beobachtungen aller anderen Modalitätsebenen darauf aufzubauen. Naheliegend ist ein solches Vorgehen, weil es die methodische Sicherheit suggeriert, über die wir bei der konversationsanalytischen Auseinandersetzung mit Gesprächen verfügen. Zudem erfolgt in vielen Fällen die voranalytische Selektion relevanter Videoausschnitte auf der Grundlage des Gesprochenen. Problematisch ist ein solches Vorgehen, weil es die theoretisch postulierte Egalität aller Modalitätsebenen bei der Analyse unterläuft.

Grundsätzlich muss sich die Entscheidung, auf welcher Ebene man in die systematische Analyse einsteigt, an dem spezifischen Erkenntnisinteresse ausrichten, und sie muss dem konversationsanalytischen Diktum folgen, sich sowohl bei der Entwicklung angemessener Analysemethoden als auch bei der Definition adäquater Kategorien von den Daten selbst leiten zu lassen. Um den rekurrenten analytischen Durchgang durch das gleiche Videosegment methodisieren und dadurch den enormen Zeitaufwand ökonomisieren zu können, wird es zudem nötig sein, systematisch die eigene „De-facto-Methodologie“ zu rekonstruieren, der man bei der Analyse von Videoaufzeichnungen folgt.

3.3.2.2 Visuelle Erstanalyse

Man kann auf die methodische Anforderung einer systematischen Mehrebenenanalyse so reagieren, dass man zur Fokussierung der körperlichen Ausdrucksformen die vokale Ebene motiviert ausschließt. In systematisierter Form führt dies zur visuellen Erstanalyse als einer Variante der Modalitätsfokussierung. Der Begriff 'visuelle Erstanalyse' verweist auf die spezifische, durch den Forschungsgang definierte Funktionalität dieses methodischen Verfahrens: Es sollte vor der Analyse des verbalen Geschehens zur Anwendung gelangen. Da die Analyse noch nicht weitgehend durch das Wissen um das verbale Geschehen beeinflusst ist, ist am Anfang vor allem ein solcher Zugang

in kategorienbildender Hinsicht besonders produktiv. Wird ein Videoausschnitt nach den Regeln der Kunst unter Ausschluss der verbalen Anteile segmentiert und in seiner Struktur beschrieben, führt das – methodisch induziert – nicht nur zwangsläufig zur Entwicklung neuer Beschreibungskategorien, sondern auch zur Konstitution neuer Struktursegmente.

Die bei einer visuellen Erstanalyse einer Videoaufzeichnung generierten Beschreibungen und ad-hoc-Kategorisierungen können dann systematisch darauf befragt werden, ob sich mit ihnen in systematischer Weise pragmatisch relevante Strukturen erfassen lassen, die wesentliche Aspekte der Interaktionsdynamik beschreiben. Ein Beispiel für diesen Zusammenhang: Im Rahmen einer visuellen Erstanalyse wurde als manifestes Verhalten deutlich, dass die Regisseurin am Filmset eine Mitarbeiterin um die Hüfte fasst und herumdreht. Nachdem man in einem folgenden Analysegang auf der verbalen Ebene eine Auseinandersetzung zwischen den beiden analysiert hat, wird deutlich, dass das Umdrehen auch im Sinne eines gesprächsrhetorischen Zuges verstanden werden kann. Es hat die pragmatischen Implikationen von „Jemanden von seiner ursprünglichen Position abbringen“. Dies ist ein instruktives Beispiel dafür, dass relevante Orientierungen der Beteiligten nicht nur verbalisiert werden, sondern auch als pragmatisch implikative Verkörperung zum Ausdruck kommen.

3.3.3 Methodologische Implikationen

Die Spezifik audiovisueller Daten führt auch im methodologischen Bereich zu interessanten Fragen. Wir werden uns diesbezüglich auf folgende Aspekte konzentrieren: die Modifikation etablierter konversationsanalytischer Konzepte, den Status von Transkripten und Video als reflexive Dokumentationsmedien und den Prozess der Datenkonstitution als aktive Herstellung.

3.3.3.1 Modifikation etablierter Konzepte

Die multimodale Perspektive auf Interaktion führt zwangsläufig zur Neukonstitution und Neubewertung etablierter konversationsanalytischer Konzepte. Mit der Fokussierung aller sichtbaren Verhaltensaspekte für die Konstitution der interaktiven Ordnung und der permanenten Gleichzeitigkeit unterschiedlicher Verhaltensmodalitäten geraten die verbal definierten Konzepte zur Beschreibung der Grundmechanismen bei der Produktion interaktiver Ordnung unter Druck.⁶

⁶ Zu den Auswirkungen auf etablierte konversationsanalytische Konzepte aus multimodaler Perspektive siehe beispielsweise Mondada (2004) und Schmitt (2004, 2005).

Eine damit zusammenhängende Implikation führt letztlich dazu, die für die theoretisch-konzeptionelle und analysefaktische Konzentration auf Verbalität zentralen Kategorien „Sprecher“/„Hörer“ aufzugeben. Sie müssen im Kontext der Analyse multimodaler Interaktion gegenüber einer Vorstellung von multimodal aktiven Interaktionsbeteiligten und dem „Interaktionsensemble“ als zentrale Handlungseinheit ersetzt werden (vgl. Schmitt i. Vorb.).

Das heißt jedoch nicht, die Kategorien „Sprecher“ und „Hörer“ gänzlich aufzugeben. Sie verlieren jedoch ihren konzeptuellen Status von Zentralkategorien und rutschen, wenn man das so sagen will, eine Stufe herab. Sie kategorisieren jetzt einen an eine spezifische Modalität gebundenen Beteiligungsaspekt im Gesamtverhalten der Interaktionsbeteiligten. Die Kategorien sind im strengen Sinne nicht der Interaktion als solche inhärent und quasi naturwüchsig zentral, sondern sie sind das Ergebnis eines spezifischen Erkenntnisinteresses und einer davon abgeleiteten Gegenstandskonstitution, wie sie beispielsweise in den Bezeichnungen *conversation* und später dann *talk-in-interaction* im konversationsanalytischen Forschungszusammenhang ihren Ausdruck findet. *Conversation* wurde zunächst als übergreifender Begriff für unterschiedliche, verbal definierte *speech exchange systems* benutzt und später durch *talk-in-interaction* ersetzt (Schegloff 1987a, 1987b und 1988).

Im multimodalen Erkenntniszusammenhang gibt es kein „Service-Verhältnis“ bestimmter Modalitätsebenen für Verbalität. Einzelne Modalitätsaspekte leisten vielmehr einen je spezifischen Beitrag bei der Interaktionskonstitution. Sie sind immer eingebettet in den übergeordneten und differenzierten Rahmen von *participation framework*, dessen Füllung eines der wesentlichen Aufgaben zukünftiger Forschung im Bereich „Multimodalität der Interaktion“ darstellt. Für die konkrete Analysearbeit bedeutet das, dass man – ausgehend vom Interaktionsensemble als zentraler Bezugsgröße – nunmehr auch die Aktivitäten derjenigen hinsichtlich ihres spezifischen Beitrags zur Interaktionskonstitution befragen muss, die aktuell nicht sprechen, sich jedoch als Teil des Interaktionsensembles multimodal verhalten.⁷

Damit sind wir bei zwei weiteren Implikationen der multimodalen Konzeption von Interaktion, die – vergleichbar dem stillschweigenden Augenzwinkern hinsichtlich der methodologischen Implikationen des Verhältnisses von dokumentierter Situation und Interaktionsdokument im Kontext der Konversationsanalyse – auch im multimodalen Erkenntniszusammenhang leicht

⁷ Siehe beispielsweise Heidtmann/Föh (2007), die „verbale Abstinenz als Form interaktiver Beteiligung“ konzeptualisieren.

übersehen werden können: Es handelt sich zum einen um *Video als reflexives Dokumentationsmedium* und zum anderen um das, was man vielleicht am adäquatesten als *Stücke-Idealisierung* im Kontext interaktionsanalytischer Untersuchungsgegenstände bezeichnen kann. Beide Aspekte sind für unser Erkenntnisinteresse an der multimodalen Herstellung fokussierter Interaktion unmittelbar relevant.

3.3.3.2 Video als reflexives Dokumentationsmedium

Einer der manifesten Aspekte, der das Nachdenken über die Spezifik des methodischen Umgangs mit den empirischen Grundlagen im audiovisuellen Forschungszusammenhang nötig macht, ist der Status der Videotechnologie. Video als Dokumentationsmedium macht nicht nur die sichtbare Seite interaktiver Ordnung der Analyse zugänglich, Video ist gleichzeitig auch eine reflexive Technologie, die auf sich selbst verweist (Mondada 2006, 2008c, 2009; Schmitt/Fiehler/Reitemeier 2007, Heath/Hindmarsh 2002, Heath/Luff 2006). Die zur Aufzeichnung eingesetzte Technik schreibt sich unabwendbar in die Aufzeichnungen der Interaktionssituationen hinein, die sie konserviert. Das, was wir hinterher bei der Analyse sehen, ist nicht die tatsächliche Situation, in der die Videoaufnahme erfolgt ist. Vielmehr handelt es sich um einen durch die Kamera und die Person hinter der Kamera festgelegten Ausschnitt und Blickwinkel, mit dem bestimmte Fokussierungen und Implikationen verbunden sind.

Damit zusammenhängend rückt auch die aktive Rolle des Kameramannes/der Kamerafrau bei der Datenkonstitution in den Fokus. Beim Betrachten von Videodokumenten wird deutlich, dass der Kameramann/die Kamerafrau unmittelbar und aktiv zur Gestaltung der von ihm/ihr dokumentierten Daten beiträgt: Dynamische Kameraführung, Schwenks und Zoomaktivitäten, die Hervorhebung einzelner Personen aus dem Gesamtzusammenhang sowie die Ausblendung ganzer Teilaktivitäten sind Ausdruck seiner/ihrer Interpretation des dokumentierten Geschehens. Diese werden als konstitutiver Bestandteil der Aufnahme mit-dokumentiert und verdeutlichen so, dass es sich bei der Dokumentation um eine aktive Herstellung handelt.

Diese nicht hintergehbare Präsenz der Videotechnologie ist jedoch nicht nur ein notwendiges Übel. Man kann daraus vielmehr auch eine Tugend machen. Unter einer solchen Perspektive wird der Kameramann/die Kamerafrau als Bestandteil des Interaktionsensembles gesehen und die Kamerabewegungen als Ausdruck seiner/ihrer relevanzgeleiteten Orientierung und als Hinweise auf seine/ihre Interpretationen des aktuellen Interaktionsgeschehens begriffen.

Folgt man also dem ethnomethodologischen Diktum, sich bei der Konstitution von Fragestellungen und der Konzeptentwicklung von den Daten selbst leiten zu lassen, werden die Kamera bzw. der Kameramann/die Kamerafrau und die Spuren, die er/sie und das Dokumentationsmedium bei der Dokumentationsarbeit im Material hinterlassen haben, selbst zum Gegenstand der Untersuchung (Broth 2004, Macbeth 1999, Mondada 2009).

Die nicht hintergehbaren Implikationen des Dokumentationsmediums haben beispielsweise zur Entwicklung des Konzepts „Fokusperson“ geführt (Schmitt/Deppermann 2007). Das Konzept trägt zum einen den interaktionsvorgängigen Ordnungsstrukturen und deren Implikationen für die Interaktion Rechnung. Dies betrifft die herausragende Position des Regisseurs – der im Zentrum kontinuierlicher Monitoring-Aktivitäten vieler Setmitarbeiter steht – für die Koordination auf dem Set. Zum andern trägt das Konzept jedoch auch den spezifischen Dokumentationsbedingungen Rechnung. In methodischer Hinsicht reflektiert das Konzept die Tatsache, dass der Regisseur gerade auch durch die Mikrofonverkabelung und durch die Kamera zum Fokus des dokumentierten Geschehens wird.

Insgesamt macht der dynamische Kameraeinsatz deutlich, dass bereits die Dokumentation unumgebar analytische Implikationen hat bzw. selbst bereits analytisch ist. Dies zeigt sich besonders dann, wenn die Kamera in Folge der Antizipation des Kameramanns/der Kamerafrau bereits einen der Interaktionsteilnehmer fokussiert und groß ins Bild holt, bevor dieser in das Geschehen eingreift. Hier kann man sehen, dass die Kamerabedienung auf der Grundlage einer intuitiven Analyse funktioniert, die dem Kameramann/der Kamerafrau selbst in der Situation nicht bewusst sein muss.

3.3.3.3 Status von Transkripten

Was ebenfalls einer systematischen Reflexion bedarf, ist der veränderte Status von Transkripten im multimodalen Forschungskontext. Bei Transkripten audiovisueller Daten handelt es sich konstitutionslogisch gesehen um „Zwitzer“. Zum einen stellt die Notation des verbalen Geschehens wie in verbalen Transkripten die Grundlage für die nachfolgende Analyse der Interaktionsstruktur und der sequenziellen Rekonstruktion der sie generierenden Mechanismen dar. Zum anderen aber sind die Notationen der sichtbaren Verhaltensweisen bereits weitgehend das Ergebnis einer durch theoretisch motivierte Erkenntnisinteressen geleiteten faktischen Analyse des Interaktionsgeschehens. Erst nach einer analytischen Beschäftigung mit der Videoaufnahme

kann die Entscheidung getroffen werden, an welcher Stelle im Transkript welche Abbildungen gebraucht werden. Hier stellt sich also die Frage, welche Analysefragen auf der Grundlage einer solchen konstitutionslogisch „zwitterigen“ Datenbasis in sinnvoller Weise verfolgt werden können. Diese Frage verweist auch indirekt auf die Notwendigkeit der Entwicklung gegenstandsadäquater Präsentationsformen für Analyseergebnisse, die auf der Grundlage audiovisueller Daten erarbeitet wurden.

Diese konstitutionslogische Differenz geht einher mit einem weiteren Unterschied, der das Verhältnis verbaler und visueller Informationen prägt. Für den Bereich der Notation des verbalen Geschehens gibt es aufgrund der grammatisch-syntaktischen Kodierung unserer Sprache eine sehr weitgehende intersubjektive Übereinstimmung zumindest im Hinblick auf das Gesagte. Da es eine vergleichbare Kodifizierung und Wiedergabe körperlicher Aktivitäten nicht gibt, ist bei diesen Informationen bei unterschiedlichen Beobachtern ein weiter interpretativer Spielraum dafür gegeben, was man beispielsweise unter einer Bewegung versteht und wie sie angemessen wiedergegeben werden soll. Allein für die objektiv wahrnehmbare Bewegung eines Lehrers von der Tafel in den Klassenraum hinein gibt es eine ganze Reihe alternativer, jedoch nicht funktional äquivalenter Beschreibungen: *von der Tafel weggehen, den Blick auf die Tafel freimachen, in den Klassenraum hineintreten, auf die Schüler zugehen* etc. Jede dieser Beschreibungen ist mit einer ganz spezifischen Implikation verbunden, die das Geschehen in unterschiedlicher Weise konzeptualisiert.

Der Einsatz von Transkriptionssoftware und des Alignments der ursprünglichen Signale und des Transkriptionstexts hat es ermöglicht, bei der Lösung der grundlegenden Problematik der Repräsentation multimodaler Zusammenhänge einen wichtigen Schritt nach vorne zu machen. Dies betrifft in erster Linie die Behandlung und Visualisierung von Zeitlichkeit, die in einer *timeline* repräsentiert wird, auf die sowohl die verbalen, gestikulatorischen, blicklichen als auch alle anderen relevanten Aktivitäten bezogen sind (Mondada 2008a). Diese Abhängigkeit von der *timeline* ermöglicht die Aufrechterhaltung der grundsätzlichen Gleichwertigkeit der verschiedenen Modalitätsebenen. Sie ermöglicht die flexible und adäquate Anordnung, die bei der jeweiligen Transkription in Abhängigkeit der Fragestellung jeweils notwendig ist. Diese Form der Repräsentation ermöglicht zudem, die Annotations Spuren ohne prinzipielle Beschränkung beliebig zu erweitern (siehe Mondada 2007a).

Die Beiträge des Bandes repräsentieren hinsichtlich der „Transkriptions-Problematik“ eine gewisse Bandbreite an Problemlösungen. Es finden sich im engeren Sinne multimodale Transkripte (Mondada, Oloff, De Stefani/Mondada), bei denen sich die Autor(inn)en am klassischen Verständnis eines Transkripts als „Analysegrundlage“ orientieren. Die Erstellung dieser multimodalen Transkripte ist durch das Bemühen gekennzeichnet, die multimodale Komplexität der Ursprungssituation – in gewisser Granularität – im Transkript grundsätzlich verfügbar zu machen. Dies führt zu einer sehr dichten und komplexen Transkription. Diese Transkripte können bei der Lektüre nicht einfach gelesen werden, sondern müssen vom Rezipienten aktiv „erarbeitet“ werden.

Es gibt daneben Beiträge (Heidtmann, Schmitt/Deppermann, Schmitt/Hausendorf), die auf die Herstellung multimodaler Transkripte verzichten. Hinter diesem Verzicht verbirgt sich folgende methodologische Überlegung: Audiovisuelle Interaktionsdokumente können nur hinsichtlich ihrer Verbalität angemessen und rezeptionsfreudig angefertigt werden. Die visuellen Aspekte hingegen sind bereits analytisch so stark aufgeladen und so komplex, dass sie den Status von Transkripten als Analysegrundlage sprengen. Ein Versuch, die multimodale Komplexität dennoch abzubilden, wird über das Einfügen von Standbildern versucht.

Es liegt nicht in der Absicht der Beiträger(innen), die Möglichkeiten und Notwendigkeiten des Einsatzes von Transkripten bei multimodalen Analysen zu homogenisieren. Es bleibt bewusst den Rezipienten überlassen, über die für ihre Bedürfnisse adäquate Präsentation der audiovisuellen Daten zu entscheiden. Der Status multimodaler Transkripte und deren vom jeweiligen Erkenntnisinteresse abhängige Granularität und Vollständigkeit ist ein zentraler Punkt, der bei der Entwicklung einer Methodologie für die Analyse multimodaler Interaktion in der Zukunft sicherlich ernsthaft diskutiert werden muss. Der vorliegende Band beschränkt sich hier lediglich darauf, ein gewisses Spektrum an Lösungen dieses Problems zu zeigen, die zum augenblicklichen Zeitpunkt zur Verfügung stehen.

4. Situationseröffnungen aus multimodaler Perspektive

Die Untersuchung von Situationseröffnungen stellt eine Reihe von Fragen, die sowohl für die Konzeptualisierung von Multimodalität und Interaktion zentral sind, als auch hinsichtlich der methodologischen Konsequenzen, die sich aus der Arbeit mit Videoaufzeichnungen ergeben.

4.1 Multimodale Dimensionen

„Späte“ **Verbalität:** Als eine grundsätzliche Erkenntnis der Analyse von Situationseröffnungen auf der Grundlage audiovisueller Interaktionsdokumente wird deutlich: Die Sprache kommt erst zu einem relativ späten Zeitpunkt ins Spiel. Werden die Interaktionsbeteiligten verbal aktiv, kann man in der Regel analytisch detailliert nachweisen, dass sie sich – anders als bei Telefongesprächen, wo dies technisch nicht möglich ist, und auch anders als bei Tondokumenten von Situationseröffnungen in Face-to-Face-Konstellationen, wo diese Informationen verloren gehen – in der Regel bereits intensiv ausgetauscht und die Möglichkeiten für den verbalen Austausch aktiv hergestellt haben.

Auch wenn man der Meinung ist, dass Sprache das zentrale Medium/die zentrale Modalität für den interaktiven Austausch ist, muss man sehen, dass der Einsatz von Sprache von den Beteiligten in der Situation selbst erst ermöglicht werden muss. Ein voraussetzungsloses Sprechen im Sinne eines unvorbereiteten interaktiven Austauschs gibt es – zumindest als unmarkierte Normalform – nicht. Das Verhältnis von Sprache und den anderen Modalitäten ist dabei ein dialektisches: Der Gebrauch von Sprache wird von den Interaktionsbeteiligten durch den Einsatz anderer Modalitäten und die Konstitution von geeigneten Interaktionsräumen systematisch vorbereitet. Die Sprache, die auf diesen interaktiv vorbereiteten Grundlagen zum Einsatz gebracht wird, reflektiert ihrerseits im Detail die für ihren Einsatz notwendigen, von den Beteiligten hergestellten Voraussetzungen und ist an diese gebunden. Wie also der erste verbale Kontakt gestaltet wird, ob es eine offizielle Begrüßung, eine namentliche Adressierung gibt, in welcher Weise auf die aktuelle Situation referiert wird etc., hängt neben den Relevanzen des Ereignisses für die Beteiligten auch ganz wesentlich von den Aktivitäten dieser Anbahnungsphase ab.

Visuelle Wahrnehmung: Die wechselseitige Kategorisierung als potenzielle bzw. zukünftige Interaktionspartner wird primär durch Prozesse der visuellen Wahrnehmung und der Wahrnehmungswahrnehmung organisiert und vorbereitet. Visuelle Wahrnehmung und die darauf basierende Blickorganisation ereignen sich – aufgrund der Qualität als Distanzmedium – zu einem sehr frühen Zeitpunkt. Beteiligte können sich schon aus großer Distanz wahrnehmen und Hypothesen über Laufrichtung, Bewegungsgeschwindigkeit und Ziele anderer bilden und – für den Fall, dass sie eine fokussierte Interaktion mit einem/den anderen Beteiligten beabsichtigen – ihr eigenes Verhalten in adäquater Weise auf der Grundlage dieser Hypothesen koordinieren.

Doch nicht nur das Verhalten der anderen wird im Hinblick auf intendierte nachfolgende Interaktion durch Monitoring-Aktivitäten verfolgt. Auch das eigene Verhalten wird im Hinblick auf die Eröffnung einer Interaktionssituation für die anderen erkennbar, d.h. in diesem Falle wahrnehmbar gemacht. Dies ist die Voraussetzung dafür, dass es nicht (nur) zu einseitig strukturierten Anbahnungen kommt, sondern dass diese Vorbereitungsphase von den Beteiligten gemeinsam organisiert werden kann.

Körperorientierung: Die gemeinsame Orientierung auf eine durch bestimmte Hinweise in der Anbahnungsphase projizierte Bereitschaft/Absicht zur fokussierten Interaktion zeigt sich im Verhalten der Beteiligten in Veränderungen ihrer Körperorientierungen, der Modifikation ihrer Bewegungsgeschwindigkeit, der Zielkorrektur, der Herstellung von Nähe, der Kopfausrichtung und der Blickorganisation als vorgehende Verdeutlichungen des Partnerbezuges. Es kommt im Laufe der Entwicklung solcher Anbahnungen zu einer immer deutlicher werdenden wechselseitigen Abstimmung dieser unterschiedlichen modalen Ausdrucksmöglichkeiten auf einen gemeinsamen Punkt (der Eröffnung) hin. Dies ist die Voraussetzung dafür, dass Außenstehende die Interaktionsabsicht der sich so gemeinsam orientierenden und koordinierenden Beteiligten mit relativer Sicherheit prognostizieren und erkennen können, dass es zu einer Situationseröffnung kommen wird. Denn das grundsätzliche Design, die Wahrnehmbarkeit des eigenen Verhaltens als Vorbereitung für die Eröffnung einer fokussierten Interaktion, erfolgt nicht nur für den zukünftigen Interaktionspartner, sondern auch für diejenigen, die nicht als Beteiligte des zukünftigen Austauschs vorgesehen sind. Auf der Grundlage ihrer Interpretation des Verhaltens als Vorbereitungsaktivitäten können sie jedoch – in bestimmten Situationen und unter bestimmten Voraussetzungen – selbst entscheiden, sich in diesen Prozess zu integrieren.

Beteiligungsrahmen: Und noch ein weiterer Punkt weist den konstitutiven Status der Anbahnung als Bestandteil von Situationseröffnungen aus: Bevor eine Situation eröffnet werden kann und die ersten Worte gewechselt werden können, müssen die Beteiligten in hinreichend erkennbarer Weise gemeinsam einen Beteiligungsrahmen etablieren und sich dadurch als potenzielle und prinzipiell willige Teilnehmer einer zukünftigen/intendierten fokussierten Interaktion wechselseitig identifizieren. In konzeptueller Hinsicht ist es eine interessante Frage, wie der Status dieser Aktivitäten im Hinblick auf die Struktur von Situationseröffnungen konkret auszuweisen und zu definieren

ist. Dass sie aus multimodaler Perspektive konstitutiver struktureller Bestandteil von Situationseröffnungen sind, steht angesichts der Ergebnisse der einzelnen Beiträge außer Frage.

Raum: Bei der Analyse der multimodalen Herstellung fokussierter Interaktion spielt dabei gerade für die Phase der Anbahnung und Vorbereitung der Situationseröffnung die Herstellung eines besonderen räumlichen Arrangements, eines gemeinsamen Interaktionsraums eine wesentliche Rolle (siehe auch Kendon (1990c) und seine Konzeption von *F-formation*). Die Art und Weise, in der ein solcher Interaktionsraum als zentrale Voraussetzung für die intendierte zukünftige Interaktion gemeinsam hergestellt wird, ist dabei projektiv hinsichtlich zentraler thematischer, pragmatischer oder sozialer Aspekte dieser zukünftigen Interaktion. Die verschiedenen Beiträge in diesem Band zeigen klar, dass und wie die Realisierung der Anbahnung und Vorbereitung der Situationseröffnung bereits zentrale Relevanzen dieser Situation verdeutlicht.

4.2 Methodologische Konsequenzen

4.2.1 Korpusaufbau

Videoaufzeichnungen verdeutlichen die grundsätzliche Bedeutung von Wahrnehmung für die Strukturierung der Situationseröffnungen in Face-to-Face-Situationen. In der Regel ist die wechselseitige Wahrnehmung und die Blickorientierung der Interaktionsbeteiligten eine Art Primärressource/-medium der Situationsanbahnung. Dies hat unmittelbare Auswirkungen auf den Korpusaufbau, vor allem aber hinsichtlich der Festlegung, wann man mit der audiovisuellen Dokumentation der Situation beginnt und Kameras und Mikrofone einschaltet.

Die zentrale Bedeutung der wechselseitigen visuellen Wahrnehmung in Face-to-Face-Situationen für die Vorbereitung von Situationseröffnungen geht einher mit ihrer für eine detaillierte und methodisch kontrollierte Analyse oft nur ungenügenden empirischen Erfassung. Aufgrund aufnahmetechnischer Beschränkungen sind viele unserer audiovisuellen Interaktionsdokumente nur in Ausnahmefällen geeignet, die interaktiv relevanten Anbahnungsaktivitäten einer systematischen Analyse zu unterziehen. Dies gilt vor allem für mimische Aktivitäten und Blickorganisation, teilweise aber auch für minimale Formen von Kopfnicken als Ausdruck der Adressatenvorauswahl. Vor allem aber gilt das für die koordinierte Abstimmung der

Anbahnungs- und Vorbereitungsaktivitäten der Beteiligten, wenn dies mit Bewegung im Raum verbunden ist. Vor allem bei einer gewissen räumlichen Distanz der Beteiligten ist dies nur mit großem technischem Aufwand dokumentier- und analysierbar.

Unter Bedingungen wechselseitiger Wahrnehmung ist es ein generelles Charakteristikum von Interaktionseröffnungen, dass in der Regel lange bevor das erste Wort zwischen in einer Situation ko-präsenten Beteiligten gesprochen wird, bereits intensiver interaktiver Austausch zwischen ihnen stattfindet. In der Phase räumlicher Annäherung beispielsweise, die eine konstitutive Voraussetzung für das „Nähe-Medium“ Verbalität ist, selektieren und identifizieren sich Beteiligte bereits mit Blicken und Körperorientierung als künftige Interaktionspartner. Sie tauschen dabei auch wahrnehmbare Vororientierungen auf einen gemeinsamen Interaktionsraum aus und kommunizieren durch ihre spezifische körperliche Präsenz im Raum und durch den Modus ihrer Bewegung, welche Qualität die projizierte Interaktion haben wird/soll.

Es ist wichtig zu betonen, dass dies nicht einfach notwendige Vorbereitungen sind, die das eigentlich zentrale verbale Geschehen ermöglichen, das in seiner Spezifik dann jedoch von der bereits stattgefundenen Interaktion in der Anbahnungsphase unabhängig ist. Es ist vielmehr so, dass solche vorgängigen Aushandlungen (Anbahnungen) einen unmittelbaren Einfluss auf die Art und Weise haben, wie die verbale Weiterführung der bereits vorbereiteten Interaktion faktisch aussehen wird (ob es beispielsweise zu einer formellen Eröffnung mit Adressierung oder Begrüßung kommt oder nicht). In diesem Sinne kann man aufgrund einer videogestützten Analyse von Situationseröffnungen nicht nur rekonstruieren, was alles an Vorarbeiten in welcher Modalität geleistet worden ist, bevor Sprache ins Spiel kommt. Dies ist ein wichtiger Aspekt bei der Gegenstandskonstitution und dem „Zuschnitt“ von Situationseröffnungen für die Analyse. Man kann umgekehrt aus der Analyse der verbalen Situationseröffnung auch rekonstruieren, was alles an Vorbereitungsaktivitäten stattgefunden haben muss, damit der Einsatz von Sprache genau in der vorliegenden Art und Weise erfolgen konnte.

Für die umfassende und systematische Untersuchung der multimodalen Herstellung fokussierter Interaktion besteht folglich die Notwendigkeit, spezielle Korpora aufzubauen, die der spezifischen Konstitutionslogik der Situationsherstellung adäquat Rechnung tragen. Das bedeutet auch, dass man dabei auf leicht zugängliche Fremddaten, wie sie uns Fernsehen und Internet zur Verfügung stellen, nicht zurückgreifen kann.

Um jedoch entscheiden zu können, wie solch ein geeignetes Korpus de facto aussehen muss, ist es nötig, zunächst einmal auf der Grundlage vorhandener Aufzeichnungen die multimodale Herstellung fokussierter Interaktion in möglichst vielen unterschiedlichen und maximal kontrastierenden Situationen detailliert zu untersuchen. Hierzu leistet der Band einen wichtigen Beitrag.

4.2.2 Ausschnitt konstituieren

Aufgrund der teilweise weit gespannten Phase der Anbahnung und Situationsherstellung und ihrer in der Regel erst späten verbalen Qualität ist es unter einer multimodalen Erkenntnisperspektive notwendig, sich bei der Analyse von Videoausschnitten die Frage nach dem für die Analyse notwendigen Ausschnitt zu stellen. Es geht dabei nicht einfach darum, in formaler Hinsicht den Umfang des Ausschnittes festzulegen. Wie weit man den Blick bei der Ausschnittkonstitution nach vorne richten muss, ist nicht leicht zu entscheiden.

Unter einem multimodalen Erkenntnisinteresse erfordert die Ausschnittbildung also bereits ernsthafte analytische Investitionen. Aufgrund der Komplexität der für die Interaktionskonstitution relevanten Aspekte und besonders wegen des mehr oder weniger dichten Netzes projektiver Verdeutlichungen in unterschiedlichsten Formen multimodaler Realisierung ist die motivierte, erkenntnisgeleitete Festlegung des relevanten Ausschnitts tatsächlich bereits ein erstes Analyseergebnis.

Schaut man sich beispielsweise Videoaufzeichnungen an, die die Vor- und Etablierungsphase von Meetings dokumentieren, dann werden zwei Dinge unmittelbar deutlich: Zum einen liefert die Art und Weise, wie diese Vor-Phase gemeinsam hergestellt wird, wie die Mitglieder des Meetings nach und nach den Raum betreten, ihren Platz suchen, miteinander reden und sich dabei nicht nur blicklich, sondern auch teilweise bereits thematisch auf das anstehende Meeting beziehen, bereits wichtige Hinweise auf das nachfolgende interaktive Geschehen und den Charakter des Meetings. Zum andern wird deutlich, dass die Frage, wann die fokussierte Interaktion als hergestellt gelten kann, nicht zu beantworten ist, ohne sich bereits analytisch mit dem Geschehen dieser Vorbereitungsphase befasst zu haben.

Andererseits wird der Beginn bereits durch die Orientierung der Person hinter der Kamera und durch situativ emergierende Relevanzen festgelegt, was oft zu einem verspäteten Einsatz des Dokumentationsmediums führt. Dadurch werden jedoch bestimmte Aspekte des Gesamtereignisses an den Rand gerückt oder gar gänzlich ausgeblendet. Zusätzlich zeigt sich häufig eine starke

Orientierung an der manifesten Eröffnung und ihrer verbalen Grundlagen, was unmittelbar die Festlegung des Dokumentationsbeginns und damit auch die Möglichkeiten der analytischen Rekonstruktion der faktischen „Tiefe“ von Anbahnungen betrifft.

Wir haben es bei Videoaufnahmen in einem doppelten Sinne mit einer vorge-deuteten Wirklichkeit⁸ zu tun: Nicht nur die Deutungen der Interaktionsbeteiligten sind inkorporiert in ihrem konkreten Verhalten in den Daten enthalten. Auch die während der Aufnahme realisierten Deutungen und Interpretationen des Interaktionsgeschehens durch den Kameramann sind Bestandteil der Daten. In Anlehnung an Bergmann (1985) und seine Überlegungen zur Flüchtigkeit sozialer Interaktion muss man hier von einer „interpretativen Konservierung“ sprechen.

Dieser konstitutive Aspekt audiovisueller Daten macht in erhöhtem Maße die methodologische Reflexion nicht nur der Datenkonstitution, sondern auch des methodisch kontrollierten und konzeptorientierten analytischen Umgangs mit diesen „aktiv hergestellten Dokumenten sozialer Praxis“ notwendig. Es ist daher erforderlich, in der detaillierten Auseinandersetzung mit konkreten Fällen der multimodalen Herstellung fokussierter Interaktion eine der konversationsanalytischen Methodologie vergleichbare Methodisierung fallspezifischer Analyseprozeduren zu entwickeln. Diese muss der spezifischen Qualität der Interaktionsdokumente Rechnung tragen, was bedeutet, dass sie neben sequenziellen Phänomenen in gleichem Maße auch die vielfältigen Aspekte multimodaler Gleichzeitigkeit berücksichtigt.

Der vorliegende Band versteht sich auch als eine Vorarbeit zu einer solchen Methodisierung.

5. Kurzbeschreibung der Beiträge

Die in den Beiträgen des Bandes analysierten Beispiele multimodaler Herstellung fokussierter Interaktionen repräsentieren bewusst eine bestimmte Bandbreite unterschiedlicher Situationen. Das Varianzspektrum umfasst Situationen, die hinsichtlich verschiedener konstitutiver Aspekte kontrastieren. Zu diesen zählen beispielsweise:

- Privat versus institutionell;
- Freizeit versus Berufswelt;
- Ritual versus Zufallskommunikation;

⁸ Zum Aspekt der immer schon vorge-deuteten Daten sozialwissenschaftlicher Untersuchungen siehe Schütz (1971).

- Dominante Kernaktivität versus thematisch-pragmatisch offen;
- Face-to-Face-Interaktion versus technisch vermittelte Interaktion;
- Variable Präsenzform versus dominante Präsenzform;
- Vordefinierte Beteiligungsstruktur versus offene Beteiligungsstruktur.

Konkret enthält der vorliegende Band folgende Situationen:

- Arbeitsmeetings, in denen in der Kooperation zwischen Filmstudenten und Dozenten Ideen für Filme entwickelt werden (Daniela Heidtmann);
- Videokonferenzen unterschiedlicher Ärzteteams, in denen Krankheitsfälle besprochen und Lösungsmöglichkeiten diskutiert werden (Lorenza Mondada);
- zufällige Treffen im öffentlichen Raum wie beispielsweise das Einholen von Wegauskünften oder das ungeplante Zusammentreffen Bekannter im Supermarkt (Elwys De Stefani/Lorenza Mondada);
- ein Raclette-Essen unter Freunden und Bekannten, bei denen die Teilnehmer(innen) nach und nach „eintröpfeln“ (Florence Oloff);
- der Gang des Regisseurs am Filmset in einer Drehpause und die dabei stattfindenden Interaktionen mit unterschiedlichen Beteiligten (Reinhold Schmitt/Arnulf Deppermann) und
- die langgedehnte Eröffnungsphase eines Gottesdienstes in einer evangelischen Kirche (Heiko Hausendorf/Reinhold Schmitt).

Dieses Varianzspektrum hängt mit dem Primärziel des vorliegenden Bandes zusammen. Angesichts der im Moment noch sehr sporadischen Untersuchungen zur komplexen multimodalen Struktur der Herstellung fokussierter Interaktion ist die detaillierte fallspezifische Rekonstruktion ausgewählter Situationen ein erster wichtiger Schritt, der noch weitgehend explorative Intentionen verfolgt.

Die fall- und situationsspezifischen Studien zur multimodalen Herstellung fokussierter Interaktion sind – angesichts der Prominenz der von der Konversationsanalyse untersuchten verbalen Gesprächseröffnungen – ein geeigneter und wichtiger Startpunkt für die Weiterentwicklung der multimodalen Konzeption von Interaktion gerade im kontrastiven Bezug zu diesem fest etablierten konversationsanalytischen Gegenstand. Dieser Band versteht sich in dieser Hinsicht auch als ein Beitrag zur grundsätzlichen Klärung der Brauchbarkeit verbal definierter Konzepte und deren empirisch basierter Reflexion, Weiterentwicklung, Modifikation oder Substitution.

6. Literatur

- Ag-Youssouf, Ibrahim/Grimshaw, Allen D./Bird, Charles S. (1976): Greetings in the desert. In: *American Ethnologist* 3, S. 797-824.
- Argyle, Michael (1975): *Bodily communication*. New York.
- Arminen, Ilkka (2002): Emergentes, divergentes? Les cultures mobiles. In: *Réseaux* 112-113, S. 79-106.
- Arminen, Ilkka (2005): Sequential order and sequence structure – the case of incommensurable studies on mobile phone calls. In: *Discourse Studies* 7, 6, S. 649-662.
- Baker, Carolyn/Emmison, Mike/Firth, Alan (2001): Discovering order in opening sequences: 6. calls to a software helpline. In: McHoul, Alec/Rapley, Mark (Hg.): *How to analyse talk in institutional settings: A casebook of methods*. London, S. 41-56.
- Bauman, Richard/Sherzer, Joel (Hg.) (1974): *Explorations in the ethnography of speaking*. Cambridge.
- Bergmann, Jörg R. (1979): *Interaktion und Exploration: Eine konversationsanalytische Studie zur sozialen Organisation der Eröffnungsphase von psychiatrischen Aufnahmegesprächen*. Univ. Diss. Konstanz.
- Bergmann, Jörg R. (1985): Flüchtigkeit und methodische Fixierung sozialer Wirklichkeit. Aufzeichnungen als Daten der interpretativen Soziologie. In: Bonß, Wolfgang/Hartmann, Heinz (Hg.): *Entzauberte Wissenschaft. Zur Relativität und Geltung soziologischer Forschung*. Göttingen, S. 299-320.
- Bergmann, Jörg R. (1995): Alarmiertes Verstehen: Kommunikation in Feuerwehrnotrufen. In: Jung, Thomas/Müller-Doohm, Stefan (Hg.): „Wirklichkeit“ im Deutungsprozeß. Verstehen und Methoden in den Kultur- und Sozialwissenschaften. 2. Aufl. Frankfurt a.M., S. 283-328.
- Birdwhistell, Ray L. (1970): *Kinesics and context. Essays on body-motion communication*. Philadelphia.
- Boden, Deirdre/Zimmerman, Don H. (Hg.) (1991): *Talk and social structure. Studies in ethnomethodology and conversation analysis*. Cambridge.
- Boden, Deidre (1994): Opening to meetings. In: Boden, Deidre: *The business of talk. Organizations in action*. Cambridge, S. 90-99.
- Bonu, Bruno (2007): Connexion continue et interaction ouverte en réunion visiophonique. In: *Réseaux* 144, S. 25-58.
- Bowles, Hugo (2006): Bridging the gap between conversation analysis and ESP – an applied study of the opening sequences of NS and NNS service telephone calls. In: *English for Specific Purposes* 25, S. 332-357.
- Bowles, Hugo/Pallotti, Gabriele (2004): Conversation analysis of opening sequences of telephone calls to bookstores in English and Italian. In: *Textus* 17, 1, S. 63-88.

- Broth, Mathias (2004): The production of a live TV-interview through mediated interaction. In: van Dijkum, Cor/Blasius, Jörg/Kleijer, Henk/van Hilten, Branko (Hg.): *Recent developments and applications in social research methodology*. (= Proceedings of the sixth international conference on logic and methodology, August 17-20, 2004, Amsterdam). Amsterdam.
- Caton, Steven C. (1986): Salam tahiyah: Greetings from the highlands of Yemen. In: *American Ethnologist* 13, S. 290-308.
- Collett, Peter (1983): Mossi salutations. In: *Semiotica* 45, 3-4, S. 191-248.
- Danby, Susan/Baker, Carolyn/Emmison, Michael (2005): Four observations on openings in calls to Kids Help Line. In: Baker, Carolyn/Emmison, Michael/Firth, Allan (Hg.): *Calling for help*. Amsterdam, S. 133-152.
- De Gaulmyn, Marie-Madeleine (1994): Appels téléphoniques d'urgence sociale: parler au nom d'un autre. In: Trognon, Alain (Hg.): *Construction interactive du quotidien*. Nancy, S. 235-261.
- Deppermann, Arnulf/Schmitt, Reinhold (2007): Koordination. Zur Begründung eines neuen Forschungsgegenstandes. In: Schmitt (Hg.), S. 15-54.
- Deppermann, Arnulf/Mondada, Lorenza/Schmitt, Reinhold (i.Dr.): Agenda and emergence in a meeting. From work to break-like activities and back to work. In: *Journal of Pragmatics*.
- Drew, Paul/Heritage, John (Hg.) (1992): *Talk at work: interaction in institutional settings*. Cambridge.
- Duranti, Alessandro (1992): Language and bodies in social space: Samoan ceremonial greetings. In: *American Anthropologist* 94, S. 657-691.
- Duranti, Alessandro (1997): Universal and culture-specific properties of greetings. In: *Journal of Linguistic Anthropology* 7, S. 63-97.
- Eibl-Eibesfeld, Irenäus (1972): Similarities and differences between cultures in expressive movements. In: Hinde, Robert A. (Hg.): *Non-verbal communication*. Cambridge, S. 297-312.
- Emmison, Michael/Danby, Susan (2007): Troubles announcements and reasons for calling: initial actions in opening sequences in calls to a national children's helpline. In: *Research on Language and Social Interaction* 40, 1, S. 63-87.
- Fornel, Michel de (1994): Le cadre interactionnel de l'échange visiophonique. In: *Réseaux* 64, S. 107-132.
- Godard, Danielle (1977): Same setting, different norms: phone call beginnings in France and the United States. In: *Language in Society* 6, S. 209-219.
- Goffman, Erving (1963): *Behavior in public places: Notes on the social organization of gatherings*. New York.
- Goodwin, Charles (1981): *Conversational organization: Interaction between speakers and hearers*. New York.

- Goody, Esther (1972): 'Greeting', 'begging' and the presentation of respect. In: La Fontaine, Jean Sybil (Hg.): *The interpretation of ritual*. London, S. 39-71.
- Greenbaum, Paul E./Rosenfeld, Howard M. (1980): Varieties of touching in greetings: Sequential structure and sex-related differences. In: *Journal of Nonverbal Behavior* 5, 1, S. 13-25.
- Hausendorf, Heiko (2001): Deixis and speech situation revisited: The mechanism of perceived perception. In: Lenz, Friedrich (Hg.): *Deictic conceptualization of space, time and person*. Amsterdam, S. 249-269.
- Heath, Christian (1981): The opening sequence in doctor-patient interaction. In: Atkinson, Paul/Heath, Christian (Hg.): *Medical work: Realities and routines*. Farnborough, S. 71-90.
- Heath, Christian (1982): The displays of reciprocity: An instance of a sequential relationship in speech and body movement. In: *Semiotica* 42, 2-4, S. 147-167.
- Heath, Christian (1986): *Body movement and speech in medical interaction*. Cambridge.
- Heath, Christian/Hindmarsh, Jon (2002): Analysing interaction: Video, ethnography and situated conduct. In: May, Tim (Hg.): *Qualitative Research in Action*. London, S. 99-121.
- Heath, Christian/Luff, Paul (2006): Video analysis and organisational practice. In: Knoblauch et al. (Hg.), S. 35-50.
- Heidtmann, Daniela/Föh, Marie-Joan (2007): Verbale Abstinenz als interaktive Beteiligungweise. In: Schmitt (Hg.), S. 263-292.
- Heritage, John/Greatbatch, David (1991): On the institutional character of institutional talk: the case of news interviews. In: Boden/Zimmerman (Hg.), S. 93-137.
- Heritage, John/Maynard, Douglas (Hg.) (2006): *Communication in medical care*. Cambridge.
- Heritage, John/Robinson, Jeffrey D. (2006): Accounting for the visit: giving reasons for seeking medical care. In: Heritage/Maynard (Hg.), S. 48-85.
- Hopper, Robert/Chen, Chia-Hui (1996): Languages, cultures, relationships: Telephone openings in Taiwan. In: *Research on Language and Social Interaction* 29, S. 291-314.
- Hopper, Robert/Doany, Neda/Johnson, Michael/Drummond, Kent (1990/1991): Universals and particulars in telephone openings. In: *Research on Language and Social Interaction* 24, S. 369-388.
- Hopper, Robert/Koleilat-Doany, Neda (1989): Telephone openings and conversational universals: A study in three languages. In: Ting-Toomey, Stella/Korzenny, Felipe (Hg.): *Language, communication and culture*. Newbury Park, CA, S. 157-179.

- Houtkoop-Steenstra, Hanneke (1991): Opening sequences in Dutch telephone conversations. In: Boden/Zimmerman (Hg.), S. 232-250.
- Hutchby, Ian (1999): Frame attunement and footing in the organization of talk radio openings. In: *Journal of Sociolinguistics* 3, 1, S. 41-63.
- Hutchby, Ian/Barnett, Simone (2005): Aspects of the sequential organization of mobile phone conversation. In: *Discourse Studies* 7, 2, S. 147-171.
- Irvine, Judith T. (1974): Strategies of status manipulation in Wolof greeting. In: Bauman/Sherzer (Hg.), S. 167-191.
- Jewitt, Carey/Kress, Gunther (Hg.) (2003) : *Multimodal literacy*. New York.
- Kallmeyer, Werner/Streeck, Jürgen (2001): Interaction by inscription. In: *Journal of Pragmatics* 33, S. 465-490.
- Kendon, Adam (1990a): *Conducting interaction – Patterns of behavior in focused encounters*. Cambridge.
- Kendon, Adam (1990b): Some context for context analysis: A view of the origins of structural studies of face-to-face interaction. In: Kendon (1990a), S. 15-49.
- Kendon, Adam (1990c): Spatial organization in social encounters: the F-formation-system. In: Kendon (1990a), S. 209-237.
- Kendon, Adam (1990d): A description of some human greetings. In: Kendon (1990a), S. 153-207.
- Kendon, Adam/Ferber, Andrew (1973): A description of some human greetings. In: Michael, Richard P./Crook, John H. (Hg.): *Comparative ecology and behaviour of primates*. London/New York, S. 591-668.
- Knoblauch, Hubert/Schnettler, Bernt/Raab, Jürgen/Soeffner, Hans-Georg (Hg.) (2006): *Video-analysis. Methodology and methods*. London.
- Kress, Gunther/Van Leeuwen, Theo (2001): *Multimodal discourse: The modes and media of contemporary communication*. London.
- LeBaron, Curtis D./Streeck, Jürgen (1997): Built space and the interactional framing of experience during a murder interrogation. In: *Human Studies* 20, S. 1-25.
- Lee, Seung-hee (2006): Second summonings in Korean telephone conversation openings. In: *Language in Society* 35, S. 261-283.
- LeVine, Phillip/Scollon, Ron (Hg.) (2004): *Discourse and technology. Multimodal discourse analysis*. Washington, DC.
- Licoppe, Christian/Dumolin, Laurence (2007): L'ouverture des procès à distance par visioconférence. In: *Réseaux* 144, S. 103-140.
- Liddicoat, Anthony J. (2000): Telephone openings in Samoan. In: *Australian Review of Applied Linguistics* 23, 1, S. 95-107.
- Lindström, Anna (1996): Identification and recognition in Swedish telephone conversation openings. In: *Language in Society* 23, S. 231-252.

- Luke, Kang Kwong (2002): The initiation and introduction of first topics in Hong Kong telephone calls. In: Luke/Pavlidou (Hg.), S. 171-200.
- Luke, Kang Kwong/ Pavlidou, Theodossia-Soula (Hg.) (2002): Telephone calls: Unity and diversity in conversational structure across languages and cultures. Amsterdam.
- Macbeth, Douglas (1999): Glances, trances, and their relevance for a visual sociology. In: Jalbert, Paul L. (Hg.): Media studies: Ethnomethodological approaches. Lanham, S. 135-170.
- Meehan, Albert J. (1989): Assessing the 'policeworthiness' of citizen's complaints to the police: accountability and the negotiation of 'facts'. In: Helm, David T./Anderson, Tim W./Meehan Albert J./Rawls, Anne W. (Hg.): The interaction order. New directions in the study of social order. New York, S. 116-140.
- Meier, Christoph (1998): Potentielle und aktuelle Präsenz: von der Interaktionseröffnung zum offiziellen Beginn in Videokonferenzen. (= Arbeitspapiere Telekooperation: Strukturen, Dynamik und Konsequenzen elektronisch vermittelter Kooperativer Arbeit in Organisation 5). Gießen.
- Mirivel, Julien/Tracy, Karen (2005): Premeeting talk: An organizationally crucial form of talk. In: Research on Language and Social Interaction 38, S. 1-34.
- Modaff, Daniel P. (2003): Body movement in the transition from opening to task in doctor-patient interviews. In: Glenn, Phillip/LeBaron, Curtis D./Mandelbaum, Jenny (Hg.): Studies in language and social interaction: In honor of Robert Hopper. Mahweh, NJ, S. 411-422.
- Mondada, Lorenza (2004): Temporalité, séquentialité et multimodalité au fondement de l'organisation de l'interaction: Le pointage comme pratique de pris de tour. In: Cahiers de linguistique Française 26, S. 169-192.
- Mondada, Lorenza (2006): Video recording as the preservation of fundamental features for analysis. In: Knoblauch et al. (Hg.), S. 51-68.
- Mondada, Lorenza (2007a): Transcript variations and the indexicality of transcribing practices. In: Discourse Studies 9, 6, S. 809-821.
- Mondada, Lorenza (2007b): Multimodal resources for turn-taking: Pointing and the emergence of possible next speakers. In: Discourse Studies 9, 2, S. 195-226.
- Mondada, Lorenza (2007c): Interaktionsraum und Koordinierung. In: Schmitt (Hg.), S. 55-93.
- Mondada, Lorenza (2007d): Operating together through videoconference: Members' procedures for accomplishing a common space of action. In: Hester, Stephen/Francis, David (Hg.): Orders of ordinary action. Aldershot, S. 51-67.
- Mondada, Lorenza (2008a): Documenter l'articulation des ressources multimodales dans le temps: la transcription d'enregistrements vidéos d'interactions. In: Bilger, Mireille (Hg.): Données orales, les enjeux de la transcription. Perpignan, S. 127-155.

- Mondada, Lorenza (2008b): L'interprétation online par les co-participants de la structuration du tour in fieri en TCUs: évidences multimodales. In: *Revue TRANEL* 48, S. 7-38.
- Mondada, Lorenza (2008c): Doing video for a sequential and multimodal analysis of social interaction: Videotaping institutional telephone calls. In: *Forum Qualitative Sozialforschung/Forum Qualitative Social Research* 9, 3. Internet: www.qualitative-research.net/ (Stand: 25.08.09).
- Mondada, Lorenza (2008d): Production du savoir et interactions multimodales. Une étude de la modélisation spatiale comme activité pratique située et incarné. In: *Revue d'Anthropologie des Connaissances* 2, 2, S. 267-289. Internet: www.cairn.info/revue-anthropologie-des-connaissances-2008-2-page-219.htm (Stand: 25.08.09).
- Mondada, Lorenza (2009): Video recording practices and the reflexive constitution of the interactional order: some systematic uses of the split-screen technique. In: *Human Studies* 32, 1, S. 67-99.
- Mondada, Lorenza (i.Dr.): Emergent focused interactions in public places: A systematic analysis of the multimodal achievement of a common interactional space. In: *Journal of Pragmatics*.
- Mondada, Lorenza/Schmitt, Reinhold (2007): Vergleichende Analysen von Situationseröffnungen/Analyses comparées d'ouvertures. Ein deutsch-französisches Kooperationsprojekt. In: *Sprachreport, Sonderheft März 2007*, S. 27-30.
- Müller, Cornelia/Bohle, Ulrike (2007): Das Fundament fokussierter Interaktion. Zur Vorbereitung und Herstellung von Interaktionsräumen durch körperliche Koordination. In: Schmitt (Hg.), S. 129-165.
- Norris, Sigrid (2004): *Analyzing multimodal interaction: A methodological framework*. London.
- O'Halloran, Kay (Hg.) (2004): *Multimodal discourse analysis: Systematic-functional perspectives*. London.
- Pallotti, Gabriele/Vacasia, Cecilia (o.J.): Service telephone call openings: a comparative study on five European languages. Internet: http://icar.univ-lyon2.fr/ecole_thematique/contacti/documents/pallotti/pallottivacasia.pdf (Stand: 25.08.09).
- Park, Yong-Yae (2002): Recognition and identification in Japanese and Korean telephone conversation openings. In: Luke/Pavlidou (Hg.), S. 25-47.
- Pillet-Shore, Danielle M. (2008): *Coming together: Creating and maintaining social relationships through the opening of face-to-face interactions*. Univ. Diss., UCLA, Dept. of Sociology, Los Angeles.
- Relieu, Marc (2002): Ouvrir la boîte noire. Identification et localisation dans les conversations mobiles. In: *Réseaux* 112-113, S. 19-47.
- Relieu, Marc (2007): La téléprésence, ou l'autre visiophonie. In: *Réseaux* 144, S. 183-224.

- Robinson, Jeffrey D. (1998): Getting down to business: talk, gaze and bodily orientation during openings of doctor-patient consultations. In: *Human Communication Research* 25, S. 97-123.
- Robinson, Jeffrey D. (2006): Soliciting patients' presenting concerns. In: *Heritage/ Maynard* (Hg.), S. 22-47.
- Ruusuvuori, Johanna (2000): Control in the medical consultation. Practices of giving and receiving the reason for the visit in primary health care. (= *Acta Electronica Universitatis Tamperensis* 16). Internet: <http://acta.uta.fi/pdf/951-44-4755-7.pdf> (Stand: 25.08.09).
- Sacks, Harvey/Schegloff, Emanuel A./Jefferson, Gail (1974): A simplest systematics for the organization of turn-taking for conversation. In: *Language* 50, S. 696-730.
- Salmond, Anne (1974): Rituals of encounter among the Maori: Sociolinguistic study of a scene. In: *Bauman/Sherzer* (Hg.), S. 192-212.
- Schefflen, Albert E. (1972): *Body language and social order: Communication as behavioral control*. Englewood Cliffs, NJ.
- Schegloff, Emanuel A. (1967): *The first five seconds: The order of conversational openings*. Univ. Diss., Berkeley, CA.
- Schegloff, Emanuel A. (1968): Sequencing in conversational openings. In: *American Anthropologist* 70, S. 1075-1095.
- Schegloff, Emanuel A. (1979): Identification and recognition in telephone conversation openings. In: *Psathas, George* (Hg.): *Everyday language: studies in ethno-methodology*. New York, S. 23-78.
- Schegloff, Emanuel A. (1986): The routine as achievement. In: *Human Studies* 9, S. 111-152.
- Schegloff, Emanuel A. (1987a): Analyzing single episodes of interaction. An exercise in conversation analysis. In: *Social Psychology Quarterly* 50, 2, S. 101-114.
- Schegloff, Emanuel A. (1987b): Some sources of misunderstanding in talk-in-interaction. In: *Linguistics* 25, S. 201-218.
- Schegloff, Emanuel A. (1988): Description in the social sciences I: Talk-in-interaction. In: *IPrA Papers in Pragmatics* 2, 1-2, S. 1-24.
- Schegloff, Emanuel A. (2002): Beginnings in the telephone. In: *Katz, James, E./ Aakhus, Mark* (Hg.): *Perpetual contact: mobile communication, private talk, public performance*. Cambridge, S. 284-300.
- Schegloff, Emanuel A./Sacks, Harvey (1973): Opening up closings. In: *Semiotica* 7, S. 289-327.
- Schmitt, Reinhold (2001): Die Tafel als Arbeitsinstrument und Statusrequisite. In: *Ivany, Zsuzsanna/Kertesz, Andras* (Hg.): *Gesprächsforschung. Tendenzen und Perspektiven*. (= *Metalinguistica* 10). Frankfurt a.M., S. 221-242.

- Schmitt, Reinhold (2004): Die Gesprächspause: „Verbale Auszeiten“ aus multimodaler Perspektive. In: *Deutsche Sprache* 32, 1, S. 56-84.
- Schmitt, Reinhold (2005): Zur multimodalen Struktur von turn-taking. In: *Gesprächsforschung – Onlinezeitschrift zur verbalen Interaktion* 6, S. 17-61. Internet: www.gespraechsforschung-ozs.de/heft2005/ga-schmitt.pdf (Stand: 25.08.09).
- Schmitt, Reinhold (2007a): Theoretische und methodische Implikationen der Analyse multimodaler Interaktion. In: Holly, Werner/Paul, Ingwer (Hg.): *Medialität und Sprache*. (= Mitteilungen des Deutschen Germanistenverbandes 54/1). Tübingen, S. 26-53.
- Schmitt, Reinhold (2007b): Von der Konversationsanalyse zur Analyse multimodaler Interaktion. In: Kämper, Heidrun/Eichinger, Ludwig M. (Hg.): *Sprachperspektiven. Germanistische Linguistik und das Institut für Deutsche Sprache*. (= Studien zur Deutschen Sprache 40). Tübingen, S. 395-417.
- Schmitt, Reinhold (Hg.) (2007): *Koordination. Analysen zur multimodalen Interaktion*. (= Studien zur Deutschen Sprache 38). Tübingen.
- Schmitt, Reinhold (2009): Bericht über das 10. „Arbeitstreffen zu Fragen multimodaler Kommunikation“. In: *Sprachreport* 1/2009, S. 27-29.
- Schmitt, Reinhold (i.Vorb.): *Zur multimodalen Konstitution interaktiver Zugehörigkeit: Das Konzept „Interaktionsensemble“*.
- Schmitt, Reinhold/Deppermann, Arnulf (2007): Monitoring und Koordination als Voraussetzung der multimodalen Konstitution von Interaktionsräumen. In: Schmitt (Hg.), S. 95-128.
- Schmitt, Reinhold/Fiehler, Reinhard/Reitemeier, Ulrich (2007): Audiovisuelle Datenkonstitution und Koordinationsprozesse. In: Schmitt (Hg.), S. 377-410.
- Schütz, Alfred (1971): *Das Problem der sozialen Wirklichkeit*. In: Schütz, Alfred: *Gesammelte Aufsätze*. Bd. 1. Den Haag.
- Sherzer, Joel (1983): *Kuna ways of speaking: An ethnographic perspective*. Austin.
- Sidnell, Jack/Stivers, Tanya (Hg.) (2005): *Multimodal interaction*. In: *Semiotica* 156, 1, 4 (Special Issue).
- Streeck, Jürgen (1996): How to do things with things: Objects trouvés and symbolization. In: *Human Studies* 19, S. 365-384.
- Taleghani-Nikazm, Carmen (2002): A conversation analytical study of telephone conversation openings between native and nonnative speakers. In: *Journal of Pragmatics* 34, 12, S. 1807-1832.
- Taylor, Maurice M. (1989): *The structure of multimodal dialogue*. Amsterdam.
- Taylor, Maurice M. (2000): *The structure of multimodal dialogue, II*. Philadelphia.

- ten Have, Paul (1988): *Sequenties en formuleringen; aspecten van de interactionele organisatie van huisartsspreekuurgesprekken*. [Sequences and formulations: aspects of the interactional organization of general practice consultations]. Dordrecht.
- Traverso, Véronique (1998): *Conversations familiaires*. Lyon.
- Turner, Roy (1972): Some formal properties of therapy talk. In: Sudnow, David (Hg.): *Studies in Social Interaction*. New York, S. 367-396.
- Varcasia, Cecilia (2003): *Chiamate di servizio in Italia e Germania: aperture a confronto*. In: Thüne, Eva-Maria/Leonardi, Simona (Hg.): *Telefonare in diverse lingue*. Milano, S. 112-132.
- Whalen, Marilny/Zimmerman, Don H. (1987): Sequential and institutional contexts in calls for help. In: *Social Psychology Quarterly* 50, S. 172-185.
- Zimmerman, Don H. (1992a): The interactional organization of calls for emergency. In: Drew/Heritage (Hg.), S. 418-469.
- Zimmerman, Don H. (1992b): Achieving context: openings in emergency calls. In: Watson, Graham/Seiler, Robert M. (Hg.): *Text in context: contributions to ethno-methodology*. London, S. 35-51.